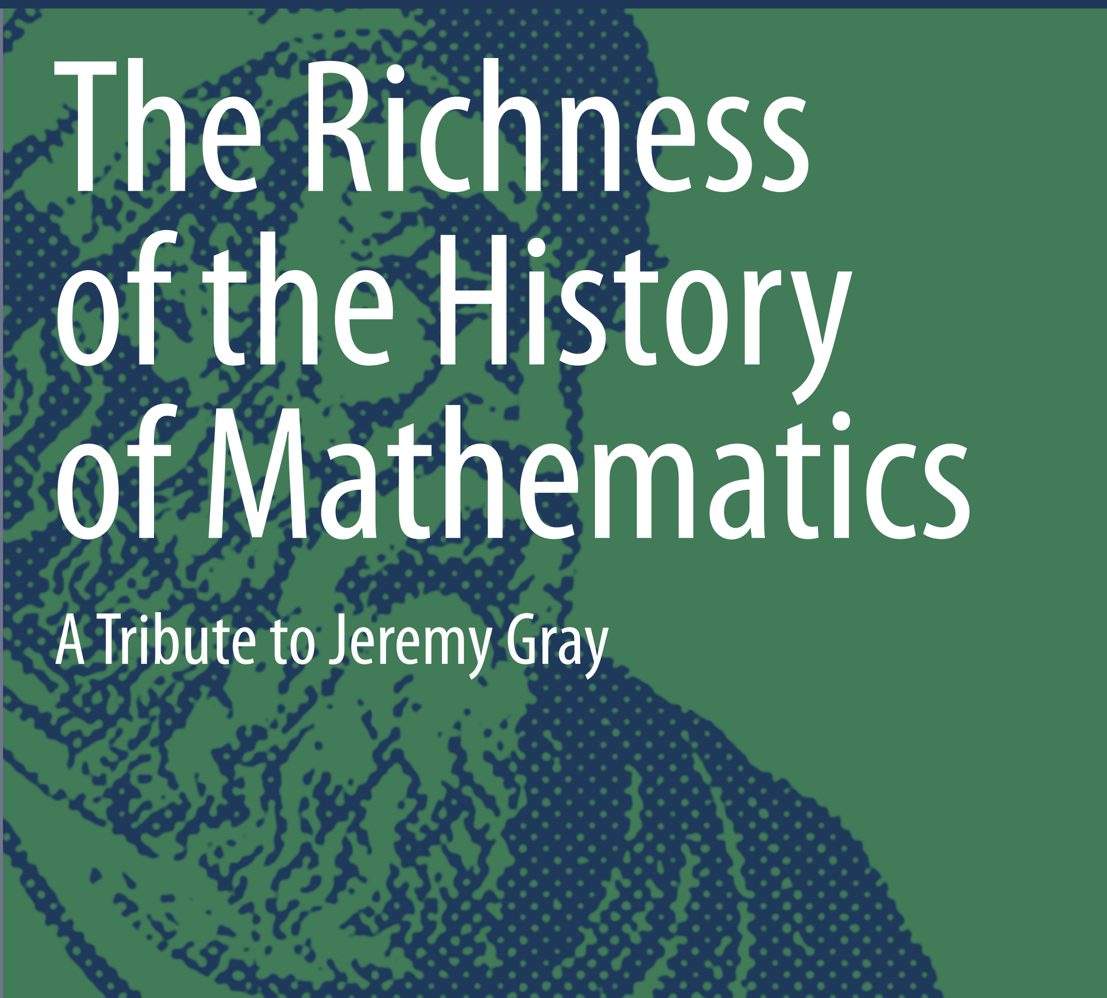
Archimedes 66 New Studies in the History and Philosophy of Science and Technology

Karine Chemla · José Ferreirós ·
Lizhen Ji · Erhard Scholz ·
Chang Wang *Editors*



The Richness of the History of Mathematics

A Tribute to Jeremy Gray

Springer

Archimedes

New Studies in the History and Philosophy
of Science and Technology

Volume 66

Archimedes has three fundamental goals: to further the integration of the histories of science and technology with one another; to investigate the technical, social and practical histories of specific developments in science and technology; and finally, where possible and desirable, to bring the histories of science and technology into closer contact with the philosophy of science.

The series is interested in receiving book proposals that treat the history of any of the sciences, ranging from biology through physics, all aspects of the history of technology, broadly construed, as well as historically-engaged philosophy of science or technology. Taken as a whole, Archimedes will be of interest to historians, philosophers, and scientists, as well as to those in business and industry who seek to understand how science and industry have come to be so strongly linked.

Submission / Instructions for Authors and Editors: The series editors aim to make a first decision within one month of submission. In case of a positive first decision the work will be provisionally contracted: the final decision about publication will depend upon the result of the anonymous peer-review of the complete manuscript. The series editors aim to have the work peer-reviewed within 3 months after submission of the complete manuscript.

The series editors discourage the submission of manuscripts that contain reprints of previously published material and of manuscripts that are below 150 printed pages (75,000 words). For inquiries and submission of proposals prospective authors can contact one of the editors:

Please find on the top right side of our webpage a link toour *Book Proposal Form.*

Karine Chemla • José Ferreirós • Lizhen Ji •
Erhard Scholz • Chang Wang
Editors

The Richness of the History of Mathematics

A Tribute to Jeremy Gray

Springer

Editors
Karine Chemla
SPHERE UMR 7219
Université Paris 7 Paris Cité
Paris cedex 13, France

José Ferreirós
Faculty of Philosophy
Universidad de Sevilla
Sevilla, Spain

Lizhen Ji
Department of Mathematics
University of Michigan
Ann Arbor, MI, USA

Erhard Scholz
School of Mathematics
University of Wuppertal
Wuppertal, Germany

Chang Wang
Institute for Advanced Studies in
History of Science
Northwest University
Xi'an, Shaanxi, China

ISSN 1385-0180 ISSN 2215-0064 (electronic)
Archimedes
ISBN 978-3-031-40854-0 ISBN 978-3-031-40855-7 (eBook)
https://doi.org/10.1007/978-3-031-40855-7

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

Paper in this product is recyclable.

This book is dedicated to our friend and colleague Jeremy Gray in honour of his contributions to the history of mathematics during the first 75 years of his life—and to all students in the world interested in getting to know how the integrated history and philosophy of mathematics may be written.[1]

[1] Drawing by Jemma Lorenat.

Preface

The historiography of mathematics has a long history. In spite of this we have to admit that there is a lack of standard introductory books to this field, which students can turn to in order to find orientation in what may be basic problems in the subject and how they can start to pursue it. Questions of this type have been posed by one of us, Lizhen Ji, a mathematician with strong interest in the history of mathematics, to many colleagues. During the last few years he discussed them with historians and philosophers of mathematics. Finally, this led to the idea of compiling a book addressing such questions in a more direct way and to dedicate it to our colleague and friend *Jeremy Gray*.

Jeremy's wide range of interests and the broadness of his research in the history of mathematics is shown by his publications; it suffices to mention the great number and thematic spectrum of the books he wrote or co-edited to prove this point. Let us here just give a list of topics for books written by him during the past four decades: ideas of space from antiquity to the twentieth century (1979), linear differential equations and group theory (1986), Desargues (with J. Field, 1987), Hilbert (2000), modernism in nineteenth-/twentieth-century mathematics (2008), complex function theory (jointly with U. Bottazzini, 2013) and Poincaré (2016). In addition, in the last few years he published a series of textbooks on the history of mathematics, which grew out of his lectures at Warwick University. They treat the history of geometry in the nineteenth century (2007), real and complex analysis in the nineteenth century (2015), abstract algebra (2018) and differential equations and variational calculus (2021). Jeremy also kept close contact with the community of mathematicians as well as that of philosophers of mathematics and was immensely active in the dissemination of historical knowledge by preparing teaching material for the Open University, Milton Keynes, Warwick University and beyond.

We thank Jeremy for contributing himself to the book by writing down some *Reflections* on his path into the historiography of mathematics (Part VII). This is a nice piece showing a not so untypical effect for newcomers to our subdiscipline of the history of science on the one hand and mathematics on the other. In relation to the weak institutional backing for the history of mathematics, the heterogeneity of backgrounds needed to work in it as well as the diversity of epistemic goals pursued, of styles and methods of work, and of the diversity of its audiences, there is often no direct ("linear") career path leading young scientists into the heart of the field. Jeremy's reflections give a smiling look back on the path he himself followed.

We tried our best to present in our book for Jeremy a collection of reading material containing information and sometimes even some kind of orientation for young researchers worldwide, who are about to enter research in the history of mathematics. Among them there is an impressively large group of young historians of mathematics at and from the Northwest University of Xi'an. We hope and expect that members of this group will soon be contributing to the international enterprise of the historical studies in the mathematical sciences. Through the mediation of one of us, Chang Wang, we had the chance to communicate with this group during the preparation of this book. All contributions to it, with the only exception of Jeremy's reflections, have been carefully read and commented by young Chinese historians of mathematics. The comments have been given to the attention of the authors in addition to the remarks of the usual peer reviews for each of the chapters. Our thanks go to the members of all three (partly overlapping) groups of authors, peer commentators and young researchers from Xi'an.[2]

Keeping the interests of our colleague Jeremy Gray in mind, the panorama of topics in this book is essentially concentrated on the history of mathematics in what is usually referred as the "modern European" area with one exception: Anjing Qu's discussion of mathematical practices in (ancient) Chinese astronomy (in Part III). This remark leads us to the task of outlining the spectrum of topics which appear here for illustrating the "Richness of the History of Mathematics". To keep the foreword as concise as possible we give here only a short overview of the structure of the book, with very selective information on the topics covered.

The book starts with a collection of essays on *methods and problems of the historiography of mathematics* (Part I). It documents how diverse the approaches to writing the history of mathematics can be. The first two chapters address the question of how one may embark—and how some researchers have indeed engaged—in historical work on mathematics. The first one by Tinne Hoff Kjeldsen gives an inviting discussion of this question by a present researcher in this field. The second one by Bruno Belhoste and Karine Chemla shows that this is not a twentieth century or even presentist question but has deep roots in the development of mathematics itself. They exemplify this point by a study of the role of history in the teaching and research carried out at the early Ecole Polytechnique. In the next two chapters we find the reflections of Lizhen Ji, a mathematician interested in the history of mathematics, and Viktor Blåsjö, a historian of mathematics with particular interest in the cooperation with mathematicians. Ji provides guidance for beginners in the history of mathematics, considering many different sources about historical thinking and the methodology of historical research. Blåsjö argues that the time is ripe for comparative interpretative work of well-researched episodes in history (his example is the calculus) and that mathematically trained researchers

[2] See the acknowledgments and the list of contributors below.

are needed for this enterprise. In the final chapter of this part, Niccolò Guicciardini shows the usefulness, perhaps even necessity, of reflexive anachronistic rewritings of technical passages in historical works, exemplified with Newton's treatment of the inverse problem for a body moving under central forces.

The contributions of Part II shed glances from different directions into the historical *practice of mathematics* and of mathematicians. With regard to geometry, David Rowe presents less known aspects of Felix Klein's early work in projective and algebraic geometry. Two chapters on the history of arithmetic follow. Catherine Goldstein shows the peculiar blend of methods that Poincaré put into play in his early works in number theory, again basically ignored in the existing literature. Nicola Oswald focuses on the correspondence between Hurwitz, Gordan and Hilbert on transcendence proofs for e. She analyses how it led to three variations of a proof of transcendence for e, which aimed to simplify the argument, and she contrasts these various proofs from the viewpoint of their later reception. Colin McLarty examines what the mathematical practices underlying various uses of the concept of function are like today and historically from the perspective of a philosopher of mathematics. The part is rounded off by two contributions on the formation of a mathematician. The first one by Jemma Lorenat deals with this issue from a gender point of view, with a case study in the history of the late nineteenth century. The second one gives an autobiographical report of John Stilwell; it tells how he himself became a mathematician deeply involved with the history of mathematics. This report may be read as a kind of parallel with Jeremy Gray's reflections on his path into the history of mathematics (Part VII).

Questions linked with the *interrelation of mathematics with the natural sciences* are addressed in Part III. In our book we can highlight again only selected topics in this vast field. The part starts with a glimpse by Anjing Qu into this relationship in an extra-European context, the only one in this book as already stated. Our author analyses the mathematical practices used in the construction of astronomical tables in *Yuanjia li*, a calendar-making system of the fifth century CE. The next chapter written by Jesper Lützen discusses the changing relationship between mathematical geometry and physics in the late nineteenth-century European context. As a result the classical identification between the concepts of the empirical space of natural science, geometry and the mathematical concept of space was broken up and new interrelation between more abstract space concepts of mathematics—like manifolds—and mechanics came into sight (several decades before the relativistic "revolution" of the early twentieth century). Jed Buchwald studies the usage of scalar and vector potentials in different theories of electrodynamics of the nineteenth century and the conditions or relations between them. He shows that the role of such relations as free choices was not clearly stated before 1903 in a paper by H.A. Lorentz. About 20 years later (after Weyl's contribution to relativistic electrodynamics, see Part IV) they were called "gauge" choices. Of course mathematical investigations in natural sciences go far beyond the classical fields of "application" in astronomy and physics. The last chapter of this part, written by June Barrow-Green—a former doctoral student of Jeremy—focuses on a collaboration between the mathematician Hilda Hudson and the medical scientist Ronald Ross. Their joint work led to an early contribution to the mathematical study of epidemics. The most important part of it took place during World War I and thus before the seminal papers of Kermack and McKenzie (1927ff.), which are influential for modelling the dynamics of epidemics until the present.

The topic of *modernism* in mathematics is taken up in Part IV. This concept was imported from cultural and social history into the history of mathematics by Herbert Mehrtens and became one of Jeremy's particular concerns. The first chapter of this part, written by Leo Corry, steps back and reconsiders the issue from the wider scope of the cultural history in which the topic of modernism was bred. Corry examines anew in which respect the question may be useful for the study of the history of mathematics. Inside mathematics, it may not be so easy to discern what it means to be "modern". Tom Archibald discusses the question for integration theory in the twentieth century. Cantorian set theory is often taken as a core feature of modernism in mathematics, standing in counter-position to other strands in twentieth-century mathematics, in particular Brouwerian intuitionism. José Ferreirós, in contrast, argues that many mathematicians were inclined to revise and weaken set theory, and therefore speaks of set theories of the twentieth century in the plural. An extreme example of this was Brouwer, who elaborated his own intuitionistic understanding of sets (another is Weyl's predicativist system). Furthermore, he finds that Brouwer's program for reforming mathematics can very well be placed under the heading of modernist approaches to mathematics. Finally Erhard Scholz takes the opportunity of this book to discuss two key figures of twentieth century mathematics, Felix Hausdorff and Hermann Weyl, which have been labelled by Mehrtens as two characteristic representatives of what he considered as, respectively, the modern and the counter-modern camp in mathematics of the last century. In spite of the rather different stance of the two protagonists with regard to modernism Scholz does not agree with labelling Weyl as "counter-modern".

The next part (Part V) addresses the question of how *mathematicians* interacted with the broader *philosophical discourse* of their time. Vincenzo de Risi unearths a debate in the eighteenth and nineteenth centuries on what was called the direction theory of parallels. After a transformation of the question posed, this discussion played a key role in the debates raised by the later emergence of non-Euclidean geometry. The related developments were also heavily intertwined with epistemological questions of early modern mathematics and, more generally, with philosophical debates between followers of, respectively, Leibniz and Kant. At the end of the nineteenth century the scene for references of mathematics to philosophical issues had diversified in various ways. In his chapter on Zeuthen's understanding of enumerative geometry, Nicolas Michel finds strong references to the Danish reception of the philosophy of Henri Bergson, supplemented by holistic psychology. Finally Umberto Bottazzini discusses Federico Enriques' epistemology of mathematics and sketches his broader view of a "scientific philosophy" which Enriques developed in reference to views held by members of the Vienna circle.

All this leads over to *issues in the philosophy* of mathematics (Part VI). Dirk Schlimm examines why Frege, Cantor and Gödel thought about the object of mathematics in Platonist terms, and why this may matter still today. James

Tappenden, for his part, argues that one may observe a strong connection between the history and the philosophy of mathematics. He addresses this topic through a discussion of the work of Riemann and Weierstrass in complex function theory and analyses their implicit philosophy in their way of forming, presenting and understanding mathematical concepts. In the last chapter of this part—a short literary piece—Jeremy Avigad reflects on "what we talk about" when we "talk" mathematics.

The closing part of the book (Part VII) presents two cases of *how* one may *become a historian of mathematics*: The first one by Snezana Lawrence—also a former PhD student of Jeremy—shows how even a twisted route may lead into this field. The second and final one contains Jeremy Gray's reflections on his own path into the history of mathematics. His recollections show that this trajectory was also far from any established career route through well-established institutions which are anyway rare in our part of the world.

Please do not overlook that, in addition to an index of names, the appendix to this book contains a selection of photographs which show our venerated colleague at different stages of his life and in different environments.

Paris, France — Karine Chemla
Sevilla, Spain — José Ferreirós
Ann Arbor, MI, USA — Lizhen Ji
Wuppertal, Germany — Erhard Scholz
Xi'an, China — Chang Wang
January 10, 2023

Books Mentioned

Bottazzini, Umberto, and Jeremy J. Gray. 2013. *Hidden Harmony – Geometric Fantasies. The Rise of Complex Function Theory*. Berlin: Springer.

Field, Judy, and Jeremy Gray. 1987. *The Geometrical Work of Girard Desargues*. Berlin: Springer.

Gray, Jeremy. 1979. *Ideas of Space: Euclidean, Non-Euclidean, and Relativistic*. Oxford: Clarendon. ^{2}1985.

Gray, Jeremy. 1986. *Linear Differential Equations and Group Theory from Riemann to Poincaré*. Basel: Birkhäuser. ^{2}2000.

Gray, Jeremy. 2000. *The Hilbert Challenge*. Oxford: University Press.

Gray, Jeremy. 2007. *Worlds out of Nothing: A Course in the History of Geometry in the 19th Century*. Berlin: Springer.

Gray, Jeremy. 2008. *Plato's Ghost. The Modernist Transformation of Mathematics*. Princeton: University Press.

Gray, Jeremy. 2013. *Henri Poincaré. A Scientific Biography*. Princeton: University Press.

Gray, Jeremy. 2015. *The Real and the Complex: A History of Analysis in the 19th Century*. Springer Nature Switzerland.
Gray, Jeremy. 2018. *A History of Abstract Algebra*. Berlin: Springer.
Gray, Jeremy. 2021. *Change and Variations. A History of Differential Equations to 1900*. Springer Nature Switzerland.

Acknowledgements

We thank our colleagues who contributed by reading and commenting on the chapters (some of them also contributors to the book): David Aubin, June Barrow-Green, Viktor Blåsjö, Umberto Bottazzini, Jenny Boucard, Leo Corry, Della Dumbaugh, David Dunning, Caroline Ehrhardt. Moritz Epple, Isabel Falconer, Craig Fraser, Paolo Fregulia, Michael N. Fried, Niccolò Guicciardini, Catherine Goldstein, Gerhard Heinzmann, Tinne Hoff-Kjeldsen, Christopher David Hollings, Anne Kox, François Lê, Sandra Lingueri, Jemma Lorenat, Jesper Lützen, Paolo Mancosu, Thomas Morel, Marc Moyon, Victor Pambuccian, Luigi Pepe, Walter Purkert, Erich Reck, Joan Richards, Tabea Rohr, David Rowe, Norbert Schappacher, Martina Schneider, Jörn Stauding, John Stillwell, Silvia de Toffoli, Peter Ullrich, Mark van Atten, Klaus Volkert, Jean-Daniel Voelke, Roy Wagner.

A special thank goes to the Chinese young historians of mathematics who commented on all the contributions to this book:

Kesheng Chen, Wei Chen, Jinze Du, Qianqian Feng, Yang Gao, Jinming Geng, Chanchan Guo, Yuwen He, Nan Jiang, Juan Li, Rui Li, Ruirui Li, Wei Li, Yaya Li, Jianxin Liu, Xi Liu, Yuanyuan Liu, Ruiping Mu, Mengqi Wang, Jiadai Xin, Kuai Xu, Baoqiang Yang, Mingli Yang, Qiang Yang, Zhongmiao Yu, Xianci Zeng, Hongxing Zhang, Qiaoyan Zhang, Yangyang Zhang, Yilin Zhang, Jiwei Zhao.

Last but not least, we thank *Kirsti Andersen, June Barrow-Green, Karine Chemla, Sue Lawrence, Chang Wang* for the photographs of Appendix A, and *Jemma Lorenat* for her beautiful picture of Jeremy on p. vi.

Contents

Contributors

Tom Archibald Department of Mathematics, Simon Fraser University, Burnaby, BC, Canada

Jeremy Avigad Department of Philosophy and Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA, USA

June Barrow-Green School of Mathematics & Statistics, The Open University, Milton Keynes, UK

Bruno Belhoste IHMC, Université Paris 1 Panthéon-Sorbonne-ENS-CNRS, Paris, France

Viktor Blåsjö Mathematical Institute, Utrecht University, Utrecht, The Netherlands

Umberto Bottazzini Mathematics, Universitá degli Studi di Milano, Milan, Italy

Jed Buchwald Caltech, Pasadena, CA, USA

Karine Chemla SPHERE, CNRS–Université Paris Cité Case 7093, Paris Cedex 13, France

Leo Corry Cohn Institute for the History & Philosophy of Science & Ideas, Tel Aviv University, Tel Aviv, France

Vincenzo de Risi SPHERE, CNRS–Université Paris Cité, Case 7093, Paris Cedex 13, France

José Ferreirós Departamento de Filosofía y Lógica, Universidad de Sevilla, Sevilla, Spain

Catherine Goldstein Institut de mathématiques de Jussieu-Paris-Gauche, UMR 7586, CNRS, Sorbonne Université, Université Paris Cité, Case 247, Paris Cedex 05, France

Jeremy Gray School of Mathematics and Statistics, Open University, Milton Keynes, UK

Niccolò Guicciardini Dipartimento di Filosofia “Piero Martinetti”, Università degli Studi di Milano, Milano, Italy

Tinne Hoff Kjeldsen Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark

Lizhen Ji Department of Mathematics, University of Michigan, Ann Arbor, MI, USA

Snezana Lawrence Department of Design Engineering and Mathematics, The Burroughs, Middlesex University, London, UK

Jemma Lorenat Mathematics, Pitzer College, Claremont, CA, USA

Jesper Lützen Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark

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

Nicolas Michel Department of Mathematics & Informatics, University of Wuppertal, Wuppertal, Germany

Nicola Oswald Department of Mathematics & Informatics, University of Wuppertal, Wuppertal, Germany

Anjing Qu Institute for Advanced Studies in History of Science, Northwest University, Xi’an, China

David Rowe Mathematics Institute, Mainz University, Mainz, Germany

Dirk Schlimm McGill University, Montreal, QC, Canada

Erhard Scholz Department of Mathematics & Informatics, University of Wuppertal, Wuppertal, Germany

John Stillwell Emeritus Professor, College of Arts and Sciences, University of San Francisco, San Francisco, CA, USA

Jamie Tappenden Department of Philosophy, University of Michigan, Ann Arbor, MI, USA

Chang Wang Institute for Advanced Studies in History of Science, Northwest University, Xi ’an, China

Part I
Practicing the History of Mathematics

Chapter 1
A Problem-Oriented Multiple Perspective Way into History of Mathematics – What, Why and How Illustrated by Practice

Tinne Hoff Kjeldsen

Abstract This chapter is written with students in mind. It introduces and describes a problem-oriented multiple perspective approach to history of mathematics, which is a methodology to history of mathematics that is based on an action-oriented conception of history. It is explained how this approach is an open approach to history of mathematics in the sense that the research is driven by a question-answer strategy where the decisive factors for the development have not been decided beforehand, and it is clarified in what sense this approach moves beyond the internal/external division in the historiography of mathematics. The approach is illustrated by three examples from the history of twentieth century mathematics. The first is focused on the invention of the concept of a general convex body, and is a case that can be seen as an exemplar of the move of mathematics into an autonomous enterprise, which is an aspect of the twentieth century mathematics. The second case is concerned with the influence of WWII in the development of mathematical programming. It is an example of how conditions, or urgencies, in society might influence the development of mathematics together with more internal motivated driving forces. The third example deals with Nicolas Rashevsky's early development of mathematical biology. This case demonstrates how conditions within the sciences, and in society, have a significant influence on what kind of research is being developed, and how mathematical modelling can function as a research tool at the frontier of science. As such, the chapter is an attempt to lay out, present and explain the theoretical perspective and methodology for a problem-oriented multiple perspective approach to history of mathematics and illustrate its strengths and versatility through the three examples.

T. H. Kjeldsen (✉)
Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark
e-mail: thk@math.ku.dk

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_1

1.1 Introduction

This chapter is written with students in mind. As the title indicates, the focus is on presenting what I call *a problem-oriented multiple perspective approach to history of mathematics* (shortened to *a p-oriented m-perspective approach* in the following) as one way into doing research in history of mathematic (Kjeldsen 2012, 2019).[1] This is one of many ways of thinking about history of mathematics, so let me begin with a disclaimer: The chapter is not a textbook-like introduction to methodology; it does not provide general guidelines on how to do e.g. archival work etc. and it does not present a survey of the historiography and methodology of history of mathematics. If one wants to become acquainted with historiography of mathematics in a more general setting, Wardhaugh (2010) and Stedall (2012) are two short introductions to the field. The conference proceedings (Remmert et al. 2016) addresses "how the historiography of mathematics has been influenced by the contexts and motivations of its practitioners" (p. 1), and is a good place to look for the history of historiography of mathematics in the nineteenth and twentieth centuries. For history of science, the book by Helge Kragh (1994) *An introduction to historiography of science* is also to be recommended for students in history of mathematics. More recent reflections on the historiography of science can be found in e.g. (Karstens 2014) where Karstens presents an overview of various turns in the history of science from the early twentieth century to the present, and (Kuukkanen 2012) in which Kuukkanen discusses the active role of the historian and calls for more 'reflexivity' in the historiography of science. This being said, what is the chapter then about? It is an attempt to present some aspects of a p-oriented m-perspective approach and to be explicit about some reflections on doing research in history of mathematics related to this approach, illustrated by concrete examples. The intention has been to move a step back and to illustrate and reflect on some issues and questions lying underneath and leading up to concrete pieces of research within a problem-oriented multiple perspective approach. That is, to illustrate and discuss the terms that such an approach to history of mathematics sets for research, as well as ways it opens into research – and as such, the paper has a subjective element, presenting primarily my own approach and sources of inspiration.

In the following, we will first look briefly at a few historians of mathematics' experiences with, and thoughts about, history of mathematics. Then the approach of a p-oriented m-perspective way into history of mathematics is presented, followed by three examples for illustration. The chapter ends with a discussion and concluding remarks.

[1] I have borrowed the term 'multiple perspective' from Bernard Eric Jensen, who is a general historian, i.e. not a historian of mathematics or science (Jensen 2003). The problem-orientation is inspired from the Roskilde Model of problem-oriented learning and project work (Andersen and Heilesen 2015), see Andersen and Kjeldsen (2015a, b) for the theoretical foundation and a review of the key concepts.

1.2 A Few Historian of Mathematics' Accounts of Experiences with Research

In 2004, Ivor Grattan-Guinness published a paper in *Historia Mathematica* on his experiences and thoughts about doing history of mathematics (Grattan-Guinness 2004). He distinguished between two ways of treating or interpreting mathematics of the past, which he called 'history' and 'heritage' respectively. He characterized 'history' as an approach to past mathematics where the actors (by which he meant the historians) are concerned with answering questions of "what happened in the past", give descriptions and (sometimes) also try to give explanations as to "why" (Grattan-Guinness 2004, p. 164). In contrast to this, he found that the actors of the 'heritage' approach (whom he named inheritors), are concerned with the question "how did we get here?" – an interpretation of past mathematics where "the present is *photocopied* onto the past" (p. 165, Italic in the original). He deemed both ways of working with past mathematics legitimate but warned against confusing them, i.e. not taking heritage for history, heritage being, he wrote, "frequently the mathematicians' view" (p. 165). So here, we see a separation of dealing with the past based on the actor's profession – not exclusively but almost, and only one of them counts as history, according to Grattan-Guinness.

Michael Fried, working also with a distinction between non-historical and historical ways of dealing with the past, called for a more elaborated, more fine-grained division, than the one proposed by Grattan-Guinness (Fried 2018). Fried is interested in distinguishing between different approaches or attitudes to the past and how they "place one in different relations to the past", as he phrased it (p. 6). Like Grattan-Guiness, Fried operates with a division between a historical and a non-historical position, which he sees as two poles. With respect to this division, he singled out three categories with various sub-categories of ways of relating to mathematics of the past, which he placed on a scale from non-historical to historical relationships. The three main categories are 'mathematicians', 'mathematician-historians' and 'historians of mathematics', with 'mathematician-historians' ranging in the middle and placed somewhere in between the two poles. Fried characterized 'mathematicians' as those who "see themselves doing mathematics, not history", when they deal with mathematics of the past; 'mathematician-historians' as being "generally mathematicians who see themselves engaged in an historical enterprise and yet to a greater or lesser degree (. . .) see a continuity between the mathematics of the past and their own mathematical work; and 'historians of mathematics' as those who "see themselves doing history and their mode of experiencing the past is historical in the sense ... [that] they relate to the past as something utterly apart from the present, the past as a problem" (p. 7). In 2014, Fried and Viktor Blåsjö had a short dispute about the historiography of mathematics, Blåsjö advocating for *critical internalism* in historiography of mathematics (Fried 2014; Blåsjö 2014). Blåsjö further elaborated on this issue as a historiography of mathematics from the mathematician's point of view, questioning the division between historical and non-historical treatments of past mathematics, basically based solely on the actors' (the authors') profession (Blåsjö 2021).

Both Grattan-Guinness's and Fried's distinctions between a 'historical' and a 'heritage' approach and a 'historical' and a 'non-historical' relation to past mathematics, respectively, are to some extent rooted in the debate about internalist and externalist approaches to history of mathematics.[2] Joan Richards described how she, at the Madison "Conference on Critical Problems and Research Frontiers in History of Science and History of Technology" in 1991, experienced hostility and an internal division in the history of mathematics group. Her impression was that historians of mathematics "were actively engaged in drawing an even sharper line to divide themselves into these two groups [intenalists and externalists]" (Richards 1995, p. 123), which she coined "*the* critical problem in the history of mathematics" at that time (p. 124, italic in the original). Instead of distinguishing between mathematicians and historians dealing with past mathematics, she placed the internalist/externalist debate in the context of new developments in history of science and philosophy of mathematics. Drawing on examples from work on history of mathematics, she advocated for exploring various approaches: "the creativity", she wrote, "of the history of mathematics lies on the fractal boundary between internalism and externalism" (p. 134). David Rowe also discussed new trends and old images in the history of mathematics at that time, pointing out that "the ever widening range of interests among scholars and students alike reflects one of the major trends now taking place in the history of mathematics. This trend is linked with a shift from a relatively narrow, Eurocentric vision of a monolithic body of knowledge to a broader, multi-layered picture of mathematical activity embedded in a rich variety of cultures and periods" (Rowe 1996, p. 4).

In his essay from 2000, 'Formen der Mathematikgeschichte', Moritz Epple discussed the meaning of causality in relation to historical work. His understanding of causality in historical work does not entail historical laws that the episodes are following. Causal relationships are understood as consisting of a web of relations that is produced by historical events themselves, and which constrain or create different domains of possibilities for the realization of individual events. He interpreted the internalism/externalism controversy as a disagreement about the decisive (he called them causal) factors that affect the development of mathematics, where internal accounts seem to adhere to the strong hypotheses that mathematics is an autonomous science developing solely according to some inner rules dictated by the rationality of mathematics, immune to influences from outside (Epple 2000, p. 146). In such purely internal representations of history of mathematics, the decisive factors responsible for the development of mathematical concepts and mathematical knowledge are sought for only within mathematics itself. As a reaction, external (to the sciences) representations developed where the decisive factors were sought

[2] Basically, internalism accounts represent the point of view that mathematics develops without any influences from outside of mathematics whereas externalism considers factors from outside of mathematics in historical accounts.

for from outside the sciences themselves. Epple advocates for a form of history of mathematics where the nature of the causal factors that influence the development of mathematics has not been decided beforehand. This approach would be open to both kinds of explanations (internal and external) and makes it possible to investigate how internal and external factors are intertwined in various events. In his work on the history of topology, he emphasized that in order to accomplish this, the historian needs to analyze "[. . .] the weave of concrete scientific action rather than an abstract life of mathematical ideas" if they wish to "adhere to the goal of producing a causally coherent account of developments like the emergence of topology" (Epple 1998, p. 307).

In the next section, I present a problem-oriented multiple perspective approach to history of mathematics – an approach, which has this "openness", that Epple is advocating for, built into it, and which moves beyond the internalism/externalism division.

1.3 A Problem-Oriented Multiple Perspective Approach to History of Mathematics

I have on several occasions talked and written about what I call a multiple perspective approach to history of mathematics, how I see this as an open approach and how it moves beyond this division between internalism and externalism, see e.g. Kjeldsen (2012, 2019). In the following I will recap the ideas of the approach.

As indicated in footnote 1, I am inspired by the general historian Bernard Eric Jensen's thoughts about history as an academic discipline, and how people work with and use history. The underlying premise is an action-oriented conception of history where people are viewed as being shaped by history and being shapers (makers) of history. History is understood as collections of processes that people make or create through their projects in life as well as the intended and unintended consequences of these projects. To understand and explain historical-social processes means to gain insights into how people have acted and thought at different times and cultures (Jensen 2003, p. 378).

If we think of mathematics as a form of knowledge that is produced and develops through activities and projects that people, mathematicians, carry out in their work and life, such a conception of history can be adapted to history of mathematics. People's activities and projects are conditioned by the past, the present and the actors' (people's) expectations for the future. Historical-social processes in the making of mathematics can then be investigated and analyzed from various perspectives or points of observation by taking points of departure in concrete episodes of people's mathematical activities and projects from where connections and relations can be studied. It can be e.g. from practices of mathematics, from interdisciplinary perspectives, from institutions, from political, power and/or financial perspectives, from personal networks, from specific views on the nature of

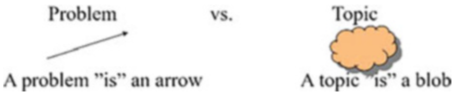


Fig. 1.1 The function of a problem formulation as a guidance for research

mathematics, from a specific culture, to name some. Such a multiple perspective approach to history of mathematics opens for exploring the interplay between developments of mathematics and the broader historical-social conditions of its development where both internal and external factors can be considered – in this sense, the "openness" mentioned above, is built into this approach.

When we do research in history of mathematics, there is always 'something', some research question, curiosity, agenda, motivation or suchlike that drives our research. This 'something' guides the investigations that the historian chooses to conduct and the analyses that are performed, whether this is articulated by the historian or not. In a problem-oriented research strategy, a question-answer process drives the research and the purpose is to find plausible answers to an articulated problem or complex of problems. A problem-formulation gives a direction for the research project of the historian and determines its boundaries – it functions as a 'steering gear' and guides the project (see Fig. 1.1).

A well-formulated problem/research question or group of research questions is an important part of all kinds of research. In history of mathematics research, it will help the historian to make decisions about what to include and what to leave out; it can function as a filter for what is important – issues that do not contribute to answering the research question, can be left out. Figure 1.1 is an attempt to illustrate the difference between having an articulated research question and a broader topic. A problem that one is trying to solve or investigate provides a direction for one's research, here represented by an arrow. A topic does not give a direction in itself to what to include and what to leave out; it is usually too broad. The blob in Fig. 1.1 illustrates this. It is not always an easy task to come up with a problem formulation, and often it is a dynamical process that sometimes continues throughout most of the historian's research and writing processes. Sometimes the process begins with a topic and a problem formulation crystalizes during the work. In other cases, the initial topic may not constrain the problem formulation since other areas that do not belong to the initial topic might turn up as important aspects to research. Sometimes the process begins with a question that crosses over a variety of topics or is not confined to a topic at all. A good place to start is with curiosity – something that strikes one as being weird or interesting, and reflections about why this is so.

In a multiple perspective approach, the question of which perspective(s) to choose, is answered or justified by the problem – the research question(s) of the historian. It raises, as all research approaches do, the question of reliability. Reliability concerns the trustworthiness of the analyses and conclusions performed and drawn by the historian. It depends on the 'evidence', so to speak, that has been available to/has been investigated by the historian, the material and sources. It also depends on the relation between the problem/the research question(s) of the historian, the methodological frameworks and tools used by the historian in

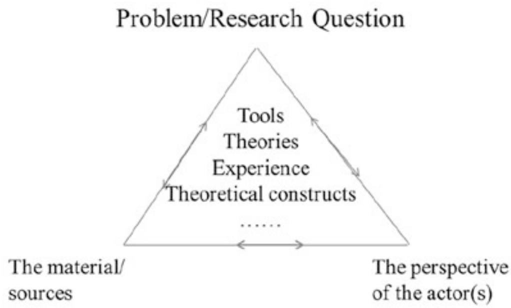


Fig. 1.2 Relation between problem – the research question, the chosen perspective of the actors and sources used. The figure is also presented in Kjeldsen (2019)

the analyses, and the chosen perspectives of the actors under investigation, the actors' motivations, goals and actions (see Fig. 1.2). The double arrows indicate the dynamics of the research process, and that all three corners might influence each other and change along the way. The initial research question, e.g., might change during the research process, which might change the choice made by the historian of which perspective of the actors, that is the lens, the vantage point, or point of observation from which to perform the investigations of the actors' actions and relations, which might cause a change in what sources are relevant and vice versa.

As I wrote, my inspiration for a m-perspective approach to history comes from the general historian Jensen's writings (see e.g. Jensen 2003). Recently, the notion of 'scientific perspectives' (Massimi and McCoy 2020) has caught attention among philosophers of science. The philosopher of science, Michela Massimi, in her book, *Perspectival Realism*, uses the notion of scientific perspective to offer an account of realism in science, which she calls a 'perspectival realism'. The focus is epistemological; it is about scientific knowledge claims as categorized by actors, taking seriously, as she phrases it, "the historically situated nature of scientific knowledge when discussing its epistemic foundation" (Massimi 2022, p. 9). Hence, it has a resemblance to the multiple perspective approach to history outlined in this chapter, in particular when the research question deals with epistemological issues in the historical development of mathematics.

Regarding the research questions (the problem-formulation), they are governed by the historian's curiosity or agenda, the processes in history of mathematics that the historian are interested in understanding and explaining – and these, again, are determined by the perspective of the historian. Hence, besides the actors' perspective, which underlines the situatedness of the historical inquiry, there is also the historian's perspective, which recognizes the role of the historian in the selections and interpretations of sources and data. This last element is related to reflectivity about historiographical practices. In the approach sketched here, with an action-oriented conception of history which takes the actors' projects, perspectives, and social and scientific practices into account, and acknowledges the perspective of the historian, there cannot be one single account – but this does not mean that "anything goes". I agree with Epple, that despite this difficulty historiography should not give up on the attempt to provide, in his framework, "causally coherent narratives" (Epple 1998, p. 308). More work on this issue needs to be done, and I will not go further into the discussion here, but end with quoting another general historian, Søren Mørch, who in paraphrasing (Jensen 2003) urges that:

> Historians should explain the position they take in relation to the subject they are writing about and the approach they are using so that the reader can better relate to the narrative that he or she is dragged into. (Mørch 2010, p. 503, my translation)

In order to illustrate and reflect on the above in a concrete and specific way, I will give short extracts in the next sections from three examples where I have used a p-oriented m-perspective approach and where I have analyzed the historical actors' mathematical activities and projects from various perspectives dictated by the problem. The three cases also exemplify various factors influencing the development of mathematics.

1.4 The Inventive Art of Minkowski and Epistemic Objects and Tools

Minkowski's 'inventive art', as David Hilbert called it in his memorial speech for Minskowski, is connected to convexity, which is a central notion in mathematics. The theory of convexity is useful in many areas of modern mathematics and in applied fields. The idea of a general convex body was crystallized in the period 1887–1897. In monographs on convexity,[3] the German mathematician Carl Hermann Brunn (1862–1939) is often credited as being the first to undertake systematic studies of general convex bodies in his thesis, with Hermann Minkowski (1864–1909) credited with further development of the theory. Both of them worked with convex bodies, in the same time period, but in two very different ways. Minkwoski's work on and with convex bodies became very successful and generated new mathematical research areas of convexity theory and geometric number theory – it was very fruitful. Brunn's work did not have the same effect. Walther von Dyck who was professor at the Technical University of Munich at that time described Brunn's thesis as containing some naïve ideas, but nothing more than that.[4]

This difference prompted several questions such as: What kinds of objects did Brunn and Minkowski study and why? How did they carry out their investigations, which methods did they use, which questions were they interested in and why? How and why was the theory of convex sets developed through Minkowski's work?[5]

[3] See e.g. Bonnesen and Fenchel (1934), Klee (1963).

[4] See Hashagen (2003, p. 245).

[5] See Kjeldsen (2008, p. 60) and Kjeldsen (2009, pp. 85–86).

These questions, among others, led to the following problem formulation (Kjeldsen 2009, pp. 85–86):

1. How and why did the concept of convex bodies emerge in the two trajectories of mathematical research of Brunn and Minkowski, respectively?
2. Why did Minkowski's strand of research lead to the development of a theory of convexity – in contrast to Brunn's? Were the differences crucial for the object they studied and for what it could be used for and lead to?

Answers to such questions, explaining differences in research practices and understanding whether such differences are significant or not, can be reached by analyzing concrete mathematical activities within the mathematical practices in which specific mathematical knowledge is produced. In this case, it can be reached by studying the mathematical activities of Brunn and Minkowski from the practice of mathematics guided by the research questions.

Historiographical tools that are well suited to dealing with history of mathematics research that is focused on mathematical activities from the perspective of the actors, their projects, motivation and wishes studied through the lens of their mathematical practices, are the notions of epistemic objects and techniques. They have been adapted into historiography of mathematics by Epple (see e.g. Epple 1999) from Hans-Jörg Rheinberger's methodological framework of epistemic things and technical objects (Rheinberger 1997). They are elements of what Rheinberger calls experimental systems, which he defines as the smallest working unit of research. Rheinberger developed his framework for investigating experimental practices of molecular biologists and, as Epple explains, even though at first sight the notion of technical object might not seem to play a role in mathematical research except perhaps for calculating (and measuring devices, surface models etc.) and thereby also might not seem to be generally applicable for historians of mathematics, that is not the case (Epple 1999, p. 15). Rheinberger's notions of epistemic things and technical objects are not meant to be taken as stable entities belonging to theory and experimental practices, respectively, but are given a functional meaning. With this in mind, Rheinberger's notion of technical object can be adapted to mathematical research where there also is a "technical machinery", as Epple explains, of various mathematical techniques involved in concrete mathematical research activities. To circumvent the association with material machinery, Epple suggests that we in history of mathematics talk about epistemic techniques instead of technical objects.

The notions of epistemic objects and epistemic techniques can distinguish between how problem-generating and answer-generating elements of particular research episodes function, interact and change in concrete episodes of mathematical research of a specific mathematician or group of mathematicians. They make processes in research practice visible. They can capture dynamics of knowledge production, and as such, they are suitable tools for providing answers to the kind of historical research questions formulated above.

Brunn wrote his inaugural thesis *About Ovals and Egg-surfaces* in 1887 at the Ludwig-Maximilian University of Munich. At that time, a transition from empirical-intuitive to formal-deductive mathematics was taking place.[6] Brunn's thesis belongs to the empirical-intuitive tradition. In his thesis, he set out to perform what he called "elementary geometrical investigations of a special kind of real curves and surfaces – oval and egg surfaces" (Brunn 1887). By which he understood what we today recognize as convex bodies in two and three dimensions. He defined an oval as a closed plane curve that has two and only two points in common with every intersecting straight line in the plane, and a full oval as an oval together with its inner points. He defined what he called egg-surfaces and egg-bodies in a similar way as the corresponding spatial objects. These, Brunn wrote in his thesis, are objects whose properties were unknown and they generate questions. He asked questions and initiated investigations of curvature, area, volume, cross-sections and extremal properties of these objects.

The questions that Brunn could ask, depended on his epistemic techniques, the tools with which he investigated the objects. This will be spelled out more clearly below in the comparison with Minkowski's work. Brunn was devoted to what he called "Steiner's methodology" of geometry, for which he gave the following reason:

> I [Brunn] was not entirely satisfied with the geometry of that time which strongly stuck to laws that could be presented as equations quickly leading from simple to frizzy figures that have no connection to common human interests. I tried to treat plain geometrical forms in general definitions. In doing so I leaned primarily on the elementary geometry that Hermann Müller, an impressive character with outstanding teaching talent, had taught me in the Gymnasium, and I drew on Jakob Steiner for stimulation. (Brunn 1913, p. 40)

Brunn is referring to a nineteenth century controversy between the synthesists and the analysists, he himself siding with the synthesists – being of the opinion that synthesis was the proper way to argue in geometry.

Minkowski came to define convex bodies from a very different direction – it came out of his work on the so-called minimum problem for positive definite quadratic forms in n variables. In a resumé of a talk "Über Geometrie der Zahlen" which he presented in 1891, he introduced the three-dimensional lattice as a collection of points with integer coordinates in space with orthogonal coordinates, in which he considered, as he wrote, "a very general category of bodies" that:

> are constructed in such a way that they circumscribe a particular lattice point – for instance the origin – in a certain way. ... [it] consists of all those bodies that have the origin as middle point, and whose boundary towards the outside is nowhere concave. (Minkowski 1891/1911, pp. 264–265)

At first, Minkowski's epistemic objects were bodies of a certain shape and they were at the outset connected to lattice points.

Two years later, in a talk where he presented an outline of his forthcoming book *Geometrie der Zahlen*, he introduced what he called the radial distance, $S(a, b)$,

[6] See Toepell (1996) and Hashagen (2003).

between two points a, and b, and its corresponding “Eichkörper”, $S(ou) \leq 1$, which we recognize as its unit ball. If the triangular inequality holds, he argued, its

> “Eichkörper” then has the property that whenever two points u and v belong to the “Eichkörper” then the whole line segment uv will also belong to the “Eichörper”. On the other hand every nowhere concave body, which has the origin as an inner point, is the “Eichkörper” of a certain “einhellig” radial distance [i.e. a radial distance for which the triangular inequality holds]. (Minkowski 1893/1911, pp. 272–273)

The manuscript for Minkowski’s talk clearly shows that he had realized by then that the essential property for his proof on the minimum theorem is the property of convexity, most notable in his famous lattice point theorem:

> If the Eichkörper for an “einhellig” and reciprocal radial distance has volume $J \geq 2^3$ then the Eichkörper contains a lattice point in addition to the origin.
>
> [...]The hereby gained theorem about nowhere concave bodies with middle point seems to me to belong to the most fruitful in the whole of number theory. (Minkowski 1893/1911, p. 274)

With the lattice point theorem, Minkowski connected the volume of nowhere concave bodies with points with integer coordinates – and in his book on geometry of numbers from 1896, he developed his theory in n dimensions. The concepts of radial distance, Eichkörper, nowhere concave bodies and the lattice, which he introduced, came out of his work on the minimum problem for positive definite quadratic forms in n variables – and enabled him to prove the minimum theorem in a simple and elegant way using geometrical intuition. In his memorial speech for Minkowski, Hilbert named this method of Minkowski’s “a pearl of the Minkowskian art of invention” (Hilbert 1909/1911, p. XI) – and no wonder, with this work, Minkowski moved number theory into a new epistemic place that led to a new research field, Geometry of Numbers, and to his beginning the further development of a theory of convex sets.[7]

Following Minkowski’s line of work that led him to crystalize the concept of convex bodies by analyzing his papers, using the historiographical tools of epistemic objects and techniques, a process becomes visible. At first, Minkowski investigated positive definite quadratic forms in n variables. These were the epistemic objects. Regarding epistemic techniques, initially he set out to investigate the minimum problem using a geometrical interpretation of positive definite quadratic forms and analytic geometry. This technique was not so well understood at the time: there was a sketch produced by Gauss for the case of two variables and a work by Dirichlet for the case of three variables. Minkowski developed the technique for n variables, introducing the n dimensional lattice. This then became a new epistemic object under investigation, leading Minkowski to introduce further objects: a nowhere concave body with middle point; the radial distance with which he generalized the notion of distance; the Eichkörper. These were fundamental objects in the geometrical technique that he was developing for his number theoretical studies – they constituted the proof-technique.

[7] For details, see Kjeldsen (2008, 2009).

Tom Bonnesen and Werner Fenchel, as they presented their subject in the introduction to their monograph on convex bodies, in 1934, considered Brunn's egg-forms and Minkowski's nowhere concave bodies to correspond to each other – as universal, de-contextualised mathematical entities (Bonnesen and Fenchel 1934). However, if we view and compare these objects and how they were treated by Brunn and Minkowski, we might ask: did they in practice investigate the same objects?

Brunn's egg-forms were abstractions of shapes of material objects in space, an interpretation that is supported by the names he gave them, which come from everyday experiences such as oval drawings and egg shapes. Minkowski's bodies which, in the beginning, were described as nowhere concave towards the outside and surrounding the origin in a certain way, were developed by Minkowski as a tool to investigate the minimum problem. They were interpreted as "measure" bodies, which he constructed through a radial distance function, and they constituted a major part of his proof-technique. Minkowski's nowhere concave bodies did not emerge from material objects – they were dictated by positive definite quadratic forms in n-variables and their ability to measure distance, hence they "lived" in n-dimensional manifolds.

Within the framework of a multiple perspective approach to history of mathematics with a focus on specific episodes of mathematicians' development of mathematical knowledge, the notions of epistemic objects and techniques are useful tools in the historiography of mathematics in the sense that they provide a framework to analyze the dynamics of knowledge production in concrete mathematical practices. The framework helps us to understand and "see" the embeddedness of mathematical, epistemic objects in the epistemic techniques used in particular research episodes, and how, in concrete mathematical research activities, the conception of a mathematical object and the way it develops and crystallizes is embedded in mathematical practice. In relation to the case of Brunn's and Minkoski's research regarding the development of general convex bodies, we saw that the questions that Brunn and Minkowski could ask depended on the epistemic techniques by which they investigated their objects. Brunn could not have asked the question of how large the volume of an egg-body must be for it to have a lattice point. He had no lattice and no tools to link the volume of an egg-body to spatial points with integer coordinates. The object, and what one can do with it, depends on the epistemic techniques that are available to us – to the specific actors in question. The ability to shed light on this and its significance for the historical development of mathematics from its practices makes epistemic objects and techniques useful tools in the historiography of mathematics.

1.5 The Significance of World War II for the Developing of Mathematical Programming

The second example is from the development of mathematical programming – or mathematical optimization, which the Society of Mathematical Programming changed the name of the society to in 2010.[8] It is concerned with theories and techniques for how to minimize or maximize a real function subject to (inequality) constraints on the variables. The so-called Diet Problem of selecting a set of foods that will satisfy specific requirements of daily uptake of vitamins, at minimum cost, is an example of a linear programming problem, where all the involved functions (the one to be optimized and the functions describing the constraints) are linear functions. Its key theoretical results circle around duality theorems. In the USA, mathematical programming developed in connection with the Second World War.[9] At a first glance, focusing on these two aspects, duality and World War II, might seem to call for an internal and external historical investigation respectively. However, using a m-perspective approach with the perspective of the actors as the point of departure and guided by the research questions formulated below, illustrates how in this case the historical investigation moved beyond the internal/external division and provided opportunities to explore interactions between the creation of mathematics and (some) conditions for its development in a broader socio-economic context.

Among others, the research questions that guided the historical investigations and analyses were:

1. How did ideas of duality emerge in linear programming? What role did they play for the development of mathematical programming?
2. What role did the military play for the emergence and development of mathematical programming and what influence did it have on the establishment of mathematical programming as a research area?

A linear programming problem is a finite dimensional optimization problem of finding the maximum or minimum of a linear function subject to linear inequality constraints. To such a problem, one can formulate a similar problem, called the dual problem, using the same data. If the original problem, called the primal problem, is a maximum problem, the dual is a minimum problem, and vice versa. The duality theorem states that if one of the two problems, the primal or the dual, has a finite optimal solution, so does the other one, and the optimal values are equal.

In the USA, George B. Dantzig (1914–2005) was one of the protagonists. He worked during the Second World War at the US Air Force's Combat Analysis Branch of Statistical Control. He worked on calculation of what they called a

[8] See the webpage of the society http://www.mathopt.org/

[9] Similar ideas were published by the Russian economist and mathematician Leonid V. Kantorovich in 1939, see Kantorovich (1939/1960) and Leifman (1990).

program. An Air Force program was a tool for logistic planning of operations and activities. According to Dantzig, it took more than 7 months to construct a consistent program (Dantzig 1982). After the war, Dantzig was re-hired by the Air Force to work on how to speed up the construction of programs. However, the advent of the computer opened up the possibility of calculating not only consistent programs, but also the least expensive one. That changed the focus of the work, and Dantzig and his group developed a mathematical model for the Air Force problem – the minimization of a linear form subject to linear inequalities. Today we recognize this as a linear programming problem.

To discuss how to solve such a programming problem, Dantzig consulted John von Neumann (1903–1957), who, besides being a professor at the Institute for Advanced Study at Princeton University in the USA, also held consulting jobs for the military in the post-war period. On one of these visits to Princeton, Dantzig met Albert W. Tucker (1905–1995), who was a professor at the mathematics department at Princeton. The Office of Naval Research, who was a major player for funding of research in the USA in the years after WWII, funded a summer project to explore the mathematics. Tucker became principal investigator of the project and two younger mathematicians, Harold W. Kuhn (1925–2014) and David Gale (1921–2008) worked initially with Tucker within the project. Their first results, which included existence and duality theorems for linear programming, were derived in 1949.

Gale, Kuhn, and Tucker's results were published in 1951 in the proceedings of a conference, which is now thought of as the "zero's" conference on linear programming (Gale et al. 1951). Their paper was the first publication of duality theorems in linear programming which explains Tucker's surprise, when Dantzig in his text book on linear programming of 1963, ascribed duality to von Neumann and not to Tucker and his group. Dantzig's answer was "because he [von Neumann] was the first to show me" (Dantzig 1982, p. 45). Dantzig further explained that von Neumann did so when they had met together in the fall of 1947. Dantzig has a longer description of what went on at that meeting in his recollections from 1982. Von Neumann had just a few years earlier finished a book on game theory together with the Austrian economist Oscar Morgenstern. According to Dantzig, at their meeting, von Neumann outlined a theory for the Air Force problem analogous to the one he and Morgenstern had developed for games, and he conjectured that the Air Force problem is equivalent to a two-person zero-sum game (Dantzig 1982, p. 45).

Dantzig's description of what went on was written two decades later and cannot be taken at face value. Recollections are reconstructions that are influenced by later events, personal developments and reflections. However, that does not mean that they are not valuable sources in historical investigations. They can reveal essential things that are relevant for the historical study at hand. In the present case, where the focus is on the duality theorem and the significance of the military involvement in the mathematicians' research on linear programming, Dantzig's description reveals several links in a concrete communicative situation, links between actors, institutions, mathematics and circumstances of the time. There is a link between two mathematicians, between a professor in academia and a mathematician working in the Air Force, which was brought about because of WWII and the involvement and organization of civilian scientists for the war effort in the USA.[10] There is a link between the Air Force problem and games, and the underlying mathematical theories of linear inequality and convexity. There was a sharing and transmission of knowledge across various institutional boundaries – all of it conditioned by WWII and the post-war organization of science support in the USA.

John von Neumann was probably the only one who could have made the connection between game theory and linear programming at that time.[11] The involvement of game theory was important both internally at the mathematically theoretical level and externally for the focus on these two subjects in the post-war military funding of mathematical research.[12]

Mina Rees (1902–1997) was another important actor for the early development of mathematical programming in the USA. She had served as a technical assistant to Warren Weaver during the war, and in the post war-period, she was head of the Office of Naval Research's mathematics program. She wrote about the establishment of a separate Logistic Branch in the Office of Naval Research because of the possibilities that some of the mathematical results of linear programming "... could be used by the Navy to reduce the burdensome costs of their logistics operations". This, she concluded "... has proved to be a most successful activity of the Mathematics Division of ONR, both in its usefulness to the Navy, and in its impact on industry and the universities" (Rees 1977, p. 111). So while the involvement of Tucker seems to have been almost a coincidence, the establishment of the Office of Naval Research logistic program and the funding of mathematical programming was not. This was rather because of a deliberate strategy on behalf of the Office of Naval Research to promote and to some extent control scientific research in the post-war period.

The research moved into academia through the project funded by the Office of Naval Research and the connection to von Neumann and his work. Tucker and Kuhn continued working on mathematical programming, funded in the beginning by the Office of Naval Research and later by the National Science Foundation (NSF) of the USA. In 1949, they tried to extend the duality result for linear programming to the nonlinear case, and the following year they presented their results in a talk entitled Nonlinear Programming, which they gave at the Second Berkeley Symposium on Mathematical Statistics and Probability and published in the proceedings (Kuhn and Tucker 1950). They did not succeed in deriving a duality result in the paper,[13] but they proved the now famous Kuhn-Tucker conditions for nonlinear programming. An analysis of their paper, however, suggests that the

[10] See for example Owens (1989), Schweber (1988), Zachary (1997).

[11] For an elaboration of that, see Kjeldsen (2000, 2001, 2003, 2009).

[12] See e.g. Leonard (1992) and Mirowski (1991).

[13] The first duality result for nonlinear programming is due to Werner Fenchel (1953), see Kjeldsen (forthcoming-a).

duality result for linear programming was in fact their motivation for extending the field on linear programming into nonlinear programming.[14]

All the way through, the historian's (i.e. my) underlying perspective of driving forces in historical developments of mathematics, studied from the perspective of the actors' motivations and opportunities to engage in this kind of mathematics, guided by the research questions above, has illustrated that internal and external factors interacted in crucial ways. These interactions were significant for the field of mathematical programming, both for its theoretical development and for its establishment as a mathematical sub-discipline.

1.6 Rashevsky and "The Power of a Superman" for Developing Mathematical Biology

The final example includes a perspective of interdisciplinarity with respect to practices of scientific research and the establishment of new interdisciplinary research fields with mathematics. It concerns the Ukrainian scientist Nicolas Rashevsky's (1899–1972) early work in bringing mathematics to biologists to create a physico-mathematical foundation for biology. Trained as a theoretical physicist, his hope was that mathematics could do for biology what it had done for physics. It turned out not to be so easy to convince biologists of the usefulness of mathematics in biology. My perspective for the piece of research presented here was interdisciplinarity, the move of mathematics as a scientific practice into other sciences and the creation of new interdisciplinary research fields – here the beginning of mathematical biology. The research was guided by questions of how Rashevsky argued for and viewed the epistemic role of mathematics in biology, and the emergence of mathematical modeling in biology. Rashevsky ran into problems in his encounters with biologists, and his early work on cell division and his philosophical ideas have been analyzed by taking his motivation and goals into account, as well as his perception of the role of mathematics as a method to gain knowledge in biology.

Rashevsky immigrated to the USA in 1924, to Pittsburg where he worked at the research department of Westinghouse Electric and Manufacturing Company. While he was there, he did research in the dynamics of colloid particles and the division of drops, drawing an analogy to cell division, but being well aware that just because there are some similarities, a biologist would never assume that the exact same phenomenon occurs in cells. However, he claimed, "the observation of such a "model" justifies the more general assumption, that some kind of variation of the surface tension of the cell, probably of a much more complicated nature, may be responsible for the cell-division" (Rashevsky 1931, pp. 143–144).[15]

[14] See Kjeldsen (2000).

[15] Rashevsky was neither the first one to suggest this analogy nor the first person to attempt to introduce mathematical and physical methods into biology, see e.g. (Israel 1993; Keller 2002).

In 1934, he received a Rockefeller Fellowship that took him to the University of Chicago where he worked on bringing mathematical methods to biology. The time was favorable for Rashevsky's vision. Warren Weaver, who was the director for the Natural Sciences of the Rockefeller Foundation, was very much in favor of the Foundation undertaking "a long-range program of support of quantitative biology – a program that would seek to apply to outstanding problems of biology some of the methods and machines that had been so successful in the physical sciences" (Rees 1987, p. 501).

Rashevsky was critical towards the experimental practices of biologists. In a paper in *Nature* in 1935, he explained why and emphasized the crucial role he thought mathematics had to play:

> ...our experiments do not and cannot reveal those hidden fundamental properties of Nature. It is through mathematical analysis that we must infer, from the wealth of known, relatively coarse facts, to the much finer, not directly accessible fundamentals. (Rashevsky 1935, p. 528).

The way to proceed, according to Rashevsky, was through what he called "paper and pencil" models. We do not have to actual build a model, he wrote in a paper in *Physics* from 1931, we can satisfy ourselves "by investigating mathematically, whether such a model is possible or not. [..] The mathematical method has a greater range of possibilities, than the experimental one, the latter being often limited by purely technical difficulties." (Rashevsky 1931, p. 144)

The thought that mathematics could or should play a role in biology was not foreign. When the biologist Reginald G. Harris took over as director of the marine biology teacher-training laboratory, the Cold Spring Harbor Laboratory at Long Island, New York, he gathered mathematicians, physicists, chemists and biologists for research symposia on quantitative biology. Rashevsky participated in 1934 with a paper on cell division. He presented his work to date on his mathematical model for cell division, which he had worked on and published in a series of papers, extending and developing the model. In the paper he presented at the symposium, he began by asking if we need to assume some special independent mechanisms to explain cell division. His opinion was, no, cell division can be explained as a direct consequence of the forces arising from cell-metabolism.

He calculated the forces, acting on a unit volume, which is produced in a cell by a gradient of concentration of a substance due to metabolism. He idealized cells as homogenous and spherical. With these assumptions, he was able to deduce that when the radius of a cell reached a critical seize, division of the cell would decrease the free energy of the system. With reference to the principle of free energy, he took this, not as a definite explanation of cell division, but as a proof of a *possible* explanation, and concluded that cell division can be explained as a direct consequence of the forces arising from cell metabolism. (Rashevsky 1934)

The discussion following the talks at the symposia is published in the proceedings, and it captures the biologists' reaction to Rashevsky's model. They were quite critical, "What is the nearest example in nature to this theoretical case?" they wanted to know, continuing with a dismissal because, in their opinion "... it doesn't help as a general solution, because a spherical cell isn't the commonest form of cell." (Rashevsky 1935, p. 198)

In the introduction to the first edition of his monograph *Mathematical Biophysics: Physico-Mathematical Foundations of Biology* from 1938, Rashevsky addressed this "problem of reality", which he was faced with from the biologists, writing in a way that indicates that he was losing his patience:

> We start with a study of highly idealized systems, which at first may even not have any counterpart in real nature. This point must be particularly emphasized. The objections may be raised against such an approach, because such systems have no connection with reality; and therefore any conclusions drawn about such idealized systems cannot be applied to real ones. Yet this is exactly what has been, and always is, done in physics. The physicist goes on studying mathematically, in detail, such nonreal things as "material points," "absolutely rigid bodies," "ideal fluids," and so on. *There are no such things as those in nature.* Yet the physicist not only studies them but applies his conclusions to *real things*. And behold! Such an application leads to practical results – at least within certain limits. This is because within these limits the real things have common properties with the fictitious idealized ones! Only a superman could grasp mathematically at once all the complexity of a real thing. We ordinary mortals must be more modest and approach reality asymptotically, by gradual approximation. (Rashevsky 1938, p. 1, italic in the original)

On one hand, there was a clash in scientific practice between Rashevsky's physico-mathematical approach with an emphasis on theory and the biologists' approach with an emphasis on collecting facts and performing experiments. These epistemological differences across the disciplinary boundaries functioned as constraints; there was no common ground for what was considered to be meaningful or useful knowledge between Rashevsky and his group and the biologists. On the other hand, there were social-economic factors at that time in the USA that promoted interdisciplinary research, especially the Rockefeller Foundation that supported Rashevsky financially.[16] Retrospectively, the analogy of droplets did not work for the phenomenon of cell division. Rashevsky's model failed to explain cell division. Michael Conrad evaluated the early work on mathematical biology in a piece on the history of the Society of Mathematical Biology as follows:

> that mathematical biology was not in a good repute [1970s], and theoretical biology was in the doghouse as well. [. . .] This atmosphere was the natural aftermath of the truly great advances in molecular biology initiated in the fifties by the discovery of the significance of DNA. [. . .] a generation of theoretical "speculation" was being sent to the graveyard in body bags. (Conrad 1996, p. 8)

Coming back to Rashevsky, his work – his activities to create a physico-mathematical foundation for mathematical biology – is historically interesting even though his model for cell division failed. It exhibits various kinds of difficulties in interdisciplinary collaboration. It illustrates the uncertainty inherent in research at the frontier where new areas are explored and/or new methods are employed. Due to the focus on promoting interdisciplinary research and the funding he

[16] For Rashevsky's relationship with the Rockefeller Foundation, see Abraham (2004). For a scientific biography of Rashevsky, see Shmailov (2016).

received, Rashevsky was able to place mathematical biology on the agenda in the scientific community regarding both research and education.[17] He also launched the first journal in the field *The Bulletin of Mathematical Biophysics*, which is now the official journal, renamed *Bulletin of Mathematical Biology*, of the Society of Mathematical Biology.

1.7 Discussion and Concluding Remarks

In history of mathematics with a problem-oriented research strategy, the product of research is a process of formulating questions and deriving answers. It goes together with a functional conception of sources in the sense that the function of a source is determined by the problem (or problem complex) the historian has decided to explore. This also means that the question of validation of such historical research hinges on the relevance of the chosen problem or problem complex. The historian must be able to justify the chosen research questions within the field of inquiry. In the problem-oriented project organized pedagogical model of teaching and education in the Roskilde Model (Andersen and Kjeldsen 2015a, b), the exemplary principle (for the sciences in the sense of Wagenschein) has been the answer. This can also function as the answer for research. The questions – or the exemplars – one chooses to answer, or investigate, must be an exemplar that shed light and gives insights into relevant issues in the field.

We will look at the three cases above one by one. The first case might in some sense be taken as representative for some aspects of twentieth century mathematics, namely the move into an autonomous enterprise. Minkowski generalized concepts that he envisioned through geometrical intuition, into n-dimensional space in order to be able to use them to solve number theoretical problems – most notably the minimum problem for positive definite quadratic forms. Minkowski conducted what we can call interdisciplinary work within mathematics, by developing a geometrical method to deal with number theory. In this endeavor, he crystalized the concept of a general convex body in a process of connecting different mathematical disciplines – this gave rise to the new mathematical research fields of geometry of numbers and the theory of convexity. The former he developed in his book *Geometry of Numbers* and the latter he began to develop in manuscripts that he did not have the time to publish before he died in 1909 at the age of 45. They were found after his death among his papers and published in his collected works. The contrast between Minkowski's and Brunn's work, and especially the very different impact

[17] See Kjeldsen (2019). For a discussion of the significance of the case of Rashevsky for teaching in mathematics education, see Kjeldsen (2017), Kjeldsen (forthcoming-b), Jessen and Kjeldsen (2021).

their various strands of research had on further developments in mathematics, sheds light on some characteristics for the growth of mathematics.[18]

The second case is an exemplar of how conditions, or one might say, urgency, in society might influence the development of mathematics, together with more internal developments of mathematics. The realization of the connection between linear programming and game theory sparked the idea of looking for duality results, so here again we see the fruitfulness of interdisciplinary research within mathematics as such. However, in this case we also see a significant influence from society outside of mathematical institutions. The connection between the two fields and their further developments were realized and promoted during the war because of the organization of civilian scientists for the war effort in the USA. After the war, the connection was maintained through the roles these scientists played in various advisory boards and the continuing military support for research in game theory and mathematical programming in the universities and in agencies such as the Rand Cooperation, funded among others by the Office of Naval Research. Through these connections and funding, the Air Force programming problem moved into academic mathematical research in the universities and became a research object of mathematicians working in academia.

The two cases of convexity theory and mathematical programming are further connected, in the sense that Werner Fenchel developed the first duality theorem for nonlinear programming during a sabbatical in the USA. He met Albert Tucker when he (Fenchel) was visiting the Institute for Advanced Study at Princeton. Fenchel was the leading expert on the theory of convex bodies, and by then the significance of convexity theory had been recognized among the people working in mathematical programming. Tucker invited Fenchel to prolong his stay in Princeton and to give a lecture course in convexity. Fenchel's famous lecture notes *Convex Cones, Sets and Functions* (Fenchel 1953) from this course, which were published in 1953, is a classic in mathematical programming literature.

The third case is an exemplar of how conditions within the sciences, what is deemed important in academia at large, and in society, have a significant influence on what kind of research is being developed. On one hand, it is an exemplar of issues that arise and become important when mathematics emmigrates, so to speak, into other sciences and tries to promote a different (mathematical) way of knowledge production, along with non-disciplinary issues of funding from the surrounding society. On the other hand, it is an exemplar of how mathematical modelling can serve as a research tool at the frontier of science as well as an explorative tool in the sense of Axel Gelfert (see Gelfert 2018).

Taken together, the three cases are meant to be in some sense exemplars for research questions addressing internal mathematical, external societal and external (to mathematics) epistemological factors. They illustrate how such factors might

[18] The case was used in connection with philosophical ideas regarding the growth of mathematics of how new objects are introduced into mathematics and how we are able to reason with new objects, see Kjeldsen and Carter (2012).

be intertwined in the emergence and developments of new mathematical objects, theories and interdisciplinary areas of research. As such, the studies presented in this paper contribute to our understanding of how mathematics, as developed and cultivated by people, and conditions of a particular time and place interfere in the history of mathematics – brought about by a problem-oriented multiple perspective approach.

References

Abraham, T. 2004. Nicholas Rashevsky's Mathematical Biophysics. *Journal of the History of Biology* 37: 333–385.

Andersen, A. S., and S. B. Heilesen, eds. 2015. *The Roskilde Model: Problem-Oriented Learning and Project Work*, Volume 12 in Innovation and Change in Professional Education. Cham/Heidelberg/New York/Dordrecht/London: Springer.

Andersen, A. S., and T. H. Kjeldsen. 2015a. Theoretical Foundations of PPL at Roskilde University. In *The Roskilde Model: Problem-Oriented Learning and Project Work*, Volume 12 in Innovation and Change in Professional Education, ed. Anders S. Andersen and Simon B. Heilesen, 3–16. Cham/Heidelberg/New York/Dordrecht/London: Springer.

———. 2015b. A Critical Review of the Key Concepts in PPL. In *The Roskilde Model: Problem-Oriented Learning and Project Work*, Volume 12 in Innovation and Change in Professional Education, ed. Anders S. Andersen and Simon B. Heilesen, 17–36. Cham/Heidelberg/New York/Dordrecht/London: Springer.

Blåsjö, V. 2014. A Critique of the Modern Consensus in the Historiography of Mathematics. *Journal of Humanistic Mathematics* 4 (2): 113–123.

———. 2021. Historiography of Mathematics from the Mathematician's Point of View. In *Handbook of the History and Philosophy of Mathematical Practice*, ed. B. Sriraman. Cham: Springer. https://doi.org/10.1007/978-3-030-19071-2_67-1.

Bonnesen, T., and W. Fenchel. 1934. *Theorie der konvexen Körper*. Berlin: Springer.

Brunn, H. 1887. Ueber Ovale und Eiflächen. Inaugural-Dissertation Munich: Akademische Buchdruckerei von F. Straub.

———. 1913. Autobiography. In *Geistiges und Künstlerisches München in Selbstbiographien*, 39–43. Max Kellers Verlag: Munich.

Conrad, M. 1996. Childhood, Boyhood, Youth. *Society of Mathematical Biology Newsletter* 9 (3): 8–9.

Dantzig, G. B. 1982. Reminiscences About the Origins of Linear Programming. *Operations Research Letters* 1: 43–48.

Epple, M. 1998. Topology, Matter, and Space, I: Topological Notions in 19th-Century Natural Philosophy. *Archive for History of Exact Sciences.* 52: 297–392.

———. 1999. *Die Entstehung der Knotentheorie. Kontexte und Konstruktionen einer modernen mathematischen Theorie*. Braunsweig/Wiesbaden: Friedr. Vieweg & Sohn.

———. 2000. Genies, Ideen, Institutionen, mathematische Werkstätten: Formen der Mathematikgeschichte. *Mathematische Semesterberichte* 47: 131–163.

Fenchel, W. 1953. *Convex Cones, Sets, and Functions*, Lecture Notes. Department of Mathematics, Princeton University.

Fried, M. N. 2014. The Discipline of History and the "Modern Consensus in the Historiography of Mathematics". *Journal of Humanistic Mathematics* 4 (2): 124–136.

———. 2018. Ways of Relating to the Mathematics of the Past. *Journal of Humanistic Mathematics* 8 (1): 3–23.

Gale, D., H. W. Kuhn, and A. W. Tucker. 1951. Linear Programming and the Theory of Games. In *Activity Analysis of Production and Allocation*, Cowles Commission Monograph, 13, ed. T.C. Koopmans, 317–329. New York: Wiley.

Gelfert, A. 2018. Models in Search of Targets: Exploratory Modelling and the Case of Turing Patterns. In *Philosophy of Science. Between the Natural Sciences, the Social Sciences, and the Humanities*, ed. A. Christian, D. Hommen, N. Retzlaff, and G. Schurz. New York: Springer.

Grattan-Guinness, I. 2004. The Mathematics of the Past: Distinguishing Its History from Our Heritage. *Historia Mathematica* 31 (2): 163–185.

Hashagen, U. 2003. *Walther von Dyck (1856–1934): Mathematik, Technik und Wissenschaftsorganisation and der TH München*. Stuttgart: Franz Steiner Verlag Stuttgart.

Hilbert, D. 1909/1911. Hermann Minkowski. Gedächtnisrede. 1909. In Minkowski, H. (1911). *Gesammelte Abhandlungen,* V–XXXI. Volume I. Leipzig, Berlin: B. G. Teubner.

Israel, G. 1993. The Emergence of Biomathematics and the Case of Population Dynamics: A Revival of Mechanical Reductionism and Darwinism. *Science in Context* 6: 469–509.

Jensen, B. E. 2003. *Historie – livsverden og fag*. Copenhagen: Gyldendal.

Jessen, B. E., and T. H. Kjeldsen. 2021. Mathematical Modelling in Scientific Contexts and in Danish Upper Secondary Education– Are There Any Relations? *Quadrante* 30 (2): 37–57.

Kantorovich, L.V. 1939/1960. Mathematical Methods of Organizing and Planning Production. *Management Science* 6: 366–422.

Karstens, B. 2014. The Peculiar Maturation of the History of Science. In *The Making of the Humanities. - Vol. 3: The Modern Humanities*, ed. R. Bod, J. Maat, and T. Weststeijn, 183–203. Amsterdam University Press.

Keller, E. F. 2002. *Making sense of life: Explaining biological development with models, metaphors, and machines*. Harvard University Press.

Kjeldsen, T. H. 2000. A Contextualized Historical Analysis of the Kuhn-Tucker Theorem in Nonlinear Programming: The Impact of World War II. *Historia Mathematica* 27: 331–361.

———. 2001. John von Neumann's Conception of the Minimax Theorem: A Journey Through different Mathematical Contexts. *Archive for History of Exact Sciences* 56: 39–68.

———. 2003. New Mathematical Disciplines and Research in the Wake of World War II. In *Mathematics and War*, ed. B. Booss-Bavnbek and J. Høyrup, 126–152. Basel/Boston/Berlin: Birkhäuser Verlag.

———. 2008. From Measuring Tool to Geometrical Object: Minkowski's Development of the Concept of Convex Bodies. *Archive for History of Exact Sciences* 62: 59–89.

———. 2009. Egg-forms and Measure Bodies: Different Mathematical Practices in the Early History of the Development of the Modern Theory of Convexity. *Science in Context* 22 (01): 85–113.

———. 2012. Uses of History for the Learning of and About Mathematics: Towards a Theoretical Framework for Integrating History of Mathematics in Mathematics Education. In *Proceedings of the International Conference on History and Pedagogy of Mathematics (HPM)*, Invited plenary paper.

———. 2017. An Early Debate in Mathematical Biology and Its Value for Teaching: Rashevsky's 1934 Paper on Cell Division. *The Mathematical Intelligencer* 39 (2): 36–45.

Kjeldsen, T. H. 2019. A Multiple Perspective Approach to History of Mathematics: Mathematical Programming and Rashevsky's Early Development of Mathematical Biology in the Twentieth Century." In Gert Schubring (Ed.) *Interfaces between Mathematical Practices and Mathematical Education*. Springer, p. 143–167.

———. forthcoming-a. From Duality in Mathematical Programming to Fenchel Duality and Convex Analysis: Duality as a Force of Inspiration in the Creation of New Mathematics. In Krömer, R., Haffner, E. & Volkert, K. (eds.). *Duality, an Archetype of Mathematical Thinking* (preliminary title). Birkhäuser Verlag, 27 p. Accepted for publication.

———. forthcoming-b. What Can History Do for the Teaching of Mathematical Modeling in Scientific Contexts? *Proceedings of the 14th International Conference of Mathematics Education* (Invited lectures). To be published.

Kjeldsen, T. H., and J. Carter. 2012. The Growth of Mathematical Knowledge – Introduction of Convex Bodies. *Studies in History and Philosophy of Science* 43 (2): 359–365.

Klee, V. ed. 1963. Convexity. *Proceedings of Symposia in Pure Mathematics*, VII, Providence: American Mathematical Society.

Kragh, H. 1994. *An Introduction to the Historiography of Science*. Cambridge: Cambridge University Press.

Kuhn, H. W., and A. W. Tucker. 1950. Nonlinear Programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 481–492.

Kuukkanen, J.-M. 2012. The Missing Narrativist Turn in the Historiography of Science. *History and Theory* 51 (3): 340–363.

Leifman, L. J. 1990. In *Functional Analysis, Otimization, and Mathematical Economics: A Collection of Papers Dedicated to the Memory of Leonid Vital'evich Kantorovic*, ed. L.J. Leifman. Oxford: Oxford University Press.

Leonard, R. J. 1992. Creating a Context for Game Theory. In *Towards a History of Game Theory*, ed. E. RoyWeintraub, 29–76. Durham/London: Duke University Press.

Massimi, M. 2022. *Perspectival Realism*, Oxford Studies in Philosophy of Science. Oxford/New York: Oxford University Press.

Massimi, M., and C. D. McCoy, eds. 2020. *Understanding Perspectivism. Scientific Challenges and Methodological Prospects*. New York: Routledge.

Minkowski, H. 1891/1911. Über Geometrie der Zahlen. In Minkowski, H. (1911). *Gesammelte Abhandlungen*, 264–265. Volume I. Leipzig/Berlin: B. G. Teubner.

———. 1893/1911. Über Eigenschaften von ganzen Zahlen, die durch räumliche Anschauung erschlossen sind. In Minkowski, H. (1911). *Gesammelte Abhandlungen,* 271–277. Volume I. Leipzig/Berlin: B. G. Teubner.

Mirowski, P. 1991. When Games Grow Deadly Serious: The Military Influence on the Evolution of Game Theory. In *Economics and National Security*, ed. D.G. Goodwin, 227–255. Durham and London: Duke University Press.

Mørch, S. 2010. *Store Forandringer*. Politikens Forlag.

Owens, L. 1989. Mathematicians at War: Warren Weaver and the Applied Mathematics Panel, 1942–1945. In *The History of Modern Mathematics, II: Institutions and Applications*, ed. David Rowe and John McCleary, 287–305. San Diego: Academic.

Rashevsky, N. 1931. Some Theoretical Aspects of the Biological Applications of Physics of Disperse Systems. *Physics* 1: 143–153.

———. 1934. Physico-mathematical Aspects of Cellular Multiplication and Development. *Cold Spring Harbor Symposia on Quantitative Biology* II: 188–198.

———. 1935. Mathematical Biophysics. *Nature* 135: 528–530.

———. 1938. *Mathematical Biophysics: Physicomathematical Foundations of Biology*. Chicago: Chicago University Press.

Rees, M. S. 1977. Mathematics and the Government: The Post-War Years as Augury of the Future. In *The Bicentennial Tribute to American Mathematics, 1776–1976*, ed. D. Tarwater, 101–116. Buffalo, NY: The Mathematical Association of America.

———. 1987. Warren Weaver, July 17, 1894–November 24, 1978. *Biographical Memoirs, National Academy of Sciences of the United States of America* 57: 492–530.

Remmert, V. R., M. R. Schneider, and H. K. Sørensen. 2016. *Historiography of Mathematics in the 19th and 20th Centuries*. Cham: Birkhäuser.

Rheinberger, H.-J. 1997. *Towards a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Standford: Standford University Press.

Richards, J. L. 1995. The History of Mathematics and 'L'esprit humain': A Critical Reappraissal. *Osiris* 10: 122–135.

Rowe, D. 1996. New Trends and Old Images in the History of Mathematics. In *Vita Mathematica. Historical Research and Integration with Teaching*, MAA Notes Series, ed. Ronald Calinger, vol. 4, 3–16. Washington, D.C: Mathematical Association of America.

Schweber, S. S. 1988. The Mutual Embrace of Science and the Military: ONR and the Growth of Physics in the United States after World War II. In *Science, Technology and the Military*, ed. E. Mendelsohn, M.R. Smith, and P. Weingart, 3–45. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Shmailov, M. M. 2016. *Intellectual Pursuits of Nicolas Rashevsky. The Queer Duck of Biology*. Birkhäuser: Springer.

Stedall, J. 2012. *History of Mathematics. A Very Short Introduction*. Vol. 305. Oxford: Oxford University Press.

Toepell, M. 1996. *Mathematiker und Mathematik an der Universität München 500 Jahre Lehre und Forschung*. Munich: Institut für Geschichte der Naturwissenschaften.

Wardhaugh, B. 2010. *How to Read Historical Mathematics*. Princeton/Oxford: Princeton. University Press.

Zachary, P. G. 1997. *Endless Frontier: Vannevar Bush, Engineer of the American Century*. New York: The Free Press.

Chapter 2
Mathematics, History of Mathematics and Poncelet: The Context of the Ecole Polytechnique

Bruno Belhoste and Karine Chemla

Abstract Jean-Victor Poncelet (1788–1867) is known as a geometer whose mathematical contributions were crucial for the development of what would later become projective geometry. In this chapter, we focus on his practice of mathematics, and notably on the fact that Poncelet systematically intertwined mathematical activity with both historical and philosophical reflections about mathematics. Indeed, many practitioners of mathematics at the time, most of whom Poncelet was in contact with, also conducted historical work on mathematics and wrote on the philosophy of mathematics. However, we argue that, in this context, Poncelet's practice of mathematics was unique, being characterized by an intimate interrelation between these three fields of inquiry. Our aim here is more specifically to shed light on the shaping of Poncelet's practice. We suggest that his training at the École Polytechnique, between 1807 and 1810, played an important role in this respect. Our argument unfolds in three main steps. We point out characteristic features of the training of students at the École Polytechnique that, in our view, left a hallmark on Poncelet's mathematical practice. We particularly bring to the fore the importance given to collective work and to reading in students' learning (Part 2). In this respect, two aspects were instrumental: the constitution of a collection of books at the library of the École—which aimed to be a collection of reference for "the arts and the sciences"—and the production of historical works that relied on this collection and were thought to be useful for the learning and the advancement of mathematics (Part 3). Lastly, we focus on the journal established at the École in 1804, *Correspondance pour l'École Polytechnique à l'usage des élèves de cette école*, which reflects several aspects of the life at the École. We argue that the journal gives clues about how, like

B. Belhoste
IHMC, Université Paris 1 Panthéon-Sorbonne-ENS-CNRS, Paris, France

K. Chemla (✉)
CNRS, Laboratoire SPHERE UMR 7219, Université Paris Cité (Campus Grands Moulins) – CNRS, Paris, France
e-mail: chemla@univ-paris-diderot.fr

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_2

teachers, students were encouraged to practice mathematics, in particular geometry, in relation to elements of their history (Part 4). This publication shows how this interest in history meshes with an emphasis on the comparison between methods. It also indicates that former students continued to practice mathematics in this way.

2.1 Introduction: A Specific Mathematical Practice

Of Jean-Victor Poncelet's (1788–1867) mathematical works, several books as well as articles and manuscripts remain. The most famous of the books is undoubtedly the *Traité des propriétés projectives des figures*, whose first edition appeared in 1822 and which arguably marks the beginning of projective geometry.[1] Poncelet continued writing works on the topic until 1833. However, thereafter, absorbed by his occupations as a military officer, he devoted himself almost entirely to mechanics and machines, returning to geometry only shortly before his death.[2] Indeed, in 1862 and 1864, Poncelet, feeling the need to document the mathematical discoveries he had made during the first half of his life, published two volumes that made public manuscripts he had composed before the publication of his 1822 treatise. As its full title makes clear, (Poncelet 1862) contains in particular the text of seven notebooks written in 1813 and 1814, while he was held in Saratoff, after having been taken prisoner during a battle fought in Russia by the Grande Armée.[3] As for (Poncelet 1864), it brings together the manuscripts of articles that Poncelet had published prior to 1822 as well as letters he had exchanged with various protagonists at the time. Poncelet added notes to all these documents, in 1862 and 1864, respectively. Later, taking the opportunity of the publication of a second edition of the *Traité des propriétés projectives des figures,* "revised, corrected and with new annotations" ("revue, corrigée et augmentée d'annotations nouvelles", Poncelet 1865), Poncelet added to the treatise a second volume that likewise made public manuscripts of articles published after 1822 as well as other unpublished documents, to all of which, again, he added annotations (Poncelet 1866). Perhaps even more important from a historical viewpoint, insofar as these sources shed light on the process of maturation that led to the 1822 treatise, (Poncelet 1864) further makes unpublished

[1] (Poncelet 1822). Gray (2005) is devoted to this book. Gray (2007: 11–78) outlines the history of projective geometry in France in the first half of the nineteenth century, focusing in particular on Poncelet's contributions.

[2] (Didion 1869: 49–59) gives a complete bibliography of Poncelet's works that shows this shift in his publications. The reader will also find in this book a biography of Poncelet.

[3] Poncelet joined the Grande Armée on June 17, 1812, and the Russian army took him prisoner on November 18, 1812. (Didion 1869: 109, 116). He arrived in Saratoff in March 1813 and began his first notebook in April 1813 (Poncelet 1862: 1). Irina Gouzévitch and Dimitri Gouzévitch (1998) deal with this period of Poncelet's life and also compare the remaining manuscript of the first notebook with the published text; (Belhoste 1998) offers a survey of Poncelet's mathematical work in these notebooks.

drafts available in which, after his return from Russia to France in September 1814, we see the mathematician honing his reflections on geometry, during the moments of leisure that his life as an officer in Metz left to him.[4]

What is striking in all these documents is what they reveal of the mathematical practice that Poncelet brought into play in his research. In particular, Poncelet's works are full of references to authors of the past, whose works he discusses.[5] Moreover, his mathematical developments are intertwined with philosophical reflections about mathematics, which are mainly related in one way or another to the issue of the relationship between geometry and analysis. These historical and philosophical reflections made a decisive contribution to the development of his theory of ideal elements, on which his entire projective geometry is based.

For sure—and we return to this issue below—, other authors who belong to the same milieu as Poncelet, such as Charles-Julien Brianchon (1783–1864), also interweave their mathematical works with historical developments. Moreover, at the time when Poncelet began to write about geometry, many practitioners were conducting their mathematical work in close connection with philosophical reflections. Suffice it to recall that, as early as 1813, a rubrique entitled "Mathematical Philosophy" appeared in the journal *Annales de Mathématiques Pures et Appliquées* that Joseph Diez Gergonne (1771–1859) launched in 1810. One of the first contributions in this rubrique was by Jacques Frédéric Français (1775–1833), who entered the École Polytechnique in 1797 and began to teach military arts at the École d'Artillerie et du Génie de Metz—precisely the École d'application in which Poncelet studied between 1810 and 1812, after his studies in Paris, at the École Polytechnique.[6] The following year, François-Joseph Servois (1767–1847), who had been one of Poncelet's teachers at the Lycée de Metz in 1806 and was a friend of Gergonne, also published a contribution to this rubrique.[7] What is more, the first article in the *Annales de Mathématiques Pures et Appliquées* whose title contained the cognate expression of "philosophy of mathematics" was published by "the editors," in the

[4] (Poncelet 1864: vi–vii). The war ended in June 1814, which put an end to Poncelet's captivity and allowed him to leave Saratoff and reach Metz in September of the same year (Poncelet 1862: 421, note date 1861; Didion 1869: 122, 124).

[5] This is less true of the Saratoff manuscripts, for reasons to which we return. The archives of the École Polytechnique keep many notes taken by Poncelet, on the basis of his reading of historical mathematical works.

[6] On Jacques Frédéric Français and his brother François (1768–1810), see (Taton 1970–1980a). Poncelet referred to the former as his teacher (Poncelet 1864: 107, 120, 593) and kept in touch with Jacques Frédéric after his return to Metz in 1814. He also commented on Français's 1813 publication (Français, 1813–1814, Poncelet 1864: 592–594, 197). Poncelet showed Français his drafts, and Français lent him precious books from the library of ancient books and manuscripts–notably mathematical–collected by Louis François Antoine Arbogast (1759–1803), which Français inherited. To get a sense of the scope of the library, see (Anonymous 1823).

[7] (Servois 1814–1815). On the relationships between the three men, see (Poncelet 1864: 32). Poncelet asserted that, as early as September 1814, he had communicated with Français about his Saratoff manuscripts, and shortly thereafter with Servois (Poncelet 1822: v–vi; 1864: 469). On Servois, see (Taton 1970–1980b).

second issue dated 1811–1812 (pp. 65–68), and it announced the publication of a book to which Poncelet would later refer: Józef Maria Hoëné de Wroński's (1776–1853) *Introduction to the Philosophy of Mathematics*, which appeared in 1811 (Wroński 1811).

However, we argue that Poncelet's mathematical practice not only articulates these three domains of inquiry—that is, mathematics, the history and the philosophy of mathematics—with each other, but also attests to a much more intimate relationship between them than is demonstrated by these other practitioners.[8] In this respect, Poncelet appears to be carrying out mathematical activity in a specific way, the closest match to it being Michel Chasles's (1793–1880) practice as attested in the 1837 *Aperçu historique sur l'origine et le développement des méthodes en géométrie*.[9]

In this essay, we aim to inquire into the shaping of this research practice. Our main focus will be to argue that the training Poncelet received at the École Polytechnique between 1807 and 1810 played a key role in the development of several facets of his mathematical practice.[10]

Indeed, Poncelet entered the École Polytechnique in November 1807. However, falling seriously ill, in May 1808 he was authorized to return home. When he could go back to Paris, in October 1808, he had to repeat the first year of study (referred to as "second division") at the École, starting the second year ("first division") only in November 1809 and eventually completing his studies there in 1810.[11] Consequently, Poncelet was exposed to the teaching of a greater number of teachers. We will outline here only those who taught him mathematics during these years.

The *instituteur* (teacher) in charge of teaching analysis to first-year students who started their studies in 1807 was Sylvestre-François Lacroix (1765–1843), whose lectures Poncelet thus followed.[12] When Poncelet resumed studying in the first year in 1808, as well as during his second year (1809–1810), the analysis lectures were given by, respectively, André-Marie Ampère (1775–1836), and then Louis Poinsot (1777–1859) (Belhoste 2003: 252). In fact, during his first year at the École, Poncelet had already had Ampère as a *répétiteur* of Lacroix's lectures.[13]

[8] Somehow, Gergonne perceived the close intricacy between mathematics and philosophy that Poncelet's work represents as well as the related depth of Poncelet's contribution (Gergonne 1826–1827). The fact that the text is part of the polemical exchanges between the authors about the priority dispute they had about duality should not overshadow Gergonne's appreciation of Poncelet's contribution. Poncelet (1866: 390–393) quotes this article *in extenso*.

[9] (Chasles 1837). As shown in (Chemla 2016), like Poncelet, Chasles developed a reflection on the reasons why analytical approaches to geometry had the virtues of generality and uniformity, aiming to find means of endowing geometrical approaches with similar virtues. His mathematical and philosophical endeavor was rooted in the historical reflection that he carried out on methods in geometry, starting from ancient Greek geometrical texts.

[10] On the establishment of the École Polytechnique, its early history as well as its international impact, see (Belhoste 2003).

[11] (Belhoste 1998). See also Poncelet's testimony about his illness in (Poncelet 1866: 406).

[12] On this *savant* and his lectures of, as well as his treatises on, analysis, see (Caramalho Domingues 2008).

As for the other main lectures on mathematics, devoted to descriptive geometry and analysis applied to geometry, for the three years, Poncelet's teacher was Jean Nicolas Pierre Hachette (1769–1834), who had assisted the creator of the discipline Gaspard Monge (1748–1818) from the very beginning of the École Polytechnique (Belhoste 1998: 12).

In what follows, we will highlight characteristic features of the training of students at the École Polytechnique that, in our views, left a hallmark on Poncelet's mathematical practice (Part 2). We will emphasize, in particular, the constitution of the collection of books at the library of the École and the production of historical works that relied on it (Part 3). We will then turn to the journal established by Hachette in 1804, *Correspondance pour l'École Polytechnique à l'usage des élèves de cette école*, which reflects several aspects of the life at the École that will prove important for us. We will argue that the journal gives clues about how students were encouraged to practice mathematics, and in particular geometry, in relation to elements of their history (Part 4). In another publication, we will argue that the lectures of analysis given at the École Polytechnique from the beginning might have encouraged a practice of mathematics in which historical and also philosophical reflections played an important role.

Our contribution can be read from two different perspectives. In a sense, it sheds light on the shaping of Poncelet's mathematical practice. Seen from another angle, it uses Poncelet as a witness to highlight the shaping, at the École Polytechnique, of a historico-mathematical culture in which Poncelet participated.

2.2 The Shaping of the Training at the École Polytechnique and the Consequences for the Students

2.2.1 Key Features of the Early École Polytechnique as an Institution

Several documents attest to the fact that Monge played a major role in the conception of the École Polytechnique, which opened its doors at the end of 1794 under the first name of "École Centrale des Travaux Publics" (Belhoste 1994). For our purpose in this chapter, it will be useful to return to some aspects of the organization of the school, as described in the project of founding decree (*projet d'arrêté*)—

[13] On the lecture courses of analysis, see (Belhoste 2003, Chap. 8). (Caramalho Domingues 2008: 403–422) reproduces several documents that shed light on the evolution of Lacroix's lectures. A partial outline of the first year during which he taught (1799–1800, after Lagrange resigned) shows Lacroix relying on his *Traité du calcul différentiel et du calcul intégral* (three large volumes, 1797–1800). Starting in 1805 for the first year and in 1806 for the second year, Lacroix was teaching according to an official programme (pp. 417–420). His actual lectures for the first year (1805–1806) and for the second year (1806–1807) were summarized by an *inspecteur des élèves* so that they could be used by Ampère, when in 1808, the latter started lecturing.

which Monge drafted and entitled "Institution de l'École Nationale des Travaux Publics"[14]—as well as to the evolution of part of the initial project due to the circumstances.

The founding decree placed the École under the orders of a director, assisted by two adjunct-directors. The teaching staff consisted of nine *instituteurs*. Altogether, with a secretary in charge of preparing the minutes of the meetings, they composed the Council, which met every 10 days, with the position of chair rotating among them every month.[15]

Of relevance for our subject, in the original plan, the staff also included *conservateurs* (curators). For mathematics, a curator was in charge of models and drawings. However, more important to us will prove to be the curator in charge of the library, whose official duties were, according to the original idea, to "distribute classic and other books to students as well as to the teachers and to oversee access into the library" (Belhoste 1994: 55, document 4, clause XVI). To begin with, as the secretary's duties fell to him, the librarian was also a member of the Council (Fourcy 1828: 63). This latter feature of the organization remained true for decades, except between *ca* April 1, 1796 and June 19, 1797, when the functions of the secretary were distinguished from those of the librarian. The first two librarians were Pierre Jacotot (1755–1821), who resigned the 30 Germinal year III (April 19, 1795), and his successor François Peyrard (1759–1822), who remained librarian between the 6 Floréal year III (April 25, 1795) and the 1 Frimaire year XIII (November 22, 1804) (Langins 1989). In the period during which the tasks of secretary were not fulfilled by the librarian, they were entrusted to Nicolas Halma (1755–1828).[16] We return below to these three figures. The position of the librarian in the Council bespeaks an importance granted to the library in the school, which is essential for our argument.

The founding decree further instituted the organization of the students, the rhythm of their days as well as the subjects they would study. Students would be divided into groups of 20 (*brigades*), led by a head (*chef de brigade*) chosen among

[14] (Belhoste 1994: 37–58) reproduces this text as document 4 and establishes (pp. 10–13) that Monge completed it on *ca.* July 8, 1794. About 10 days later, a revision of the end of the text of the founding decree was put forward by Jacques-Élie Lamblardie (1747–1797), the then director of the École des ponts et chaussées (see document 6, pp. 66–70). Our page numbers for this article refer to the online version posted here: https://www.sabix.org/bulletin/b11/belhoste.pdf (accessed August 15, 2022). The final decree was promulgated on the 6 Frimaire year 3 (November 26, 1794). In what follows, we rely on these documents. For a more nuanced account, we refer the reader to this article.

[15] At the time when Poncelet was a student, the Council had been divided into three different Councils. The institution corresponding to a similar composition as the initial Council was the Instruction Council (*Conseil d'Instruction*), responsible for "everything related to the teaching and the students' study" (Belhoste 2003: 50).

[16] For the composition of the Council, see (Belhoste 2003: 52). Halma signed minutes of the Council as early as 12 Germinal year 4 (April 1, 1796) (Dooley n.d., vol.1: 189–190). (Dooley 1994) publishes extracts of the minutes of the Council. However, we draw on (Dooley n.d.) in two volumes, which reproduce the minutes extensively. (Fourcy 1828: 94, 125–126), respectively, outlines the function of the secretary and explains the circumstances and the date when both positions were merged.

the 20 oldest students of the whole promotion. To each of these groups, a specific study room would be allocated. This organization of the space corresponded to a conception of the learning process. To begin with, plenary lectures would be given to students at precise hours of the day. Immediately afterwards, specific times of the day would be set aside for the brigades to meet in their respective rooms and repeat the lessons just learnt, in a collective and active way, that is, by putting into practice the operations taught during the general course. In other words, a form of mutual teaching was implemented, the collective organization encouraging students to form small groups to engage with the subjects taught and to go beyond what they had seen in class. Poncelet (1862: 457) provides evidence of how he and the students who belonged to the same group in study room number 6 discussed better ways of solving questions and of making drawings in descriptive geometry than those published by the teacher.[17] Below, we return to the type of work Poncelet conducted with other students.

As for the topics, with respect to mathematics, the founding decree instituted descriptive geometry and analysis as the pillars of the training. The original plan distinguished between teaching the methods for an initial 3-month period and working on their application in the subsequent years. In the first 3 months, a complete overview of the branches taught during the curriculum would be dispensed, analysis and descriptive geometry being given in separate sets of lectures. Monge had thought of Joseph-Louis Lagrange (1736–1813) for this analysis course.[18] After this 3-month period, analysis would only be taught as applied to geometry and mechanics, whereas descriptive geometry would be applied to various domains of the arts. The impossibility of applying this scheme in practice quickly led to a rethinking of the program. An elementary lecture course in algebraic analysis was created while, for the first 5 years, Lagrange actually taught lessons in analysis, but as a complementary topic, during the holidays, and precisely in the library. Indeed, his lectures were optional, being in fact mainly followed by more advanced students and other teachers.[19] What matters for us is that this teaching would lead to the publication of two books, *Théorie des Fonctions Analytiques* (1797) and *Leçons sur le Calcul des Fonctions* (1801), which would have a key influence on the subsequent shaping of the course of analysis at the École Polytechnique, when the topic was given more weight in the curriculum. On the other hand, as early as 1795, the lectures devoted to analysis applied to geometry were placed in the hands of those teaching descriptive geometry, with a significant consequence: students were thereby confronted with the differences between the geometrical and the analytical ways of dealing with the same topics—a theme dear to Monge's heart, as shown from his lectures at the École Normale de l'an III (Belhoste and Taton

[17] See also (Poncelet 1862: 460–461).

[18] What follows draws on (Belhoste 2003: 235–247). For a more detailed analysis, see also (Wang Xiaofei 2017, 2020).

[19] (Belhoste 2003: 237). See Jean-Guillaume Garnier's (1766–1840) testimony, in an autobiographical manuscript, published by Adophe Quetelet (1839: 171)

1992). The centrality of this issue for Poncelet's reflections on geometry is difficult to underestimate.

2.2.2 *Focus on the Management of the Library*

As we have just mentioned, the presence of the librarian as a member of the Council suggests the significance attached to the library. In fact, many more clues indicate that in the original organization of the school, the library was given special attention.

Already in the first text outlining the organization of the new school in 1794 (document 2, Belhoste 1994: 29–30), the suggestion was made that a specific space be devoted to the library. This was an innovation in comparison with the École des ponts et chaussées,[20] however, not one that would be specific to the École Polytechnique: after the Revolution, the École des Mines would adopt a similar institution, creating a library to supply teachers and students with the books needed for the training.[21] The same holds true more generally for the *Écoles Centrales*—the new type of school that replaced the former colleges of the Old Regime (schools of secondary level).

The importance that the library had for the Council can be seen from the following episode: In April 1796, when some of the workers constructing the school buildings stopped work on account of lack of funding, the Council decided that the remaining workers would concentrate on the three rooms for drawing and the library.[22]

Issues related to the library were brought to the Council, and the minutes of its meetings highlight several facets of the use of the library. In March 1796, the *chefs de brigades* put forward a request to the Council, asking "to extend the days when the library was open".[23] In response, in a new regulation approved by the Council on November 18, 1798, new opening hours for the library were set.[24] In the meantime, in July 1796, the Council allowed those among the *chefs de brigade* deserving the privilege to borrow books at 9 am—that is, outside the opening hours—, if and only if they gave them back before 2 pm so that all students could use them.[25] The testimony of the Danish astronomer Bugge who visited the École Polytechnique

[20] (Laboulais 2014: 54, n. 16), based on Nathalie Montel, *Une revue des savoirs d'État. De la genèse à la fabrique des* Annales des ponts et chaussées au XIXe siècle, Habilitation, 2008, p. 534.

[21] On the process of the constitution of the collection and the various uses of the library thereby constituted that actors developed, see (Laboulais 2014).

[22] In what follows, we will refer to the minutes of the Council meetings using the transcription carried out in (Dooley n.d.). Here, see the minutes of 28 Germinal year 4 (April 17, 1796), in (Dooley n.d.: 73).

[23] Minutes of the 22 Ventôse year 4 (March 12, 1796), in (Dooley n.d.: 187).

[24] (Dooley n.d.: 238).

[25] See the minutes of the 2 Thermidor year 4 (July 30, 1796), in (Dooley n.d.: 200).

in 1798 further suggests that many students consulted books in the library.[26] The borrowing of books appears to have been active, since the Council regularly required that all books be returned.[27] What is more, in 1797, we see the Council inviting readers to borrow books in small quantities and to avoid those that were of daily use in the library.[28]

We might interpret this extensive use of the library as an application of a general pedagogical idea, put forward during the French Revolution. It consisted in replacing the students' copying of the master's courses by their reading textbooks—which motivated the establishment of school libraries. This project was the starting point of the École Normale de l'an III. Monge adopted this idea for the École Polytechnique. This appears clearly in a statement he put forward with respect to how books should be employed in the training provided at the École Polytechnique. On the 20 Pluviôse year 3 (February 8, 1795), he asserted during a Council meeting:

> We should make [students] copy no manuscript that can be printed or handed out. However, the students may be required to carry out, and even carry out correctly what the book teaches him to do, and the use he has made of the book may be judged by the manner in which he operates. The same holds true of certain drawings . . . .[29]

One of the students who entered the École Polytechnique in 1801, Georges de Chambray (1783–1848), has left us an interesting testimony of some of the ways students used books. For him, oral lessons in mathematics were useless—by which he clearly meant lessons in analysis rather than in descriptive geometry. In his view, for this topic knowledge was thus acquired through an extensive use of books, sometimes requiring the help of the *chef de brigade* or discussions with other students of his *brigade* (Chambray 1836: 45–47). Chambray went as far as stating this:

> However, the distinguished students of the École Polytechnique derived a particular advantage from this excess of work that they had given themselves when studying transcendental mathematics, and that was to have acquired the conviction that they could learn almost anything without a teacher, with the help of books alone.[30]

Chambray entered the school a couple of years after Lagrange had finished teaching there, and yet he still remembered hearing that Lagrange's lectures were

---

[26] Quoted in Bradley 1976: 166).

[27] See, for instance, the minutes for the 8 Nivôse year 5 (December 28, 1796) and of the 8 Nivôse year 6, in (Dooley n.d.: 100, 112–113), respectively.

[28] Minutes of the 8 Nivôse year 6 (December 28, 1797), in (Dooley n.d.: 113).

[29] "Il ne faut faire copier aucun manuscrit que l'on peut faire imprimer ou distribuer; mais on peut éxiger (sic) que l'Elève éxécute (sic), et éxécute (sic) bien ce que le Livre apprend à faire, et l'on juge de l'usage qu'il a pu faire du Livre par la manière dont il opère. Il en est de même de certains [dessins] . . . " (Dooley: 148)

[30] "Toutefois, les élèves distingués de l'École Polytechnique retiraient un avantage particulier de ce travail excessif, auquel ils s'étaient livrés en étudiant les mathématiques transcendantes, c'était d'avoir acquis la persuasion qu'ils pouvaient apprendre presque tout sans maître, avec le seul secours des livres." (Chambray 1836: 54).

meant for the "advancement of science," being (as a result, we would add) "not compulsory" (Chambray 1836: 22–23). The books which Chambray needed to understand the lectures on analysis that were part of the regular program were textbooks. Poncelet leaves a strikingly different and yet similar testimony about the way he uses books while he was a student at the École Polytechnique. Indeed, he recalls how together with "M. Guillebon, an excellent friend, an honest and innocent mind under a sickly shell (a weak body)," who belonged to the same study room as he did, he "worked, during recreation hours, reading Lagrange's *Calcul des Fonctions* and *Mécanique Analytique*, and so on."[31] In his engagement with books, like Chambray, Poncelet thus adopted the method of collective work encouraged by the institution. However, instead of replacing oral lessons with books, he looked to research books—and precisely Lagrange's books—for further development. Poncelet could most probably have found these books in the library.[32]

The thesis that, at the École Polytechnique, Poncelet made, more broadly, an extensive use of the library is supported by several pieces of evidence in his writings. To begin with, in his foreword to the second volume of *Applications d'analyse et de géométrie*, we read

> It was above all during the winters of (. . .) 1815 to 1820 (. . .) that I was finally able to resume the course of my geometrical ideas and *read the scientific books and journals of which I had been entirely deprived since I left the École Polytechnique*.[33]

Moreover, when, in the first volume, he gives an account of how he conducted his research at Saratoff, Poncelet makes his interest in scientific libraries explicit and further gives important information about what he read before the Russian campaign, notably at the École Polytechnique. His account, in which he speaks of himself in the third person, reads as follows:

> . . . he had to painstakingly redo, so to speak, one by one, the elements indispensable to mathematical studies, *deprived as he was of any book*, of any precision instrument, both of which were difficult to obtain in this city of Saratoff, which was, moreover, *devoid of scientific libraries* at the time. One should *not* therefore *expect to encounter here the reflection or the distant echo* of the *profound analytical works* of scholars such as Euler, Bernoulli, Huygens, Newton, d'Alembert, etc., nor even of the more recent and no less admirable works of [scientists] such as Lagrange, Legendre, Laplace, Monge and their disciples—*works that had left no trace in his memory* in the midst of the perils and anguish of such an unfortunate beginning to his military career.

---

[31] " . . . M. Guillebon, excellent ami, esprit droit et naïf sous une chétive enveloppe, avec lequel je piochais, pendant les heures de récréation, le *Calcul des fonctions* et la *Mécanique analytique* de Lagrange, etc. . . . " (Poncelet 1862: 457).

[32] (Bradley 1976:173–175 and related footnotes) suggests that in 1807, the École expected students to own their books and that it bought books in large numbers so that it could sell them to students at a reduced price. She reproduces lists of books thus bought. We find the *Mécanique Analytique*, but no other book by Lagrange.

[33] (Poncelet 1864: VII), our emphasis. "Ce fut surtout pendant les hivers (. . .) 1815 à 1820 (. . .) que je pus enfin reprendre le cours de mes idées géométriques et prendre connaissance *des livres et des journaux scientifiques dont j'avais été entièrement privé depuis ma sortie de l'École Polytechnique*."

> The Author, who left the École Polytechnique in November 1810, (...) had little thought, in such an agitated life, of devoting himself to the abstract sciences. Reduced to his recollections of the Lycée de Metz and *the École Polytechnique*, *where he had cultivated with predilection the works of Monge, Carnot and Brianchon*, it must be acknowledged that he was unable to borrow anything from the last writings published before his return to France in September 1814.[34]

In the notes added to his 1862 book—for instance in (Poncelet 1862: 143, 311)—, Poncelet returns to the Saratoff period to express regret for the lack of correct acknowledgement of his predecessors' works in his notebooks. He associates this problem with the conditions under which he was working at the time and refers to his 1822 treatise, in which he considers that he has settled the issue. This sheds light on one of the uses Poncelet associated with a library.

These observations invite us to examine the sort of book that Poncelet and the other students could find in the library.

2.3 Collecting Books for the Library and Working on the History of Mathematics

2.3.1 The Constitution of the Library's Book Collection

In Monge's draft for the founding decree, one of the essential duties of the school Council was to examine the books that would be used for the students' learning (Belhoste 1994: 38, document 4, clause 25). Lamblardie's revised write-up repeats this clause with slight modifications (Belhoste 1994: 70, document 6, clause 24). However, it adds emphasis on the Council's responsibility with respect to books in a way that widens the use intended for them. Indeed, in the enumeration that this time lists the school Council's *primary* duties, we read that one of its key attributions would be precisely "to choose the books and models of all kinds that could contribute most efficiently to the [students'] *advancement* and thus to order

[34] (Poncelet 1862: IX), our emphasis. "... il dut refaire péniblement, et pour ainsi dire un à un, les éléments indispensables aux études mathématiques, *privé qu'il était de tout livre*, de tout instrument de précision, difficiles à se procurer dans cette ville de Saratoff, d'ailleurs *dépourvue alors de bibliothèques scientifiques*. On ne doit donc *pas s'attendre à rencontrer ici comme le reflet ou l'écho lointain des profonds travaux analytiques* des Euler, des Bernoulli, des Huygens, des Newton, des d'Alembert, etc., ni même des travaux plus récents et non moins admirables des Lagrange, des Legendre, des Laplace, des Monge et de leurs disciples, *travaux qui n'avaient laissé aucune trace dans sa mémoire* au milieu des périls et des angoisses d'un aussi malheureux début dans la carrière des armes. L'Auteur, sorti en novembre 1810 de l'École polytechnique, (...) ne songeait guère, dans une vie aussi agitée, à s'occuper des sciences abstraites. Réduit à ses souvenirs du lycée de Metz et de *l'École polytechnique, où il avait cultivé avec prédilection les ouvrages de Monge, de Carnot et de Brianchon*, on doit reconnaître qu'il n'a rien pu emprunter aux derniers écrits publiés avant sa rentrée en France, en septembre 1814." In effect, the references that we find in the Saratoff notebooks are mainly to Monge, Carnot and Brianchon.

them or to approve them."[35] These documents confirm the emphasis the institution placed on the library.

In application of the latter regulation, the minutes of the Council meetings regularly mention discussions about books and the opportunity to acquire them for the library. On 8 Frimaire year 4 (December 9, 1795), the Council approved the acquisition of a "list of several elementary books" put forward by the curator of the library Peyrard.[36] It also validated the funding that Peyrard suggested using to this effect: In 1795, the school established the *Journal de l'Ecole Polytechnique*, whose sales generated profits that could be employed for the purchase of books. On 12 Frimaire year 4 (December 3, 1795), at the request of its author, the Council nominated two referees to examine a textbook for artillery by Durtubie, and twenty days later, on 2 Nivôse year 4 (December 23, 1795), the referees presented a report, in which they emphasized the book's appropriateness for instruction and suggested "sending it to the competition for elementary books."[37]

Books given as gifts likewise had to receive the approbation of the Council. For example, on 8 Vendémiaire year 5 (September 29, 1796), the Council accepted Lesage's gift of his brochure on elasticity, asking that "it be deposited in the library."[38] Similarly, it expressed its gratitude when, on 18 Brumaire year 7 (November 8, 1798) and on 12 Floréal year 6 (May 1st, 1798), the curator Peyrard and the teacher Labey presented to the library, respectively, "a new edition of Bezout's algebra, with additions" and "the part 2 of [Labey's] translation of Euler's *Differential Calculus*."[39] On the second complementary day of year 3 (September 18, 1795), Monge presented a report on a miscellanea devoted to stereotomy, on the basis of which "the Council decided that it be deposited in the library."[40] Even La Peyrouse's travel book that was sent by the Ministry of the Navy for the school library went through the Council, on 8 Vendémiaire year 5 (September 29, 1796).

[35] Our emphasis. "... sur le choix des ouvrages et des modèles en tout genre qui pourront contribuer le plus efficacement à leur *avancement* et qui seront à ordonner ou à approuver." (Belhoste 1994, document 6: 67, clause 4).

[36] (Dooley n.d.: 51).

[37] (Dooley n.d.: 175, 177, respectively). The book is Théodore D'Urtubie's (1741–1807) *Le manuel de l'artilleur,* which had many editions—it has proved difficult to determine the date of the first edition, but the second dates from 1787.

[38] (Dooley n.d.: 95). This is probably a work by Georges-Louis Lesage (1724–1803), however we could not identify the actual title.

[39] (Dooley n.d.: 233, 220, respectively). The former book is part of *Cours de mathématiques, à l'usage de la marine et de l'artillerie, par Bézout. Édition originale, revue et augmentée par Peyrard et renfermant toutes les connaissances mathématiques nécessaires pour l'admission à l'*École Polytechnique. Its publication was carried out "chez Louis" in Paris, in 1798. The latter work is the second volume of Jean-Baptiste Labey's (1752–1825) *Introduction à l'analyse infinitésimale par Léonard Euler, traduite du latin en français, avec des Notes et des Éclaircissements par J. B. Labey*, Professeur de Mathématiques aux Écoles Centrales du Département de la Seine. It was published "chez Barrois aîné," in Paris and in 1796.

[40] (Dooley n.d.: 166).

On the same day, the Council further decided on the acquisition from its bookseller Bernard of a book on acoustics,[41] and later "for the library, of the latest volumes from the Academy of Sciences."[42] Similarly, on 18 Frimaire year 7 (December 8, 1798), the purchase of other journals was again decided for the library.[43] In the Spring of 1796, the Council projected other costly acquisitions, such as Euler's complete works, and on three occasions discussed how to secure this purchase from a financial viewpoint.[44] An inventory of the collection and lists of books that were lacking were a part of reports that the Council would send to the Directoire as argument justifying a required budget.[45]

On 12 Thermidor year 6 (July 30, 1797) the procurement of books had extended to Adrien-Marie Legendre's (1752–1833) book "on the properties of numbers," Lagrange's "forthcoming memoir," "Bossut's Calcul différentiel et integral" and *Le langage du calcul* by Étienne Bonnot de Condillac (1714–1780).[46] On 28 Nivôse year 7 (January 17, 1799), a long list of books was approved.[47] The fact that it includes, in addition to textbooks in multiple copies and journals, Condorcet's *Manière de compter*,[48] Hougthon's *Voyage en Afrique* and Linné's *Œuvres* indicates that the library was no longer thought of as simply a support for the students' learning.

This is a conclusion on the evolution of the library that Luigi Pepe (1996) has drawn from an analysis of other channels—different from the commercial channels discussed above—through which books were sent to the library. Indeed, the bulk of the books that entered into the library had wholly different origins.

When the École Polytechnique opened at the end of 1794, its library was for its main part constituted with a large number of books received from the library of the École du génie de Mézières, to which the revolutionary *Comité de Salut*

---

[41] (Dooley n.d.: 220). Similarly, on 8 Pluviôse year 5 (January 27, 1797), the Council decided the acquisition of "a copy of Brissot's physics" (p. 101), of "a copy of [Baltard's] book," after the latter had presented his *Études à l'usage des ingénieurs civils, militaires et géographes* (p. 231). On 8 Brumaire year 7 (October 29, 1798), the Council also decided the acquisition of *Le tir de l'artillerie* by Lombard d'Auxonne (we do not know to which title exactly this book refers), of Durtubie's book on artillery (that is, the artillery textbook examined in 1795), as well as of Gassendi's *Aide-Mémoire d'artillerie* (p. 233) This list is not exhaustive. On Bernard as the approved bookseller of the École, see (Bradley 1976: 173–174).

[42] (Dooley n.d.: 111).

[43] (Dooley n.d.: 242). On the periodicals in the library, see (Bradley 1976: 167, 169–170).

[44] 8 Floréal year 4 (April 27, 1796), as well as 12 Floréal year 4 (May 1, 1796) and 22 Floréal year 4 (May 11, 1796), see (Dooley n.d.: 73, 193).

[45] See the minutes of the 2 Ventôse year 4 (January 21, 1796), in (Dooley n.d.: 184).

[46] (Dooley n.d.: 127).

[47] (Dooley n.d.: 256–257).

[48] This title refers to a book by Marie Jean Antoine Nicolas de Caritat, marquis de Condorcet (1743–1794), posthumously published in 1798 under the title *Moyens d'apprendre à compter avec facilité* (Paris: Moutardier, 1798).

Public had put an end.[49] The first librarian Pierre Jacotot, who gave lectures on mathematics in the first months of the Ecole and would resign from his position of librarian in April 1795, established the first inventory on 2 Nivôse year III (January 19, 1795).[50] At the time, the library amounted to 564 volumes, 76 of which were listed under mathematics.[51] Among the latter, we find eighteenth-century textbooks on mathematics and engineering. The books also included some older titles by Leonhard Euler (1707–1783), Daniel Bernoulli (1700–1782), and Pierre Varignon (1654–1722). Of note are Isaac Newton's (1642–1727) *Philosophie naturelle*,[52] Gottfried-Wilhelm Leibniz' (1646–1716) *Oeuvres philosophiques*, and Condorcet's *Essai d'analyse.* Interestingly, the oldest volume of all was the *Examen des oeuvres* that the inventory attributes to Girard Desargues (1591–1661), but that refers to Jacques Curabelle's 1644 diatribe against the Lyon mathematician's work on stonecutting and other methods.[53] The other main rubriques were physics, chemistry, hydraulics and military architecture, as well as civil architecture, history and travel, to which we must add the *Encyclopédie* and memoirs published by various Academies.

The Revolution led to the confiscation of books from *émigrés* as well as from ecclesiastic institutions and former scientific societies. These books were kept in depots to which the librarians of the École Polytechnique went to select books for the library (Fourcy 1828: 16–17; Bradley 1976: 166–168). This represented another essential channel through which the collection was constituted. Jacotot started this work, which was actively continued by his successor François Peyrard.[54]

---

49 (Anonymous 1892) outlines the history of the various types of acquisitions carried out for the library at different stages and provides useful documents.

50 On Jacotot's biography, see (Barbier 1999). On his actions as a librarian, and also the reproduction of his inventory, see (Pepe 1996: 157–160 and 174–178 (appendix 1)), whose analysis we summarize here. We are grateful to Enrico Giusti for having helped us gain full access to this article. However, (Pepe 1996) has the incorrect name of Jacolot. See also (Langins 1989).

51 The figure seems to refer to physical volumes and not to titles. We cannot analyze the classification. However, what counts as mathematics in this inventory awaits discussion. (Pepe 1996: 158) establishes that of these 76 mathematical volumes, 65 came from Mézières.

52 In other words, Newton's *Philosophiae Naturalis Principia Mathematica*, first published in London in 1687.

53 (Poncelet 1866: 312; 1865: 409) refer to this book by Curabelle (1585–16?) using in both cases the word "diatribe." In the latter reference, Poncelet mentions reading this book in detail and keeping the manuscript of excerpts he had copied. For a translation and an analysis of Desargues's work on conics, see (Field and Gray 1987).

54 For Peyrard's biography, and his actions as a librarian, we draw on (Langins 1989) as well as on (Pepe 1996). (Pepe 1996: 163–166) deals with Peyrard as a librarian. (Pepe 1996: 180–185) reproduces an extract—concentrating on mathematical books — of the library's inventory completed by Peyrard on 27 Germinal year 4 (April 16, 1796). (Anonymous 1892: 128–129, fn 2) outlines a biography of Peyrard and gives a substantial bibliography of his works. (Anonymous (P. L. B.) 1893) offers a biography that reproduces important documents on François Peyrard, in particular the public lecture courses in mathematics that Peyrard gave at the end of the 1780s. We could not identify the author to whom the initials P.L.B. refer.

Peyrard, who originated from Haute-Loire and had been trained in the humanities and philosophy, enrolled in the "Gardes Françaises" to go to Paris. After having followed public lectures and taught mathematics—notably in public lecture courses—, he took an active role in the revolution, becoming in particular in 1793 a member of the Committee of Public Instruction of the Department of Paris, in which Monge and Lagrange were also members. In April 1795, Peyrard was nominated librarian at the École Polytechnique, a position from which he would be expelled in 1804, becoming thereafter a teacher of mathematics and astronomy at the Lycée Bonaparte. His action as a librarian at the École Polytechnique is generally considered to have radically changed the size and the nature of the library.

The assessment is clearly correct with respect to the size: Indeed, the catalogue that Peyrard established one year after having become librarian already lists 3400 books, among which 270 volumes belong to "mathematics and mechanics". His 1801 inventory, classified by author, refers to 7555 books (Bradley 1976: 169), and Langins (1989) mentions an evaluation that the library held 10,000 books when Peyrard left his position. Peyrard continued looking for books in the depots. However, the library also benefited from requisitions carried out by French forces in conquered countries. After the annexation of Belgium to the French Republic, a set of eighty books on mathematics and architecture was sent from Belgian collections to the library of the École Polytechnique (Pepe 1996: 159–163). The list of these books, established by Peyrard on October 11, 1795, is reproduced in (Pepe 1996: 178–180).

As Pepe rightly emphasizes, the list points out a key shift in nature with respect to the books that had been catalogued by Jacotot, since it includes many sixteenth- and seventeenth- century mathematical treatises of international provenance. Among these books, let us point out the 1703 edition of Euclid's complete works, published by David Gregory (1659–1708) in Oxford; the 1710 edition of Apollonius' *Conics*, also published in Oxford by Edmund Halley (1656–1742); and, last but not least, the 1621 edition of Diophantos' *Arithmétiques*, by Claude-Gaspard Bachet de Méziriac (1581–1638). In fact, the catalogue produced by Peyrard in April 1796 shows that the librarian had worked along the same lines in his choice of books from the depots. Indeed, the library now further included other editions of Apollonius' *Conics* and of Euclid's works, as well as, for Archimedes' works, the 1544 *editio princeps* published in Basel and the 1615 Paris edition. It also contained a large number of classic scientific works of the early modern period, and in particular the works of authors such as "Euler, Bernoulli, Huygens, Newton, d'Alembert", as well as "Lagrange, Legendre, Laplace, Monge," which, as we have seen above, Poncelet mentions in the opening pages of *Applications d'Analyse et de Géométrie* (see footnote 34). For Pepe (1996:163, 166), these acquisitions are more broadly what turned a library that would serve as a support for teaching into an institution that would become a library of reference for scientific works and beyond, since Peyrard did not limit himself in his choice of books.[55]

[55] See (Langins 1989) and (Pepe 1996: 165).

This evaluation is consistent with Peyrard's perception of the book collection he had established. Indeed, in the autobiographical defense signed by Peyrard on September 23, 1804, when he was threatened with being fired from the École—a memoir on which (Langins 1989) mainly draws—, Peyrard asserts about the time before he took the position: "it is true to say that there was not yet the beginning of the library that exists today and that contains *almost all the good books, both ancient and modern, dealing with the sciences and the arts*."[56] Quite significantly for our purpose, a few months earlier, on 28 Frimaire year 11 (December 19, 1803), Peyrard had written the following to Fourcroy: "All the works in the Ecole Polytechnique library are classified according to format, contents and chronological order. I should like to see an inventory in conformity with this classification . . . This would provide a *catalogue by order of contents and by chronological order of the best books*, and would subsequently *provide a brief summary of human knowledge*."[57] The expression of his wish testifies to two essential points. First, in his view, the physical arrangement of the books on the shelves was historically meaningful. Moreover, the library was rich enough to offer in abridged form an overview of human knowledge.

As early as 1798, the Danish astronomer Thomas Bugge (1740–1815) had also noted the value of the library for scholarship in the domains taught at the École (Bradley 1976: 166). In the same year, as a member of the commission who had been dispatched to Italy in 1796 to collect books and art works for French libraries and museums, Monge sent to the library of the École about 200 precious and valuable works that Napoleon's victories had enabled him to acquire for France (Pepe 1996: 166–173, 1997).[58] On January 26, 1801, Peyrard signed an inventory of books from Italy—which is reproduced in (Pepe 1996: 192–195). These works would further expand the scope of the library by enriching it with rare and precious books. However, Monge's work in Italy would not simply concern the library of the École Polytechnique. Most importantly for us, in his capacity as the officer in charge of the requisition of books and artworks, between the beginning of 1796 and July 1797, Monge was also in charge of choosing manuscripts from the Vatican library and of sending them to the French national library. He selected, in particular, several codices of great significance for the history of science. Among them, let us mention the *Vaticanus Graecus* 190, which contained an important ancient edition of

[56] (our emphasis) "il est vrai de dire que l'on n'avait pas encore le commencement de la bibliothèque qui existe aujourd'hui et qui renferme *presque tous les bons livres anciens et modernes relatifs aux sciences et aux arts*." Quoted in (Langins 1989).

[57] Bradley's translation (Bradley 1976: 169) and our emphasis. The original is quoted in note 21, p. 177: "Tous les ouvrages de la Bibliothèque de l'Ecole Polytechnique sont classés par ordre de format, par ordre de matière et par ordre chronologique. Je désirerais qu'on en fit (sic) dresser un inventaire conforme cette classification... un catalogue par *ordre de matière et par ordre chronologique des meilleurs ouvrages*... offrirait par conséquent *le tableau abrégé des connaissances humaines.*"

[58] On 28 Vendémiaire year 7 (October 19, 1798), the minutes of the Council meeting report that books and objects coming from Italy were given to the school and distributed to the appropriate services in the École, including the library.

Euclid's *Elements* that would contribute to Peyrard's fame as well as the *Vaticanus Graecus* 184 and 1038—which included editions of Ptolemy's *Almagest* that would soon be useful for Halma (Pepe 1996: 169).

These last remarks raise a key question: how can we better appreciate the impact of a library of this kind at the École on both the staff and the students, and, in particular, on Poncelet?

2.3.2 *Editing Anew and Translating Greek Works of Antiquity for the Present*

Langins (1989) provides several pieces of evidence showing that, from the beginning of the nineteenth century at the latest, for the staff as well as for other *savants*, the library of the École Polytechnique was a place where they came to carry out research. At least in hindsight, Poncelet shared the idea that the library was rich enough to allow one to deal with any question extensively. Writing, in 1862, about a memoir Jules de la Gournerie (1814–1883)—professor of descriptive geometry at the École Polytechnique—had published in 1851 in the school journal, Poncelet wrote:

> It is regrettable that this skillful teacher did not take advantage of his position at our alma mater to elucidate the historical aspect of this interesting question.[59]

However, the first person who drew on the library's collection for his works was none other than Peyrard himself. In 1804, Peyrard published a translation into French entitled *Les élémens de géométrie d'Euclide, traduits littéralement*...[60] The title page, which refers to Peyrard as the librarian of the École Polytechnique, does not mention that the translation contains only books I, II, III, IV, VI, XI, XII—a point that the table of contents makes explicit. Only in the second edition (Peyrard 1809) does Peyrard—now referred to as "professeur de mathématiques et d'astronomie au Lycée Bonaparte"—announce the additional translation of book V, thereby still leaving aside Euclid's arithmetical books.

The preface to the first edition emphasizes the relationship between the project of translating ancient Greek mathematical works and the environment of the École Polytechnique. It opens as follows:

> When I was appointed Librarian of the École Polytechnique, I formed the project of giving to the public a literal translation of the works of Euclid and Archimedes, the two greatest

[59] (Poncelet 1862: 457). "Il est regrettable que cet habile professeur n'ait pas mis à profit sa position à notre mère École, pour élucider la partie historique de cette intéressante question."

[60] (Peyrard 1804). We return to this edition and the following ones below. See (Aujac 1990, 2007, Xiaofei 2017), on which we draw for the description of the editions.

> geometers of antiquity. I thought that it was in some way my duty to *devote my spare time to works that were similar to those of the École Polytechnique*.[61]

All such references were removed from the preface to the second edition of 1809—a preface written a few years after Peyrard's dismissal from the Ecole and which is much more historical in nature, featuring a different approach to history than the first. In the context of our chapter, it is important to emphasize that the edition of Euclid's works on which Peyrard drew for his 1804 translation was the one published in Oxford by Gregory in 1703 (Aujac 1990: 396). As we have seen above, this was precisely an edition that was among the books sent from Belgium that Peyrard completed listing on October 11, 1795. This remark sheds light on the connection between the library and the research Peyrard was carrying out.

In the same 1804 preface, Peyrard announced for the following year the publication of a "literal translation of Archimedes' complete works" and indicated that a subscription would be open until 1 Vendémiaire year XIII. Money for this should be sent to him at the École Polytechnique or to the publisher F. Louis. Monge and Gaspard de Prony (1755–1839) subscribed to this Archimedes volume (Langins 1989). The book in question was in fact only published in 1807, without a list of subscribers and with another publisher.[62] In 1804, Peyrard explained that, Archimedes' works not being "elementary enough", he had instead decided to compose a "supplement," in which he dealt with the same topics as Archimedes, following "Archimedes' principles" (Peyrard 1804: x–xi). The supplement was included in the 1804 volume of Euclid's *Elements*.

2.3.3 The Viewpoint of the School Council on the Editions and Translations of Greek Works of Antiquity

Interestingly, Langins (1989) notes the following paragraph in the minutes of the Council meeting of 19 Prairial year XII (June 8, 1804):

> The deliberation on purchasing a copy of the Oxford edition of Archimedes for the library (an action suggested by Peyrard) gave several members the opportunity to discuss the advantages that would result for the *students' benefit*, for the *advancement of the exact sciences* and for the *glory of the École*, if, following a regular and well-designed plan, it succeeded in obtaining *translations of the ancient Greek and Latin authors* in which the *teachers* of the *École* would have *worked together with the most learned men of letters*

[61] (Peyrard 1804: ix), our emphasis. "Lorsque je fus nommé Bibliothécaire de l'Ecole Polytechnique, je formai le projet de donner au public une traduction littérale des Œuvres d'Euclide et d'Archimède, les deux plus grands Géomètres de l'antiquité. Je pensois qu'il étoit en quelque sorte de mon devoir de *consacrer mes momens de loisir à des travaux qui fussent analogues à ceux de l'Ecole Polytechnique.*"

[62] (Peyrard 1807). Joseph Delambre contributed a memoir to this edition (see below). The second edition (Peyrard 1808) was revised by Delambre (Aujac 1990: 396; Aujac 2007: 234).

> to *establish the genuine text*, and to *render it faithfully* by *adding* to it the necessary or interesting *notes* that the most direct study of these authors would provide them.[63]

This paragraph contains multiple pieces of evidence that are important for our argument.

First, Peyrard suggests that the Oxford edition of Archimedes' works be purchased—a decision approved by the Council. Interestingly, Peyrard's foreword to the 1807 translation of Archimedes' works—which is comparable in its historical approach with the preface to his second edition of Euclid's *Elements*—ends with an overview of the translations and commentaries published since the sixteenth century (Peyrard 1807: xxxv–xxxvii). This overview mentions the 1544 Basel edition—which Peyrard dates from 1545—as well as the 1615 Paris edition: we have indicated above that both editions were already available in the library of the *École*. The overview is concluded by a reference to the 1792 edition published by Giuseppe Torelli (1721–1781) in Oxford—precisely the one whose purchase is under consideration—, Peyrard praising the elegance and the faithfulness of its Latin translation as well as its critical notes, to which his text refers.[64] It would be interesting to clarify how Peyrard had access to this edition after he left the École Polytechnique.

Second, the deliberation about the purchase leads the Council to broaden the discussion in a way that shows the members establishing a link between the publication of such books and the work students and staff carry out at the École Polytechnique. To their eyes, translations of the kind Peyrard had begun planning would be useful for the students. We will return to the evidence we have that, in the subsequent years, historical texts were used in this way. Let us simply note for now that Peyrard's translation of Archimedes' works as well as the second edition of his Euclid's *Elements* both bear the mark that the government approved of their use in *lycées.*

Third, the Council imagines that teachers at the school could cooperate with men of letters to produce editions of such Greek texts. Does this suggest an implicit criticism of Peyrard? We cannot tell. However, interestingly, the Council sees an

---

[63] Quoted following Langins (1989) and adding emphasis. "La délibération d'acheter un exemplaire de l'édition d'Archimède d'Oxford pour la bibliothèque (action suggérée par Peyrard) fournit à plusieurs membres l'occasion de développer les avantages qui résulteraient pour *l'utilité des élèves*, pour *l'avancement des sciences exactes* et pour la *gloire de l'École* si d'après un plan régulier et bien combiné elle parvenait à se *procurer les traductions des anciens auteurs grecs et latins* dans lesquelles les *Professeurs de l'École auraient concouru* avec les plus *savants littérateurs* pour *établir le véritable texte*, *le rendre fidèlement* en y *ajoutant les notes* nécessaires ou intéressantes que leur fourniraient (sic) l'étude plus immédiate de ces auteurs."

[64] It seems that in the 1807 edition, Torelli is mentioned only through a letter by Delambre, which Peyrard reproduces. However, in the notes to the second edition of Archimedes' works in 1808, Peyrard refers to Torelli's critical notes on several occasions. One would need to study on the basis of which edition Peyrard carried out his translation and how the 1808 edition differs from the 1807 one.

operation of this kind as useful for the "glory of the École"—as useful to this end, perhaps, as the library.

Fourth, the Council outlines a scheme that should be followed for such publications, and it is noteworthy that this scheme describes precisely Peyrard's work in the subsequent years. Indeed, the Council pointed out the necessity of establishing a genuine text before translating. Peyrard had not carried out this operation for his first publications, since we have seen that he had relied on Oxford editions. However, his third edition of Euclid's works (Peyrard 1814–1818) would fill this gap precisely. We have mentioned above that Monge had sent to the French national library, from the Papal manuscript collection in Rome, the *Vaticanus Graecus* 190. For his subsequent edition, Peyrard offered a critical edition of Euclid's text, in addition to a French as well as a Latin translation. For this, Peyrard relied on the manuscripts gathered in the "Imperial"— soon to become "Royal"— "Library," and especially on the *Vaticanus Graecus* 190—to which philologists now refer as "Codex P" in the honor of Peyrard's work on it. Similarly, and during the same years, Nicolas Halma, whom we have met above as secretary of the École between 1796 and 1797 and who, around 1810, with Lagrange's support, became librarian at the École des ponts et chaussées,[65] produced, in addition to a French translation, a critical edition of Ptolemy's *Almagest* (Halma 1813–1816): For this work, he, too, drew on the manuscripts of the Imperial Library taken from Rome.

Echoing with the scheme outlined by the Council of the École, both Peyrard and Halma emphasized that their translations were literal. Moreover, Peyrard's 1804 edition of Euclid's *Elements* and the 1807 rendering of Archimedes' works had explanatory notes to the translations.[66] Similarly, Halma completed his edition and translation of the *Almagest* with notes which were composed by Joseph Delambre (1749–1822). As already mentioned, the same Delambre had also been closely involved in Peyrard's publication on Archimedes, for which he composed a memoir titled "De l'Arithmétique des Grecs," to which we return shortly.[67]

In conclusion, these remarks support the idea that the École Polytechnique contributed to the conception—and hence to the shaping—of how a corpus of ancient Greek authors in mathematics should be—and actually was—produced at the beginning of the nineteenth century.

A final point from the Council's discussion deserves emphasis: the idea that translations of this kind would serve the advancement of science. Interestingly, in the historical introduction to his edition and translation, Halma raises the same question, and his answer is interesting for us to consider. Indeed, Halma writes:

> What benefit can we derive from Ptolemy's *Mathematical Composition*, [when we know] the degree of perfection that astronomy has now reached? What will be the usefulness of a

[65] (Guoyt de Fère and François-Fortuné 1858).

[66] For Euclid, see (Peyrard 1804: 559–573) and for Archimedes, see (Peyrard 1807: 445–536).

[67] (Wang Xiaofei 2017: 44–55 and Wang Xiaofei 2022) analyze this memoir and the very unusual approach that Delambre takes to the history of arithmetic.

> new translation of this book, after the two Latin versions that we have had for a long time? And is it not a retrograde step for science to take it back, as it were, to its cradle?[68]

A few pages later, Halma offers the following answer to this rhetorical question:

> (. . .) as a result, instead of slowing down the progress of science, on the contrary, it is a way of *enlightening* it in its march, to publish a translation—free from the errors rightly reproached to these [previously discussed] two versions—of the work that presents *its first steady steps*, or the *first operations guided by the spirit* of *method* and *calculation* that reigns there. (...). And since science is interested in *finding*, in an *accurate interpretation* of the meaning of our author, the *observations* he reports and the *methods* he employs, *that of the modern languages* to which the treasures of its literature have ensured the *universality* that, in the past, the Greek language had in the East, and the Latin language in the West, was the most suitable to spread everywhere the knowledge of these *observations* that cannot be dispensed with and of these *methods* that are still fruitfully studied.[69] (pp. V–VI)

We note the emphasis that Halma places on the history of the methods in his answer. This will appear to be a common theme for many teachers who used historical material in the context of the École.

Lastly, let us add that the Council nominated a committee in which Monge, Prony, Lacroix, Labey, Hachette, and Siméon Denis Poisson (1781–1840) took part, to think about the opportunity of its project about ancient classical texts (Langins 1989). In other words, several teachers who would have Poncelet as a student after he entered the École 3 years later, had been involved in the reflection on such uses of ancient Greek mathematical texts. In fact, years later, Poncelet remembered that one of his own personal inquiries in geometry, at the École, involved reflections on Greek mathematicians of antiquity. He formulated this memory in a note that he added in 1864 to a notebook written in the winter of 1815–1816. In this notebook, the young Poncelet asserted, about the line described by a point, that "the simplest

68 (Halma 1813–1816: I) "Quel fruit peut-on retirer de la *Composition Mathématique* de Ptolémée, au degré de perfection où l'astronomie est aujourd'hui parvenue? Quelle sera l'utilité d'une nouvelle traduction de ce Livre, après les deux versions latines que nous en avons depuis longtemps? Et n'est-ce pas faire rétrograder la science, que de la ramener, pour ainsi dire, à son berceau?"

69 (Halma 1813–1816: V-VI, our emphasis). Needless to say, the modern language to which Halma refers as the language enjoying universality is French. "C'est donc, au lieu de rallentir (sic) les progrès de la science, *l'éclairer* au contraire dans sa marche, que de publier, de l'ouvrage qui en expose les premiers pas assurés, ou les *premières opérations dirigées par l'esprit de méthode et de calcul* qui y règne, une traduction exempte des fautes justement reprochées à ces deux versions. (. . .). Et puisque la science est intéressée à *trouver* dans une *interprétation exacte* du sens de notre auteur, les *observations* qu'il rapporte et les *méthodes* qu'il emploie, *celle des langues modernes* [that is, French] à laquelle les trésors de sa littérature ont assuré *l'universalité* qu'avoient eue autrefois la langue grecque en orient, et la langue latine dans l'occident, étoit la plus propre à répandre partout la connoissance de ces *observations* dont on ne peut se passer, et de ces *méthodes toujours étudiées avec fruit*."

among these lines had already been studied by the Ancients."[70] Poncelet's 1864 note reads as follows

> This very extensive class includes not only Nicomedes' *conchoid*, Diocles' *cissoid*, Conon's and Archimedes' *spiral*, but a large number of other lines with simple or multiple convolutions. During my stay in 1808 and 1809 at the École Polytechnique, I dealt with these curves with a kind of exclusive perseverance, in connection with the research on... (here, Poncelet returns to the piece of work that he developed through discussions with students of the same study room as him and that we have mentioned in Sect. 2.1).[71]

We have already mentioned hints suggesting that the historical work undertaken on Greek mathematical texts had an impact on the teaching of mathematics. For instance, from 1807 onward, Peyrard's editions were published with reference to the government's adoption of the books for the libraries of the *lycées*. Moreover, in his *Traité d'arithmétique,* Jean-Guillaume Garnier included a chapter on the "arithmetic of the Greeks" that was based on Delambre's memoir.[72] What clues can we gather, more specifically, on the way in which mathematical texts of the past were used to train students at the École Polytechnique?

2.4 Hachette and the *Correspondance sur l'Ecole Polytechnique*

One way of approaching the latter question is to examine a journal that Jean-Nicolas-Pierre Hachette established in 1804 under the title *Correspondance pour l'École Polytechnique à l'usage des élèves de cette école*.[73] Indeed, in addition to reflecting most aspects of the institutional life of the École, this bulletin published works by students, former students and teachers that bespeak interesting features of their practice of mathematics.

As we have seen, Hachette assisted Monge in teaching descriptive geometry and analysis applied to geometry, and he was Poncelet's teacher in these fields during the three years that Poncelet spent at the École. Poncelet does not seem to have harbored a great admiration for Hachette, describing methods the latter taught as "long,

[70] (Poncelet 1864: 255). "... dont les plus simples avaient déjà occupé les Anciens (*)."

[71] (Poncelet 1864: 255). Italics in the original text. "Cette classe fort étendue comprend non-seulement la *conchoïde* de Nicomède, la *cissoïde* de Dioclès, la *spirale* de Conon et d'Archimède, mais un grand nombre d'autres lignes à circonvolutions simples ou multiples, dont je me suis occupé avec une sorte de persévérance exclusive, pendant mon séjour en 1808 et 1809 à l'École Polytechnique, à propos des recherches sur ...".

[72] The first edition of Garnier's treatise was published in 1803 and did not include a chapter of this kind. However, as early as 1808, the second edition presented this chapter of the history of arithmetic. See (Wang Xiaofei 2017: 160; and Wang Xiaofei 2022).

[73] Published under the form of booklets appearing at regular intervals, the *Correspondance* was republished as three books, with slightly different titles (Hachette (ed.) 1808); (Hachette (ed.) 1814); (Hachette (ed.) 1814–1816). We refer to its articles through these three volumes.

complicated and cumbersome" (Poncelet 1862: 457). The expression "estimable professor" (Poncelet 1822: 399) by which Poncelet refers to Hachette seems much less enthusiastic than the adjectives he uses for Lacroix and Ampère. That said, Poncelet appears to have been an avid reader of the *Correspondance*, to which his writings refer constantly.[74]

2.4.1 *Problems and Methods Taken from the Library for Collective Use*

As the title of the *Correspondance* makes explicit, the journal was primarily meant for use by the students. To begin with, it published, in particular, students' works, thereby reflecting the students' personal research. We will find Poncelet among them.

In fact, Hachette and other professors put forward problems and theorems, on which the students worked together, submitting solutions and demonstrations, some of which Hachette edited and published in the *Correspondance*. For example, in July 1806, Hachette published a solution by Louis Puissant (1769–1843) for a problem in analytic geometry, and he concludes the article with the statement of a theorem, "invit[ing] the students of the École Polytechnique to provide a proof of this theorem."[75]

Interestingly, some of these problems were explicitly attached to a mathematical text of the past. For instance, the first issue of the *Correspondance*, in April 1804, presents a "note about activities carried out at the *École* during the years XI-XII"—which covered the academic years 1802–1803 and 1804–1805. These activities include a rubrique—"On the Contact between Spheres"—that begins with the quotation of a problem by Pierre de Fermat (160?–1665) in his "treatise *De contactibus sphericis*." The problem consists in finding a sphere tangent to four other spheres whose centers and radii are known.[76] Hachette does not only translate the quote by Fermat, but he also mentions the year and place of the publication of Fermat's selected works, from which he is quoting, as well as the number of the question in the treatise.[77] We will see that such precision characterizes the references that Hachette makes to the books that most probably he consulted—

[74] We give a single example: (Poncelet 1822: 78).

[75] (Hachette (ed.) 1808: 191–193, in particular p. 193).

[76] As early as 1794, Monge had put forward this problem to the "students-instructors," whom he trained to become the *chefs de brigade* of the first promotion (Belhoste 2009).

[77] The book in question is in the old collection of the library of the École Polytechnique, but Hachette does not give its title: *Varia Opera mathematica D. Petri de Fermat, senatoris Tolosani. Accesserunt selectae quaedam ejusdem epistolae, vel ad ipsum a plerisque doctissimis viris Gallicè, Latine, vel Italice, de rebus ad Mathematicas disciplinas, aut physicam pertinentibus scriptae*, 1679, Toulouse: J. Pech. The quotation is in (Hachette (ed.) 1808: 8).

and invited students to consult—in the library. The *Correspondance* thus provides evidence of one way the library was used in training the students.

In the same note, Hachette mentions other related problems tackled by the first students trained at the École, such as the problem of finding a circle touching three other given circles. The following issue, in August–September 1804, begins with an article titled "Activities carried out at the *École*: Memoir on the contact between spheres," in which Hachette brings together several solutions given to seven related problems in the context of "research carried out at the École Polytechnique."[78] The solution to the first problem makes use of a property for which a footnote refers to a precise paragraph of Monge's *Géométrie descriptive.* Another footnote attributes the solution to the fifth problem explicitly to a former student, Charles-André Dupin (1784–1873). For the sixth problem—precisely the problem of the three circles—, a footnote refers the reader to analytical solutions given by, respectively, Newton in his *Arithmetica Universalis* as well as Euler and Nicolas Fuss (1755–1826), through the indication of, respectively, a precise book with page number and a precise volume of the *Mémoires de l'Académie de Pétersbourg.* All these publications could be found in the library of the École.

Problems were also formulated by other contributors, such as the former student Louis Poinsot, at the time a "teacher at the Lycée Bonaparte." In the January 1807 issue, Hachette published part of a letter Poinsot sent him on January 6, 1807, which ends with the statement of a problem.[79] In the subsequent issue of May 1807, Hachette published several solutions that he "received" for the "issue raised by Poinsot."[80]

Students frequently returned to problems that had been raised earlier, with new solutions. For example, in the same issue of January 1807, Hachette published another solution to the problem relating to the three circles that he had received from Augustin-Louis Cauchy (1789–1857), a student at the École between 1805 and 1807.[81] Moreover, in 1811, it was again to this problem as well as to the problem relating to Poinsot's question that Poncelet devotes the only article of his published in the *Correspondance*.[82] Let us note that a practice of this kind encouraged students to compare different methods for the solution of the same problem, including methods that had been put forward by authors of the past.

In his "Memories from the École Polytechnique (1809-1810)," Poncelet (1862: 443–446) reproduces *verbatim* the solution published in 1811 to the problem of the three circles. On this occasion, he gives evidence about the students' interaction with Hachette in the context of the publication of their solutions. Indeed, Poncelet asserts that he had given Hachette the article two years earlier, in 1809, and that the

[78] (Hachette (ed.) 1808: 17–28).

[79] (Hachette (ed.) 1808: 245–246).

[80] (Hachette (ed.) 1808: 305–307).

[81] (Hachette (ed.) 1808: 193–195).

[82] (Hachette (ed.) 1814: 271–274). On the relation between Poinsot's problem and the problem solved, see Poncelet (1862: 446).

editor "had *demanded* that several passages that he considered useless or foreign to the question be deleted" (p. 443, our emphasis).

In the same document, Poncelet (1862: 447–457) published a note that derives from a manuscript that he dates from his second year at the École Polytechnique (1809–1810) (Poncelet 1862: 456–457). Interestingly, the manuscript points out that it makes use of a method that Gilles Personne de Roberval (1602–1675) put forward for tangents, which seems to echo an article also published in January 1809 in the *Correspondance* by Louis Gaultier (1776–1852), a former student who had entered the school in 1798 and was, when he wrote, a teacher at the Conservatoire des arts et métiers.[83] We thus see how the historical inquiry of some of the students made ancient ways of approaching problems available to all, which then became part of a collective culture.

2.4.2 *History of Mathematics as Part of the Students' Activities at the École*

The articles published in the *Correspondance* by Hachette point out other uses of history as well as other uses of the library.

In the first issue of 1805 (Pluviôse year XIII), in the rubrique "Activities carried out at the *École*," for descriptive geometry, Hachette published a "Complete solution of the triangular pyramid."[84] The first paragraph mentions, with careful dating, the contributions to spherical trigonometry—a subpart of his topic—prior to Neper, that is, in his view, those of Hipparchus, Theodosius, Menelaus, Gebert (sic) and Regiomontanus. His historiography emphasizes a break in the treatment of the topic, with the application of algebra to geometry, and points to the works of Euler and Lagrange—with precise references—as the pinnacle of what could be achieved through "taking this new route open to modern geometers" (p. 41). However, Hachette continues, even though the formulas thereby produced give the "simplest arithmetic operations" to solve the problems that can be raised, "they do not indicate the geometric constructions that lead the most directly from the lines given to the lines sought-for." And, here, Hachette returns to the historiography of the subject, with another, parallel, line of development.

He grants that "the inventors of the art of stereotomic drawing (*art-du-trait*)"—that is, the carpenters and stonecutters, on whose methods Monge had relied to shape his descriptive geometry— had undoubtedly "solved these problems in this way."[85] But he laments that they did not produce writings testifying to this fact. Hachette further indicates that "the Arabs" and "the Goths" "frequently made use of this

[83] (Hachette (ed.) 1814: 24) mentions Gaultier's work in relation to Roberval's method, and gives an outline on pp. 27–28 and 87–93.

[84] (Hachette (ed.) 1808: 41–51).

[85] On stereotomic drawing, see (Sakarovitch 1998).

art," lamenting here too that they "transmitted neither the names of the inventors, nor the principles of geometry on which this art was based." Hachette could again assign names and refer to precise books only starting from the end of the sixteenth century: he points to the architectural treatise by Philibert de l'Orme (1514–1570), the treatise on stonecutting by Mathurin Jousse (*ca.* 1575–1645), and a treatise on stone- and wood-cutting by Amédée François Frézier (1682–1773) as well as to the mathematical treatise by the Jesuit priest "Deschalles" (Claude François Milliet Dechales (1621–1678)). Hachette further refers to François Derand (1591–1644) and Desargues's treatises on stonecutting, with, however, no precise bibliographic indication. Among these works, Jousse's book features in Jacotot's inventory, and *Mundus Mathematicus* by "Deschales", in Peyrard's 1796 catalogue (number 45, Pepe 1996). Moreover, we have already mentioned, in the library, Curabelle's 1644 diatribe— especially on stonecutting— against Desargues's work. Derand's book on the art of vaults as well as de l'Orme's and Frézier's treatises are all in the old collection of the library of the École. Hachette concludes this second part of his historiography with Monge's descriptive geometry, presented as the pinnacle of this line of development, and he situates his article in line with his master's work, adding to it precisely for the purpose of his lectures at the École Polytechnique.

In sum, Hachette writes a mathematical article for the students, and includes in it a historiographic treatment of the subject. His historiography situates Monge's contribution in one of the two traditions of approach to geometrical problems that he outlines: Hachette, thereby, projects backwards on the history of mathematics the two types of method, the comparison of which Monge emphasized in his lectures. In addition, the history of mathematics that Hachette offers emphasizes the contributions of scholars who were not from Europe as well as those of scholars who worked in the context of practical activities. All these features characterize the type of historiography that Poncelet and later Chasles would develop.

A last point deserves emphasis: Hachette's historical treatment was made possible thanks to the book collection available to him at the time. This remark holds true for another article that Hachette published in the January 1809 issue of the *Correspondance* on spherical epicycloids.[86] He begins with a reference to Charles Etienne Louis Camus's (1699–1768) *Traité de statique* (Peyrard's 1796 catalogue number 34, Pepe 1996: 182) and explains where spherical epicycloids occur in the theory of machines, before turning to his treatment of epicycloids in general and tackling the problem of the geometric construction of their tangents. Note that Gaultier's article that we have mentioned above, and on the basis of which Poncelet would compose a manuscript, applies Roberval's method precisely to the determination of the tangent of epicycloids. After having introduced the definition of spherical epicycloids, Hachette refers the students to volume III of Johann Bernoulli's (1667–1748) works, using a precise bibliographic reference. This is precisely a book that the library acquired in the set sent from Belgium and that features in Peyrard's catalogue (Pepe 1996: 179, 181, respectively).

[86] (Hachette (ed.) 1814: 22–27).

Hachette seems to have spent time on these volumes of Johann Bernoulli's works. Indeed, in November 1805, in an article devoted to the computation of the shortest twilight, he refers to Johann Bernoulli's research on it and quotes extensively the excerpt of a letter published in January 1693 and included in volume I of Bernoulli's works (p. 64).[87] Hachette further refers to "the Portuguese geometer" Pedro Nuñes' (1502–1578) solution to this problem by means of spherical trigonometry—yet another book possessed by the library. Hachette then turns to Monge's geometrical solution, which he invites students to prove geometrically, while he himself establishes it analytically in the last pages of the article. This publication thus illustrates the mixture of history and mathematics to which students were exposed and in which they were encouraged to participate, being thereby invited to proceed using different methods and comparing them.[88] We find exactly the same mixture in an article that Hachette composed on the double refraction of light, in which the author quotes extensively Christiaan Huygens' (1629–1695) treatise on light.[89]

Interestingly, the *Correspondance* also sheds light on an entirely different way in which students were invited to think about the sciences on the basis of their history. Indeed, since 1804, a course of "grammar and belles-lettres" had been established for both first-year and second-year students, and it was given by François Andrieux (1759–1833) (Belhoste 2003: 176–179). In the issue of July 1805, the *Correspondance* gives an outline of the content of these classes and publishes the topics given for students to compose a written essay.[90] One of the topics given to first-year students reads as follows:

> Hiero, king of Syracuse, writes to the geometer Archimedes, his relative and friend, to urge him not to make geometry a purely intellectual and speculative science, but to apply it to useful inventions, for example, to build war machines to defend [Syracuse] against the Romans who threaten Syracuse.
>
> Archimedes agrees, and he promises the king machines whose effect will be sure and prodigious *Hiero*'s *Letter* and *Archimedes' answer*.
>
> (See *Plutarch*, *Titus-Livius*, *Polybius*, etc.)[91]

The composition had to offer an outline of the two elements of the list italicized. Moreover, the teacher gave three books in which the students could find information, and the authors' names were also italicized. It is noteworthy that valuable editions of Plutarque's *Vitae Illustrium virorum* and of Tite-Live's works were among the books Monge sent from Rome to the library of the Pantheon (Pepe 1996: 196). We

[87] (Hachette (ed.) 1808: 148–151).

[88] (Hachette (ed.) 1808: 193) mentions yet other types of solution to the problem.

[89] (Hachette (ed.) 1814: 281–289).

[90] (Hachette (ed.) 1808: 86–88).

[91] "Hiéron, roi de Syracuse, écrit au géomètre Archimède, son parent et son ami, pour l'engager à ne pas faire de la géométrie une science purement intellectuelle et spéculative, mais à l'appliquer à des inventions utiles, par exemple, à construire des machines de guerre pour se défendre contre les Romains qui menacent Syracuse. Archimède y consent, et promet au roi des machines dont l'effet sera sûr et prodigieux *Lettre d'Hiéron et réponse d'Archimède.* (Voyez *Plutarque, Tite-Live, Polybe, etc.*)" (p. 88, italics in the original).

are fortunate enough to have a composition of a first-year student on this topic: Charles-Hippolyte de Paravey (1787–1871), who was at the École between 1803 and 1806.[92] The manuscript further bears the marks of the teacher.

To the second-year students, Andrieux proposed another topic touching the history of science, which the *Correspondance* reproduced, the italics likewise indicating the task and the sources to use:

> Viviani, a student of Galileo's, defends his master before the inquisition in Florence in 1633. Galileo was accused of heresy for having taught and maintained the movement of the earth around the sun.
>
> *Speech*.
>
> (See *Laplace, Exposition du système du monde*, book 5; *Histoire de l'astronomie moderne*).[93]

The reflection on the history of science to which the students were invited thus went beyond the simple search for various historical methods to solve some problems or the correct attribution of results to their true inventors.

2.4.3 The Pursuit of this Form of Practice After School Years and the Formation of a Network

Through the *Correspondance*, students were thus incited to work on the same problems and hence encouraged to practice joint research of the kind Monge had aimed to foster. They were also put in contact with former students, entering into a network of scholars who would remain in contact with one another and devote time to mathematical research (Belhoste 1998: 12). Among the former students, we have evoked Poinsot and Gaultier. Many more sent articles to the *Correspondance*, and, interestingly for our purpose, these contributors, like Gaultier, regularly displayed an interest in mathematics mixed with an interest in the history and also the philosophy

[92] See the digitized version of *Papiers divers du chevalier de Paravey* (1871). V. *Documents sur l'École polytechnique* at the link: https://gallica.bnf.fr/ark:/12148/btv1b10090440m/f7.item.r=catalogue%20de%20la%20bibliothèque%20de%20l'Ecole%20polytechnique (accessed on August 5, 2022). The letter attributed to Hiéron is on pp. 14–15 of the pdf, and the answer attributed to Archimedes is on pp. 15–16. After the École Polytechnique, Paravey entered the ponts et chaussées École d'application, and there too, he was asked to compose texts on topics relating to the history of science. One such topic reads (see the same set of documents on p. 23): "Ecole des ponts et chaussées. Concours du 28 mars 1809: de la manière d'écrire l'histoire des sciences et des arts, soit dans les ouvrages où elle est réunie à l'histoire politique des peuples, soit dans les ouvrages dont elle constitue l'objet spécial. Nota. Un des points importants de la question proposée consiste à examiner comment une histoire bien faite des sciences et des arts est un moyen de les propager, d'en préparer et d'en hâter les progrès." See his composition on pp. 23–25.

[93] "Viviani, élève de Galilée, défend son maître devant l'inquisition, à Florence, en 1633. Galilée étoit accusé d'hérésie pour avoir enseigné et soutenu le mouvement de la terre autour du soleil. *Discours*. (Voyez *Laplace*, *Exposition du système du monde*, liv. 5; *Histoire de l'astronomie moderne*)." (p. 89, italics in the original).

of mathematics. They thus seem to have acquired a way of practicing science at the École and to continue with a practice of this kind through their exchanges.

Among them, we have mentioned Jacques Frédéric Français as the author of one of the first articles of the "philosophie mathématique" rubrique, in the *Annales de Mathématiques pures et appliquées,* (Français 1813–1814) and also as someone to whom Poncelet would show his first Saratoff results and who would lend Poncelet books from Arbogast's library. Français, having entered the École Polytechnique in 1797, continued contributing to the *Correspondance* after he left the school.[94]

Another former student who sent letters and contributions to Hachette and was quite influential on Poncelet, Charles-Julien Brianchon, illustrates quite nicely the nature of the engagement with history to which these mathematicians had been exposed and which they continued in their work.

In a letter sent from Metz, on January 3, 1807, to Hachette—who reproduced it in the *Correspondance*[95]—, Brianchon put forward and proves properties of curves of the second degree. Brianchon hesitantly attributes the first of these properties to Colin Maclaurin (1698–1746): one might assume that he had read Maclaurin's work and relied on his memory to suggest this attribution. Brianchon further emphasizes the ease with which the concept of transversal—to which he associates the name of Lazare Carnot (1753–1823) and Servois—enables one to derive properties that would require much longer treatments through analysis and even descriptive geometry. We recognize the constant comparison carried out between methods and their relative merits. Brianchon adds: "It is undoubtedly curious to find in several ancient authors, notably in Pappus, some traces of research of this kind."[96]

In the case of the article just examined, a mathematical text is interspersed with historical remarks. In January 1813, Hachette published another article of exactly the same type by Brianchon, titled "Géométrie de la règle." This article addresses problems on conics that Brianchon suggests solving using only a ruler. Interestingly, at the end, Brianchon adds historical remarks.[97] There, he first emphasizes that the solution to some of the problems was given in lecture courses on architecture—a point that he proves by referring to the work by François Blondel (1618–1686), using the following reference: "*Résolution des quatre principaux*

[94] After a letter to Hachette on April 17, 1807, which the latter published (Hachette (ed) 1808: 320–321), Français sent an article that would be published in the issue of January 1808 (Hachette (ed) 1808: 337–349). The last pages of the article return to a problem raised earlier by Hachette. In the issue of January 1810, Français has two articles published. The first offers yet another solution to Fermat's problem (Hachette (ed) 1814: 63–69). The second (Hachette (ed) 1814: 69–70) is devoted to the problem for which Poncelet claims to have given, in 1809, with Guillebon, a much better solution to the students of their study room. This is the topic of the third manuscript published in his "Memories from the École Polytechnique (1809–1810)," Poncelet (1862: 456-461). In it, Poncelet criticizes explicitly the methods published by Hachette and Français, among others.

[95] (Hachette (ed.) 1808: 281–289).

[96] (Hachette (ed.) 1808: 309). "il est sans doute curieux de retrouver dans plusieurs auteurs anciens, notamment dans *Pappus*, quelques traces de ce genre de recherches."

[97] (Hachette (ed.) 1814: 386–387).

Problèmes d'Architecture; au Louvre, 1673, in-folio." The first catalogue, by Jacotot, already mentions this book as possessed by the library of the École (Pepe 1996:177). What is more, this reference shows that Brianchon's historiography had points in common with Hachette's.

In the same 1813 article, Brianchon further adds remarks on Blaise Pascal's (1623–1662) works on conics in a way that shows the depth and the accuracy of his knowledge of the historical evidence. In particular, Brianchon refers the reader to Pascal's works and to "a letter written by *Leibnitz* and placed at the end of the fifth or last volume of the 1779 edition." Strikingly, this edition of Pascal's works is number 112 in Peyrard's catalogue of the library established in April 1796 (Pepe 1996: 184). These remarks cogently suggest that Brianchon, probably when he was a student, made extensive use of the library.

However, it also seems that Brianchon continued this practice of working with books after he left the École. Indeed, the next letter that Brianchon sent—from Toledo, this time—on April 8, 1810 and that Hachette published in January 1811 is more purely historical in nature.[98] Brianchon had had in the hands "by chance" a Latin edition of Ptolemy's *Almagest*—whose reference he gives with all due details—and had identified that this book in fact contains "the fundamental principle" of the "theory of transversals"—what is now generally called Menelaus theorem. Brianchon thus translates the related pages of the *Almagest*, and in italics establishes a relationship between Ptolemy's formulation and Carnot's statements. Similarly, the following year, in 1812, Hachette published in the *Journal de l'Ecole Polytechnique* a translation of Fermat's Latin text from which he had selected the problem of the four spheres.[99] In these ways, just as Peyrard and Halma had sought to translate Greek classical texts of antiquity into French in order to broaden their accessibility, Brianchon and Hachette were producing the means of shaping a culture of mathematics that could be shared.

The preceding remarks shed light on why Poncelet's introduction to his 1822 treatise, precisely after mentioning Brianchon's discovery about Ptolemy's *Almagest,* refers to Brianchon as someone to whom we "owe a lot for the history of projective properties".[100]

The last of these former students worth mentioning with respect to the *Correspondance* is Olry Terquem (1782–1862), who became a student at the École in 1801 and who was also in close contact with Poncelet as the latter was developing the research that led to the publication of the 1822 *Traité*. In January 1816, Terquem, at the time a teacher in artillery schools, published an article titled "Histoire de l'algèbre. Sur l'algèbre des Indiens", which he had written on the

[98] (Hachette (ed.) 1814: 257–260).

[99] Hachette, "Du Contact des sphères par Fermat, traduit par M. Hachette", *Journal de l'Ecole Polytechnique*, septième et huitième cahier, second volume, 1812, pp. 279–289.

[100] (Poncelet 1822: xxxvii–xxxviii).

basis of recent English publications on the topic.[101] In fact, Terquem would be more broadly one of the most active proponents of the history of mathematics in the nineteenth century. In particular, in the context of the journal he established in 1842 with Camille-Christophe Gerono (1799–1891) under the title *Nouvelles Annales de mathématiques*, from 1855 and until his death, he published what can be considered as the first journal devoted to the history of mathematics, that is, a yearly supplement, with a separate pagination, titled *Bulletin de bibliographie, d'histoire et de biographie mathématiques*. This publication would leave the enduring imprint of a practice of mathematics shaped to train students at the École Polytechnique on the history of mathematics, notably thanks to the library to which Peyrard contributed in such a remarkable manner.

2.5 Conclusion

In this chapter, we have focused on certain publications linked with the École Polytechnique. They have allowed us to observe the activity in the history of mathematics that developed in this context, in relation to the way students were encouraged to learn. A key point has emerged: the interest in the history of mathematics, which was made possible essentially thanks to the library, was intimately connected with the practice of comparing methods, in which students were trained. This interest was also encouraged by the fact that teachers and former students selected problems to which students were invited to find new solutions and for which the literature of the past offered other methods. The publications under consideration were mainly about geometry. We will turn to analysis, which presents a different case, in another publication that we will devote to the practice of philosophy of mathematics in the same context.

To date, we have shown that for geometry, a mathematical culture took shape, in which students and former students participated, Poncelet among them. As a student, as we have seen, he contributed a solution to the problem of finding a circle tangent to three given circles. Moreover, with a friend from the same study room he read extensively books which they probably found in the library. Later, he continued studying ancient treatises, such as what he referred to as Maclaurin's posthumous treatise of algebra (Bruneau 2011). The notes to his 1822 treatise show the extent of his historical knowledge.

In our view, this context sheds interesting light on several features of the *Traité des propriétés projectives des figures*.

To begin with, the book, as Poncelet defines it, centers precisely on methods. Indeed, in (Poncelet 1822: vi), we read:

> The purpose of this book, however voluminous it may seem, is not so much to multiply the number of these properties as to indicate the path that we must follow. In brief, I

[101] (Hachette (ed.) 1814–1816: 4–17, 275–283 + Plate 271).

> have sought, above all, to perfect the method of proving and discovering using simply geometry.[102]

The specificity of the "path" to be followed could be best illustrated by returning to old problems and showing how the new method allowed him to deal with them. This is what Poncelet does when, in his *Traité des propriétés projectives des figures*, he returns precisely to the problem of the three circles (Poncelet 1822: 141). First, in a footnote, he outlines the history of the treatment of this problem as well as the related problem about spheres. Three points deserve our attention in this historical account.

First, when he mentions research of the past, he lists: "Appollonius (sic), Viète, Fermat, Newton, Euler, and Fusse (sic), etc.": if we set aside Apollonius and Vieta, these are precisely the names of those whose contributions were mentioned in the various articles in the *Correspondance.* In other words, a historiography had taken shape, which Poncelet inherited.[103]

Second, when Poncelet mentions modern and "general" approaches to the problems, he adds comments to the list of those he refers to by adding that they are, "for the most part, former students or professors of the École Polytechnique."

Third, interestingly, in this footnote, Poncelet goes deeper into Pascal's comments on these two problems, and he quotes Pascal—who refers to Apollonius and Vieta—from volume 4 of the same 1779 edition that had been used by Brianchon. Did Poncelet study this edition in the library of the École? Did he have access to it in Arbogast's library that Français kept (Anonymous 1823: 494)? We do not know. However, the mere fact that he refers to it explicitly as well as the precision of the bibliographic reference both show that his use of books was similar to Hachette's and to Brianchon's.

In these pages, Poncelet makes explicit his reason for returning to the three circles problem: he aims to show that the methods he has introduced lead to the properties in a mostly "simple" and "natural" way, which in addition will be generalizable. Comparison between methods would remain central to his mathematical work.

For this, the "path" followed by others was crucial, a task for which historical research was important. We cannot thus be surprised to read Poncelet stating in 1843:

> ... Since, as it is worth repeating here, *what interests us most* in the *history and philosophy of science* is the *path* through which the human mind has arrived at the *discovery of fundamental truths*.[104]

---

[102] "Le but de ce livre, quelque volumineux qui il paraisse, est moins de multiplier le nombre de ces propriétés que d'indiquer la route que l'on doit suivre. En un mot, j'ai cherché, avant tout, à perfectionner la méthode de démontrer et de découvrir en simple Géométrie."

[103] Gaultier (1813: 126)—also a former student of the École—repeats exactly the same references.

[104] "... car, il est bon ici de le répéter, *ce qui intéresse le plus dans l'histoire et la philosophie des sciences*, c'est la *route* par laquelle l'esprit humain est parvenu à la *découverte des vérités fondamentales*." "Extrait des *Comptes rendus de l'Académie des Sciences*, t. XVI, 1843, p. 947 à 964," quoted in (Poncelet 1866: 346). Our emphasis.

Acknowledgements It is a tremendous pleasure for us to dedicate this chapter to our friend Jeremy, whose affection and constant support throughout the years have enlightened our lives. Our most sincere thanks go to the Fondation des Treilles, thanks to whom we could write this chapter during a research stay (September–October 2022). We are also grateful to Qianqian Feng, Richard Kennedy, as well as to the editors of the book and to the two readers, whose comments helped us improve our argument.

References

Anonymous. 1823. Notice d'une bibliothèque de mathématiques, de philosophie, de physique et de chimie, à vendre de gré à gré. In *Bulletin général et universel des annonces et nouvelles scientifiques*, ed. M. le Baron de Férussac, 493–495.

———. 1892. La Bibliothèque de l'École Polytechnique. *Les Nouvelles de l'Intermédiaire* 34 (10 décembre 1892): 123–131.

Anonymous (P. L. B.). 1893. François Peyrard. *Les Nouvelles de l'Intermédiaire* 6 (28 février 1893): 43–46.

Aujac, Germaine. 1990. Science grecque et révolution française. *Bulletin de l'Association Guillaume Budé : Lettres d'humanité* 49: 395–409. https://doi.org/10.3406/bude.1990.1757. Link: http://www.persee.fr/web/revues/home/prescript/article/bude_1247-6862_1990_num_49_4_1757.

———. 2007. La Science grecque revisitée aux entours de la révolution. In *Les Autorités. Dynamiques et mutations d'une figure de référence à l'Antiquité*, ed. Didier Foucault and Pascal Payen, 227–250. Grenoble: Jérôme Millon.

Barbier, Paul. 1999. Pierre Jacotot (1756-1821), Professeur de Collège à Dijon, Bibliothécaire de l'Ecole centrale des Travaux publics. *Bulletin de la SABIX (Société des Amis de la Bibliothèque et de l'Histoire de l'École Polytechnique)* 20: 17–38. https://www.sabix.org/bulletin/b20/jacotot.html; https://journals.openedition.org/sabix/872.

Belhoste, Bruno. 1994. De l'Ecole des ponts et chaussées à l'Ecole centrale des travaux publics. Nouveaux documents sur la fondation de l'Ecole polytechnique. *Bulletin de la SABIX (Société des Amis de la Bibliothèque de l'Ecole Polytechnique)* 11: 1–29. https://doi.org/10.4000/sabix.617. http://journals.openedition.org/sabix/617;https://www.sabix.org/bulletin/b11/belhoste.pdf. References are made to the latter digitization, since the former is incomplete.

———. 1998. De l'École polytechnique à Saratoff, les premiers travaux géométriques de Poncelet. *Bulletin de la SABIX (Société des Amis de la Bibliothèque de l'Ecole Polytechnique)* 19: 9–29. http://www.sabix.org/bulletin/b19/belhoste.html.

———. 2003. *La Formation d'une technocratie. L'Ecole polytechnique et ses élèves de la Révolution au Second Empire*. Paris: Belin.

———. 2009. Charles Dupin et l'héritage de Monge en géométrie. In *Charles Dupin (1784-1873). Ingénieur, savant, économiste, pédagogue et parlementaire du Premier au Second Empire*, ed. Carole Christen and François Vatin, 81–97. Rennes: Presses Universitaires de Rennes.

Belhoste, Bruno, and René Taton. 1992. Leçons de Monge. In *L'Ecole Normale de l'An III. Leçons de mathématiques. Laplace-Lagrange-Monge*, ed. Jean Dhombres, 266–459. Paris: Dunod.

Bradley, Margaret. 1976. An Early Science Library and the Provision of Textbooks: The Ecole Polytechnique, 1794-1815. *Libri* 26: 165–180.

Bruneau, Olivier. 2011. Le *De Linearum* de MacLaurin: Entre Newton et Poncelet. *Revue d'histoire des mathématiques* 17: 9–39.

Caramalho Domingues, João. 2008. *Lacroix and the Calculus*, Science Network Historical Studies. Vol. 35. Basel: Birkhäuser.

Chasles, Michel. 1837. *Aperçu historique sur l'origine et le développement des méthodes en géométrie, particulièrement de celles qui se rapportent à la géométrie moderne, suivi d'un* *mémoire de géométrie sur deux principes généraux de la science : la dualité et l'homographie*. Bruxelles: M. Hayez.

Chemla, Karine. 2016. The Value of Generality in Michel Chasles's Historiography of Geometry. In *The Oxford Handbook of Generality in Mathematics and the Sciences*, ed. Karine Chemla, Renaud Chorlay, and David Rabouin, 47–89. Oxford: Oxford University Press.

Curabelle, Jacques. 1644. *Examen des oeuvres du Sieur Desargues*. Paris: M. et I. Henault.

de Chambray, Marquis Georges. 1836. *Sur l'École Polytechnique*. Paris: Anselin Libraire.

Didion, Monsieur, and le Général. 1869. *Notice sur la vie et les ouvrages du Général J.-V. Poncelet. Lue à l'Académie impériale de Metz dans la séance du 18 mars 1869*. Paris: Gauthier-Villars, Imprimeur-Libraire.

Dooley, E.L. 1994, November. Procès verbaux des Séances du Conseil de l'Ecole polytechnique de l'an III (1794) à l'an VII (1799). Transcription des registres à partir de microfiches, par le Colonel E. L. Dooley (Virginia Military Institute). Recension par Emmanuel Grison, Ecole polytechnique, mise en page et révision à la Bibliothèque de l'Ecole polytechnique. *Bulletin de la SABIX* 12: 8–64. https://doi.org/10.4000/sabix.703. http://journals.openedition.org/sabix/703.

———. n.d. Transcription des procès-verbaux des séances du conseil d'instruction et d'administration de l'Ecole Polytechnique. De l'An 3 à l'An 7. Transcription par E.L. Dooley, Virginia Military Institute. Recension par E. Grison, Ecole Polytechnique. Vol. 1: Transcription. Vol. 2: Index.

Field, J.V., and Jeremy J. Gray. 1987. *The Geometrical Work of Girard Desargues*. New York: Springer.

Fourcy, Ambroise. 1828. *Histoire de l'École Polytechnique*. Paris: chez l'auteur, à l'École Polytechnique.

Français, Jacques Frédéric. 1813–1814. Philosophie mathématique. Sur la théorie des quantités imaginaires. *Annales de Mathématiques pures et appliquées* 4: 222–227. http://www.numdam.org/item?id=AMPA_1813-1814__4__222_1.

Gaultier de Tours, Louis. 1813. Mémoire sur les moyens généraux de construire graphiquement un cercle déterminé par trois conditions, et une sphère déterminée par quatre conditions. *Journal de l'École polytechnique* 16e cahier (tome IX): 124–214.

Gergonne, Joseph Diez. 1826–1827. Réflexion sur le précédent article. *Annales de Mathématiques pures et appliquées* 17: 272–276. http://www.numdam.org/item?id=AMPA_1826-1827__17__272_1.

Gouzévitch, Irina, and Dimitri Gouzévitch. 1998. La Guerre, la captivité et les mathématiques. *Bulletin de la SABIX (Société des Amis de la Bibliothèque et de l'Histoire de l'Ecole Polytechnique)* 19: 31–68. https://www.sabix.org/bulletin/b19/gouzevitch.html.

Gray, Jeremy. 2005. Jean Victor Poncelet, *Traité des propriétés projectives des figures*, first edition (1822). In *Landmark writings in Western mathematics, 1640-1940*, ed. Ivor Grattan-Guinness, 366–390. Amsterdam: Elsevier.

———. 2007. *Worlds Out of Nothing. A Course in the History of Geometry in the 19th Century*. London: Springer.

Guyot de Fère, François-Fortuné. 1858. L'abbé Nicolas Halma. In *Nouvelle Biographie Générale depuis les temps les plus reculés jusqu'à nos jours*, ed. Le Hoefer, 200–203. Paris: Firmin Didot.

Hachette, Jean Nicolas Pierre, ed. 1808. *Correspondance sur l'Ecole Impériale Polytechnique à l'usage des élèves de cette école. Tome premier: 1804 (April)-1808 (March)*. Paris: Chez Bernard.

———, ed. 1814. *Correspondance sur l'Ecole Polytechnique à l'usage des élèves de cette école. Tome Second: 1809 (January)-1813 (January)*. Paris: Me Veuve Courcier.

———, ed. 1816. *Correspondance sur l'Ecole Royale Polytechnique à l'usage des élèves de cette école. Tome troisième: 1814 (January)-1816 (January)*. Paris: Me Veuve Courcier.

Halma, Nicolas. 1813–1816. *Composition Mathématique de Claude Ptolémée, traduite pour la première fois du grec en français, sur les manuscrits originaux de la Bibliothèque Impériale* *de Paris par M. Halma et suivie des notes de M. Delambre*: Vol. 1: chez H. Grand. Vol. 2: imprimerie de J.M. Eberhart, imprimerie du collège royal de France.

Laboulais, Isabelle. 2014. La bibliothèque de l'École des mines, lieu de savoir et lieu de mémoire pour les ingénieurs. *Revue de la BNU* 10: 44–55. https://doi.org/10.4000/rbnu.1627. http://journals.openedition.org/rbnu/1627.

Langins, Janis. 1989. Histoire de la vie et des fureurs de François Peyrard, bibliothécaire de l'Ecole polytechnique de 1795 à 1804 et traducteur renommé d'Euclide et d'Archimède. *SABIX* 3: 2–12. https://www.sabix.org/bulletin/b3/peyrard.html;https://journals.openedition.org/sabix/556.

Pepe, Luigi. 1996. La formazione della biblioteca dell' Ecole Polytechnique. Il contributo involontario del Belgio e dell'Italia. *Bollettino di storia delle scienze matematiche* 16: 155–198.

———. 1997. Gaspard Monge: un matematico nella storia delle grandi biblioteche italiane (1796-1798). *Bollettino di Storia delle Scienze Matematiche* 17: 155–187.

Peyrard, François. 1804. *Les élémens de géométrie d'Euclide, traduits littéralement et suivis d'un Traité du cercle, du cylindre, du cône et de la sphère, de la mesure des surfaces et des solides, avec des notes, par F. Peyrard, Bibliothécaire de l'Ecole Polytechnique. Ouvrage approuvé par l'Institut National*. Paris: F. Louis.

———. 1807. *Œuvres d'Archimède, traduites littéralement, avec un commentaire par F. Peyrard, Professeur de Mathématiques et d'Astronomie au Lycée Bonaparte, suivies d'un Mémoire du traducteur sur un nouveau miroir ardent et d'un autre Mémoire de M. Delambre sur l'arithmétique des Grecs. Ouvrage approuvé par l'Institut et adopté par le gouvernement pour les bibliothèques des lycées. Dédié à Sa Majesté l'Empereur et Roi*. Paris: F. Buisson.

———. 1808. *Œuvres d'Archimède, traduites littéralement, avec un commentaire par F. Peyrard, Professeur de Mathématiques et d'Astronomie au Lycée Bonaparte, suivies d'un Mémoire du traducteur sur un nouveau miroir ardent et d'un autre Mémoire de M. Delambre sur l'arithmétique des Grecs. Ouvrage approuvé par l'Institut et adopté par le gouvernement pour les bibliothèques des lycées. Dédié à Sa Majesté l'Empereur et Roi. Seconde édition, 2 volumes*. Paris: F. Buisson.

———. 1809. *Les élémens de géométrie d'Euclide, traduits littéralement et suivis d'un Traité du cercle, du cylindre, du cône et de la sphère, de la mesure des surfaces et des solides, avec des notes, par F. Peyrard, professeur de mathématiques et d'astronomie au Lycée Bonaparte. Seconde édition augmentée du cinquième livre. Ouvrage approuvé par l'Institut et adopté par le gouvernement pour les bibliothèques des lycées.* 2nd edition. Paris: F. Louis.

———. 1814–1818. *Les œuvres d'Euclide, en grec, en latin et en français, d'après un manuscrit très-ancien qui était resté inconnu jusqu'à nos jours, par F. Peyrard, traducteur des œuvres d'Archimède. Ouvrage approuvé par l'Institut de France. Dédié au Roi*. 3 volumes. Paris: Patris.

Poncelet, Jean-Victor. 1822. *Traité des propriétés projectives des figures; ouvrage utile à ceux qui s'occupent des applications de la géométrie descriptive et d'opérations géométriques sur le terrain*. Paris: Bachelier, libraire, quai des Augustins.

———. 1862. *Applications d'analyse et de géométrie qui ont servi de principal fondement au Traité des propriétés projectives des figures, comprenant la matière des sept cahiers manuscrits rédigés à Saratoff dans les prisons de Russie 1813 à 1814 et accompagnés de divers autres écrits, anciens ou nouveaux, annotés par l'auteur et suivis d'Additions par MM. Mannheim et Moutard, anciens élèves de l'École polytechnique. Tome 1*. Paris: Mallet-Bachelier.

———. 1864. *Applications d'analyse et de géométrie qui ont servi de principal fondement au Traité des propriétés projectives des figures, avec Additions par MM. Mannheim et Moutard, anciens élèves de l'École Polytechnique. Tome deuxième et dernier*. Paris: Gauthier-Villars.

———. 1865. *Traité des propriétés projectives des figures; ouvrage utile à ceux qui s'occupent des applications de la géométrie descriptive et d'opérations géométriques sur le terrain. Tome premier. Deuxième édition, revue, corrigée et augmentée d'annotations nouvelles*. Paris: Gauthier-Villars, Imprimeur-libraire.

———. 1866. *Traité des propriétés projectives des figures; ouvrage utile à ceux qui s'occupent des applications de la géométrie descriptive et d'opérations géométriques sur le terrain.*

Tome second. Deuxième édition, revue par l'auteur et augmentée de sections et d'annotations nouvelles ou jusqu'ici inédites. Paris: Gauthier-Villars, Imprimeur-libraire.

Quetelet, Adolphe. 1839. Notice sur Jean-Guillaume Garnier. *Annuaire de l'Académie royale des sciences et belles-lettres de Bruxelles* 5: 161–207.

Sakarovitch, Joël. 1998. *Epures d'architecture, de la coupe des pierres à la géométrie descriptive, XVIe-XIXe siècles*. Basel: Birkhäuser.

Servois, François-Joseph. 1814–1815. Philosophie mathématique. Réflexions sur les divers systèmes d'exposition des principes du calcul différentiel, et, en particulier, sur la doctrine des infiniment petits. *Annales de Mathématiques pures et appliquées* 5: 141–170.

Taton, René. 1970–1980a. Français, François (Joseph) and Français, Jacques Frédéric. In *Dictionary of Scientific Biography*, ed. C. C. Gillispie, 110–112.

———. 1970–1980b. Servois, François-Joseph. In *Dictionary of Scientific Biography*, ed. C. C. Gillispie, 325–326.

Wroński, Józef Maria Hoëné de. 1811. *Introduction à la philosophie des mathématiques, et technie de l'algorithmie*. Paris: Courcier.

Wang Xiaofei. 2017. The Teaching of Analysis at the École Polytechnique 1795-1809. Thèse de doctorat, Université Paris Diderot Paris 7 (Université de Paris).

———. 2020. Greek Texts and the Rigorization of Analysis: An Inquiry into J. L. Lagrange's Work on the History of Mathematics. *Chinese Annals of History of Science and Technology* 4: 139–165. https://doi.org/10.3724/SP.J.1461.2020.01139.

———. 2022. How Jean-Baptiste Delambre Read Ancient Greek Arithmetic on the Basis of the Arithmetic of "Complex Numbers" at the Turn of the 19th Century. *Historia mathematica* 59: 146–163. https://doi.org/10.1016/j.hm.2020.12.002.

Chapter 3
Advice to a Young Mathematician Wishing to Enter the History of Mathematics

Lizhen Ji

Abstract In this chapter, I will try to explain some basic issues in doing the history of mathematics, which seem almost self-evident though often not easy to be followed or followed consciously (or conscientiously), and then illustrate the importance of some of these points by examining several well-known historical studies on the famous testamentary letter of Galois written on the eve of his fatal duel, showing some damaging consequences for violating these basic points. We conclude with the notion of the spacetime of mathematics which can provide one common framework for multiple approaches to the study of the history of mathematics and illustrate its use by two examples concerning some works of Poincaré and Hilbert.

3.1 Introduction: Reasons for Writing This Chapter

After spending many years in the world of mathematics, I want to understand better some global and historical aspects of mathematics by doing some research on the history of mathematics.[1] Several years ago, I started to look for books and papers which can serve as helpful guides to young people and mathematicians who want to learn proper ways to do some good research on the history of mathematics. After a period of relatively extensive search through various sources and asking multiple historians of mathematics whom I could have a chance to meet or interact, I could

[1] Partly motivated by the metaphor: if one has lived for life in a big city with a long history such as London, Paris, Beijing, and Rome or Xi'an, it is reasonable to expect that this person should know something about the major historical landmarks and main roads connecting them in the city, and some important events in the history of the city and their impacts on the development of the city, in particular, their current impacts. Now replace the big city by mathematics.

L. Ji (✉)
Department of Mathematics, University of Michigan, Ann Arbor, MI, USA
e-mail: lji@umich.edu

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_3

not find many such writings on the historiography of mathematics, or rather the methodology of the history of mathematics. Partly because of this, during multiple meetings of the editors of this book *The Richness of the History and Philosophy of Mathematics*, I tried to emphasize the idea of asking the experts on the history and philosophy of mathematics who are contributing to this book to share their experiences and explain their perspectives on "what, why and how" about the history of mathematics. I strongly believe that this should be helpful to several groups of people, especially the young historians and mathematicians who are interested in history, since many problems about the history of more recent mathematics, for example, the late nineteenth century and twentieth century mathematics, need to be studied or understood better, hence contributions from the younger generations and mathematicians are needed. Naturally, to make it easier for these people to enter the field, it is probably helpful, or even definitely desirable, for them to learn from the experiences, not only the works, of the experts in the history of mathematics.

Probably because of my persistence on the above point, José Ferreirós suggested that I could try to write one such introductory paper first, which could then be improved by experienced experts. Though this is a challenging task, writing such an article is a good learning experience for me: not only the benefit from the demanding process of finding out what have been written already about basic and important problems in the historiography of mathematics but also a good opportunity to interact with the experts in history with concrete questions in mind. So I took the challenge and wrote an article titled *Some Perspectives on the History of Mathematics: What, Why, and How* (Ji 2022). Since there are not many books and papers discussing basic issues in the methodology of the history of mathematics,[2] I tried to look up as many books as possible on the methodology of general history and related philosophy books. After all, the history of mathematics should have shared some similarities with the general history about basic questions and big issues.[3] Both for my own education and for the convenience of the possible readers, I quoted extensively from various writings in order to convey more balanced views of many experts, for example, to avoid the danger of distorted meanings of chopped quotes. After working on it intensively and exclusively over a period of several months, I made the article definitely too long to be included in this book (for example, the references occupy over 17 pages). Several people suggested to pick out only certain parts and produce a much shorter and more focused paper. This paper is the outcome of this shortening process. In some sense, it is a different and improved paper.

[2] See some illuminating quotes from several influential writers on historiography to explain why working historians often do not want to talk too much about "what, why, and how" of their subjects in Ji (2022, §1.2).

[3] Maybe the following description of the historian of mathematics is reasonable: a historian of mathematics is a historian who studies the history of mathematics. Therefore he/she needs to know basic techniques in the study of history, and also has a good command of mathematics topics under study. Other subjects such as philosophy are also needed, and he/she should be able to write up his/her research results in an authentically historical, yet accessible, way which are of historical value to the readers.

José Ferreirós originally suggested a title *Advice to a young mathematician wishing to enter the history and philosophy of mathematics* for my article, which was also endorsed by Erhard Scholz. In this reduced version, I took their suggested title except dropping one key word "philosophy", since I do not know enough about philosophy even after some struggles with it and hence could not say much about it which might be useful.

After writing this lengthy article (Ji 2022), thinking over various issues discussed in it and reading more books and articles on the methodology of history, I had a revelation recently: Many points I studied and quoted in Ji (2022) are all reasonable and almost self-evident, even though some of them may have not been discussed so explicitly or naively by experts in history. But *What is difficult* is *how to make them a part of working habit*[4] *and to apply them to do research on some concrete and solid problems in the history of mathematics.*[5]

One point emphasized in Ji (2022) is the dialectic perspective: instead of choosing sides between opposing views, one needs to take all of them into consideration without going to extremes in following either of them and to *combine* them in fruitful ways. To achieve this, it is certainly important to learn and use various basic points in the historiography of mathematics, while keeping in mind that though having some guiding principles is important, one should not be bounded by them absolutely or pursue the extreme cases allowed by them. For example, to describe a mathematical result or an event adequately and in a satisfying way, one needs to have a global or overall understanding of many relevant mathematics topics and related social, cultural, political, and philosophical factors. On the other hand, one also needs to analyze some particular things in minute detail. This involves the dialectic pair of global studies versus local studies, and how their interaction can produce a fair and complete picture of the issue under discussion. Probably it is safe to say that very few historians will disagree with the above points. On the other hand, how many people including the leading mathematicians in the world can claim to have a good or perfect overall understanding of the whole of mathematics, and how many specialists can have absolute and very detailed command of their special subjects or even topics? The above dialectic perspective suggests that we do not need to go to such extremes. It also suggests that even if we cannot achieve either of them, it does not mean that we should not try as much as possible to move in the directions suggested by them (i.e., pursuing both breadth and depth) and to *combine them in suitable ways* (for example, how different topics or subjects interact with each

[4] Maybe one meaningful comparison can be made to driving. A beginning driver needs to take some driving lessons in order to start and to learn some basic rules. But continuing driving is essential, and after a while, the driver will not even need to think about how one should drive anymore, since it has become a habit.

[5] I am also more convinced than before, when I wrote Ji (2022), that a beginning historian should spend some time to think about the basic and big issues about the historiography of mathematics, but probably not too much of it. Is it similar to our life? A moderate amount of religious and philosophical meditation will probably help one to live a more balanced and satisfying life. But the key to life is to live it, ideally an enlightened life.

other, and how some details in a proof of a theorem turn out to have a big impact on mathematics, while keeping in mind multiple famous lemmas in mathematics which outshone major works of their creators!) in order to understand better the problem under study.

Since the dialectic perspective is crucial to this chapter, we will need to explain more precisely what we mean by it. It contains, but is more than, "the *Golden Mean*" of Aristotle, i.e., finding a balance between two extremes, which is equivalent to the *Doctrine of the Mean* (*Zhong Yong Zhi Dao*), one of the most basic principles of the ancient Chinese philosophy. One *additional* crucial aspect of the dialectics is the *combination* or *synthesis* of two extremes (the thesis and antithesis). This will also be seen in other examples of this chapter.

Although I have very little experience in the study of history, I wish that some people, books or papers had given me the following pieces of advice when I first started to learn how to do research in the history of mathematics:

1. Have some solid understanding of *What is the history of mathematics.*
   The history of mathematics is about *the development of mathematics*. Good research works in the history of mathematics can and should provide better understanding about this development. Yes, there are many important aspects in describing the development of mathematics: mathematical, social, cultural, political, and philosophical etc; and there are also many ways to increase our knowledge about this development.
   In spite of the relevance and importance of all other factors, mathematics itself should be one essential part. Mathematical discussions can be simplified and discussed tangentially, but what is provided should be correct, and hopefully clear and relatively easy to be understood too. Each historical research project and writing can focus on some aspects (mathematical, social, cultural, political, or philosophical), but in the total sum of works on the history of mathematics, mathematical discussions should have the leading role. Otherwise, it is probably not the proper history of mathematics, at least not the complete picture of it, or as my colleague Jamie Tappenden told me, the right English phrase to describe the situation would be: "Hamlet without the Prince".[6]
2. Think about *Why one wants to do some research project in the history of mathematics*.
   Every person has different motivations for doing each particular research project and for spending time and effort to write up their discoveries or results. But several obvious questions are still important and should be kept in mind all the time: How can it contribute to expanding our knowledge of the history of mathematics in some important ways, for example clarifying the evolution process of a particular theory or a subject? Or how does it fill in some gaps in our knowledge, especially crucial gaps, about the development of mathematics?

[6] In a conversation with Chang Wang, I described the situation through a cake and its toppings: even though rich toppings greatly enhance a cake, the cake itself is still the most essential part which makes it a cake.

How does it give better explanation than existing ones of the historical event under discussion? Why do people want to read your papers or books? Or to a certain extent, why is it interesting for others to read them? More importantly, how can people benefit from reading them?, and so on and so forth.

Of course, all these questions should be taken in a dialectic perspective. If a problem is really interesting or justifiable to yourself, the other factors and considerations[7] probably do not matter too much, though they should not be completely ignored either.

3. Take a *Dialectic Perspective* towards both the development of mathematics and various aspects of and approaches to the study of the history of mathematics.

   Indeed, the dialectic pairs such as discrete versus continuous, local versus global, provide illuminating guides to understand essential parts of the development of mathematics. For example, the evolution of calculus manifests well the dialectic interaction between the local (differentiation) and the global (integration), and their synthesis (the fundamental theorem of calculus). This dialectic pair of local and global also clarifies the structure or nature of a majority of works in the modern and contemporary global differential geometry: the local (curvature assumptions) and the global (topological invariants and global geometric properties), and the interaction between them (most significant results such as the Gauss-Bonnet theorem and the more recent Atiyah-Singer theorems).

   More globally, people often say that mathematics studies numbers and shapes, in particular it has two important aspects: quantitative and qualitative. Of course, each part is crucial to mathematics, but the interaction between them is probably the more interesting and essential driving force behind the rich development of mathematics through out its long history.[8]

   In the methodologies of the history of mathematics, there are also dialectic pairs of the internal history versus the external history,[9] and of the view from the

[7] See Ji (2022, Section 2) for many uses of the history of mathematics with concrete examples.

[8] The ubiquity and importance of dialectic processes can be seen through the most basic and important process in nature and in the world: the life reproduction. The synthesis of the male and female produces a new dialectic pair of the male and female, and this process repeats itself with possible improvement for their best chances of survival and reproduction. It is a bit surprising that such a fundamental dialectic process is not mentioned explicitly in many writings about dialectics, for example, in the famous book *Dialectics of Nature* by Frederich Engels (1940).

[9] There are many examples which show the influence of the interaction between the internal and external factors on the development of mathematics. As it is well-known, Riemann's habilitation lecture *On the hypothesis which lies at the bases of geometry* initiated the subject of Riemannian geometry and had also a huge impact on Einstein's general theory of relativity. But this was the third, or the last, topic which Riemann presented to the committee chair, Gauss, and Gauss picked the least expected one in the eyes of Riemann, or the least prepared topic by Riemann. One can analyze the cultural and social contexts for the custom of the requirement at that time that each candidate needed to provide three possible topics for the habilitation lecture. Though it might look like a fruitful accident, accidental causes are one important part of history (Carr 1961, Chap. IV). Of course, Gauss must have his own mathematical reasons for making this choice. In any case, all these considerations, both mathematical and non-mathematical, make the history of Riemannian

present versus the view from the past itself. They all need to be taken into consideration and combined in fruitful ways. Otherwise, bad consequences might arise.

I hope that these obvious points can be relevant and interesting to others too. As in life, people often overlook the most basic and obvious things. We will make some of the above discussions more precise and describe a few more concrete points which might be considered as basic requirements in doing research on the history of mathematics. Whenever possible, relevant examples are provided. After summarizing comments by mathematicians and historians on various aspects of history in Sect. 3.2, we examine some historical studies on Galois' testamentary letter to highlight the importance of using the primary sources in Sect. 3.3. Then in Sect. 3.4, we explain the key components of historical thinking. Finally we introduce the notion of spacetime of mathematics in Sect. 5 and illustrate its use to get new perspectives on works of Poincaré and Hilbert.

3.2 Comments by Mathematicians and Historians on What, Why, How of History

Given that the history of mathematics is a long one, it is expected that some major mathematicians and historians of mathematics have written about several basic issues in the study of the history of mathematics. It is not surprising that each has different perspectives and some of them have expressed their opinions strongly and sharply. It is not surprising either that mathematicians and historians of mathematics have vastly different or even seemingly conflicting opinions. Similarly, even for some historians of mathematics, there is a strong division between the so-called internal history and external history, though this division is not clearly defined. From the dialectic perspective emphasized in this chapter, we will try to present an overview of them by making long quotes from major representatives so that the readers can understand better the precise meaning of opinions of the people from whom we have quoted.[10]

geometry richer and more interesting. More importantly, they can also explain why something happened in the ways it did.

[10] It seems that the general advice given to young historians in books on the methodology of history is to quote sparingly or briefly by incorporating others' opinions into the text through either summarizing or paraphrasing others' writings. In this chapter, I am going against this instruction because short quotes and paraphrasing will often distort the original meanings and hence can cause misunderstanding, which can be seen in the discussions about Galois' testamentary letter in Sect. 3.3 below. This is also consistent with the basic principle in the study of history of using the primary sources or the original writings as much as possible. I do not wish to express my opinions either on these quotes representing different perspectives about history, since it seems important that the readers need to understand all of them and then make up their own decisions about the basic issues in the history of mathematics and how to make use of the suggestions in the quotes. Of course, this is the dialectic view that we want to emphasize in this chapter.

Before quoting from writings of these major figures, it is helpful to give a *definition*, or rather *a description, of the history of mathematics*, which is more concrete and detailed than what we mentioned briefly in the introduction. Since I could not find a reasonable definition of the history of mathematics in books and papers which I have checked, I will quote a definition of the general history. Among all the books I have checked, it seems that the following from *Encyclopedia Britannica* in 1780 is probably still the most appropriate: *concise and to the point, yet rich in content.* Maybe this is why it was selected for inclusion in a recent well-respected book *The Modern Historiography Reader: Western Sources* (Budd 2010, 72):

> History, in general, signifies an account of some remarkable facts which have happened in the world, arranged in the true order in which they actually took place, together with the causes to which they were owing, and the different effects they have produced, as far as can be discovered.

The reader can translate this into a suitable definition for the history of mathematics. Though different people will probably make different translations, their basic contents should be similar. It is also important to emphasize that such a definition can only serve as a general guide, providing some helpful directions and reminding people of what are important when working on problems in history. Hence one should *take this definition dialectically* in the sense explained in the introduction.

Views of Mathematicians

Among mathematicians with very keen interest in history, Andre Weil is one major representative. Some of his opinions are expressed in Weil (1980, 1975). In Weil (1980, 228–230), Weil wrote:

> as Moritz Cantor observed, one may, in dealing with mathematical history, regard it as an auxiliary discipline, meant for providing the true historian with reliable catalogues of mathematical facts, arranged according to times, countries, subject-matters and authors. It is then a portion, and not a very significant one, of the history of techniques and crafts, and it is fair to look upon it entirely from the outside....
>
> From another point of view, mathematics may occasionally provide the cultural historian with a kind of "tracer" for investigating the interaction between various cultures. With this we come closer to matters of genuine interest to us mathematicians; but even here our attitudes differ widely from those of professional historians. To them a Roman coin, found somewhere in India, has a definite significance; hardly so a mathematical theory....
>
> Now, leaving the views and wishes of laymen and of specialists of other disciplines, it is time to... consider the value of mathematical history, both intrinsically and from our own selfish viewpoint as mathematicians... we may say that its first use for us is to put or to keep before our eyes "illustrious examples" of first-rate mathematical work. Does that make historians necessary? Perhaps not....
>
> The historian can help in still another way. We all know by experience how much is to be gained through personal acquaintance when we wish to study contemporary work... to us their biographies are of no small value in bringing alive the men and their environment as well as their writings....
>
> Mathematical strategy is concerned with long-range objectives; it requires a deep understanding of broad trends and of the evolution of ideas over long periods. This is almost indistinguishable from what Gustav Eneström used to describe as the main object of mathematical history, viz., "the mathematical ideas, considered historically", or, as Paul Tannery put it, "the filiation of ideas and the concatenation of discoveries". There we have the core of the discipline we are discussing, and it is a fortunate fact that the aspect towards

> which, according to Eneström and Tannery, the mathematical historian has chiefly to direct his attention is also the one of greatest value for any mathematician who wants to look beyond the everyday practice of his craft.

For Weil, the task of historians of mathematics is clear (Weil 1980, 231, 234):

> once we have agreed that mathematical ideas are the true object of mathematical history, some useful consequences can be drawn... the ability to recognize mathematical ideas in obscure or inchoate form, and to trace them under the many disguises which they are apt to assume before coming out in full daylight, is most likely to be coupled with a better than average mathematical talent....
> One important task of the serious historian of mathematics, and sometimes one of the hardest, is precisely to sift such routine from what is truly new in the work of the great mathematicians of the past.

Although Weil was a major mathematician in the twentieth century, he is often considered as a major historian of mathematics too, which can be testified by his writings on the history of mathematics. Then one question is: What is the opinions of some other major mathematicians on the history of mathematics?

In an interview (Chong and Leong 1986, 12), Serre said:

> I am already interested [in the history of mathematics]. But it is not easy; I do not have the linguistic ability in Latin or Greek, for instance. And I can see that it takes more time to write a paper on the history of mathematics than in mathematics itself. Still, history is very interesting; it puts things in the proper perspective.[11]

Perhaps this represents one view towards history of a definite percentage of working mathematicians: respect and appreciation for the history of mathematics.

The following quote of Milnor from an interview Raussen and Skau (2012, 406) shows that he was fully aware of the bigger scope of history than the so-called Whig history:

> I certainly enjoy trying to track down just when and how the ideas that I work with originated. This is, of course, a very special kind of history, which may concentrate on obscure ideas which turned out to be important, while ignoring ideas which seemed much more important at the time. History to most scientists is the history of the ideas that worked. One tends to be rather bored by ideas that didn't work. A more complete history would describe how ideas develop and would be interested in the false leads also. In this sense, the history I would write is very biased, trying to find out where the important ideas we have today came from, who first discovered them. I find that an interesting subject.

In the preface to Atiyah's collected works published in China, which was arranged by Shiing-Shen Chern for the sake of Chinese students and mathematicians, Chern wrote:

> No matter how refined or improved a new account is, the original papers on a subject are usually more direct and to the point. When I was young, I was benefited by the advice to read Henri Poincaré, David Hilbert, Felix Klein, Adolf Hurwitz, etc. I did better with

[11] The question from the interviewers is: "Do you think that you will ever be interested in the history of mathematics?", which followed the earlier question and answer: "Q: You mentioned papers which have been forgotten. What percentage of the papers published do you think will survive? A: A non-zero percentage, I believe. After all, we still read with pleasure papers by Hurwitz, or Eisenstein, or even Gauss."

> Wihelm Blaschke, Elie Cartan and Heinz Hopf. This has also been in the Chinese tradition, when we were told to read Confucius, Han Yu in prose, and Tu Fu in poetry.[12]

The reader might have noticed that in the above quotes we have tried to mix up on purpose two aspects of the history of mathematics: writings by past great mathematicians and historical studies about them.

Since the history of mathematics is often compared with the history of science, in particular, the history of physics, it will be interesting to compare the above opinions of mathematicians with views of some major physicists on history.

The distinguished physicist, Nobel Laureate, Steven Weinberg had a distinct view of the history of science (Weinberg 2004, 5), probably different from many scientists, but very close to that of many hisotrians:

> I have been asked to review the history of the formation of the Standard Model. It is natural to tell this story as a sequence of brilliant ideas and experiments, but here I will also talk about some of the misunderstandings and false starts that went along with this progress, and why some steps were not taken until long after they became possible. The study of what was not understood by scientists, or was understood wrongly, seems to me often the most interesting part of the history of science. Anyway, it is an aspect of the Standard Model with which I am very familiar, for as you will see in this talk, I shared in many of these misunderstandings.

In an article titled “Keeping an Eye on the Present” (Weinberg 2018, 58), Weinberg wrote:

> The historian of science Bruce Hunt recalls that when he was in graduate school in the early 1980s, “whiggish” was a common term of abuse in the history of science. To avoid that charge, people turned away from telling progress stories or giving “big picture” stories of any kind, and shifted to accounts of small episodes, tightly focused in time and space.
> Nevertheless, in teaching courses on the history of physics and astronomy, and then working up my lectures into a book, I have come to think that whatever one thinks of whiggery in other sorts of history, it has a rightful place in the history of science.

This quote makes one wonder whether the “whiggish” history of more recent mathematics has “a rightful place” in the history of mathematics.

On more general related issues, Einstein wrote in a letter (Einstein 1944):

> I fully agree with you about the significance and educational value of methodology as well as history and philosophy of science. So many people today–and even professional scientists–seem to me like somebody who has seen thousands of trees but has never seen a forest. A knowledge of the historic and philosophical background gives that kind of independence from prejudices of his generation from which most scientists are suffering. This independence created by philosophical insight is–in my opinion–the mark of distinction between a mere artisan or specialist and a real seeker after truth.

[12] It is probably well-known that Confucius’ teachings collected in the *Analects* and his writings have greatly influenced the culture of China and nearby countries such as Japan and Korea. Han Yu is much less known in the West, but his writings have had a huge impact on the Chinese literary tradition, and he can be compared to Dante, Shakespeare or Goethe in the West. Tu Fu is often considered as one of the two greatest Chinese poets in history, the other being Li Bai.

This quote raises one question: Is there a similar "mark of distinction between a mere artisan or specialist and a real seeker after truth" in mathematics?

Views of Historians

Next, we take a look at opinions of known historians of mathematics on the history of mathematics. According to Hawkins (Hawkins 1987, 1642),

> The challenge to the historian is to depict the origins of a mathematical theory so as to capture the diverse ways in which the creation of that theory was a vital part of the mathematics and mathematical perceptions of the era which produced it.

This quote describes clearly one main task of historians of mathematics. But several questions arise naturally. Which kind of "mathematical theory" should historians consider in terms of subject matters? Furthermore, a past one or one in the contemporary mathematics? Or both? This is an important question, since it is very relevant to the question of its usefulness to the current working mathematicians.

Since "depict the origins of a mathematical theory" is going back in time, another question is whether we should go forward in time and consider impacts, both positive and negative, of "a mathematical theory". Such a view is natural when we consider the whole mathematics across all the time, and also across subject matters.

According to Grattan-Guinness, different choices can make a big difference. In the abstract of (Grattan-Guinness 2004a), Grattan-Guinness wrote:

> Mathematics shows much more durability in its attention to concepts and theories than do other sciences: for example, Galen may not be of much use to modern medicine, but one can still read and use Euclid. One might expect that this situation would make mathematicians sympathetic to history, but quite the opposite is the case. Their normal attention to history is concerned with heritage: that is, how did we get here? Old results are modernized in order to show their current place; but the historical context is ignored and thereby often distorted. By contrast, the historian is concerned with what happened in the past, whatever be the modern situation. Each approach is perfectly legitimate, but they are often confused.

The earlier quotes of Serre and Milnor shows that not all mathematicians belong to the category described by Grattan-Guinness in the above quote. There are probably mathematicians similar to Weinberg too with regard to the "history" of their subjects.

Grattan-Guinness explained his points in more details by introducing two key words: *history* and *heritage* (Grattan-Guinness 2004a, 164–165):

> By "history" I refer to the details of the development of N [which stands for "a theory (or definition, proof-method, technique, algorithm, notation(s), whole branch of mathematics, . . .)"]: its prehistory and concurrent developments; the chronology of progress, as far as it can be determined; and maybe also the impact in the immediately following years and decades. History addresses the question "what happened in the past?" and gives descriptions; maybe it also attempts explanations of some kinds, in order to answer the corresponding "why?" question... History should also address the dual questions "what did not happen in the past?" and "why not?"; false starts, missed opportunities..., sleepers, and repeats are noted and maybe explained. The (near-)absence of later notions from N is registered, as well as their eventual arrival; differences between N and seemingly similar more modern notions are likely to be emphasized.
>
> By "heritage" I refer to the impact of N upon later work, both at the time and afterward, especially the forms which it may take, or be embodied, in later contexts. Some modern

> form of N is usually the main focus, with attention paid to the course of its development. Here the mathematical relationships will be noted, but historical ones in the above sense will hold much less interest. Heritage addresses the question "how did we get here?," and often the answer reads like "the royal road to me." The modern notions are inserted into N when appropriate, and thereby N is unveiled (a nice word proposed to me by Henk Bos): similarities between N and its more modern notions are likely to be emphasized; the present is photocopied onto the past.
>
> Both kinds of activity are quite legitimate, and indeed important in their own right; in particular, mathematical research often seems to be conducted in a heritage-like way, whether the predecessors produced their work long ago or very recently. The confusion of the two kinds of activity is not legitimate, either taking heritage to be history (frequently the mathematicians' view–and historians' sometimes!) or taking history to be heritage (the occasional burst of excess enthusiasm by a historian); indeed, such conflations may well mess up both categories, especially the historical record.
>
> A philosophical difference is that inheritors tend to focus upon knowledge alone (theorems as such, and so on), while historians also seek motivations, causes, and understanding in a more general sense. The distinction sometimes made by historians of science between "internal" and "external" history forms part of this difference.

It does not seem obvious what is the percentage of mathematicians who are "inheritors" tending "to focus upon knowledge alone (theorems as such, and so on)" when they think about the development of mathematics.

It is perhaps helpful to note that there are also "motivations, causes, and understanding" inside the "internal" history of mathematics. This important point is consistent with what Wussing wrote in (Wussing 1991, 66):

> Without studying the history of problems and ideas, the picture of the history of mathematics would remain incomplete and basically incorrect.

Hawkins, Grattan-Guinness and Wussing can be considered as major representatives of recent historians of mathematics. It is perhaps helpful to quote from some earlier historians too. Many people will agree with the importance of knowing the mathematics topics involved in historical projects under study. But what Paul Tannery wrote in (Tannery 1930, 164) may sound somewhat too strong to some historians:

> The first unfavorable condition is that the history of a science can only be truly treated by a man really possessing this science as a whole, or, at the very least, able to go further by himself into all the scientific questions which he has to deal with during this history.... Moreover, a scientist can possess or acquire all the abilities necessary for the composition of an excellent history of science to which he has devoted himself, and furthermore, the more this scientist has talents, the more the value of his historical work will shine through to all eyes, this is a point that should not be doubted at all.

Such an opinion was supported by Weil (1980, 231) to a certain extent, though it is important to take these opinions dialectically:

> How much mathematical knowledge should one possess in order to deal with mathematical history? According to some, little more is required than what was known to the authors one plans to write about; some go so far as to say that the less one knows, the better one is prepared to read those authors with an open mind and avoid anachronisms. Actually the opposite is true. An understanding in depth of the mathematics of any given period is hardly ever to be achieved without knowledge extending far beyond its ostensible

> subject-matter. More often than not, what makes it interesting is precisely the early occurrence of concepts and methods destined to emerge only later into the conscious mind of mathematicians; the historian's task is to disengage them and trace their influence or lack of influence on subsequent developments. Anachronism consists in attributing to an author such conscious knowledge as he never possessed; there is a vast difference between recognizing Archimedes as a forerunner of integral and differential calculus, whose influence on the founders of the calculus can hardly be overestimated, and fancying to see in him, as has sometimes been done, an early practitioner of the calculus.

Note also an interesting description, or rather an example, of "Anachronism" by Weil in the above quote.

The importance of mathematics, hence also of mathematicians, was also explained by some historians of mathematics. For example, Grabiner wrote in (Grabiner 1975, 442, 443):

> We expect the historian to know the general history of a particular time as well as the mathematics of that time. He should have a sense of what it was like to be a person, not just a mathematician, at that time. Sometimes such knowledge has great explanatory value...
> Historians of mathematics should certainly know the mathematics whose history they are writing. But mathematicians are still needed – and not just because they know the mathematics better. Historians need the mathematician's point of view about what is mathematically important. The mathematician's work determines what it is that most needs a historical explanation. Only the mathematician can tell us which of a half-dozen contemporary concepts is really the crucial one, and which older concepts are worth looking into again – infinitesimals are one example... Furthermore, the mathematician has a better idea of the logical relationship between mathematical ideas, and can suggest connections to the historian which might not be apparent from the historical record alone.

Some Views of External History

The above discussions deals with some differences between mathematicians and historians. But even among historians of mathematics, there is often some difference between the so-called internal history and external history. Of course, this division is not accepted or liked by all historians. One concern is that there is probably no precise definition of internal or external histories. They need to be taken dialectically too, i.e., how the internal and external factors need to be both considered and how their interaction with each other needs to be taken into account, as we emphasized in this chapter. After all, some historical studies involve both aspects at the same time.

The above quotes from historians represent some views of the internal history of mathematics. Some opinions of the external history are expressed by Stedall (Stedall 2012, 110):

> Historians of mathematics have increasingly moved away from a purely 'internalist' view in which mathematical developments are seen to come about of their own accord, regardless of outside influences. As has now been shown over and over again in this book, mathematical activity has for centuries manifested itself in a variety of ways, all of them socially and culturally determined. We should not throw out the baby with the bathwater, however: mathematicians often devote themselves to a particular problem not because it might be useful or because anyone requires them to do so, but because the problem itself catches their imagination. This was precisely the case for Newton and Leibniz with the calculus, Bolyai and Lobachevskii with non-Euclidean geometry, or Wiles with Fermat's Last Theorem. In

> such cases, progress depends first and foremost on deep and concentrated engagement with the mathematics, and in that sense mathematical creativity can be said to be an internal process. But the mathematical questions that are considered important at a particular time or place, the way they have come to be there, the way they are understood and interpreted, are all influenced by a multitude of factors outside the mathematics itself: social, political, economic, and cultural. Context has become as important to the historian as content.
> Another significant change in recent years has been the growing recognition that the mathematics done by a small number of famous mathematicians has not reflected (though it has built on) the diversity of mathematical activity and experience at other levels of society... Historians of mathematics, like scholars in many other disciplines, have also become much more sensitive to questions of gender and ethnicity... Consequently, the mathematics of the past is no longer regarded simply as a precursor to the mathematics of the present but as an integral part of its own contemporary culture.

Some non-mathematical factors and their influence on mathematics are also described by Chemla in the abstract of Chemla (2018):

> This contribution argues that history of mathematics should take as its object not only knowledge, but also ways of doing mathematics that are collectively shared (what I call 'mathematical cultures'), and additionally the connections between the two. I provide evidence showing that there is a history of ways of doing mathematics, and this history suggests that mathematical knowledge takes shape at the same time as practices do. Indeed, ways of doing mathematics do not fall out of the sky. They are shaped and transformed by actors in the process of working out some problems and addressing some issues. They represent one of the outcomes of mathematical research. I further argue that attending to the mathematical culture in the context of which actors worked is essential for interpreting their writings.

Dialectic Views of Doing History

If we take a dialectic perspective towards these different views of history, we should say that they are all important and should be combined in suitable ways, and people from different groups should work together, or at least trying to understand and appreciate approaches different from their own favorite ones, instead of emphasizing their differences. This is exactly the point expressed by Weinberg in a recent book *Third Thoughts* Weinberg (2018, 58):

> Historians who have not themselves worked as scientists may feel that they cannot match the working scientist's understanding of present science. On the other hand, it must be admitted that a scientist like myself cannot match the professional historian's mastery of source material. So who should write the history of science, historians or scientists? The answer seems to me obvious: both.

Actually such a point of view was advocated by several distinguished historians of mathematics a long time ago. For example, Grabiner wrote in Grabiner (1975, 444):

> We have seen that the mathematician and the historian bring different skills and different perspectives to their common task of explaining the mathematical present by means of the past. Therefore, ... collaboration between mathematicians and historians can be fruitful. The value of such a collaboration will be enhanced if each collaborator understands the unique contributions which can be made by the other. The importance of the common task, I think, makes it well worth the collective efforts.

In (May 1975, 453), May wrote:

> Clearly in historical work the danger in missing the mathematical point is matched by the symmetric hazard of overlooking a historical dimension. The mathematician is trained to think most about mathematical correctness without a time dimension, i.e., to think ahistorically. Of course it is interesting to know how a historical event appears when viewed by a twentieth century mathematician. But it is bad history to confuse this with what was meant at the time. The historian concentrates on significance in the historical context and on the historical relations between events. And this is equally interesting to the mathematician who wishes to understand how mathematics actually developed.
> One could continue indefinitely, but the essential point is that the best history requires sensitivity to both mathematical and historical issues, a respect for good practice of the crafts of both the historian and the mathematician. It may even be that the best mathematical research is aided by an appreciation of historical issues and results. I know of many instances and hope that the work of historians may contribute to increasing their frequency.

Therefore, it is perhaps better not to contrast or differentiate historians and mathematicians too much. In an email on May 24, 2019, Joseph Dauben wrote:

> I think you will find that a consensus emerges that the best history of mathematics is probably produced by mathematicians and historians of mathematicians working together. This is what has happened in my own work.

Conclusion The above discussions show that even though mathematicians and historians have different opinions about the history of mathematics, they are fully aware of the values of the other sides. On the other hand, they are convinced of the importance of their own approaches and perspectives, and often prefer to follow their own ways. The ideal solution is that they can be combined in fruitful ways which are beneficial to both parties. This suggestion was raised almost half-century ago by historians of mathematics. Hope that it can be more widely accepted by future generations of the people who are interested in the history of mathematics.

3.3 Reexamination of Some Known Historical Studies of Galois' Testamentary Letter

In the previous section, we have seen different opinions of mathematicians and historians about the history of mathematics and how research on history should be done. We strongly believe that all these contrasting approaches should be balanced and combined, and descriptions of the development of mathematics itself should be an essential part. We also believe that philosophy plays an important role in the study of history, as expressed in the famous aphorism attributed to Imre Lakatos: "Philosophy of science without history of science is empty; history of science without philosophy of science is blind."[13] Otherwise, the outcome will very likely be incomplete, unfair or even misleading. To show the above points, we will reexamine some well-known historical studies of Galois' testamentary letter written on the eve of his fatal duel.

Before commenting in some detail on these studies, we need to think about and keep the following questions in mind:

1. *What are some really important historical questions about Galois' testamentary letter?*
2. *Why do we want to understand better Galois' testamentary letter?*
3. *How can we properly understand Galois' testamentary letter?*

These are three basic philosophical questions. Given the importance of Galois' contributions to mathematics, in particular the notion of groups and the Galois theory, his unusually short life and tragic death, it is not surprising that there have been many writings about him, for example, multiple books about his life. It is probably very fair to say that if Galois' mathematical works did not have the impacts on the development of mathematics as we know it now, not many people (historians, mathematicians, or the public) would be interested in his life and stories about him, no matter how tragic and unusual his life was. These reasons explain why people want to understand Galois and his works better. They are related to Question (2) above, but the relation depends on what is contained in the letter and why Galois wrote this letter. We will make some comments on these issues, and also on Questions (1) and (3) near the end of this section.

Among all the writings about Galois, probably the most famous and widely read is the story about him by Bell in the famous book *Men of Mathematics* (Bell

[13] This sharp statement was quoted in a paper of Hanson (1963, 458). It is perhaps helpful to explain more concretely what is philosophy. If a person works very hard on something for a long time, say pulling a heavy cart through a forest or onto the top of a mountain in a hot day, it seems reasonable, important, or even necessary, that he/she should make some stops periodically, resting for a while, looking around himself (both in space and time), and also try to understand what he is pulling, think about why he is doing it, where he/she came from and where he/she is going etc. This is philosophy!

There are two kinds of philosophies which are closely related to the history of mathematics: (1) the philosophy of mathematics, and (2) the philosophy of the history of mathematics.

It seems that (1) is parallel to the history of mathematics and hence at the same level. When one writes about the history of mathematics, the philosophy of mathematics often comes in naturally, since it has influenced the thinking and working of some mathematicians, usually great or reflective mathematicians. Hence this philosophy is particularly important for histories dealing with such mathematicians. On the other hand, the philosophy of the history of mathematics is in some sense one level above the history of mathematics and provides some guidance to historians of mathematics, for example, what kinds of problems should be considered, what is important in the history of mathematics, and why something should be written. Of course, there is often overlap between the above two kinds of philosophies.

The above comments help explain the first half of Lakatos' aphorism. The second half can also be seen through how mathematics and its history have been used in a substantial way in the long development of philosophy.

1937). On the other hand, many historians of mathematics have criticized Bell for his story about Galois. It seems that the most authoritative, or well-regarded, critic is Tony Rothman, whose original writings appeared in Rothman (1982a) and Rothman (1982b), which were later expanded in Rothman (1989, 148–200). The paper (Rothman 1982b) was reviewed by the distinguished mathematician with strong interest in the history of mathematics, Dieudonné, in MathSciNet and highly praised by him:

> this excellent article, in presenting with admirable simplicity and conscientiousness everything that is known about Galois's life from the documents we possess... Without diminishing the depth of Galois's work, nor disguising in what ways it was ahead of its time, the author demonstrates to what extent an incredible succession of misfortunes and setbacks could have exacerbated Galois's unyielding and touchy disposition, and shows that one has to soften a bit the summary judgment that his tragic fate is chiefly due to the fact that the scientific personalities of the era were persecuting a misunderstood genius. May this exemplary work give pause to amateurs of "novelistic" lives and incite them to exercise their critical faculties a little more frequently.

The updated version in Rothman (1989, 148–200) is the basis for the biographic exposition about Galois in the recent book Barrow-Green et al. (2022, 554), where on page 567, the authors commented: "the one [essay] on Galois [by Rothman] ... is a fine antidote to the hyperbole that surrounds him." In probably the most popular textbook on the history of mathematics (Katz 2009, 762) in the past few decades, Katz wrote: "The best recent article on the life of Galois is T. Rothman," (Rothman 1982b). This paper (Rothman 1982b) is also cited in Neumann (2011, 383) and seems to be endorsed there too.

These works of Rothman and his conclusions are also used by other experts in history, for example by Kragh in (1987). As mentioned earlier, there are not many writings about the historiography of mathematics, i.e., the methodologies of the history of mathematics. In fact, there are not many books about the historiography of science either, and the book Kragh (1987) is the only systematic exposition about it, which seems to include the historiography of mathematics as one part of it. Its discussion of Galois' testamentary letter is one of the few places in this book where it deals with the history of mathematics and appears in Chapter 15: *The biographic approach* (Kragh 1987, 168–170):

> Biographies of eminent, individual scientists are one of the oldest forms of history of science... Biographical works are still, however, an important part of history of science and they will remain so... As to romance, few cases in the history of science can equal the tale of the death of the French mathematician Évariste Galois... Even in the prison did Galois continue to develop his mathematical ideas, later known as group theory.

Then Kragh quoted from Rothman (1982a, 136), where Rothman in turn quoted from Bell (1937, 375):

> The night before the duel Galois 'spent the fleeting hours feverishly dashing off his scientific last will and testament... What he wrote down in those last desperate hours before the dawn will keep generations of mathematicians busy for hundreds of years'.

After this, Kragh drew his conclusion (Kragh 1987, 170):

> Unfortunately for romantics, the story is largely a myth... Galois's alleged 'scientific last will and testament' is a legend: the night before the duel Galois was indeed occupied with mathematics but actually of a rather trivial sort, viz. making editorial corrections on manuscripts. The destruction of the Galois myth yields a more authentic history without diminishing the scientific originality of Galois. If it makes his biography a little less exciting, this is a cost which should not be regretted.

This seems to be influenced by what Rothman wrote in Rothman (1982a, 149):

> Galois's mathematical writings the night before the duel were actually confined to making editorial corrections on two manuscripts and to summarizing the contents of these and one other paper in a long letter to Chevalier. The first paper was the one rejected by Poisson; the second was a fragmentary version of an article that had already been published in Ferussac's *Bulletin*. The third has not been found and its content is known only from the summary in the letter; it apparently concerned integrals of general algebraic functions.

Note the very certain, but harsh, judgement by Kragh: "the night before the duel Galois was indeed occupied with mathematics but actually of a rather trivial sort, viz. making editorial corrections on manuscripts." We would like to argue that the violation of the *basic standing principle in history of checking against the primary sources* by both Rothman and Kragh caused more damage to the "authentic history" than other aspects of the myth caused by Bell and others, and is unfair to Galois to a serious degree, as far as the history of mathematical ideas is concerned.

It is probably a pity that Kragh quoted too little from Rothman (1982a, 136), and Rothman also quoted not sufficiently from Bell (1937, 375) either. To compare them, let us see what Rothman quoted in Rothman (1982a, 136) and similarly in a related paper Rothman (1982b, 84):

> All night long he had spent the fleeting hours feverishly dashing off his scientific last will and testament, writing against time to glean a few of the great things in his teeming mind before the death which he saw could overtake him. Time after time he broke off to scribble in the margin "I have not time; I have not time," and passed on to the next frantically scrawled outline. What he wrote in those last desperate hours before the dawn will keep generations of mathematicians busy for hundreds of years. He had found, once and for all, the true solution of a riddle which had tormented mathematicians for centuries: under what conditions can an equation be solved?

Rothman chopped off the *next crucial sentences* in Bell (1937, 375–376):

> But this was only one thing of many. In this great work, Galois used the theory of groups (see chapter on Cauchy) with brilliant success. Galois was indeed one of the great pioneers in this abstract theory, today of fundamental importance in all mathematics.

All these sentences are full of real meaning, for example, the sentence: "But this was only one thing of many" can make a big difference. A proper reading and understanding of these sentences will provide a strong rebuke to Kragh's claim above: "Galois's alleged 'scientific last will and testament' is a legend: the night before the duel Galois was indeed occupied with mathematics but actually of a rather trivial sort, viz. making editorial corrections on manuscripts."

To make things clearer, let us summarize the main points claimed by both sides:

1. From the perspective of Rothman and Kragh, (1) Galois was only "occupied with mathematics ... of a rather trivial sort, viz. making editorial corrections on manuscripts." (2) Consequently, what Galois wrote in his letter does not contain any new things. (3) Therefore, Galois' letter is not so important except for adding some minor improvements to his memoirs already written, and "Galois's alleged 'scientific last will and testament' is a legend", and Bell's claim about the impact of Galois' letter is just a myth created by Bell.
2. According to Bell, (1) Galois "spent the fleeting hours feverishly dashing off his scientific last will and testament, writing against time to glean a few of the great things in his teeming mind before the death." (2) Galois found "the true solution of a riddle: ... under what conditions can an equation be solved?", but "this was only one thing of many." (3) "What he wrote in those last desperate hours before the dawn will keep generations of mathematicians busy for hundreds of years."

From the above two lists, it seems clear that Bell, Rothman & Kragh were talking about different things: Bell was talking about the content of the whole letter of Galois, and its impacts on the later development of mathematics, while Rothman & Kragh focused on new mathematical results or ideas contained in the letter only, but not in the earlier memoirs, hence considering only the impact of new things contained the letter.[14] Since Rothman & Kragh claimed that there was nothing

[14] Bell's sentence "writing against time to glean a few of the great things in his teeming mind before the death" and the more dramatic sentence "Time after time he broke off to scribble in the margin 'I have not time; I have not time,' and passed on to the next frantically scrawled outline" could lead one to think that Galois was writing out new results which were only contained in his brain. On the other hand, it can also mean that Galois was summarizing results in papers which were not published yet. It can even mean that Galois was writing a summary of his main mathematical results up to that point since he was trying to preserve his scientific legacy. Or it can be combination of both. For the last point, one crucial but confusing word is "new" in the sentence "I have made some new discoveries in analysis". Does it mean new things he did which were not written down before? Or what he did in analysis, whether he wrote it up already or not, was new. Therefore, there could be multiple different ways to read and understand these few words of Galois. But are such hair-splitting arguments important? For which purposes do they serve?

Maybe it is helpful to quote what three distinguished historians of mathematics, Boyer, Cajori and Kline, wrote about this event in their famous books on the history of mathematics: "The night before the duel, with forebodings of death, Galois spent the hours jotting down, in a letter to a friend name Chevalier, notes for posterity concerning his discoveries" (Boyer 1989, 527), "The night before the duel he wrote his scientific testament in the form of a letter to Auguste Chevalier, containing a statement of the mathematical results he had reached and asking that the letter be published, that 'Jacobi or Gauss pass judgement, not on their correctness, but on their importance' " (Cajori 1991, 351), "The night before his death Galois drew up a hastily written account of his researches which he entrusted to his friend Auguste Chevalier" (Kline 1990, 756).

One can compare them with what Bell wrote in terms of content and style. It is also interesting to quote from Ehrhardt (2011, 14–15): "Commnly baptized 'Testamentary Letter', the letter addressed to his friend Auguste Chevalier, in which Galois summarized his research on the eve before the duel which had caused his death, is ... considered as one prophetic document, announcing the results which were only rediscovered many years later."

new in the letter, they were talking about different things, and hence Rothman & Kragh could not really destroy the arguments and conclusions of Bell, or rather the "legend" created by Bell.[15] We note that Rothman would agree with Bell if we consider the impacts of the whole letter of Galois, since Rothman wrote in Rothman (1982a, 147): "The theory of groups is one of the most fruitful areas of mathematical research; Bell is correct when he writes that it will keep mathematicians busy for hundreds of years."

To understand the whole situation better, we need to consider the following questions:

1. What are contained in the earlier memoirs of Galois according to Rothman & Kragh? This is an important question since the memoirs determine the contents of the letter, according to them.
2. What could Bell mean by "But this was only one thing of many"? The answer can show whether Rothman & Kragh's assessment is correct and fair. This question reminds one of what Galois wrote near the end of his letter (see the quote below): "You know my dear Auguste that these subjects are not the only ones that I have explored."
3. Did Galois write down anything new beyond corrections on manuscripts? This is crucial to the whole issue under consideration.
4. What are the main impacts of Galois' works contained in the memoirs according to Rothman & Kragh? How about other results of Galois, in particular those that Galois wrote down *only* in his letter? Do they have impacts which have lasted hundreds of years? These questions are also important for the discussions here.

In such a controversial case, *the most natural thing*, or *the only correct thing*, to do is to *read and check Galois' letter carefully* in order to have a reasonable knowledge about it. It does not mean that one needs to understand precisely and clearly the meanings of what Galois wrote in the letter, which is very difficult as we shall see. But it is probably not too much to demand that serious historical scholars need to read the entire letter *carefully* to get at least an idea of what is contained in the letter, for example, the scope of topics discussed, instead of just listening to standard stories about Galois' mathematical contributions. This is especially important when Rothman set out to destroy myths about Galois, especially the "legend" created by Bell. One simply cannot destroy a standard myth by other standard stories.

In the writings of Rothman and Kragh about Galois, we can observe the following:

1. In the quote from Bell by Rothman (1982a, 136), Rothman (1982b, 84), the only sentence containing specific mathematical content is: "He [Galois] had found,

[15] Viktor Blåsjö suggested to make explicit these differences. He also suggested to distinguish between (1) "direct causal influence of Galois' writings", and (2) "Galois being prescient about developments that would unfold independently".

once and for all, the true solution of a riddle which had tormented mathematicians for centuries: under what conditions can an equation be solved?”

2. Rothman emphasized that Galois wrote about or rather summarized the group theory and the Galois theory for algebraic equations in the testamentary letter. For example, Rothman wrote in Rothman (1982b, 89): “Bell also does not make it clear that the papers listed above [published papers: ‘An Analysis of a Memoir on the Algebraic Resolution of Equations’, ‘Notes on the Resolution of Numerical Equations’, ‘On the Theorem of Numbers’] (plus a later memoir), constitute what is now called Galois theory. If this point had been clarified, the claim that Galois had written the theory down on the eve of the duel would be difficult to substantiate or even to suggest.” We note that the last sentence implies that Bell claimed that Galois wrote down the Galois theory on the eve of the duel.
   Rothman wrote in Rothman (1982a, 140): “[Galois’] articles make it clear that in 1830 he had progressed beyond all others in the search for the conditions that determine the solvability of equations, although he did not yet have the complete answer in hand. By January, 1831, however, he had reached a conclusion, which he submitted to the Academy in a new memoir, written at the request of the mathematician Simeon Denis Poisson. The paper is the most important of Galois’s works, and its existence more than a year before the duel makes nonsense of the story that all Galois’s work on the theory of groups was written down in a single night.” The emphasized topic here is “the theory of groups”.
3. He wrote in Rothman (1982b, 104): “After hearing of my investigation, physicists and mathematicians all open conversations with me with the same question: ‘Did Galois really invent group theory the night before he was killed?’ No, he didn’t.” The main topic in question is the group theory.
4. In both Rothman (1982b, 84) and Rothman (1989, 149), Rothman wrote: “As with all legends the truth has become one of many threads in the embroidery. E.T. Bell has embroidered more than most, but he is not alone. James R. Newman, writing in *The World of Mathematics*, notes: “The term group was first used in a technical sense by the French mathematician Évariste Galois in 1830. He wrote his brilliant paper on the subject at the age of twenty, the night before he was killed in a stupid duel.” Note that only the notion of group is mentioned here in connection with the “legend” created by Bell.
5. On the other hand, many mathematical results Galois tried to put down in the middle and the latter part of the letter were not mentioned by Rothman in any of his writings on Galois in Rothman (1982a), Rothman (1982b), Rothman (1989). Rothman’s discussion mainly dealt with the first part and the very end of this testamentary letter. For example, after quoting the very beginning of the letter, he wrote (Rothman 1982b, 102, 103): “Galois then goes on to describe and elucidate the contents of the memoir which was rejected by Poisson, as well as subsequent work. Galois had indeed created a field which would keep mathematicians busy for hundreds of years, but not ‘in those last desperate hours before the dawn.’ During the course of the night he annotated and made corrections on some of his papers. He comes across a note that Poisson had left in the margin of his rejected memoir ...” After describing Galois’ reaction to Poisson’s note, Rothman quoted

the very end of the letter and concluded his description of this testamentary letter with: “And that was the end.”

6. Rothman did note in Rothman (1982a, 149) that Galois mentioned three memoirs in his letter, and wrote that the third one “apparently concerned integrals of general algebraic functions.” This is essentially what Galois said at the beginning of his letter, and Rothman did not give any hint of the nature of Galois’ results on integrals, while he gave a reasonably accessible exposition of the group theory and the Galois theory of algebraic equations in Rothman (1982a, 142–148). We note that such a brief statement on the third memoir does not appear in his later papers Rothman (1982b), Rothman (1989).
7. In Rothman (1989, 191), Rothman reprinted a *New York Times* editorial in 1982, which states: “Now this [Galois] story [created by Bell] has been a trifle deflated by an article in *Scientific American* [i.e., Rothman’s article] which holds that Galois was not writing out new theories but merely redrafting an already written paper. The manuscript, an account of his celebrated theory of groups (invaluable for solving Rubik’s cube), had been returned to him for revision before publication. Historical accuracy is a fine thing, but what a niggling correction to so haunting a story.” Since this editorial was inserted by Rothman in his more complete and probably the final writing about Galois’ letter, it suggests that Rothman agreed with the points expressed here.
8. Kragh probably drew all his conclusions about Galois’ letter and its mathematical contents from those of Rothman, as he wrote in Kragh (1987, 170): “Recent scholarship has argued that Galois’ alleged ‘scientific last will and testament’ is a legend... T. Rothman, who has contributed to the undermining of the Galois myth...”, and made several quotes in Kragh (1987, 170) from Rothman.

To understand better our questions about Galois’ letter and related issues in dispute, we need to *check what Galois wrote in his letter exactly*. We will use the translation of it in the book Neumann (2011, 85–97) and quote only several paragraphs from it, since the letter contains many striking results, some of which are still mysterious, at least to me.

Quotes from Galois’ Testamentary Letter

> I have done several new things in analysis. Some concern the theory of equations, others integral functions.
>
> In the theory of equations I have looked for the circumstances under which equations were soluble by radicals; this has given me occasion to deepen this theory and to describe all possible transformations on an equation even in case it is not soluble by radicals.
>
> Three memoirs could be made from all this.
>
> The first is written, and in spite of what Poisson has said about it I stand by it with the corrections that I have made in it.
>
> The second contains some pretty interesting applications of the theory of equations. Here is a summary of the most important things. (Neumann 2011, 85)
>
> ....
>
> The third memoir concerns integrals.
>
> It is known that a sum of terms of the same elliptic function always reduces to a single term plus algebraic or logarithmic quantities.
>
> There are no other functions for which this property holds.

> But absolutely analogous properties replace it in all integrals of algebraic functions.
> Let us treat at one and the same time all functions integrals of which the differential is a function of the variable and of an irrational function of the variable, whether this irrational is or is not a radical, whether or not it may be expressed by radicals. (Neumann 2011, 91–93)
> ....
> You know my dear Auguste that these subjects are not the only ones that I have explored. For some time my main thinking was directed towards the application to transcendental analysis of the theory of ambiguity. It was concerned with seeing a priori in relations between transcendental quantities or functions what exchanges one could make, what quantities one could substitute for the given quantities, without the relation ceasing to hold. (Neumann 2011, 95)
> That makes immediately recognisable the impossibility of many expressions that one could look for. But I do not have the time, and my ideas are not yet well enough developed in this area, which is immense.
> But all that I have written here has been in my head for almost a year and it is not in my interest to make a mistake so that one could suspect me of having announced theorems of which I did not have the complete proof.
> You will publicly ask Jacobi or Gauss to give their opinion not on the truth but on the importance of the theorems. After that there will, I hope, be people who will find profit in deciphering all this mess. (Neumann 2011, 97)

It should be emphasized that these quoted passages only describe a part of results which Galois stated in his letter.

Now we compare them with what Rothman (1982a,b, 1989), Kragh (1987), and Bell (1937) have written, as quoted above.

As explained above, it seems that Kragh only, and Rothman mainly, noticed the part on the Galois theory of algebraic equations, which are stated in the first part of Galois' letter, which is concerned with the first memoir. To be fair with Rothman and to give him credit for his writings about Galois, Rothman spent so much time and energy to study and check all the writings about the life of Galois and Galois' non-mathematical writings to make the record straight. Unfortunately, it seems that he did not check or read mathematical writings about Galois' testamentary letter by some major mathematicians such as Klein and Picard (see below for references and some details). Given that mathematics itself is *the issue under discussion*, to say nothing that it is also the essential part which makes the Galois story so special and attractive, and Rothman and Kragh tried to correct what they believed to be seriously misleading claims of Bell about what Galois wrote in the letter, it seems *mandatory* that they should read the whole letter carefully in order to gain at least a rough idea of what is inside it.

There is no question that the Galois theory and the related group theory is the best known part of Galois' works, maybe also the most important part. In any case, this is also the most clearly described part in Galois' letter. As explained above, the descriptions of Rothman and Kragh were exclusively devoted to it. But Rothman and Kragh seemed to overlook that Bell had another crucial sentence: "But this was only one thing of many". For example, this was chopped out of their quotes.

What are the many?

Bell did not say or explain. Naturally, *the only reliable way is to check the primary sources: the testamentary letter in this case!*

According to the above quotes from the letter, "the many" include at least some general results on *abelian integrals*, which is contained in the third memoir, and more importantly "*the theory of ambiguity*" mentioned near the end of the letter.

How about *Bell's claim: "What he wrote in those last desperate hours before the dawn will keep generations of mathematicians busy for hundreds of years" ?*

Yes, Bell's claim is true in several senses. First, the first part of Galois' letter was devoted to the Galois theory and the group theory, and every mathematician will agree that they have kept mathematicians busy for hundreds of years.

How about some other results mentioned by Galois, especially those which were only contained in the letter? We will show that the answer is also positive by explaining these two things of "the many" together with one result on modular equations.

The general theory of abelian integrals which generalizes the theory of elliptic integrals, a substantial part of which was stated explicitly in Galois' letter, was developed systematically a few decades later by Riemann, in particular in his famous paper *Theory of abelian functions* published in 1857, the most celebrated paper of Riemann in his lifetime, which has been influencing the development of mathematics up to now and into the foreseeable future. It is not entirely clear if Riemann's work on abelian integrals was influenced by Galois' work on them.[16] But whether or not Riemann was influenced by Galois' work, there is no question that the mathematics Galois described kept generations of mathematicians busy.

The Theory of Ambiguity

The mysterious "theory of ambiguity" stated in Galois' letter has fascinated many great mathematicians ever since Galois' letter was published. The list includes Felix Klein, Sophus Lie (Lie 1895), Émile Picard, Jules Drach, Teiji Takagi, George Birkhoff, Alexander Grothendieck, Jean Dieudonné, Alan Connes, and Hiroshi Umemura who spent the major portion of his life to establish "the theory of ambiguity" (Ramis 2020; Okamoto and Ohyama 2020). We will briefly mention some of these people's comments. More details will be explained on a future occasion (Ji 2023). It should be stressed that we cannot be certain of what exactly Galois meant by the theory of ambiguity. But the discussions below will show that this notion or idea suggested by him has inspired many good mathematicians in a period of more than 100 years to produce theories which can fit into the theory of ambiguity.

[16] So far, people have not found definite records that Riemann had read Galois' testamentary letter. On the other hand, this letter was published before Riemann went to University of Göttingen in 1846, and the other papers of Galois were also published by Liouville in 1846. Since some people in Göttingen such as Dirichlet and Dedekind were probably aware of Galois' works, it is a natural guess that Riemann might have heard something about Galois' works. See the paper Neuenschwander (2022).

One early comment was given by Picard in the preface to Galois' writings (Picard 1897, IX):

> We could almost guess what he [Galois] means by this [the application of the theory of ambiguity to transcendental analysis], and in this field, which, as he says, is immense, there still to this day remain discoveries to make.[17]

Picard did not explain further what his guess was, but we can guess that it was related to the Picard-Vessiot theory for linear differential equations, i.e., the differential Galois theory.

In his famous history book Klein (1979, 81), originally published in 1926, Klein speculated:

> Galois spoke of investigations into "ambiguity of functions"; it is possible that this referred to the idea of Riemann surfaces and multiple connectivity.

Later on the same page, Klein wrote:

> Galois theory does answer important questions of the theory of equations in the most general way; but it also opens the door to a new, vast and still largely unknown region, whose full range of problems we cannot yet foresee. Following a tradition arising from personal communication with Gordan, I would like to call this field "hyper-galois".

The endorsement of Klein on ideas and results contained in Galois' letter can also be clearly seen in Klein (1979, 83–84). For example, Klein wrote: "In claiming that he had left behind nothing false, Galois was right."

Another great mathematician, Teiji Takagi, also wrote a book on the history of mathematics and gave some stimulating comments.[18] For the theory of ambiguity, he suggested the idea of monodromy.[19]

Since we have used the translation of Galois' testamentary letter in Neumann (2011), it will be interesting to mention some comments on the theory of ambiguity by Neumann himself (Neumann 2011, 386):

> The tantalising phrase *théorie de l'ambiguité* ... appears only [in] ... the *Lettre testamentaire*. There is nothing beyond the context (that of abelian integrals and functions) and the following sentence to tell us what Galois had in mind. And those tell us little. I have seen

[17] This sentence was translated and included in the source book Smith (1959, 284).

[18] It was written in Japanese and became very popular. Later it was translated into Chinese, but there is no English translation of it. In one version of Chinese translation, Takagi claimed that Galois' work on abelian integrals indeed influenced Riemann, but not in a newer translation.

[19] It may be that there are even deeper connections that Galois discerned, however vaguely, that we have still to explore. In an email on November 16, 2020, Pierre Deligne told me that he agreed with Takagi's interpretation of the theory of ambiguity in terms of monodromy, and said that the application of the theory of ambiguity to "transcendental quantities" might be related to some Grothendieck's conjectures on periods of motivic Galois groups, which is certainly consistent with what Galois wrote about abelian integrals, which give rise to periods. He also added that the transcendency of π can be explained through this conjecture, since π arises as a period, the length of a unit circle, and a combination of the period conjectures and the equality $\log e = 1$ can also help explain the transcendency of e.

> speculation that Galois had in mind a theory like that of Riemann surfaces, but I find that doubtful; for me it does not chime with the sentence following the phrase in question.

In defense of Rothman & Kragh, reading and understanding Galois' letter in full detail is very difficult, if not impossible.[20] This difficulty was recognized by many people a long time ago. For example, Hermann Weyl made the following comments on this letter in his famous book on symmetry (Weyl 1952, 138):

> Galois' ideas, which for several decades remained a book with seven seals but later exerted a more and more profound influence upon the whole development of mathematics, are contained in a farewell letter written to a friend on the eve of his death, which he met in a silly duel at the age of twenty-one. This letter, if judged by the novelty and profundity of ideas it contains, is perhaps the most substantial piece of writing in the whole literature of mankind.

The last sentence in this quote from Weyl is interesting and especially relevant to our discussion here. Weyl made it clear that Galois' letter is not a legend as Kragh described in Kragh (1987, 170): "Galois's alleged 'scientific last will and testament' is a legend".

On the other hand, if people who want to write about Galois letter have tried to read it, even casually, from the beginning to the end, they probably will not fail to notice that the *well-known Galois theory of algebraic equations is only one part* of the treasures in it, and Galois did try to put down many new ideas and results into his testamentary letter.[21] If they compare what Galois wrote with later development of mathematics,[22] they will probably also notice the tremendous scope and originality of Galois's achievements and ideas, beyond the Galois theory, contained in the letter, and how unfinished some of Galois' ideas were at that night before the duel.[23]

---

[20] In an email on June 5, 2020, J.P. Serre told me that Galois was clearly aware of the problem of nth division points of the Jacobian variety of an algebraic curve, based on what was written in his letter.

[21] Another result of Galois concerning modular equations has also far-reaching consequences for the contemporary mathematics. Galois announced that the modular equation of degree 6 for the prime 5 can be reduced to an equation of degree 5. This reduction was proved and used by Hermite to solve the general quintic equation by elliptic functions and elliptic modular functions (Goldstein 2011, 230–236). This result of Hermite partially motivated Klein's work in his famous book Klein (1884) on the icosahedron and the solution of equations of fifth degree (Brechenmacher 2011, §3.2). Klein's geometric approach leads to his work in the theory of modular functions and automorphic functions with respect to $SL(2, \mathbb{Z})$ and its subgroups, for example, his famous four extensive books on such topics written jointly with Robert Fricke, which in turn was developed or perfected by Hecke and generalized by Siegel, Selberg, Langlands and others to the elaborate and powerful theory of automorphic forms and automorphic representations for general Lie groups, which is the foundational part of the celebrated Langlands program, probably the greatest theory in mathematics in the 20th program and maybe also in the twenty-first century. See also Smith (1906, 17) for some related comments on the modular equations and discrete transformation groups.

[22] This is consistent with the sound advice of Popper (1968, 16): "Among the many methods which he [a philosopher] may use–always depending, of course, on the problem in hand–one method seems to me worth mentioning. It is a variant of the (at present unfashionable) historical method. It consists, simply, in trying to find out what other people have thought and said about the problem in hand: why they had to face it: how they formulated it: how they tried to solve it."

Probably it is also important to note another point. Bell made it clear that his writings in Bell (1937) were not history at all. In fact, in the first and second paragraphs of the introduction of the book Bell (1937, 3), Bell wrote:

> I should like to emphasize first that this book is not intended, in any sense, to be a history of mathematics, or any section of such a history.
>
> The lives of mathematicians presented here are addressed to the general reader and to others who may wish to see what sort of human beings the men were who created modern mathematics. Our object is to lead up to some of the dominating ideas governing vast tracts of mathematics as it exists today and to do this through the lives of the men responsible for those ideas.

As we have seen above, Rothman and Kragh have harshly criticized Bell's book as a history book. On the other hand, since they claim to have done historical research up to the highest standard and are writing scholarly papers about Galois, they should uphold the basic requirement of doing history: *always check carefully and use the primary sources if they are available.* Another basic principle for the history of mathematics is probably also relevant: *when writing about mathematical issues in the history of mathematics, mathematics is a central part and should be handled with care.*

But Rothman and Kragh are not the only historians who have harshly criticized Bell's book as a history book. To be honest, this is one curious, or bothersome, question to me.[24] Since there are many such complaints, I will just quote from two prominent historians of mathematics and science. In Grattan-Guinness (1971, 350), Ivor Grattan-Guinness wrote:

> ... the popular but unscholarly works of E.T. Bell ... especially his *Men of Mathematics* [is] perhaps the most widely read modern book on the history of mathematics. As it is also one of the worst, it can be said to have done a considerable disservice to the profession.

Similarly, Truesdell wrote (Truesdell 1984, 423–424):

> The only course Bell taught was abstract algebra; while he did little to excite the students in that subject, he was admired for his science fiction and his *Men of Mathematics*. I was shocked when, just a few years later, Walter Pitts told me the latter was nothing but a string of Hollywood scenarios; my own subsequent study of the sources has shown me that Pitts was right, and I now find the contents of that still popular book to be little more than rehashes enlivened by nasty gossip and banal or indecent fancy.

It seems both Grattan-Guinness and Truesdell treated Bell's book as a *scholarly history book* and measured it against the ensuing strict, or high, standard for books in the history of mathematics. They probably also chose to ignore the mathematical contents in this book, which are one of the goals of Bell for the book as the above quote shows.

---

[23] It is perhaps fitting here to quote from the first paragraph of the preface of Neumann (2011, vii): "The *Lettre testamentaire* ... is an extraordinary summary of what he [Galois] had achieved and what he might have achieved had he lived to develop and expound more of his mathematical ideas."

[24] According to José Ferreirós, "The main reason is the impact Bell had on so many mathematicians and scientists, I think. They thought it was history."

I wonder whether Bell was aware of the possibility of such misunderstanding by so many historians and hence put the *disclaimer* at the beginning of his book as quoted above, and whether the critics of Bell's book have noticed the important disclaimer stated so explicitly and prominently by Bell and taken it into consideration when they made harsh judgements on him and his book.

The above comments on the mathematical content of Galois' letter give some supporting evidence to Bell's claim about the "object" of his book: "lead up to some of the dominating ideas governing vast tracts of mathematics", considering that this is a popular book for "the general reader" and "others who may wish to see what sort of human beings" mathematicians are. Maybe this explains why Bell did not explain the meaning of the somewhat mysterious sentence: "But this was only one thing of many" quoted above.

What Are Some Really Important Historical Questions About Galois' Letter? After spending so much time and space to understand better the differences between the understandings and writings by Bell, Rothman and Kragh about Galois' letter, and merits of their works, it is probably a good place here to go back and address Questions (1) and (3) raised at the beginning of this section.

Before that, let us ask a general question: *What would a person on the death bed do if he were going to write the last letter about his scientific works in order to keep a legacy for him?* A reasonable answer seems to be: *He would try to summarize all his main contributions, whether they were written up already in some papers or not, and he wanted people to appreciate what he had done and benefit from them.* In other words, he wanted people to read and understand his letter carefully and be inspired by them, even though he knew that this might not be an easy undertaking. That was what Galois wrote at the end of the letter: "ask Jacobi or Gauss to give their opinion not on the truth but on the importance of the theorems. After that there will, I hope, be people who will find profit in deciphering all this mess."

Based on this, it seems that the issue about the differences and merits of the writings of Bell, Rothman and Kragh is *not* important. Instead, the following perspectives and questions are really important:

1. Try to understand Galois' testamentary letter as a whole, in particular, try to understand how different parts or mathematical results in it are related. Though the Galois theory and the theory of groups are absolutely important, other parts should not be ignored. (One intriguing question to me is: How could such a young person do such original and foundational works with long lasting impacts in so many subjects in such a short time? Of course, Galois was a genius! But this is a simple answer, an easy way to avoid a difficult and important question.)
2. Understand or interpret more precisely what Galois wrote down in the letter, including all theorems, claims, suggestions and speculations etc., and how they arose and influenced the development of mathematics: for example, their motivations and impacts, the status of the subjects discussed before his works, unfinished parts envisioned by him, applications and further developments of his theories etc.

3. Understand Galois' letter in different contexts: (1) in the context of all his writings, (2) in the context of mathematics of his time across different subjects since Galois' letter indicated his broad scope already. For example, since he specifically asked for opinions of Gauss and Jacobi, it might be interesting to compare his works with those of Gauss and Jacobi, and naturally also with the works of Abel and Riemann. The total works of these five great mathematicians open one special window into the world of mathematics.
   Since Galois and his testamentary letter (or rather the mathematics contained in it) are so unique in the whole history of mathematics, we need to understand their uniqueness from various perspectives in mathematics.
4. Understand Galois and his mathematics also in the broader social, cultural, and political contexts etc. Since Galois' letter contains his most important contributions to mathematics, this question can also be considered as a question about his letter.

As Leo Corry commented on an earlier version of this chapter, no historical study on Galois can neglect the existing major studies in Ehrhardt (2011, 2012), Brechenmacher (2011). It seems that Ehrhardt (2011, 2012), Brechenmacher (2011) are mainly concerned with Question (4), though including some parts of Questions (2) and (3). For example, the main mathematics topic in the books Ehrhardt (2011, 2012) is the Galois theory, and the main writing of Galois is the first memoir, which deals with the Galois theory, while Galois' testamentary letter is mentioned (Ehrhardt 2011, 14–15) but not discussed or emphasized. They do not mention the theory of ambiguity explicitly. The discussions in Ehrhardt (2012, 201–210) about applications of the Galois theory to differential equations mention some works of Picard, Vessiot, Drach, and they are related to our concern in this section about the theory of ambiguity. The focus of the paper Brechenmacher (2011) is to examine a collection of books and related writings which involve or refer to some parts of the works of Galois, in comparison with Jordan's famous book *Traité des substitutions et des équations algébriques*, published in 1870.

Conclusion

Given the above long discussions about Rothman's papers and Kragh's book about Galois' letter, it is worthwhile to draw some conclusions. There is no question that Rothman did a very careful and scholarly job about the social, political, cultural and biographic aspects of Galois' life. Probably this is the reason why his papers were highly regarded and cited by multiple distinguished historians and mathematicians as mentioned at the beginning of this section. But he seems to ignore, or rather without paying enough attention to, the mathematical part of Galois' life, even when the issue was about the mathematical content of his testamentary letter. In terms of the so-called internal and external histories of mathematics, Rothman's papers were an excellent piece of work in the sense of the external history, but a poor, unfair and misleading writing in the sense of internal history of mathematics. Yes, his papers on Galois probably have mislead many people about what Galois wrote in the last scientific letter of his life. Yes, he was unfairly critical to Bell as far as mathematics is concerned. Probably more seriously, he was unfair to Galois, though unintentionally, if we recall what Galois wrote near the end of his letter.

I wonder how would Galois react if he were to read what Rothman and Kragh had written about the mathematical content and purpose of his testamentary letter, his last contribution to mathematics, by misunderstanding and distorting his intention, in particular when he read the sentence in Kragh (1987, 170): "The destruction of the Galois myth yields a more authentic history without diminishing the scientific originality of Galois."

Therefore, one conclusion is that it is important to balance or rather to *combine* the external and internal histories in doing research in the history of mathematics, since both are important and needed. Otherwise, good intentions might do much harm!

3.4 What Is Historical Thinking and Some Basic Points in the Methodology of History

After the above lengthy discussion about Galois' letter, we return to some general discussions on historiography. There are many books and writings about what, why and how of the general history, some of which seem helpful to the history of mathematics too. We will mention only a few points here. Many more detailed discussions with quotes from multiple sources are given (Ji 2022), which also contains a rather extensive list of books on the historiography of general subjects.

If one wants to do research in history, one needs to think like a historian. Though history is a natural part of everyone's life, the ability of thinking historically needs to be learned and cultivated. When I first started to read reviews or comments by historians about mathematicians' works related to the history of mathematicians, I learnt immediately that one common sin committed by mathematicians is *anachronism*, in particular the *presentism*. More generally, a basic conclusion, or a serious difficulty, is that many mathematicians do not, or cannot, *think historically*.

But what does it mean to think historically? I have tried to search for it among writings about the history of mathematics rather extensively. Unfortunately I can not find a simple, clean and clear answer. Then I searched among books on the general historiography and found a satisfying answer in a successful textbook on history *The Methods and Skills of History* by Furay and Salevouris (2015, 28–30):

> A recent book on the subject is titled *Historical Thinking and Other Unnatural Acts*, dramatizing the unsettling truth that thinking historically is not something that comes naturally. Historical reality is complicated, and becoming historically "literate" can be a challenging enterprise. But it is well worth the effort. Let's begin by looking at three essential components of effective historical thinking.
>
> 1. Sensitivity to multiple causation.
> All too often our popular culture favors quick, simplistic answers to the challenges of the day... Historians, and students of history, however, know that it is a mistake to look

for a single cause. Every situation or event is the product of multiple "causes" or factors, some short-term and some long-term. To think like an historian you will need to consider the wide range of factors and conditions that have led to the events you are investigating.

2. Sensitivity to context, or how other times and places differ from our own.
   How can twenty-first-century observers understand cultures from a strange and distant past?... historians must make a serious effort to bridge the cultural and temporal gap, even though they know it won't be easy.
3. Awareness of the interplay of continuity and change in human affairs.
   Every situation is an amalgam of the old (continuity) and the new (change). "The mixture is never the same; history does not repeat itself; the new is new. But the old persists alongside it and in time the new is grafted on to the old and continuity with the past–continuity, not identity–is not disrupted but restored." The challenge for the historian is to figure out how this balance manifests itself at any particular point in the past....
   There can be a "history" only when there is change. In essence, history is the story of change.... Most changes take place in the overall context of continuance of many of the old ways of doing things, and are often no more than patchwork alterations of the existing system.

This definition of historical thinking should be taken dialectically too. It does not mean that every piece of historical writing needs to enjoy all three qualities, and every historian will and can do things in such ways at all times. But they can serve as helpful guides when one is involved in historical researches. This is especially important for young historians to pay attention to them and develop such habits so that they can think historically in a subconscious way.

The above definition of historical thinking is consistent with some basic recommendations or rather requirements from the methodology of the general history. We select three from Ji (2022), adapting them to mathematics and adding some brief comments and examples.

3.4.1 Use Primary Sources Whenever They Are Available

Probably no historians will disagree with the importance of using the primary sources, and one should always check and use properly the primary sources even when "excellent secondary sources" are available. It is *the Golden Rule* in doing history.

But it is also very hard to follow this golden rule for several reasons: besides problems with languages, it is often difficult to understand the meaning of a piece of writing from a long time ago, since many things have changed and the reader really needs to be familiar with the appropriate contexts and have the required knowledge about the subject matter to read it and then understand it. (For example, definitions of mathematical concepts and notations change with time, and implicit standard assumptions and results may not be standard anymore and hence may become difficult to guess and identify. There are also cultural and social factors which change with time.) This is especially hard when secondary sources, which are assumed to be good and reliable, or having good reputations, are available. Then the temptation to skip the original writings and instead read others' interpretations is great, often too great. Probably this is the reason why most mathematicians do not read classical writings in their subjects, even though they still have great impacts on the subjects, leaving it to historically oriented people or historians to do the work to interpret classical writings for them. This fits well the saying from the Bible: "the spirit is willing but the flesh is weak".

The example about Rothman and Kragh's writings about Galois' letter discussed in the previous section amply shows this phenomenon. When the problem is concerned with the mathematical content of this letter, the letter itself is without any question the most important *primary source*. It seems that those historians who criticized Bell's book *Men of Mathematics* as a bad history book did not read (or rather discarded) Bell's disclaimer at the beginning of the book and also ignored the primary sources. Therefore, one natural conclusion is that it is not only mathematicians who sometimes commit this sin in doing history, though for historians, it is probably more inexcusable.

My first somewhat surprising experience with the violation of this golden rule occurred during working on the paper Ji and Wang (2020) on the motivations for Poincaré to systematically develop analysis situs (or topology). I first searched for all writings about this question by both historians and mathematicians (i.e., secondary sources). Due to the fame of Poincaré and his extensive and important works in topology, in particular the Poincaré conjecture and the publicity of or chaos around its recent solution by Perelman, there were quite a few of them. They usually give one or two reasons, which are rather consistent, as Poincaré's motivations for working on topology. Later on, I checked the original writings of Poincaré and compared Poincaré's stated motivations with what others have written. It was a big surprise, since Poincaré gave at least 7 explicitly stated reasons across many different subjects, which are consistent with Poincaré's universal contributions to mathematics. Naturally, understanding clearly Poincaré's stated motivations is a different story. More details are explained in Ji and Wang (2020).

The uses of the primary and secondary sources provide a good example of a dialectic perspective. Naturally everyone's work is built upon the works of the predecessors, and hence making use of the secondary sources about any research project in history is natural and important.[25] They are definitely different from the primary sources, but can be used to help understand primary sources of the project under study. Conversely, the primary sources can also clarify, confirm or disprove secondary sources. This dialectic process can be repeated multiple times.[26]

[25] Reading secondary sources is also one effective way to find research problems in history, since they often expose gaps in the existing historical knowledge and expositions.

[26] Since we just quoted from the Bible, it may be illuminating to use reading the Bible as an example. Going to the church and listening to sermons, and reading commentaries and other books on the Bible goes hand to hand with reading the Bible itself. They are both important and needed, and this process can often last a long time.

3.4.2 Study Problems in Appropriate Contexts

Since the history of mathematics studies the development or evolution of mathematics, no problem should be studied in isolation. Instead it should be viewed as a part of a process over time and in suitable contexts. As the Golden rule explained above, the importance of this requirement is also self-evident, since it is a part of the definition of history. But actually doing it and doing it well may not be so easy.

In my very limited experience with the history of mathematics, it took a lot of effort and time for me to realize the importance to putting things into suitable historical contexts. At the beginning of a project on the history of Poincaré conjecture, I was surprised to read and learn that what some famous mathematicians have written about it was too simplified or incorrect versions of the history of the Poincaré conjecture, which one could easily check from the published papers of Poincaré on topology. When I told this to Jeremy Gray, he said that correcting mistakes of mathematicians, even famous ones, about history of mathematics is not so valuable from the perspective of historians. Later on, I found that some historians of mathematics also made similar incorrect statements about the Poincaré conjecture. This additional finding still did not convince me of the importance of writing a paper to correct such errors. I thought over this for a long time, and one day it suddenly occurred to me that I should view the history, or rather the evolution, of the Poincaré conjecture in bigger contexts:

1. In the context of the total sum of Poincaré's works in topology. This will allow us to see how Poincaré valued this conjecture, and what kind of role it plays in his relatively long and extensive period to establish (algebraic) topology as a separate subject.
2. In the context of the development of topology, in particular, the problem of classification of manifolds in terms of algebraic invariants. This can serve several purposes: (1) the Poincaré conjecture is a special case, or maybe the simplest case, of the classification of general manifolds, and hence works on this case can contribute to the general case, (2) this can allow us to see the importance of the Poincaré conjecture to the general theories of topology.
3. In another context of how various mathematicians viewed the classification problem, in particular the Poincaré conjecture, relative to other problems and results in topology, and how their views were reflected in the major books and papers on topology.

In this way, a proper understanding of the history of the Poincaré conjecture is not limited to the conjecture itself. Instead it provides one way to understand the history of topology. Once I realized this, the rest was an extensive, though relatively standard, task in studying history. For example, found as many as possible books and papers related to the above issues, extracted useful information, drew various conclusions, and then organized everything into a historical narrative, which eventually appeared in Ji and Wang (2022). Looking back at this experience, it should have been obvious if I knew more about the proper methodology of working on the history of mathematics.

As a continuation of this project, I wanted to understand better Poincaré's works in topology in the context of all Poincaré's scientific works. This looks like an interesting problem for me since in his self-analysis of scientific works (Poincaré 1921, 100), Poincaré wrote: "As for me, all the various ways in which I have engaged successively led me to Analysis Situs." It seems that this context is also natural, given that even though Poincaré is known and remembered for works in many different subjects, his pioneering work in algebraic topology is one of his most important contributions to mathematics, and definitely most known to general mathematicians and the educated public in the last decade.

To carry out this project satisfactorily, it seems necessary to have a good understanding of all of Poincaré's scientific works. To do this, it seems necessary to read the eleven volumes of the collected works of Poincaré, on top of many secondary writings about Poincaré's works. It seems that even a superficial reading of them (for example, the beginning and ending of each paper and the statements of main results in each paper) is probably too much to ask for, and a good understanding of the overall works of Poincaré is probably beyond the capability of many people, at least me. So the requirement of putting a problem into the context and of following the golden rule of reading the primary sources seems almost impossible to be satisfied. On the other hand, one still needs to try, in particular in combination with or with the help of good secondary writings, since it is an interesting and important problem due to the fact that Poincaré did greatly influence the development of mathematics at least up to now. If we take a dialectic perspective, we can still do some researches in this direction without trying to finish the project fully or ideally: one can do what one can in the right direction.[27]

There are also other problems and contexts to understand the works of Poincaré, for example, to understand the works of Poincaré in the context of the mathematics community of Paris, and the bigger context of the mathematics community of Europe. One can also study relations and interactions between his research works and his popular lectures and writings.

There is another problem which probably fits well the perspective of external history. One could also try to understand better Poincaré's works by analyzing and understanding what Klein wrote about Poincaré in Klein (1979, 356):

> Without doubt, a portion of this versatility [of Poincaré] was due to his thorough education by the firmly articulated French educational system, in which the traditional parts of all mathematics are grasped from all sides in early years–quite otherwise than in Germany, where the growing mathematician is quite glad to attach himself to a single master, which is good for a first, ad hoc production, but often goes no further.

[27] According to José Ferreirós, there are multiple layers to be considered in doing research in history, and one does not need to be absolutely competent regarding every layer. Instead, one can combine detailed knowledge of some, good knowledge of other layers, with serious study of what happened in others (perhaps second-hand). One can get a good grasp of the general views of a scientist by judicious employment of some of his own writings, plus secondary literature.

Klein's comments raised several questions to me. If what he said is true, one may wonder why there was not another contemporary French mathematician comparable to Poincaré, without belittling the great tradition of French mathematics? In terms of broad knowledge of mathematics across subjects, it seems that Klein surpassed Poincaré at the early stage of Poincaré career as shown in their correspondences (see for example de Saint-Gervais (2016, 385–414)). This is also confirmed by Hilbert's speech on the 60th birthday event of Klein when he compared Klein with Poincaré in the presence of both of them (Rowe 1986, 76):

> If I should select, secondly, a special mathematical area, we only need to hear the names Poincaré and Klein together and what mathematician would not be reminded of the automorphic functions, whose main theory was founded by Poincaré, but whose rich design is due to you [Klein].

One concrete problem is to understand, maybe more objectively than Klein, the influence of the earlier education of Poincaré on his latter research works.[28]

Each different context will shed different light on Poincaré and his works. Putting a historical study into an appropriate context is also important for another purpose, as explained by Arnold (2000, 8):

> In the late nineteenth and early twentieth centuries some historians did work in this way, collecting and translating interesting pieces of evidence they thought might appeal to a wider readership. Such books are useful treasure troves, and have led to detailed work by other historians. They can be a pleasure to read, infecting readers with their enthusiasm for the past. But for most modern historians, this is not enough. We need to *interpret* the past, not simply present it. Finding a larger context for the story is an attempt to say not just 'what happened' but what it meant.

The above discussion of various contexts shows how they can allow us to understand better Poincaré's works.

3.4.3 Understand Motivations and Impacts

Description of the development of mathematics is certainly a central part of the history of mathematics. But it is not the only part. Motivations for and impacts of the results under discussion are also important. In other words, we need to interpret or explain what we have found and described.

In the classic book *What is History* (Carr 1961, 112–113), Carr wrote:

> The historian, like any other scientist, is an animal who incessantly asks the question, Why? .... The study of history is a study of causes. The historian ... continuously asks the question, Why?; and, so long as he hopes for an answer, he cannot rest. The great historian – or perhaps I should say more broadly, the great thinker – is the man who asks the question, Why?

[28] Some German mathematicians told me jokingly that if their educational system can produce mathematicians such as Gauss and Riemann, it is probably not too bad!

Naturally, the simple word “Why” is very complicated. For example, some causes are accidental or casual, while others are rational (Carr 1961, 140); and they are sometimes difficult to distinguish. One should not always believe in and try to find a unique, or “the definite”, cause. As explained above in §3.4 about historical thinking, one should always keep in mind the fact that there are usually multiple causes for every situation or happening. It should also be emphasized that though important, understanding causes is only one aspect, but not the whole, of the study of history.

Similar opinions were expressed by Kuhn (1977, 8):

> [historians’] first concern was to discover what each one had thought, how he had come to think it, and what the consequences had been for him, his contemporaries, and its successors.

Hawkins also wrote more directly about the history of mathematics in Hawkins (2000, vi):

> My experience is that the understanding of a theory is deepened by familiarity both with the considerations that motivated various developments and with the less formal, more intuitive manner in which they were initially conceived. Hence in addition to stressing the motivational, historical context of the mathematics, I have tried to expound the original mathematical conceptions and reasoning clearly and fairly extensively.

All the above quotes confirm the importance to understand motivations and impacts in historical studies. As with the previous two basic requirements in doing history, analyzing correctly and clearly motivations and impacts is also much easier said that done. One reason is that this problem is closely related to the issue of choosing and understanding appropriate contexts. We will take a brief look at two examples.

As it is known, there have been many writings about the so-called competition between Klein and Poincaré on automorphic functions, in particular the uniformization theorem of Riemann surfaces. Given the importance of the mathematical theories and the fame of the two competitors, this is not surprising. Even Klein himself openly described it in reasonable details in the book (Klein 1979, 326): “[The subject of automorphic functions] is the subject in which I competed with Poincaré in 1881, 1882”. One important and interesting question concerns who did better at which points and who won this competition. It seems that the general consensus is that Poincaré won the competition while Klein surpassed Poincaré at one part of the proof of the uniformization theorem. This is one conclusion of the well-known paper by Freudenthal (1955, 218), and this conclusion has been cited by multiple important books such as Gray (2000a, 199), Hadamard (1999, 8), de Saint-Gervais (2016, 144).

On the other hand, if we think of the uses and impact of the uniformization of Riemann surfaces as an equivalence between the following three categories:[29]

[29] Actually there are more categories of spaces and structures which can be identified with Riemann surfaces. But these three are probably the most notable ones.

1. Compact Riemann surfaces of genus greater than 1,
2. Plane algebraic curves of genus greater than 1,
3. Fuchsian groups, or more precisely Fuchsian groups acting the upper half-plane whose quotients are compact Riemann surfaces of genus greater than 1,

and realizing that it is algebraic curves which are important for arithmetic geometric problems and how difficult it is to write down algebraic equations for a given Fuchsian group, Freudenthal's conclusion on Klein's advantage in the proof of the uniformization theorem becomes not so conclusive or convincing anymore. More details will be given on future occasions.

Another important concept in mathematics whose motivations, development and impacts are worthy of detailed analysis is the notion of moduli space of Riemann surfaces, or equivalently the moduli space of algebraic curves.

Ever since the definition of, or rather the word, *moduli* was introduced by Riemann in his classic paper on Abelian functions in 1857, the moduli space of Riemann surfaces has been intensively and extensively studied by many mathematicians. It is still one of the most important spaces in algebraic geometry and is currently under active investigation by algebraic geometers, topologists and mathematical physicists. It will continue to be studied by generations of mathematicians in the foreseeable future.

On the other hand, the original motivations, various approaches to understand it, whether successful or failed ones, new concepts emerging in the process (for example, the major concepts such as the Teichmüller theory and the mapping class group of surfaces), and many applications in different subjects all contribute to the rich history of moduli spaces, though a detailed historical exposition is still lacking. In some sense, one difficulty comes from its richness. For example, many mathematicians and historians know about Riemann's famous count for the dimension of the moduli space of Riemann surfaces, but the Riemann's original motivation for introducing the notion of equivalence classes of Riemann surfaces and the ensuing moduli space may not be so well-known.

The history of moduli spaces also needs to be studied in the contexts of elliptical integrals, elliptical modular functions and more general modular and automorphic functions. For example, Riemann's use of the word of moduli came from its use in the pamametrization, though not exactly the classification, of elliptical integrals, and transformations of elliptic integrals in connection with the moduli parameters led to elliptic modular functions and more general modular functions. Therefore, the study of moduli was intimately connected with modular functions right at the beginning. On the other hand, this aspect of automorphic functions seems to be lacking, or not sufficiently addressed, in the study of the moduli spaces of Riemann surfaces of genus greater than 1 in comparison of the moduli space of elliptic curves, i.e., Riemann surfaces of genus 1. Is it because of the intrinsic difficulties, the lack of direct applications of such functional theoretic studies, or the lack of historical knowledge or interest about the history of moduli and modular functions?

3.5 One Common Framework Through the Spacetime of Mathematics

In the previous sections, we discussed various perspectives on and approaches to the history of mathematics. In this section, we propose a common framework of the spacetime of mathematics which can allow us to unify these different approaches. More importantly, it can also provide a new perspective or platform to study the history of mathematics, for example how to think of and identify some important issues in the history of mathematics such as whether there are revolutions in mathematics as counterparts of Kuhn's famous concept in the history of sciences (Kuhn 2012), how to judge the importance of problems and results in the study of the history of mathematics. Some of these points will be illustrated through two examples to understand better certain aspects of Poincaré's and Hilbert's works.

Since the history of mathematics is concerned with the evolution of "objects" in the world of mathematics with respect to time and time is crucial in history (think of the issue of anachronism), it occurred to me one day to combine the time factor with the world of mathematics into the *spacetime of mathematics* so that *both factors can be and have to be studied together*. Then the study of the history of mathematics is reduced to the study of this spacetime of mathematics, and maybe can be carried out more effectively in this framework; or rather it provides one convenient way to view the history of mathematics. For example, in this way, the two different ways of studying the history of mathematics, "history" and "heritage" promoted by Grattan-Guinness (2004a,b), as we quoted extensively in Sect. 3.2, becomes obvious. One can also see that even the combination of "history" and "heritage" is rather limited, since they only represent two directions in the spacetime of mathematics. But there are so many other directions! For example, when view things in the spacetime of mathematics from the perspective of time, there are also large time scales and short times scales, both of which are important in historical descriptions. But the framework of spacetime suggests even more balanced ways to understand history. Just think of our daily life experiences: If one wants to take a 3D-picture of an object, one needs to take pictures from the front, back, both sides, and also from the top and bottom. Therefore, a complete description of a historical event requires multiple perspectives. Note that both the time factor and the space factor are always, but only, parts of them.[30] Another important point is that when we view a mathematical event or result as *an embedded point* in the spacetime of mathematics, this embedding itself suggests the importance of the whole for the purpose to understand each individual point.

[30] The importance of multiple perspectives on the same historical event has been convincingly illustrated in a well-known history book *History in Three Keys: The Boxers as Event, Experience, and Myth* (Cohen 1997). In this book, the same event was viewed and examined from different vantage points.

A more academic motivation for the introduction of the spacetime of mathematics comes from the Minkowski spacetime in special relativity, which is the combination of the three-dimensional Euclidean space and time into a four-dimensional manifold, and was a great contribution of Hermann Minkowski in 1908. The introduction of the Minkowski spacetime revolutionized the theory of special relativity of Einstein by providing the right framework, and also provided a particularly important example of more general spacetimes in the theory of general relativity. Though the intermingled relation between space and time was known earlier to Einstein and others, for example, Poincaré, this simple formulation by Minkowski of combining two naturally separate factors of space and time into a common spacetime changed people's perspectives and understanding of both the space and the time.

There are several things we can learn from the theory of relativity in physics, for example, how the combination of the space and the time makes it easier to understand the interaction between and motion of objects in space, for example, the causal structure. Since understanding causes is one major issue in the history of mathematics, putting the study of the history of mathematics in the setting of the spacetime of mathematics can help. Another useful lesson from the spacetime is the important insight that the gravity can be interpreted in terms of the curvature of the spacetime. Therefore, an adequate understanding of the geometry of the spacetime of mathematics will allow people to understand the motion of objects in the spacetime, in particular their trajectories. This suggests that a sufficient knowledge of the spacetime of mathematics is important for the purpose to understand the evolution of single mathematical result or event.

Another reason for the introduction of the spacetime of mathematics is related to mathematics itself. Since the history usually deals with "remarkable facts" according to the definition we cited at the beginning of Sect. 3.2 and often emphasizes views of events from the perspective of long time, though some detailed short time period and local studies are important and emphasized too, and topology in mathematics also studies large scale features of topological spaces, in particular manifolds, I wondered if some ideas of topology can also provide some guidance to the study of the history of mathematics. The particularly relevant mathematical theory is the Morse theory in differential topology which identifies special points of a given manifold, the so-called *critical points*, and shows how these critical points interact and determine the global shape of the space. In some sense, this gives one important concrete example illustrating the fruitfulness of the *dialectic pair of local versus global*. For example, the south pole and the north pole are the only two critical points of the sphere from a perspective of the Morse function given by the height function. Starting at these poles, one can gradually expand each of them into a disk which eventually becomes a hemisphere, and the two hemispheres meet at the equator, giving rise to the whole sphere. This example explains how the *global properties* of the sphere are explained by the *local properties* of the two poles.

Once we have the notion of the spacetime of mathematics, then the Morse theory comes into play. Major historical results or events correspond to critical points of the spacetime of mathematics. As in the Morse theory, we need to understand the *local structures* of the spacetime around these points. Then there is also a question of how to put these local studies into a global understanding of the development of mathematics. They correspond to the micro and macro studies in the history.

One important lesson from the Morse theory is that *there are different types of critical points*, as we have seen in the case of a torus in footnote.[31] Therefore, the standard *single* model of a scientific revolution in Kuhn (2012) seems to be too simple minded. Instead, we should look for critical points in the spacetime of mathematics, and how they have affected the geometry, or rather the structure of the spacetime of mathematics, i.e., the development of mathematics. If one thinks globally about mathematics and its development, several examples of critical points would come to our mind:

1. Gauss' book *Disquisitiones arithmeticae*. One detailed study of the impacts of this book was done in Goldstein et al. (2007).
2. Riemann's works in complex analysis, in particular his paper on abelian functions, and their influences in related subjects such as topology and algebraic geometry, besides the complex analysis itself. One comprehensive study was done in Bottazzini and Gray (2013).
3. Hilbert's list of 23 problems. One systematic overview was given in Gray (2000b).
4. Probably to a lesser degree, Poincaré's use of the hyperbolic geometry in his works on Fuchsian functions, which put hyperbolic geometry into the main stream of mathematics,[32] before that the main interest was more philosophical, as Poincaré explained in his extensive self-analysis of his scientific works (Poincaré

[31] For a more general manifold, the process is similar, though the result or its shape is more complicated. One good example is the torus, i.e., a donut. Make the donut stand up. Then one can imagine how to built it from four critical points: top, bottom, and two middle saddle points. Note that there are different types of critical points in this case.

[32] This use of the hyperbolic plane by Poincaré was a starting point of many great theories and major results in mathematics. For example, the hyperbolic plane is a special case of symmetric spaces introduced by E. Cartan, which are Riemannian manifolds whose curvature tensor has zero covariant derivative and include and are more general than Riemannian manifolds with constant curvature. One major insight of Cartan is that symmetric spaces are closely connected with Lie groups, and this connection between symmetric spaces and Lie groups has opened up a new frontier for differential geometry, the theory of Lie groups and harmonic analysis. For example, the theory of Fuchsian groups and Fuchsian functions have led to the current Langlands program through the extensive theory of automorphic functions; and the uniformization of Riemann surfaces through surfaces of constant curvature, in particular hyperbolic surfaces, has also led to Thurston's geometrization program in geometry and topology, which was proved by Grigori Perelman in 2002–2003.

1921, §II);[33] and his works in topology which have changed the landscape of mathematics in the twentieth century.

In some sense, the question whether each of the above is a revolution or not is not so important. What is important is to understand how they have influenced the development of mathematics. This brings up one point which we have emphasized at the beginning of this chapter. Even if the notion of spacetime of mathematics does provide a convenient and helpful framework to think about and understand the history of mathematics, the difficult and important thing is to do it. For any of the above four items, which of them will not take a huge amount of efforts in order to achieve a satisfactory comprehension?

When we only talk about the spacetime of mathematics above, it might give the impression that it only represents the so-called internal history of mathematics. But the external history of mathematics is also important and crucial for the purpose to obtain a complete understanding of the history. In this case, some mathematical thinking can help too. We can either enlarge the scope of the spacetime of mathematics, since it is often difficult to separate mathematics from other subjects. Or we can introduce a bigger spacetime: *the spacetime of everything*, and embed the spacetime of mathematics into the spacetime of everything as a *subspace*.

Now imagine a surface S such as the sphere S^2 embedded in the space $\mathbb{R}^3$. There are two kinds of geometry of S. The extrinsic (or rather external ?) geometry of the surface S is concerned with the geometry and shape of S such as the bending or curvature with respect to the ambient space $\mathbb{R}^3$, and has probably been studied since antiquity (think of conic sections in the Greek mathematics). It was Gauss who initiated the *intrinsic geometry* (or should we say "internal geometry"?) of the surface S. Then Riemann vastly generalized the intrinsic geometry to higher dimensional abstract manifolds in his famous lecture "On the hypotheses which lie at the foundations of geometry". The rest is history, as people would like to say.

Using this description of the external and internal geometries of surfaces, we can apply it to the embedding of the spacetime of mathematics in the spacetime of everything, and understand this subspace both internally and externally. Consequently, for the history of a mathematical question, we also need to view it from the ambient space, the spacetime of everything including the culture, social, philosophical, political perspectives and more. Hence the external history of mathematics is an essential part of the history of mathematics. Naturally, the devil is in the details.

How to Make Use of the Spacetime of Mathematics

After reading the earlier version of this chapter, Xiaofei Wang suggested to illustrate the use of the framework of spacetime of mathematics through some concrete problems in the history of mathematics in order to see what kind of guides it can

[33] Here we are mainly talking about the specific non-Euclidean geometry given by the hyperbolic geometry, but not the general non-Euclidean geometries. Otherwise, as José Ferreirós pointed out, it will be necessary to enter into extensive discussions related to philosophical issues and applications of non-Euclidean geometries in mathematical physics.

provide. We will discuss two examples: (1) Poincaré and his scientific works, (2) Comparison between the open problems proposed by Hilbert in 1900 and Poincaré in 1908.

(1) A Framework to Understand Poincaré and His Scientific Works

Given the importance and reputation of Poincaré scientific works, it is not surprising that there have been many books and papers about him and his works, in particular two recent outstanding biographies Gray (2013), Verhulst (2012). One question for us is whether the notion of the spacetime of mathematics can be used to find another framework or a perspective to understand Poincaré and his scientific works.

When we look at the development of mathematics, or imagine a global view of the spacetime of mathematics, it is probably clear to many people that Poincaré's works occupy a very distinguished subset, since they have had a big impact on mathematics up to now. The problem is to understand how this special *Poincaré subspace* has affected or fitted into the shape of the whole spacetime of mathematics and also the spacetime of everything. The study of history often considers and emphasizes the evolution of events, i.e., the time factor. One important suggestion from the spacetime of mathematics is that *the view across different subjects and the unity between them are as important as the view through time.* As an extension of this, it is also natural and important to consider any problem under study in the history of mathematics by taking into consideration many subjects and factors outside mathematics, since they are living together inside the spacetime of everything.

We want to stare at this Poincaré subspace and nearby places and look for several kinds of things: global features or major themes running through them, and distinguished places or special spots in this Poincaré subspace, and how they are connected. Maybe it is helpful to compare it to how to look at a mountain both from afar and close by, and then summarize the impressions of what we have seen or experienced.

After thinking for a while, it seems the following images and questions may arise, among other things:

1. Poincaré contributed substantially to several fields including mathematics, celestial mechanics, mathematical physics, physics, philosophy of science, popular writings on science. But his most important contribution lies in mathematics, and the richness and unity of mathematics is well reflected in his works and other activities.
2. In the overall picture of his mathematical contributions, several major themes appear right from the beginning of Poincaré's research in mathematics: (a) Differential equations and (b) Number theory.
   These two trees have grown up in the soil of geometry, which is broadly interpreted and includes group actions or the theory of transformation groups, and topology (or analysis situs). They have branched out respectively in multiple directions. For the first, it includes both local and global studies of differential equations, striking applications to complex analysis, and differential equations arising in celestial mechanics and mathematical physics. For the

second, it includes the geometry and reduction of quadratic forms, the analytic number theory and the arithmetic geometry.

These trees or themes also interact at several places. For example, Fuchsian groups occur in both: more general Fuchsian groups arising from the monodromy groups of linear differential equations with algebraic coefficients, and the arithmetic Fuchsian groups from the automorphism groups of indefinite ternary quadratic forms.

3. Topology first appeared in the background of Poincaré works. Later it came to the front and emerged as one of his most important contributions, if not the most important, in mathematics.
4. In the above two major themes, there are special spots. For the first, we can spot Poincaré theory of Fuchsian functions and its application to the uniformization of algebraic curves, the qualitative study of the three body problem, and eigenvalues and eigenfunctions of linear differential operators. For the second, we can spot his pioneering work on the group structure of rational points of elliptic curves.
5. If we look more closely, there are many other special regions in the Poincaré subspace: complex analysis of both one and several variables, algebraic geometry, the Lie theory etc.
6. Outside the spacetime of mathematics, there are many lecture notes and papers on diverse topics in subjects including physics, mathematical physics, celestial mechanics and dynamics, and probability etc. There are also his popular lectures and writings on sciences and philosophy.
7. If we look beyond Poincaré's own works, we need to examine the works before him and works after him, the latter being huge.
8. Now we need to figure out connections between all the above things. For example, among many other questions, here are several: (a) how the French educational system affected Poincaré's formation as a mathematician, in particular the education at École Polytechnique. (b) How the teaching at universities influenced Poincaré's research interests and the choice of topics. (c) Which culture and social factors determined the teaching systems at both the lower and higher levels. (d) How Poincaré's status as the leading French scientist at his time impacted his works outside pure mathematics, and how they were reflected in Poincaré's popular writings. (e) How to understand the impact and legacy of Poincaré both in short and long terms, in particular, in comparison with Hilbert.
9. How did Poincaré's philosophy and philosophical writings influence and reflect his works in mathematics and related subjects?
10. We need to compare Poincaré's works in mathematics with all the major topics in his time, i.e., compare the Poincaré subspace with the spacetime of mathematics, to either substantiate his fame of being "the last universalist in mathematics", or point out some important gaps in his extensive works. For example, Weyl wrote in Weyl (1951, 531): the "most fascinating branch of mathematics [is] ... class field theory", but Poincaré never worked on or contributed to this subject.

11. Last but not the least, we need to understand the unity of the whole scientific works of Poincaré, and positive, and maybe also negative, impacts of the whole on each part, or one part on another.[34]

We hope this outlook, or the map, based on the spacetime of mathematics can provide one way to organize questions and answers to understand better Poincaré and his works. Of course, the hard part is to do detailed work and understand the points and questions mentioned above.

(2) A Framework to Understand and Compare Open Problems Posed by Hilbert and Poincaré

If we think of lists of open problems in mathematics, the first list which comes to our mind is very likely Hilbert's famous list of 23 problems given in Paris in 1900 (Gray 2000b). Many individual or groups of mathematicians have tried to do similar projects. If we want to find another list and compare it with Hilbert's list, the open problems proposed by Poincaré in 1908 in Rome might be a good choice. Comparison between them has indeed be emphasized and made in Davis and Mumford (2008) and Gray (2012). (A complete new English translation of Poincaré's open problems is contained in Gray (2012).)

We want to take another look at the choice of this comparison and some related questions through the framework of spacetime of mathematics.

1. Both Hilbert and Poincaré left strong traces in the spacetime of mathematics. Though their open problems are not limited by their own works, they are certainly related. Hence this provides one concrete way to compare the mathematical works of Hilbert and Poincaré. Both lists of open problems cover broad ranges of topics in mathematics and show the unity of mathematics across different subjects, and a careful study of them provides a special way to understand the richness and some global features of mathematics. These are basic views in the framework of spacetime of mathematics.
2. Both the contents and styles of the mathematical works of Hilbert and Poincaré are different. In some sense, they complement each other. Therefore, their combinations can cover more parts of the spacetime of mathematics. In some sense, they cover almost the whole spacetime of mathematics up to now. The separation and blending of these two extensive mountains in the spacetime of mathematics can convey and make us feel the dynamics in the history of mathematics, for example, the rise and fall of some subjects and problems, and high hopes in some problems before their solutions and disappointment after their solutions.

[34] As Gray pointed out in Gray (2012, 15–16), one outcome of simplifying Poincaré's more technical papers for his expository books and of the easy access and popularity of such writings might have "contributed to a misleading picture of Poincaré's famous lecture" on the open problems under discussion here.

3. The spacetime of mathematics also suggests ways to understand them. For example, it is natural to understand each in connection with the whole spacetime. Specifically, the following questions and problems seem relevant:

   (a) Compare Hilbert's problems with the mathematical works of Hilbert to understand better why Hilbert chose these problems. How do they reflect Hilbert's view and understanding of mathematics?
   (b) Compare Hilbert's problems with the bigger space of the total of all major works in mathematics around that time to learn how they reflected the status of mathematics at that time, and to identify possible major missing areas.
   (c) Usually attention is paid to each individual problem. But it is also important and fruitful to understand connections, especially the implicit ones, between seemingly different and unrelated problems. Unexpected connections which surfaced only later are particularly interesting.
   (d) Consider the history and motivations of each individual problem, or rather related problems and subjects, before Hilbert's formulation of the problem. We can also consider the history of interactions between different problems.
   (e) Do the same things for Poincaré, and compare similarities and differences between Hilbert's problems and Poincaré's problems. For example, how the missing parts of one person are filled in by the other.
   (f) We can move forward in the spacetime and look at the development of these problems of Hilbert and Poincaré and related subjects, up to now, including their current status. This may not be a typical history problem, but certainly is important and interesting to many people. In any case, this will be one important part of a complete description of the traces of these special Hilbert and Poincaré sites in the spacetime of mathematics.

   The above represents different views in the spacetime of mathematics.
4. It seems that Hilbert's problems have been very influential, while Poincaré problems have been relatively unknown. There are explanations about this and various other issues in Gray (2012) and Davis and Mumford (2008). We can also try to understand their difference in terms of several aspects both inside and outside the spacetime of mathematics for this difference: (1) Is it because of the contents of the problems? This also raises the question of what is important in mathematics. (2) Is it because of the ways the problems were formulated and presented?[35] (3) How much did institutional factors, academic networks, political and other factors contribute to this?

[35] Comparison between the problems of Poincaré and Hilbert makes one wonder if they are similar to two kinds of books on the nature of history as we will mention in the last footnote: one gives higher level and general guidances or suggestions while the other gives more concrete and detailed descriptions. For example, Lie groups are discussed in both Poincaré's Problem VII (or Part VII) (Gray 2012, 22) and Hilbert's Fifth Problem (Gray 2000b, 254–257), and the above differences between them seem quite visible in this case.

As in the previous example about Poincaré and his works, this framework can only provide one way to raise and organize questions and answers about the open problems of Hilbert and Poincaré. Detailed understanding of the mathematics and other non-mathematical factors involved in their formulations and development will be the key to any historical project about these problems.

Due to the length limit, we refer the reader to Ji (2022) for some other details and aspects about the use of the spacetime of mathematics in the study of the history of mathematics.

3.6 Conclusion

Given the assigned title of this chapter, I need to conclude this paper with several general remarks, or do a bit of philosophizing.

Before doing history, one needs to know what is history, which is a proper and relevant philosophical question.[36] While methodologies in history[37] are needed and helpful, especially for beginners, one can learn them better in combination with doing it: the hands-on learning philosophy.

The key point of doing history is to develop a habit of thinking historically by doing it: working on problems which can allow people to understand better the development of mathematics, i.e., understanding the structure of the spacetime of mathematics through its critical points and its embedding in the spacetime of everything.

One central part of historical thinking is the *multiplicity*: There are multiple types of problems, approaches, and results. They are all important. The question is how to appreciate them and combine them. Some historians like to say: the house of history is big enough, and there is room for everyone. Probably this is also true with the history of mathematics. There is something to learn from others.

Take a dialectic view towards everything. Walking in the right direction is probably more important than walking fast in order to reach the final destination. On the other hand, no amount of thinking about the history of mathematics will amount to much unless sound methodologies are put into practice and one is willing to look into details and to do hard work, for example, reading and trying to understand

[36] According to Collingwood (1994, 348, 359), the "central question" of the philosophy of history is "what is history". See also Ji (2022, §1) for some comments and various quotes on relations between philosophy and history. Socrates said famously: "The unexamined life is not worth living". Can this be adapted to the history of mathematics? And how?

[37] It seems that there are two kinds of books which explain the nature of history. One kind is more philosophical and provides guidance at the higher level or the macro guidance. One example is the book Collingwood (1994). Another kind is more at the practical level or the micro level, giving concrete and detailed instructions on how to do history. One example is the book Furay and Salevouris (2015). Though different, both kinds of books are important and needed. They can be used in a dialectic way.

difficult primary sources, but one should not be buried in minute details without keeping a big picture in mind either.

While doing history is probably enjoyable and rewarding to the researcher, it is also important to keep others in mind: Why your results and writings are interesting and beneficial to others. Of course, the ultimate test or judgement is given by history.

Acknowledgments I would like to thank the following people for reading the long paper (Ji 2022): José Ferreirós, Jamie Tappenden, Viktor Blåsjö, Leo Corry, Wenlin Li, Chang Wang and Yuanyuan Liu, Chan-chan Guo, Erhard Scholz. Their valuable comments and suggestions have influenced the writing of the much shortened and hopefully improved version in this chapter. In particular, Jamie's suggestion for emphasizing the discussion about Galois' letter and his teaching of the English idiom "Hamlet without the Prince", José's suggestion for focusing on different ideas about history by mathematicians and historians, Erhard's suggestion for focusing on the idea of the spacetime of mathematics, which is one of few new contributions in Ji (2022), Viktor's and Leo's very encouraging comments on the whole paper and specific suggestions on several points including the discussion about Galois' letter all have influenced the choice of topics included in this chapter. José's initial suggestion and later correspondences are crucial to the writing of both Ji (2022) and this chapter. His emphasis on philosophy encouraged me to learn a bit about philosophy, which led me to the important point of view of dialectics, which is probably the most important thing I have ever learned about the nature and practices of the history of mathematics. If I were to include "philosophy" in the title of this chapter as suggested by José and Erhard and were to give the beginning student in the history of mathematics an one-sentence advice, I would say: "Think dialectically in the sense as explained above in order to have a productive and joyful adventure in the history of mathematics!"

After the first version of this chapter was finished, many valuable comments and constructive suggestions have been made on it by the following people: a referee, Viktor Blåsjö, Leo Corry, José Ferreirós, Chan-chan Guo, Wenlin Li, David Rowe, Erhard Scholz, Jamie Tappenden, Xiaofei Wang, Jing Yang, and Chang Wang and a team of 6 young historians organized by him: Jinze Du, Yuwen He, Juan Li, Yuanyuan Liu, Qiang Yang, Zhongmiao Yu. I am grateful for their help, especially for the careful reading of Viktor, Leo, José and Jamie.

Finally, I would like to dedicate this chapter to Jeremy Gray for his 75th birthday, thanking him for his many helpful email correspondences about the history of mathematics in the past few years and his very valuable and encouraging comments on an earlier version of this chapter, and wishing him many healthy and productive years to come!

References

Arnold, J. 2000. *History: A Very Short Introduction*. Oxford University Press.

Barrow-Green, J., J. Gray, and R. Wilson. 2022. *The History of Mathematics: A Source-Based Approach*, vol. 2, 687 pp. AMS/MAA Press.

Bell, E.T. 1937. *Men of Mathematics: The Lives and Achievements of the Great Mathematicians from Zeno to Poincaré*. Simon and Schuster.

Bottazzini, U., and J. Gray. 2013. *Hidden Harmony–Geometric Fantasies. The Rise of Complex Function Theory*, xviii+848 pp. Springer.

Boyer, C.B. 1989. *A History of Mathematics*, 2nd ed. Edited and with a preface by Uta C. Merzbach, xviii+762 pp. John Wiley & Sons.

Brechenmacher, F. 2011. Self-portraits with Évariste Galois (and the shadow of Camille Jordan). *Revue d'histoire des mathématiques* 17: 273–371.

Budd, A. 2010. *The Modern Historiography Reader: Western Sources*. Routledge.

Cajori, F. 1991. *A History of Mathematics*, 5th ed., xi+524 pp. AMS Chelsea Publishing.
Carr, E.H. 1961. *What is history?* Vintage Books.
Chemla, K. 2018. How has one, and how could have one approached the diversity of mathematical cultures? In *European Congress of Mathematics*, 1–61. Zürich: Eur. Math. Soc.
Chong, C.T., and Y.K. Leong. 1986. An interview with Jean-Pierre Serre. *The Mathematical Intelligencer* 8 (4): 8–13.
Cohen, P. 1997. *History in Three Keys: The Boxers as Event, Experience, and Myth*. Columbia University Press.
Collingwood, R.G. 1994. *The Idea of History*, revised edition. Oxford University Press.
Davis, P., and D. Mumford. 2008. Henri's crystal ball. *Notices of the American Mathematical Society* 55: 458–466.
de Saint-Gervais, H.P. 2016. *Uniformization of Riemann surfaces. Revisiting a Hundred-year-old Theorem*, xxx+482 pp. European Mathematical Society.
Ehrhardt, C. 2011. *Évariste Galois. La fabrication d'une icone mathématique*, 301 pp. Paris: Éditions de l'École des Hautes Études en Sciences Sociales.
Ehrhardt, C. 2012. *Itinéraire d'un texte mathématique: Les réélaborations des écrits d'Evariste Galois au XIXe siècle*, 298. Paris: Hermann.
Einstein, A. 1944. A letter to R. A. Thornton, unpublished letter dated 7 December 1944 (EA 6-574), Einstein Archive, Hebrew University, Jerusalem.
Engels, F. 1940. *Dialectics of Nature*, xvi + 383 p. New York: International Publishers.
Freudenthal, H. 1955. Poincaré et les fonctions automorphes. In *Le livre du centenaire de la naissance de Henri Poincaré: 1854–1954*.
Furay, C., and M.J. Salevouris. 2015. *The Methods and Skills of History: A Practical Guide*, 4th ed.
Goldstein, C. 2011. Charles Hermite's stroll through the Galois fields. *Revue d'histoire des mathématiques* 17: 211–270.
Goldstein, C., N. Schappacher, and J. Schwermer. 2007. *The shaping of arithmetic after C. F. Gauss's Disquisitiones arithmeticae*, xii+578 pp. Springer.
Grabiner, J. 1975. The mathematician, the historian, and the history of mathematics. *Historia Mathematica* 2: 439–447.
Grattan-Guinness, I. 1971. Towards a biography of Georg Cantor. *Annals of Science* 27: 345–391.
Grattan-Guinness, I. 2004a. The mathematics of the past: distinguishing its history from our heritage. *Historia Mathematica* 31: 163–185.
Grattan-Guinness, I. 2004b. History or heritage? An important distinction in mathematics and for mathematics education. *The American Mathematical Monthly* 111: 1–12.
Gray, J. 2000a. *Linear differential equations and group theory from Riemann to Poincaré*, 2nd ed. Boston: Birkhäuser, xx+338 pp.
Gray, J. 2000b. *The Hilbert Challenge*, xii+315 pp. Oxford University Press.
Gray, J. 2012. Poincaré replies to Hilbert: on the future of mathematics ca. 1908. *Mathematical Intelligencer* 34: 15–29.
Gray, J. 2013. *Henri Poincaré. A Scientific Biography*, xvi+592 pp. Princeton University Press.
Hadamard, J. 1999. *Non-Euclidean Geometry in the Theory of Automorphic Functions.* With historical introduction by Jeremy J. Gray, xii+95 pp. American Mathematical Society.
Hanson, N.R. 1963. Commentary in *Scientific Change*, ed. A.C. Crombie, 458–466. London: Heinemann.
Hawkins, T. 1987. Cayley's counting problem and the representation of Lie algebras. In *Proceedings of the International Congress of Mathematicians* (Berkeley, 1986), vol. 2, 1642–1656. American Mathematical Society.
Hawkins, T. 2000. *Emergence of the Theory of Lie Groups. An Essay in the History of Mathematics 1869–1926*, xiv+564 pp. Springer.
Ji, L. 2022. Some perspectives on the history of mathematics: what, why, and how, preprint, April 2022, 145 pages.
Ji, L. 2023. *The Theory of Ambiguity in Galois' Testamentary Letter*, in preparation.

Ji, L., and Wang, C. 2020. Poincaré's stated motivations for topology. *Archive for History of Exact Sciences* 74: 381–400.
Ji, L., and Wang, C. 2022. Poincaré's works leading to the Poincaré conjecture. *Archive for History of Exact Sciences* 76: 223–260.
Katz, V. 2009. *A History of Mathematics. An Introduction*, 3rd ed. HarperCollins College Publishers.
Klein, F. 1884. *Lectures on the Icosahedron and the Solution of Equations of the Fifth Degree.* With a new introduction and commentary by Peter Slodowy. Higher Education Press, Beijing, 2019. xiv, XI+306 pp. Originally published in German in 1884.
Klein, F. 1979. *Development of Mathematics in the 19th Century* Mathematics Sci Press, ix+630 pp.
Kline, M. 1990. *Mathematical Thought from Ancient to Modern Times*, vol. 2. 2nd ed., i–xx, 391–812, i–xxii. The Clarendon Press/Oxford University Press.
Kragh, H.S. 1987. *An Introduction to the Historiography of Science*. Cambridge University Press.
Kuhn, T. 1977. *Essential Tension*. The University of Chicago Press.
Kuhn, T. 2012. *The Structure of Scientific Revolutions: 50th Anniversary Edition*, 4th ed. University of Chicago Press.
Lie, S. 1895. Influence de Galois sur le développement des mathématiques. In *Le cententenaire de lÉcole normale* (1795–1895) (Paris, 1895), 481–489.
May, K. 1975. What is good history and who should do it? *Historia Mathematica* 2: 449–455.
Neuenschwander, E. 2022. Die Ausleihjournale der Göttinger Universitätsbibliothek: eine bisher kaum benutzte Quelle zur Analyse von Riemanns bahnbrechenden Ideen, Preprint N°512, the Max Planck Institute for History of Science in Berlin, 2022.
Neumann, P. 2011. *The Mathematical Writings of Évariste Galois*, xii+410 pp. European Mathematical Society.
Okamoto, K., and Y. Ohyama. 2020. Mathematical works of Hiroshi Umemura. *Annales de la Faculté des sciences de Toulouse: Mathématiques* 29: 1053–1063.
Picard, É. 1897. *Oeuvres Mathématiques d'Évariste Galois, avec une introduction par Emile Picard*, 8vo, x + 63. Paris: Gauthier-Villars et Fils.
Poincaré, H. 1921. Analyse des travaux scientifiques de Poincaré faite par lui-même. *Acta Mathematica* 38: 1–135.
Popper, K. 1968. *The Logic of Scientific Discovery*. Harper & Row, Publishers.
Ramis, J.-P. 2020. Hiroshi Umemura et les mathématiques francaises. *Annales de la Faculté des sciences de Toulouse: Mathématiques* 29: 1007–1052.
Raussen, M., and C. Skau. 2012. Interview with John Milnor. *Notices of the American Mathematical Society* 59 (3): 400–408.
Rothman, T. 1982a. The short life of Évariste Galois. *Scientific American* 246: 136–149.
Rothman, T. 1982b. Genius and biographers: the fictionalization of Évariste Galois. *The American Mathematical Monthly* 89 (2): 84–106.
Rothman, T. 1989. *Science à la mode: Physical Fashions and Fictions*, 207 pp. Princeton University Press.
Rowe, D. 1986. David Hilbert on Poincaré, Klein, and the world of mathematics. *The Mathematical Intelligencer* 8 (1): 75–77.
Smith, D.E. 1906. *History of Modern Mathematics*, 4th ed. J. Wiley & Sons.
Smith, D.E. 1959. *A Source Book in Mathematics*, 2 vols. Dover Publications.
Stedall, J. 2012. *The History of Mathematics. A Very Short Introduction*, xviii+123 pp. Oxford University Press.
Tannery, P. 1930. *Mémoires Scientifiques*, vol. X. Gauthier-Villars.
Truesdell, C. 1984. Genius and the establishment at a polite standstill in the modern university: Bateman. *An idiot's fugitive essays on science: methods, criticism, training, circumstances*, 423–424. Springer.
Verhulst, F. 2012. *Henri Poincaré. Impatient Genius*, xii+260 pp. Springer.
Weil, A. 1975. Book Review: Leibniz in Paris 1672–1676, his growth to mathematical maturity. *Bulletin of the American Mathematical Society* 81 (4): 676–688.

Weil, A. 1980. History of mathematics: why and how. In *Proceedings of the International Congress of Mathematicians* (Helsinki, 1978), 227–236.

Weinberg, S. 2004. The making of the Standard Model. *The European Physical Journal C* 34: 5–13.

Weinberg, S. 2018. *Third Thoughts*. Harvard University Press.

Weyl, H. 1951. A Half-Century of Mathematics. *The American Mathematical Monthly* 58 (8): 523–553.

Weyl, H. 1952. *Symmetry*, viii+168 pp. Princeton University Press.

Wussing, H. 1991. Historiography of mathematics: Aims, methods, tasks. In *World Views and Scientific Discipline Information*, ed. W.R. Woodward and Robert S. Cohen, 63–73. Springer.

Chapter 4
Why Historical Research Needs Mathematicians Now More Than Ever

Viktor Blåsjö

Abstract Using the history of the calculus as an example, I identify some trends in recent scholarship and argue that the time is ripe for a "new internalism" in the historiography of mathematics. The field has made steady progress in the past century: mathematicians have provided clear expositions of the technical content of past mathematics, and historians have produced meticulous editions of textual sources. These contributions have been invaluable, but we are reaching a point where the marginal utility of further works of these types is diminishing. It is time to shape a paradigm of historical scholarship that goes beyond the factual-descriptive phase of the past century. Comparative interpretative work is now feasible thanks to the gains of the past century. Cognitive questions about mathematical practice provide a fascinating and underexplored avenue of research that we now have the tools to tackle. Mathematically trained researchers are needed for this enterprise.

4.1 Introduction

Today is a golden moment to reunite historians and mathematicians. Their acrimonious divorce some decades ago is proving increasingly detrimental to both. Mathematicians sit on technical expertise and are as interested as ever in history, but they are spinning their wheels with repetitive expository accounts, since no historiographical framework helps them mobilise their skills for historical research purposes. Historians have shut themselves off from mathematicians to avoid anachronism, but forget that, while this asceticism may once have been a sound cleanse, it is unsustainable as a permanent diet.

The work that the two divorcees have done while apart is a perfect foundation for their reunion. Retreating to their individual comfort zones, scholars perfected the state of local scholarship in those domains. But we cannot keep tinkering in

V. Blåsjö (✉)
Mathematical Institute, Utrecht University, Utrecht, The Netherlands
e-mail: V.N.E.Blasjo@uu.nl

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_4

fragmented niches forever. With the accumulation of detailed studies, we are now in a position to take up new lines of research based on synthesising and comparative perspectives.

Let me take an area I have worked on—the early history of the calculus—as a case in point to highlight the fruitful circumstances that make a new internalist historiography more opportune than ever.

4.1.1 Opportunity: Re-engage Mathematicians in History of Mathematics

The history of the calculus remains highly relevant to the mathematician's world-view, as seen for instance in recent high-profile books where the history of calculus features prominently, such as Strogatz (2019)—a *New York Times* Bestseller—or Bressoud (2019)—an interweaving of historical and educational aspects of calculus by a former President of the Mathematical Association of America. But, regrettably, in terms of historical content, these books are less groundbreaking, relying in large part on a rather limited set of historical set-pieces that are often repeated in one popular work after another. The title of one very successful book of this type—"The Calculus Gallery" (Dunham 2005)—inadvertently hints at the limitations of this approach: historical mathematics is reduced to a canonised collection of iconic snapshots, briskly toured under fluorescent lights; seen only on their aloof pedestal rather than in the creator's workshop. Left unanswered are questions about how the technical details of particular mathematical masterpieces were organically embedded and functioned in broader research practice.

All of these authors are highly qualified mathematicians, yet their competence is wasted on repetitive re-exposition of known material because mathematically inclined authors lack—and do not find in recent historical scholarship—any sense of direction in which history of mathematics as a research field could evolve through the kind of analysis that a mathematician can provide. "The early history of the calculus of variations is a well-beaten track" (Giaquinta and Freguglia 2016, vii) another recent book apologetically admits, before beating the same track once again. There is a wealth of fascinating and unexplored historical questions that mathematicians could very fruitfully address, but mathematicians do not know how to do new and valuable scholarship by asking novel questions about the technical substance of past mathematical practice. We need a new historiography to provide this lacking impetus, and thus rejuvenate history as a mathematical research field.

4.1.2 Opportunity: Recent Historical Scholarship Abundant in Details but Lacking in Global Vision

Current scholarship in the history of calculus is lopsided toward specialised source studies. The Newton Project and *Akademie-Ausgabe* of the works of Leibniz are epicenters of expertise in the field. By providing comprehensive and meticulous editions of sources, these projects are invaluable. But their success inherently contain the germ of a fresh start in a different direction: the excellent state of specialised source work opens the way for synthesising perspectives.

This is timely, as for the history of the calculus (as for many other historical topics) no comprehensive and accessible survey that synthesises the insights of recent research and points the way to future research has been written for generations. Highly dated books such as Edwards (1979) are still in print and widely used; the antiquated Boyer (1959) is still Amazon's top hit for "history of the calculus" and no up-to-date alternative is available. This is doubly unfortunate. For historical scholarship itself, it shows that increasing specialisation has left the field lacking in big-picture vision. Furthermore, for students and mathematicians, the lack of accessible overview is a gatekeeping barrier that makes it very difficult to keep up-to-date with recent historiography and enter the field of historical research. This blocks mathematically talented people from contributing to the field, and hence the sense that modern historical scholarship is divorced from mathematics becomes a self-fulfilling prophecy.

4.1.3 Opportunity: Join Forces with the "Practice Turn" in Recent Philosophy of Science and Mathematics

Not only mathematicians can profitably be re-invited into historical scholarship, but also philosophers. Again the timing is just right. In the twentieth century, much philosophy of mathematics was fixated on logical rigour. In the case of the history of the calculus, this meant for example many papers on the relation between Robinson's nonstandard analysis and classical infinitesimal calculus—an anachronistically motivated debate that is orthogonal to the concerns of historical mathematicians. But with the more recent "mathematical practice" movement in philosophy of mathematics, philosophers have turned to questions regarding the motivation, methodology, heuristics, and research choices of historical mathematicians, as well as cognitive-historical questions such as for instance how visual and notational devices shape styles of thought. Hence the interests of the historian and the philosopher are more favourably aligned now than in the past century.

4.1.4 Opportunity: Current Societal-Educational Questions Turn on Calculus

A new historiography has the opportunity to be inclusive in another important direction as well. Again the history of the calculus provides a case in point. Calculus has an image problem, now more than ever. It was never a crowd-pleaser to begin with, but the old student refrain "when will I ever use this?" has lately been gaining considerable traction among senior academics as well. An October 2019 *Freakonomics* episode joined a growing chorus that would be happy to see dusty old calculus yield space in the curriculum for more "twenty-first-century skills" such as "data fluency." As books such as Strogatz (2019), Bressoud (2019), and Orlin (2019) indicate, history is one of the mathematician's best tools for conveying the relevance and excitement of calculus amid such assaults. What Heilbron (1987, 559), says of science is true for mathematics as well: as "applications threaten to suffocate the traditional core of the subject"—a core "informed by the humanistic ideal"—"partnership with history may be the most promising course by which science may save itself from being crushed by its technological successes."

We historians must find a way to build on all of these opportunities constructively, rather than isolating ourselves to uphold a puritanical ideal of our subject.

4.2 Example: Huygens's Proto-Calculus

Figure 4.1 outlines a mathematical argument from the works of Huygens that is quite typical of its time. The first thing that strikes a modern reader is the geometricity of the proof. Indeed, one may say that "Huygens actually thinks geometrically, he sees the relations in the figures, formulas are secondary to him," as Bos (1980, 132), observed in a different context. But this is merely a descriptive observation. We want to dig deeper and consider the ramifications of this point of view for the mathematical practice of the time. Compared to calculus proper, what were the cognitive possibilities and limitations of this style of proto-calculus mathematics?

Many aspects of Huygens's argument can be matched with analogous notions within the calculus: geometrical properties of tangents play the role of derivatives; inferring "global" properties of the system from a characterisation of all its instantaneous local states plays the role of integrating a differential equation; using a circle as a reference figure plays the role of using trigonometric functions to express the quantities and relationships involved. To what extent were these proto-calculus analogs functionally equivalent to their calculus counterparts? In some respects they could do everything the calculus can do; in other respects not. What respects are these exactly?

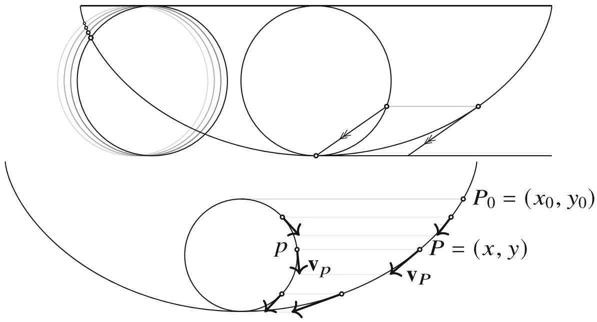


Fig. 4.1 Top: Definition of cycloid as the curve traced by a fixed point on a rolling circle, and the tangent of the cycloid expressed in terms of its generating circle in its middle position. Bottom: Huygens's proof that the period of cycloidal motion is independent of amplitude. The particle P is descending under gravity along a cycloid, starting from rest at P_0. Consider the horizontal projection p of this point onto an associated circle as shown. From physics we know that $|\mathbf{v}_P| = \sqrt{2g(y_0 - y)}$. From the tangent result shown on the left we know how to decompose this into vertical and horizontal components. By definition, $\mathbf{v}_p$ has the same vertical component, and is tangent to the circle. This determines the magnitude $|\mathbf{v}_p|$, which turns out to be constant throughout the descent and proportional to y_0. Hence the time of descent of P has been expressed in terms of the arc length of the circle. From there it immediately follows that the time of descent is the same for any P_0

In other words, what exactly did the calculus add that was new compared to these existing practices? For example:

- Did the calculus remove the need for the geometrical ingenuity exhibited by Huygens, and replace it with routine applications of symbolic-computational rules? Leibniz often stressed the value of his calculus in such terms, but he is a biased witness. Did arguments such as that of Huygens truly rely greatly on geometrical ingenuity and imagination, or is that merely how it appears to someone without a working knowledge of this style of mathematics?
- Does a calculus solution to a particular problem carry over more easily to a similar problem while Huygens-style geometrical proofs are sui generis? Euler (1736, 2), thought so. Can his opinion be validated by a comparison of pre-calculus and calculus historical sources, or did Euler only feel this way because he was more familiar with calculus methods?
- Did the calculus provide the tools to state general theorems about, say, entire classes of functions, whereas Huygens-style methods are limited to specific, concrete cases? An argument against this hypothesis, perhaps, is for example Huygens's completely general proof that the evolute of any algebraic curve is itself algebraic (Huygens 1673, III, Prop. XI).
- Is Huygens's approach damagingly dependent on working with "global" properties of entire figures and systems, whereas the calculus can successfully operate in the dark with local (differential equation) information and only need to interpret the final solution globally at the end, if at all?

- Did the calculus facilitate these kinds of problems primarily by brute-force power ("crunching the formulas"), or as a conceptual heuristic and a way of thinking about how to even formulate the problem in the first place? The latter point of view is perhaps what is captured by the paraphrase by Arnol'd (2012, vii), of a Newtonian maxim as "it is useful to solve differential equations."

4.2.1 Mathematical-Practice Historiography

Questions like those above may be called cognitive, to distinguish them from what is purely textual or factual. Cognitive questions concern how certain ideas functioned in the minds of historical thinkers, and what overall role they played in their mathematical thought: What could these ideas do, and what not? What was the lay of the land of mathematical research as seen through the lens of these ideas? How did outlooks such as that of Huygens and that of Leibniz differ in how they drew the boundaries and the infrastructure connections between the well settled, the active frontiers, and the aspirational *terra incognita* on the research landscape map?

Cognitive questions cut to the heart of what makes history relevant for many audiences. Mathematicians are drawn to these questions because they concern reconstructing past mathematics as it appeared through the eyes of active researchers. Mathematics teachers and students, because these questions point to a path of hands-on examples from which a mature view of the field gradually crystallises. Philosophers, because these questions trace the formation of fundamental concepts. Historians, because these questions target precisely what was idiosyncratic and uniquely situated about past ways of thinking.

But cognitive questions are elusive, since they try to get at thought processes that are beneath the surface. They cannot be answered by a purely textual analysis of source documents (the expertise of historians), nor by a purely formal analysis of the mathematical content (the expertise of mathematicians). Tackling cognitive questions therefore requires new historiographical methods that go beyond established practices of historians and mathematicians, but build on the strengths of both.

4.3 Need to Move Beyond "Photorealism" Historiography

The historiography of mathematics is stuck in a binary that for the past decades have pitted mathematicians and historians against each other in cartoonish terms. In what is by now a tired cliche, historians condemn the mathematicians' practice of utilising modern mathematics to analyse and illuminate historical works. Portraying this as the root of all evil, historians prided themselves on banning mathematical paraphrase and restricting themselves entirely to literal scrutiny of textual sources. This was in some ways a corrective in the right direction at the time, but it should not be mistaken for the endpoint and perfection of historiographical method. The simplistic good-versus-evil self-fashioning of the present consensus has become such a dominant narrative that past generations of mathematically oriented historical scholarship is now routinely dismissed as "at best anachronistic" (Imhausen 2021, 80).

Ultimately this point of view is as sterile and one-dimensional as taking hyperrealistic still lives to be the endpoint and perfection of art. Breaking free from the stifling ideal of photorealism allowed artists to better see the organic soul of their scenes and capture their human significance with greater emotive force. The same will happen in the history of mathematics. Liberating ourselves from the moribund "still life" textualist historiography, we will bring out what is less tangible but more vividly alive.

To be sure, reconstructing past mathematical thought is a tightrope walk that has long been difficult to get right. But Chang (2017) is right to subdivide internalist history of science into an "orthodox" and a "complementary" mode. "Orthodox" internalism is subservient to the current values of the field whose history is investigated, whereas "complementary" internalism is pursued precisely because it complements current orthodoxy in the field and regards it critically. Traditional internalist-mathematical approaches have been too often carelessly dismissed by historians based on arguments that in reality strike only against orthodox internalism.

It is true that mathematicians can be too cavalier in projecting modern notions onto past mathematics, as when Arnol'd speaks of Huygens investigating "the manifold of irregular orbits of the Coxeter group H_3" (Arnol'd 1990, 8). Of course, such approaches are likely to be insensitive to historical thought, and to bulldozer over its nuances with predetermined ideas and unwarranted extrapolations. (Examples of episodes from the history of the calculus that have been misunderstood for such reasons are discussed in Blåsjö (2015), Blåsjö (2017a), and Bell and Blåsjö (2018).) This is why it has become *de rigueur* among historians to insist—as for example the most prominent historical monograph on Huygens's mathematics immediately does—that "historical accuracy and insight are lost when results are couched in modern terms" (Yoder 2004, 7). This may be called the photorealism axiom of modern historiography.

The insistence on "photorealism"—or exact adherence to the surface form of the written text—has been a blessing and a curse for the historiography of mathematics. This ban on paraphrase has cleansed the field of many a naïve anachronism, as intended. But less widely recognised are its unintended knock-on effects. Photorealism effectively precludes comparative, synthesising studies, and hence forces a fragmentation of historical scholarship into narrowly specialised studies. This is an overreach of the photorealism axiom that goes beyond its originally intended scope and justification. New vistas for progress would open up if comparative and synthesising analyses could be rehabilitated as historical methods without reversing the gains made by the photorealism phase of the past decades.

4.3.1 Consequence of "Photorealism" Historiography: Microscopic Focus on Minor Sources

Photorealism historiography predictably steers the field into an arms race of hyperspecialisation. Comparative, synthesising perspectives necessarily go beyond the textual surface and are hence at odds with photorealism, whereas celebrating previously neglected textual specifics is the bread and butter of this historiography. Thus, predictably, the recent literature is heavily lopsided toward detailed studies of such things as the unpublished views on infinitesimals of a historical figure whose English Wikipedia page consists of two sentences (Domingues 2004), or how Leibniz once made a computational slip (using nine leading zeroes instead of eight in the decimal representation of a fraction) on a piece of scrap paper when he was trying to estimate e numerically from its power series (Probst and Raugh 2019). Such an increasingly microscopic focus has greatly improved local precision and expertise in historical scholarship, but with an exclusive focus in this direction the field is left without purposeful vision on a more global scale.

4.3.2 Alternative: Global Cognitive Contextualisation

Let us take the example of Leibniz's calculation of e and consider what new questions we would ask about this episode from a practice-oriented cognitive perspective.

As Probst and Raugh (2019) observe, Leibniz's manuscript appears to have been the first explicit occurrence of the numerical value of e. But what was the significance of this to the mathematical practice at the time? The natural logarithm and e eventually became fundamental in mathematics, but what did this enable mathematicians to do that they couldn't do before? In the context of the calculus, $\ln(x)$ is "natural" by virtue of being the logarithm function with the simplest derivative, but seventeenth-century calculus often reasoned in terms of proportionality and dimensional homogeneity, which arguably meant that there was no marked preference for $1/x$ over a/x. Is the modern canonisation of $\ln(x)$ and e merely cosmetic, or does it have cognitive import? If so, in what way, and did Leibniz see it that way?

Leibniz used power series for his computation, but sophisticated computational techniques for logarithms had been around since before Leibniz was even born. Already in 1622, Speidell gave a table of genuine natural logarithms for all integers from 1 to 1000, agreeing with the modern $\ln(x)$ to six decimal places, though the table omits the decimal point (Cajori 1991, 153; Speidell's work is now available at *Early English Books Online*). How does the calculus-based power series paradigm compare with earlier computational practices such as those for logarithms? Did the new paradigm excel compared to earlier techniques by efficiency, extension, unification, or simplification? How easily could earlier mathematicians such as

Speidell have repurposed their algorithms to compute e if they had wanted to? Since base-10 and base-e logarithm tables were available in print half a century before Leibniz, and since calculating $e = 10^{\lg(2)/\ln(2)}$ is readily reduced to such tables, the numerical value of e could in theory have been looked up in five minutes in a good library in Leibniz's time. Would contemporaries of Leibniz well-versed in established logarithmic practices have regarded computing e (once defined) as routine? More generally, how were calculus innovations parsed in relation to established proto-calculus practices, and how does the significance of key calculus concepts differ from modern perceptions when read through such a lens?

To answer such questions we must be attentive not only to specifics of individual documents such as Leibniz's e manuscript; we must also understand the overall scope of the know-how of logarithmic functions established in the mathematical practice of the time. Only a synthesising, comparative perspective could ever answer to this purpose.

The Leibniz e episode also raises intriguing questions about the relation between concrete calculations and curve plotting on the one hand and abstract theory on the other. As Probst and Raugh (2019) observe, Leibniz's most immediate purpose with computing the numerical value of e was to plot the graph of the shape of a hanging chain, the catenary $y = (e^x + e^{-x})/2$. Was this a mere "after the fact" pragmatic way to draw a curve already found theoretically, or did visual and numerical checks play a role of verification that removed lingering doubts that the theoretical derivation may not have been correct? More generally, were these kinds of calculations and drawings used as an integral part of research itself, for verification or explorative purposes, rather like a modern mathematician may use a computer?

Again, these are issues that cannot be answered by zooming in on the details of isolated cases but only by a comprehensive analysis of patterns of thought. Fortunately, philosophers have recently been very interested in issues such as diagrammatic reasoning (e.g. Giaquinto 2007; Hanna and Sidoli 2007), so we are better equipped than ever with conceptual tools to help us in such an analysis.

4.3.3 Consequence of "Photorealism" Historiography: Overdependence on Novel Sources

Two of the most high-profile interpretative innovations in the recent literature on the history of the calculus are the claim that previously unpublished manuscripts by Leibniz suggest "a complete transformation of the prevailing view on the position Leibniz held on the foundation of infinitesimal techniques" (Rabouin 2020, 19), and the claim that when modern imaging techniques revealed some previously unreadable words on an ancient Archimedes manuscript, this was "a major discovery" that "made us see, for the first time, how close Archimedes was to modern concepts of infinity" (Netz and Noel 2007, 29).

In both cases, revisionist interpretations are based on previously unpublished documentary sources. Indeed, it could hardly be otherwise, given the exclusively textual focus of photorealism historiography. This may seem right and proper: it simply shows that our field is evidence-driven. Yet if this is the only game in town then the consequences will be predictably detrimental.

If publishing new sources is a precondition for publishing new interpretations, then the field closes itself off from the analytic insights of mathematicians and philosophers, who have unique expertise that they would not have been able to develop if they had devoted the bulk of their time to editing unknown texts. Meanwhile, historians who do excellent work on editions of sources automatically have their voice greatly amplified also on interpretative questions: a conflation of credibility in one domain with authority in another.

Furthermore, an addiction to extracting ever more from the increasingly depleted potential of hitherto unpublished sources, and a lack of alternative ways of making scholarly progress, dooms us to keep mining the archives for novelty with diminishing returns. One may say that a historiography that refuses to go beyond the directly textual is bound to enter a "fracking" stage toward the end of its life cycle, as researchers are pushed to find ways of extracting new discoveries from documents that previous prospectors had treated as unpromising.

4.3.4 Alternative: Rigorous Historiography for Evaluating Other Types of Interpretative Hypotheses

Just as physics thrives on an interplay of theorists and empiricists, so historical scholarship would benefit from a wide range of interpretative thought rather than regarding as exclusively legitimate that of those who personally work on publishing ever more novel source material. For this, we need new standards of assessing hypotheses that go beyond direct one-to-one correspondence with textual evidence. We need to shift the focal point of historical research from the microtextual to a more overarching level of patterns of thought.

4.3.5 Consequence of "Photorealism" Historiography: Overemphasis on Surface Form

The axiom that fidelity to historical actors' modes of expression is the same thing as fidelity to the conceptual essence of their underlying thought entails that differences between geometric and algebraic styles are ipso facto profound. Hence, predictably, modern historians have placed considerable emphasis on the work of the second generation of calculus practitioners, who opted for a more algebraic approach than the geometrical style of people like Huygens and Newton. This was "a monumental conceptual shift" (Shank 2018, 234), "a major step that cannot be overestimated" (Speiser 2008, 108), modern historians assert. But it is hard to escape the impression that these claims are driven more by historiographical commitments than by analyses of mathematical practice. For example, Shank (2018) opens with a long and detailed chapter on the historiography of mathematics but has no technical discussion of actual mathematics in the entire book.

4.3.6 Alternative: Practice-Based Assessments of Importance

We need a new approach that neither erases differences between geometrical and algebraic approaches through anachronistic translation into modern mathematical terms, nor assumes that such differences are necessarily conceptually profound. I suggest that to answer these kinds of questions is to build up a comprehensive picture of what for instance Huygens's methods discussed above could and could not do. Only through such an overall sense of what it was like to wield these tools as research weapons can we understand the significance of the technical details of an argument such that by Huygens. And only through a detailed, comparative study of many specific examples can we build up such a general picture.

4.3.7 Conclusion on "Photorealism" Versus Cognitive History

All-out war on anachronistic paraphrase has not only eliminated the intended culprits but also inflicted additional casualties: comparative perspectives and mathematically insightful commentaries face a hostile climate under the new regime, and mathematicians and philosophers are alienated from the field. Left are historians playing it safe in the wake of this ideological purge, limiting themselves to the most directly textual domains of scholarship: narrowly specialised studies and archival work on unpublished sources. The field is losing vigour, like a person who avoids food poisoning only at the cost of suffering severe malnutrition.

Cognitive history shows how digging into mathematical practice *in media res* and asking new questions can be both mathematically and historiographically exciting. It reconstructs the living, bustling research scene of the time—the hopes and dreams and conundrums and technical obstacles that these historical mathematicians wrestled with in their daily practice. Consequently, instead of having to dig ever more deeply into the archives for fresh kicks, cognitive history shows that scholarship on canonical sources is far from "done" merely by mathematicians having given technically accurate, local descriptions of their content, and historians having traced their documentary network. The field is ready to graduate beyond this basic descriptive phase of scholarship and to dare to pursue new comparative and interpretative perspectives.

4.4 Example: Newton's Unusual Quadrature Manipulations to Describe Inverse-Cube Orbits

Figure 4.2 shows how Newton described the orbits that result from an inverse-cube force law in his *Principia* (1687). As throughout the *Principia*, Newton's style is classical and geometrical. However, because this is one of the more complicated technical problems of the *Principia*, Newton had relied on calculus "behind the scenes" to arrive at these solutions, using the steps outlined in Fig. 4.3. For this reason, this is an interesting case study for clarifying the boundary between classical and calculus methods.

4.4.1 Historiographical Lessons of the Newton Example

The existing literature on this episode is very typical. As far as excellent technical commentaries and explanations of the steps of Newton's derivation are concerned, there is an abundance—or, one might almost say, oversaturation—of literature, including for instance Erlichson (1994), Brackenridge (2003), and Guicciardini (2016). The last of these declares itself "deeply indebted" (210) to the earlier ones, and indeed largely consists of re-exposition rather than novelty as far as technical analysis of mathematical content is concerned. With multiple articles showing so much overlap and rapidly converging to a consensus, one is bound to get the impression that this mathematical style of historical scholarship is effectively "done" and has little more to contribute. No wonder, then, that recent scholarship,

"If with centre C and principal vertex V any conic VR . . . is described, and from any point R of it the tangent RT is drawn so as to meet the axis CV . . . at . . . T; and . . . there is drawn the straight line CP, which is equal to . . . CT and makes an angle VCP proportional to the sector VCR; then, if a centripetal force inversely proportional to the cube of the distance of places from the centre tends towards that centre C, and the body leaves the place V with the proper velocity along a line perpendicular to the straight line CV, the body will move forward in the trajectory VPQ which point P continually traces out."

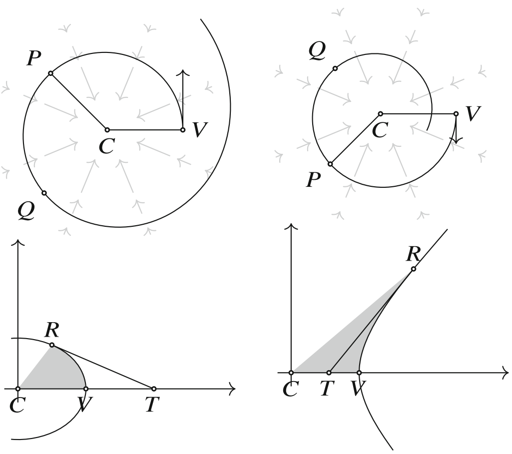


Fig. 4.2 Newton's description of trajectories in an inverse-cube force field

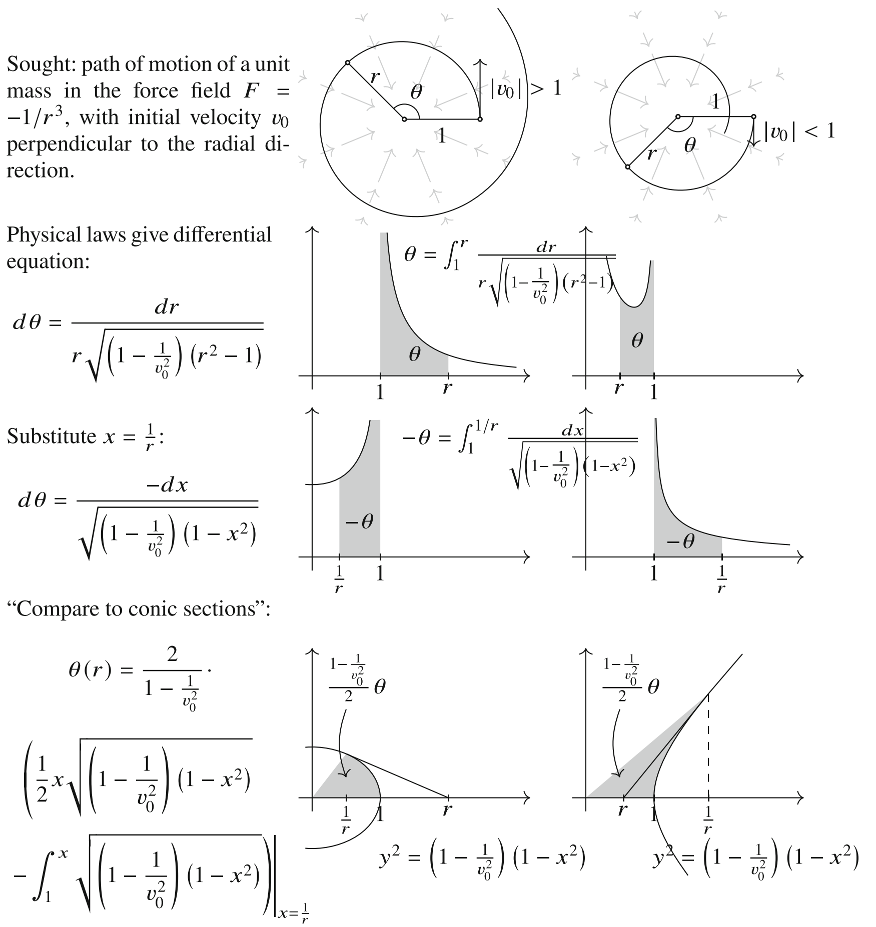


Fig. 4.3 Paraphrase of the steps Newton used to derive his description of trajectories in an inverse-cube force field

such as Guicciardini (2016), is instead turning its attention more to textual aspects, such as the context of this episode in Newton's manuscripts and correspondence.

But mathematically oriented historical scholarship has not stagnated because it has exhausted its potential or become obsolete. It is spinning its wheels at the moment, but by reorienting it in a new direction we will be able to harness its power in new ways.

The mathematical commentaries on Newton's orbit derivation are "done" only because they set themselves too limited a task. Very typical is the conclusion by Erlichson (1994) that "the ultimate key to the mystery ... is Newton's practice of expressing abstract quadratures by concrete visualizations" (154). Note the word *ultimate*, as if there was nothing further to be explained! From a cognitive point of view, the postulation of a particular type of geometrical predilection in Newton's style is not an "ultimate" brute fact but merely the beginning of what is to be explained. *Why* was this Newton's practice?

Cognitive questions pick up precisely where purely formal analyses left off. Erlichson stopped when he reduced the matter to a particular disposition in Newton's style, because that's where objective mathematics ends and subjective preferences begin. We need to break down this common barrier that isolates technical mathematical analyses from the less tangible but equally crucial cognitive considerations that drive mathematical research. Technical details of mathematical arguments on the one hand, and broader stylistic and philosophical attitudes on the other, co-evolved and were intimately intertwined.

Today there is agreement on what "the answer" is to integrals such as those involved in the Newton example above (in that case the solutions can be expressed in terms of trigonometric and exponential functions). But, as that example shows, the situation was much more fluid at the time. It is indeed a non-trivial question (that the early practitioners of the calculus wrestled with extensively) to decide even what kind of thing "the" answer should be. What do we want the solution to a differential equation to do? Should it be numerically tractable, visualisable, or qualitatively illuminating? By what such standards is, for example, Newton's reduction to convoluted conic measurements superior to other possibilities that were readily available to him, such as the quadratures of higher-degree curves obtained as intermediate steps in his own derivation, or power series methods? Indeed it is striking that in his final step Newton is able to express everything in terms of conics only by making the shapes of the area segments more complicated and even by completely dropping the geometrical representation of the proportionality constant from the visualisation altogether. So Newton opted for this particular type of geometrical interpretation even though it came at a notable cost.

Newton's choice was hardly mere conservatism, because comparable preferences for qualitative, geometrical characterisations of integrals are commonplace in the early calculus. This includes for instance the prominent seventeenth-century practice of "rectifying quadratures" (Blåsjö 2012), that is to say, expressing insoluble integrals as arc lengths, such as elliptic integrals in terms of the arc length of the lemniscate. Equally odd to modern eyes is the recurrent theme in this period of finding when a generally transcendental problem could be expressed without reference to transcendental quantities (such as trigonometric functions or π). For instance, Huygens, Leibniz, and Johann and Jakob Bernoulli all tried to find classes of segments of the cycloid whose area is "squarable" in this sense. The Bernoullis investigated this in depth and to this end were led to developing what is today known as Chebyshev polynomials (Henry and Wanner 2017). To name another example, Huygens's solution to the catenary problem fell short precisely because it failed to fully reduce the necessary quadratures (Bos 1980, 142), showing that this was a complicated matter that stumped even the best minds. Similarly, Leibniz solved the brachistochrone problem but failed to recognise from his own solution formula that it was the well-known cycloid (Blåsjö 2017b, 185).

Altogether, the varied ways in which seventeenth-century mathematicians chose to transform and "solve" integrals were based on deliberate choices and priorities that were mathematically and philosophically rich but are poorly understood today. The only way to illuminate such questions is through a comparative perspective. If we look at episodes like the Newton orbit example in isolation then we have little choice but to leave it at the weak non-explanation that Newton preferred to express the solution geometrically rather than by a formula. But by taking a comprehensive view and reconstructing the overall state of calculus research at the time, we will be in a much better position to situate his precise choices in a coherent context.

For example, Guicciardini (2016) has shown what calculus manuscripts Newton relied on in his solution, so by studying the guiding motivation implicit in the structure of that treatise, and the uses Newton made of those ideas in other works, we will be able to say much more about what guided Newton's choice of representation in the orbit example than we could have by looking at that episode alone.

In the same way, comparing Newton's approach with those of his contemporaries will also illuminate what he saw as the particular strengths of his chosen integration methods and means of representing curves. Guicciardini (2016) is right that "interesting questions remain open" such as "in what sense do Newton's methods differ from those deployed by Leibniz, Varignon, Johann Bernoulli, and Euler?" (234) It is no coincidence that these questions remain open, since the photorealism axiom penalises comparative research. A historiographical rethink is needed to make progress in these directions. A cognitive turn will redress precisely this problem and thereby revitalise the field and show how the expertise of mathematically trained researchers can be mobilised in new ways to reach new kinds of insights about history.

4.5 Conclusion

Let me summarise the invitation to mathematicians that I have proposed. In a historical text we find a mathematician using a particular technique. For example, Huygens finding the motion of a particle sliding down a cycloid by relating it to the geometry of an associated circle. Or Newton finding the orbit in an inverse cube force field by relating it to arcs and areas of conics.

We want to know what the broader cognitive significance of this technique is. Typically, isolated cases are insufficient to say anything conclusive about this. So we form several interpretative hypotheses consistent with the case at hand.

For example, we may hypothesise that Huygens's use of circle geometry to solve a dynamical problem is effectively equivalent to using the calculus of trigonometric functions. Or in the Newton case we may hypothesise that Newton preferred his convoluted expression in terms of conics, rather than the obvious alternatives such as power series, because it better illuminates the qualitative properties of the orbit.

Such hypotheses entail testable predictions. If the hypothesis correctly puts the finger on a key aspects of the mathematical thought of that author, then in comparable cases that author ought to act in accordance with that hypothesis. So to test our hypotheses we then turn to other works.

For instance, among the various problems that Huygens solved geometrically that we today would solve using the calculus of trigonometric functions, which utilise a reference circle and its properties such as theorems about tangents to effectively go from the same premisses to the same conclusion as the modern calculus proof? Are there cases where Huygens's approach has demonstrable drawbacks compared to an approach based on the calculus of trigonometric functions? Are there cases where Huygens failed to solve a problem that the next generation could solve by the calculus or trigonometric functions?

In the Newton case, his reduction to conic areas and arcs is based on an unpublished catalogue of such reductions—effectively a "table of integrals." What other uses did Newton make of this catalogue? Are those uses consistent with our hypothesis? Is the theoretical structure of the catalogue consistent with our hypothesis?

These kinds of questions can only be answered by a mathematical analysis that goes beyond what is explicit in the texts, and by a comparative perspective that looks at the mathematical practice of the time comprehensively. Answering such questions requires understanding that only mathematicians are likely to posses.

It is easier than ever for mathematicians to enter the field and do this kind of work. Recently published sources and specialised studies have made the field more accessible and easier to navigate, and it has made comparative and interpretative work drastically more feasible. The old hostility to mathematicians among professional historians of mathematics that had its brief heyday is nowadays sooner the subject of historical study itself (Schneider 2016) than a force in the present that anyone needs to fear. Mathematicians turning to history are likely to find a warmer reception today than in those "cold war" decades.

Thus mathematicians can help historians, but there will be benefits in the opposite direction as well. A cognitively reorientation of historical scholarship will make it more mathematically exciting. Our questions about the thought and practice of Huygens and Newton, for example, are precisely the kind of history that is directly relevant to teaching and thoughtful understanding, not as decorative anecdotes but as insights deeply intertwined with content and substance.

References

Arnol'd, V. I. 1990. *Huygens and Barrow, Newton and Hooke: Pioneers in Mathematical Analysis and Catastrophe Theory from Evolvents to Quasicrystals*. Birkhäuser.

Arnol'd, V. I. 2012. *Geometrical Methods in the Theory of Ordinary Differential Equations*, 2nd ed. Springer.

Bell, J., and V. Blåsjö. 2018. Pietro Mengoli's 1650 proof that the harmonic series diverges. *Mathematics Magazine* 91 (5): 341–347.

Blåsjö, V. 2012. The rectification of quadratures as a central foundational problem for the early Leibnizian calculus. *Historia Mathematica* 39 (4): 405–431.

Blåsjö, V. 2015. The myth of Leibniz's proof of the fundamental theorem of calculus. *Nieuw Archief voor Wiskunde* 16 (1): 46–50.

Blåsjö, V. 2017a. On what has been called Leibniz's rigorous foundation of infinitesimal geometry by means of Riemannian sums. *Historia Mathematica* 44 (2), 134–149.

Blåsjö, V. 2017b. *Transcendental Curves in the Leibnizian Calculus*. Elsevier.

Bos, H. J. M. 1980. Huygens and mathematics. In *Studies on Christiaan Huygens*, ed. H. J. M. Bos, 126–146. Routledge.

Boyer, C. 1959. *The History of the Calculus and Its Conceptual Development*. Dover Publications.

Brackenridge, J. B. 2003. Newton's easy quadratures "omitted for the sake of brevity". *Archive for History of Exact Sciences* 57: 313–336.

Bressoud, D. M. 2019. *Calculus Reordered: A History of the Big Ideas*. Princeton University Press.

Cajori, F. 1991. *A History of Mathematics*, 5th ed. AMS Chelsea.

Chang, H. 2017. Who cares about the history of science? *Notes and Records: The Royal Society Journal of the History of Science* 71: 91–107.

Domingues, J. C. 2004. Variables, limits, and infinitesimals in Portugal in the late 18th century. *Historia Mathematica* 31: 15–33.

Dunham, W. 2005. *The Calculus Gallery: Masterpieces from Newton to Lebesgue*. Princeton University Press.

Edwards, C. 1979. *The Historical Development of the Calculus*. Springer.

Erlichson, H. 1994. The visualization of quadratures in the mystery of corollary 3 to proposition 41 of Newton's Principia. *Historia Mathematica* 21 (2): 148–161.

Euler, L. 1736. *Mechanica sive motus scientia analytice exposita*.

Giaquinta, P., and M. Freguglia. 2016. *The Early Period of the Calculus of Variations*. Birkhäuser.

Giaquinto, M. 2007. *Visual Thinking in Mathematics*. Oxford University Press.

Guicciardini, N. 2016. Lost in translation? Reading Newton on inverse-cube trajectories. *Archive for History of Exact Sciences* 70 (2): 205–241.

Hanna, G., and N. Sidoli. 2007. Visualisation and proof: a brief survey of philosophical perspectives. *ZDM Mathematics Education* 39: 73–78.

Heilbron, J. L. 1987. Applied history of science. *Isis* 78: 552–563.

Henry, P., and G. Wanner. 2017. Johann Bernoulli and the cycloid: a theorem for posterity. *Elemente der Mathematik* 72 (4): 137–163.

Huygens, C. 1673. *Horologium oscillatorium sive de motu pendulorum ad horologia aptato demonstrationes geometricae*. Paris.

Imhausen, A. 2021. Quo vadis history of ancient mathematics: who will you take with you, and who will be left behind? *Historia Mathematica* 57: 80–93.

Netz, R., and W. Noel. 2007. *The Archimedes Codex*. Da Capo Press.

Orlin, B. 2019. *Change Is the Only Constant: The Wisdom of Calculus in a Madcap World*. Black Dog & Leventhal.

Probst, S., and M. Raugh. 2019. The Leibniz catenary and approximation of e — an analysis of his unpublished calculations. *Historia Mathematica* 49: 1–19.

Rabouin, D. 2020. Exploring Leibniz's Nachlass at the Niedersächsische Landesbibliothek in Hanover. *EMS Newsletter* (116): 17–23.

Schneider, M. R. 2016. Contextualizing Unguru's 1975 attack on the historiography of ancient greek mathematics. In *Historiography of Mathematics in the 19th and 20th Centuries*, ed. V. Remmert, M. Schneider, and H. K. Sørensen, 245–267. Birkhäuser.

Shank, J. B. 2018. *Before Voltaire: The French Origins of "Newtonian" Mechanics 1680–1715*. University of Chicago Press.

Speiser, D. 2008. *Discovering the Principles of Mechanics 1600-1800*. Birkhäuser.

Strogatz, S. 2019. *Infinite Powers: How Calculus Reveals the Secrets of the Universe*. Mariner Books.

Yoder, J. G. 2004. *Unrolling Time: Christiaan Huygens and the Mathematization of Nature*. Cambridge University Press.

Chapter 5
Further Thoughts on Anachronism: A Presentist Reading of Newton's *Principia*

Niccolò Guicciardini

Abstract This chapter is an exercise in anachronism. I read a hitherto unnoticed corollary in Newton's *Principia* availing myself of concepts and results that were obviously unavailable to the author. I contend that even extreme forms of anachronism can be helpful in the historical interpretation of past mathematical texts. I also claim that such anachronistic readings must be followed by an attempt to capture the differences between past and present mathematics. The text I analyze is Proposition 41, Book 1, and its three corollaries (especially, the second one), which are part of a set of propositions where Newton deals with the motion of a body accelerated by a central force.

5.1 Newton on Graphical Methods

It is well known that Isaac Newton deployed graphical methods in order to construct orbits traced by a body accelerated by a central force. When Newton wrote the *Philosophiae Naturalis Principia Mathematica* (the *Principia* for short), in the years 1684–1687, the methods of the calculus were still very rudimentary: mathematicians were not used to writing differential equations. In order to study the motion of bodies accelerated by central forces several methods for graphically approximating the orbits were devised.

There is a vast literature dealing with Newton's graphical methods. So far, scholars have studied Newton's use of the parallelogram rule in order to construct orbits via vector composition of motions, according to Corollary 1 to the laws of motion. Attention has also been paid to the geometrical methods employed by Newton in Sections 2 and 3, Book 1, of the second edition of his *Principia* in

N. Guicciardini (✉)
Università degli Studi di Milano, Milano, Italy
e-mail: niccolo.guicciardini@unimi.it

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_5

which he resorted to the radius of curvature[1] (Nauenberg, 1994; Brackenridge, 1995, pp. 79–85). Lately, Nauenberg (2018) and, more thoroughly, Chin have studied Prop. 1, Book 1, of Newton's *Principia* as a "a geometric embodiment of a numerical algorithm, the first symplectic integrator devised 300 years before its time," (Chin 2022, p. 13) where a symplectic integrator is a numerical integration scheme for Hamiltonian systems, "a canonical transformation which seeks to integrate Hamilton's equation to obtain the system's coordinate and momentum as a function of time." (Chin 2022, p. 4).

What I will do in this paper is show that in Corollary 2, Prop. 41, Section 8, Book 1 of the *Principia* (Corollary 2, for short), Newton devised a method—to allow ourselves the use of anachronistic language from the very beginning! — for associating a *slope field* to the differential equation of a mass-point accelerated in a central force field. Thus, Corollary 2 paves the way for a graphical construction of the orbit traced by the mass-point. Newton's construction thus defines a concept, that of slope field, that is important for methods for the graphical solution of differential equations developed later, most notably by Leonhard Euler, Karl Heun, Carl Runge, and Wilhelm Kutta (Butcher and Wanner 1966).

5.2 Propositions 39–41, Book 1

These propositions have often been commented on and for very good reasons.[2] They are a masterpiece of physical and mathematical insight. In their broad outlines, they are still taught in courses on classical mechanics, even though Newton's language and conceptions are peculiar to his times and present many divergences from our own conceptions. They deal with the motion of a point mass in a spherically symmetric potential $V(r)$. The corresponding central force is

$$\mathbf{F} = -\frac{dV}{dr}\hat{\mathbf{r}} \tag{5.1}$$

where $\hat{\mathbf{r}}$ is the unit vector pointing away from the centre of force. Without loss of generality we can assume that the mass is $m = 1$. Indeed, this is a one-body problem (a body attracted by a centre of force) and no use is made of the third law of motion.

5.2.1 Proposition 39

In Proposition 39 Newton considers a "body" moving in a straight line ADE under the action of a centripetal force (see Fig. 5.1). Thanks to a very simple reasoning, he

[1] These methods are based on the fact that the normal component of force is proportional to the speed squared and inversely proportional to the radius of curvature.

[2] A crystal-clear treatment is Newton (1999, pp. 334–345).

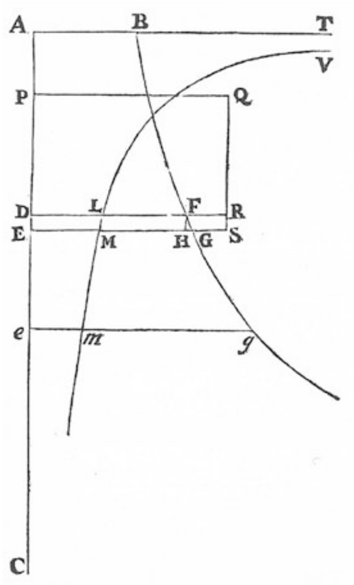


Fig. 5.1 Diagram for Proposition 39, Book 1, of the third edition of the *Principia* (1726). Source: Newton (1726, p. 120). The curve VL in the first two editions (1687) and (1713) was wrongly drawn

shows that the square of the speed is proportional to the area subtended under the curve that represents the variation of the force's intensity as a function of distance. Let's provide a translation into algebraic symbols of Newton's text.

In Fig. 5.1 a body is falling from rest in A towards the centre of force C. The curve BF represents the intensity of the centripetal force F.[3] The force is directed towards C. I choose as a sign convention that for a centripetal force $F < 0$. Newton writes that the ordinate DF is "proportional to the centripetal force in that place [at point D] tending towards the centre C" (Newton 1999, 525). Since in this paper I translate Newton's proportions into equations, I need to make a proportionality constant explicit. I choose:

$$DF = 2F. \tag{5.2}$$

The curve VLM represents the inverse of the speed v of the body, thus the speed of fall at D is proportional to $1/DL$. Let's put:

$$\frac{1}{DL} = v. \tag{5.3}$$

[3] It might well be that Newton chooses "F" to denote that point because it is the end-point of the ordinate "DF" measuring the intensity of force.

Newton proves that

> The speed of the body in any place D will be as the straight line whose square is equal to the curvilinear area $ABFD$ (Newton 1999, p. 525).[4]

Newton's statement is framed in terms of the theory of proportions. I attempt a first translation

$$\frac{1}{DL} \propto \sqrt{ABFD} \tag{5.4}$$

where $ABFD$ is the area subtended to the curve BF, and I use $\propto$ as an abbreviation for "is proportional to." While Newton does not use the symbol $\propto$, he is happy to use the square root symbol: he does write $\sqrt{ABFD}$ in Prop. 41.

What Newton writes in the language of proportion theory in the *Principia* can easily be translated into our familiar equation:

$$v(r)^2 = 2\int_{r_0}^{r} F dr, \tag{5.5}$$

where we set $CA = r_0$, $CD = r$, where $r_0 > r$. According to the proposed algebraic translation, the speed squared is equal to the area of $ABFD$: or $v^2 = ABFD$.

Of course, after this translation, the historian should pay great attention to Newton's original language, whose "equivalence" with our calculus translation is a very subtle interpretative matter (see Sect. 5.5). Most notably, I have transformed a proportion into an equation, which means that I have made proportionality factors explicit.

Further, even more anachronistically, once written as above, Newton's proposition will immediately be read by a modern reader as stating the law of conservation of mechanical energy, whereby the sum of the potential energy and kinetic energy must remain constant during the time evolution of an isolated system. In a sense, it is legitimate to say that Newton stated, or maybe anticipated, the principle of conservation of mechanical energy. On the other hand, it is also true that Newton did not isolate the concept of energy, and even less those of kinetic and potential energy, and that our formulation is conceptually richer (we use the concepts of closed system, dissipative vs conservative force, and potential for example).

5.2.2 Proposition 40

In Proposition 40, Newton considers a body orbiting in a plane orbit under the action of a central and isotropic force after having been fired laterally (not towards the force

---

[4] Lettering modified for consistency. For example, Newton uses $ABGE$ rather than $ABFD$. The two expressions are equivalent, since DE is infinitesimal. Cohen and Whitman translate "velocitas" as velocity.

centre) from a given position with a given velocity.[5] He proves that the speed is a function of the distance from the force centre only. Thus, in order to determine the speed of a body moving in an orbit in a central force field, one can apply the result of Proposition 39.

5.2.3 *Proposition 41: The Statement and Result*

The statement of Proposition 41 is:

> Supposing a centripetal force of any kind and granting the quadratures of curvilinear figures, it is required to find the orbits in which bodies will move and also the times of their motions in the orbits so found (Newton 1999, p. 529).

This is a problem (in Newton's times known as the "inverse problem of central forces") that we still teach to our students in courses on Newtonian mechanics.

5.2.3.1 An Anachronistic Rendering

Today, we ask to determine the motion of a point mass in a central force field, given as initial conditions the position and velocity at time $t = 0$. I remind the reader that I consider the mass unitary: $m = 1$.

We ask to reduce this problem to two differential equations, a choice of polar coordinates r and θ being the best for reasons of symmetry. The equations that can be found in our treatises on Newtonian mechanics, such as De Lange and Pierrus (2010, 217–219), are

$$dt = \frac{\pm dr}{\sqrt{v_0^2 + 2\int_{r_0}^{r} F dr - h^2/r^2}}, \tag{5.6}$$

and

$$h = r^2 \frac{d\theta}{dt} \tag{5.7}$$

where h is the angular momentum, $r_0 = r(0)$ is the initial position, and $v_0 = v(0)$ is the initial velocity. Inversion of

$$t(r) = \int_{r_0}^{r} \frac{dr}{\sqrt{v_0^2 + 2\int_{r_0}^{r} F dr - h^2/r^2}}, \tag{5.8}$$

[5] The orbit is planar, as Newton proves in Propositions 1 and 2, Book 1.

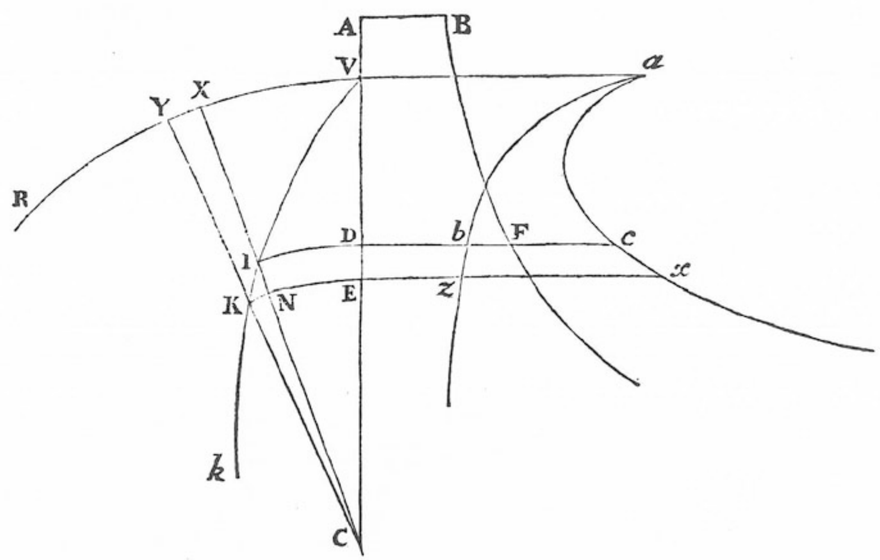


Fig. 5.2 Diagram for Proposition 41, Book 1, of the *Principia*. Source: Newton (1726, p. 125). The diagram for the first and second editions was compromised by a few mistakes

gives the radial position $r(t)$ in function of time.[6] Substituting this in Eq. (5.7) allows to express the angular position in function of time as

$$\theta(t) = h \int_0^t \frac{dt}{r^2(t)} + \theta_0, \tag{5.9}$$

where $\theta_0 = \theta(0)$. Equations (5.8) and (5.9) determine the time-dependent orbit. The geometric orbit is instead given by

$$\theta(r) = h \int_{r_0}^{r} \frac{dr}{r^2\sqrt{v_0^2 + 2\int_{r_0}^{r} F dr - h^2/r^2}} + \theta(r_0). \tag{5.10}$$

5.2.3.2 Newton's Approach

In a way, Newton's procedure is similar to ours. Such an 'equivalence' is what—indeed—is discussed in this paper.

Newton considers a "body" setting out from V with a given velocity under the action of a given centripetal force (see Fig. 5.2). The force's centre is C. Note that $VXYR$ is circle with radius CV.

[6] Note the sign ambiguity in Eq. (5.6). Here we choose the positive root.

The orbit, to be found, is $VIKk$.

The point A is specified by the property that when a falling body reaches point V, falling vertically from rest at an initial position A, the magnitude of its velocity will be the same as the magnitude of the initial velocity of the body moving along the curved path $VIKk$.[7]

As coordinates of the body's position at a point I of the orbit Newton uses the distance from the force's centre $CI = \mathfrak{A}$ and the area of the circle sector VCX. This is very nearly what we would do nowadays by using polar coordinates.

As in Prop. 39, the curve BF represents the intensity of the force as a function of distance from the force's centre.[8] Thus, the area of the surface $ABFD$ subtended to this curve measures what we would call mechanical work. For the conventions that I adopted in Sect. 5.2.1, $ABFD$ is equal to the double of the work.

It should be noted that Newton avails himself of two properties of central force motion that he proved in previous pages.

In Newton's terms, the first property is that the area law holds if and only if the force is central (Propositions 1 and 2, Book 1) (Newton 1999, p. 444). Thus the motion is planar and the areal velocity is constant. This first property allows Newton to geometrically represent time as the area of the surface swept by the radius vector. For example, the area VCI is proportional to the time taken by the body to traverse the arc VI. Newton denotes the constant areal velocity as $Q/2$ and sets $Z = Q/\mathfrak{A}$.[9] Using our algebraic symbols $Z = h/r$, where h is the angular momentum.

The second property is proved, as we have just seen, in Propositions 39 and 40 (see Sects. 5.2.1 and 5.2.2). This property implies that the square of the speed at point I is proportional to the area of the surface $ABFD$. Using our algebraic symbols $v^2 = AFBD = v_0^2 + 2\int_{r0}^{r} F dr$.

In modern terms, we would understand these two properties as the law of conservation of angular momentum $h = Q$ and the law of conservation of mechanical energy E.

Starting from these two properties, Newton obtains two curves abz and acx that must be "squared" in order to determine the dependence of time (measured by the area $VCI = (Q/2)t$) and the dependence of angle (measured by the area of the circular sector $VCX = (CX^2/2)\theta$) from distance. "Squaring" a curvilinear surface meant calculating its area.

Newton proves that the two curves have ordinates Db and Dc given by the following equations:

$$Db = \frac{Q}{2\sqrt{(ABFD - Q^2/\mathfrak{A}^2}} = \frac{Q}{2\sqrt{ABFD - Z^2}}, \tag{5.11}$$

---

[7] I am here deeply indebted to Bernard Cohen's translation into symbols of Prop. 41. See Newton (1999, p. 145).

[8] In Newton's words "[the ordinate] DF is proportional to the centripetal force in that place [$CI = CD = \mathfrak{A}$] tending towards the centre C." (Newton 1999, p. 525).

[9] Edmond Halley suggested this abbreviation to Newton.

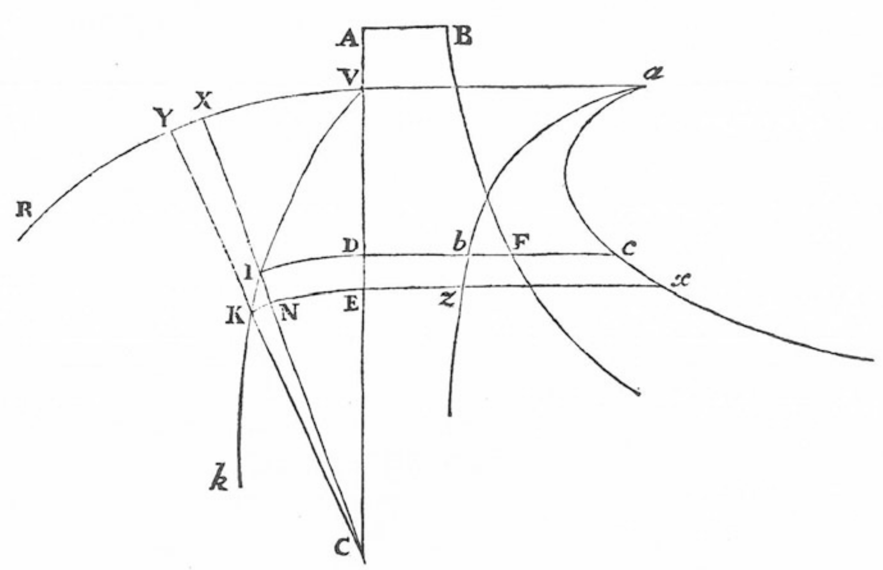


Fig. 5.3 Diagram for Proposition 41, Book 1. The corrected version in the third edition. Source: Newton (1726, p. 125)

and

$$Dc = \frac{Q \times CX^2}{2\mathfrak{A}^2\sqrt{ABFD - Q^2/\mathfrak{A}^2}} = \frac{Q \times CX^2}{2\mathfrak{A}^2\sqrt{ABFD - Z^2}}. \tag{5.12}$$

5.2.4 Proposition 41: The Demonstration

Let us now take a closer look at Proposition 41, for the purpose of familiarizing ourselves with Newton's notation and demonstration.

5.2.4.1 Newton's Approach

For the readers's convenience, I will repeat the diagram for Proposition 41 (see Fig. 5.3), in which VIk is the sought orbit, and IK is an infinitesimal arc traversed by the body during an infinitesimal interval of time.[10] Further, KN is drawn perpendicularly to CI.

[10] "Let the points I and K be very close indeed to each other." (Newton 1999, p. 530).

Newton's demonstration can be subdivided into three steps. We parse the crucial passage as follows:

1. "the line-element IK, described in a given minimally small time, will be as the velocity
2. and hence as the straight line whose square equals the area $ABFD$,
3. and the triangle ICK proportional to the time will be given; and therefore KN will be inversely as the height IC, that is, if some quantity Q is given and the height IC be called $\mathfrak{A}$, as $Q/\mathfrak{A}$." (Newton 1999, p. 530).

This is a bit confusing for the modern reader, because Newton does not use the concept of function, as we would do, and employs the language of proportion theory. Thus, for example, he cannot say that the speed is equal to the infinitesimal arc divided by the infinitesimal time, since this would be a ratio between two inhomogeneous magnitudes. Rather, Newton employs geometrical diagrams, such as the curve BF, that represent the functional dependence of different continuously varying kinematic and dynamical magnitudes.

Moreover, Newton subdivides time into an infinity of infinitesimal equal increments and states that the instantaneous speed is proportional to the infinitesimal arc IK. This is step 1 above.

Because of propositions 39 and 40, IK is proportional to $\sqrt{ABFD}$, that is, IK is proportional to the square root of the area subtended under the curve BF. This is step 2.

Since the area law holds for central force motion (as proved in Propositions 1 and 2, Book 1), Newton can state that the infinitesimal increments ICK (triangles since IK is equated to a straight infinitesimal segment) of the surface swept by the radius vector in equal infinitesimal intervals of time have equal areas. This implies that KN is inversely proportional to $IC = \mathfrak{A}$, or $KN \times IC = KN \times \mathfrak{A} = Q$. Thus, $KN = Q/\mathfrak{A}$. This is step 3.

Putting together the above three steps, one immediately obtains (bearing in mind that $Z = Q/\mathfrak{A}$):

$$\frac{IK}{KN} \propto \frac{\sqrt{ABFD}}{Q/\mathfrak{A}} = \frac{\sqrt{ABFD}}{Z}. \tag{5.13}$$

The formula (5.13) is the basic result that allowed Newton to tackle the inverse problem of central forces in Proposition 41. It is the formula that will be applied in Corollary 2.

It is straightforward to deduce from Eq. (5.13) the curves (5.11) and (5.12), which must be squared in order to solve the inverse problem of central forces.

5.2.4.2 An Anachronistic Rendering

Instead of following Newton's manipulation of symbols, such as IK and KN, representing infinitesimal segments, I will adopt an anachronistic mode of expression and translate the formula (5.13) in modern calculus symbols.[11] As we shall see, Newton's formula is 'equivalent' to a differential equation in polar coordinates.

Let's rewrite Eq. (5.13) as:

$$\frac{ds}{rd\theta} = \frac{r\sqrt{v_0^2 + 2\int_{r_0}^{r} F dr}}{h}, \tag{5.14}$$

Then:

$$\frac{dr^2 + r^2 d\theta^2}{r^2 d\theta^2} = \frac{r^2\left(v_0^2 + 2\int_{r_0}^{r} F dr\right)}{h^2} \tag{5.15}$$

Rearranging, one gets:

$$\frac{dr^2}{r^2 d\theta^2} = \frac{r^2\left(v_0^2 + 2\int_{r_0}^{r} F dr\right) - h^2}{h^2} \tag{5.16}$$

Thus:

$$\frac{d\theta}{dr} = \frac{\pm h}{r^2\sqrt{v_0^2 + 2\int_{r_0}^{r} F dr - h^2/r^2}} \tag{5.17}$$

a differential equation (see (5.10)), which is familiar to our students of Newtonian mechanics.

Proposition 41 has often been seen as the high point of Newton's ability to apply his method of fluxions (more precisely, his quadrature techniques) to the science of motion and force (in this case, to central force motion). Yet, as we shall see in the next Sect. 5.3, an intuitive, non algebraic, graphical method for approximating orbits is at work in the group of propositions we are examining. This is spelled out in Corollaries 1 and 2, Proposition 41.

5.2.4.3 Corollary 3: A Brief Look

For the sake of completeness, I should note that in Corollary 3, Proposition 41, Newton squares the curves *abz* and *acx* in order to find the orbits traced in an inverse-cube force field (see Fig. 5.4). He limits his treatment to the case in which the initial velocity is orthogonal to the radius vector. Thus, Newton indeed solves the inverse problem of central forces in the inverse-cube case "granting the quadrature of curvilinear figures," as he claims in the statement of Proposition 41. However,

[11] For a discussion respectful of Newton's original language, see Newton (1999, pp. 334–345).

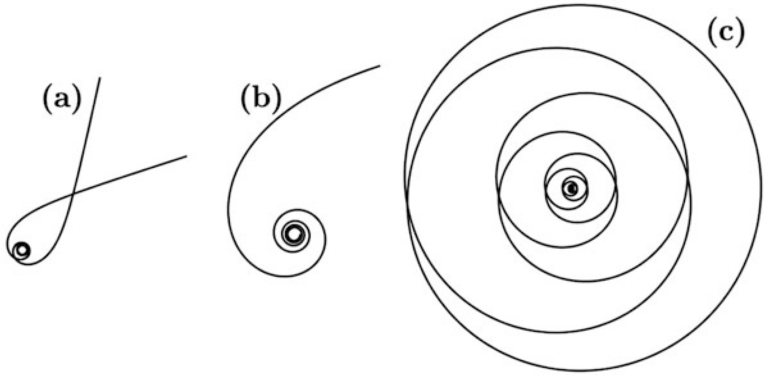


Fig. 5.4 Cotesian spirals in an inverse-cube force field. Cases (**a**) and (**c**) are those obtained by Newton in Corollary 3, Prop. 41, Book 1. Case (**b**) is constructed in Fig. 5.6 : it was discovered by Roger Cotes. The equations in polar coordinates are as follows: $1/r = \cos k\theta$, $1/r = \sinh k\theta$, $1/r = \cosh k\theta$. Other possible orbits are the logarithmic spiral ($\log r = k\theta$), identified by Newton in Proposition 9, Book 1, and the inverse Archimedean spiral ($1/r = k\theta$), identified by Johann Bernoulli in 1710. Circular orbits are also possible, but they are unstable: a tiny radial impulse will cause the body to spiral either to the centre of force or to infinity (Guicciardini 2016)

"for the sake of brevity," he leaves the quadrature as an exercise to be done by his readers, and explains the details only to some of his correspondents.[12]

It is unclear if Newton was able to square curves abz and acx for forces other than inverse-cube ones. Most notably, we have no evidence that he was able to solve the inverse problem for inverse-square forces by quadratures. This notable result was published in print for the first time in 1710 by Jacob Herman and Johann Bernoulli. As a matter of fact, as we would say today, for a central force $\mathbf{F} = kr^n\hat{\mathbf{r}}$, elementary integrations (in terms of trigonometric or hyperbolic functions) are possible only for $n = 1$, $n = -2$, and $n = -3$. In Newton's times the notion of function and integration were still *in nuce*. Newton thought in terms of "squaring of curves." By "squaring" he meant calculating the area of the surface bounded by a curve.

5.3 Corollaries 1 and 2, Proposition 41, Book 1

5.3.1 Corollary 1

Corollary 1 reads as follows:

> Corollary 1. Hence the greatest and least heights of bodies (that is the apsides of their orbits) can be found expeditiously. For the apsides are those points in which the straight line IC

[12] On Corollary 3, see Guicciardini (2016) and the literature cited there.

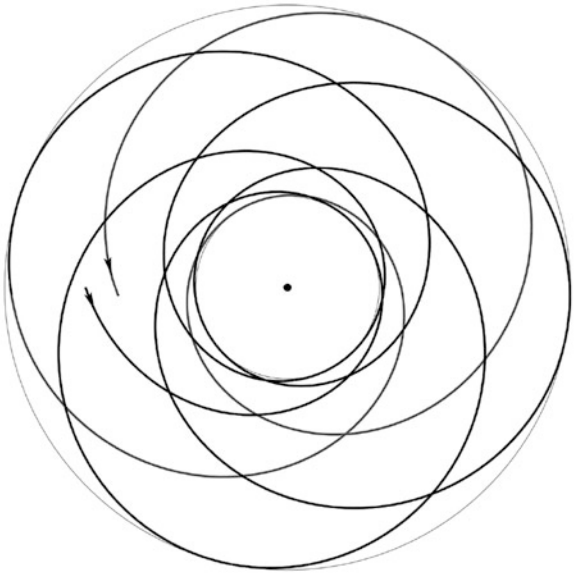


Fig. 5.5 A case of bounded motion. The apsidal distances r_1 and r_2 are the radii of the inner and outer circles. The orbit displays a large apsidal precession. For a generic potential, in the case of bounded orbits there are two distinct periods: one is that of the oscillation of the radius; the second is that of revolution. The orbit closes if and only if the relationship between the periods is a rational number. If the relationship is irrational the orbit never closes: one speaks of quasi-periodic motion, in the sense that after a sufficiently long time the orbiting body will return arbitrarily close to the initial condition. In general, the ratio between the two periods is a function of energy and angular momentum (i.e., of the initial conditions), therefore with small variations of the two parameters one passes continuously from irrational to rational values, so that one gets a rosette-shaped orbit. The two cases of the harmonic and the Keplerian potential are singular because the relationship between the two periods is independent of both the energy and the angular momentum (that is, the initial conditions). In the harmonic case the period of oscillation of the radius is half of that of revolution, and the orbit is an ellipse with a centre in the centre of force. In the Keplerian case the two periods are equal, and the orbit (if bounded) is an ellipse with the centre of force in a focus

> drawn through the centre falls perpendicularly upon the orbit VIK, which happens when the straight lines IK and KN are equal, and thus when the area $ABFD$ is equal to Z^2 (Newton 1999, p. 531).

In the first corollary Newton determines the possible existence of apsidal distances r_1 and r_2 (see Fig. 5.5). Such turning points occur for $\dot{r} = dr/dt = 0$ or $dr/d\theta = 0$ (see our Eqs. (5.6) and (5.10)).[13]

In Newton's terms, the "apsides" occur at distances from the force's centre C in which IK and KN are equal (see Fig. 5.3).

[13] I wish to thank George Smith for an enlightening discussion on this corollary.

If one considers Newton's formula (5.13), the apsides occur for values of the distance $CI = \mathfrak{A}$ for which

$$ABFD - Z^2 = ABFD - (Q/\mathfrak{A})^2 = 0, \tag{5.18}$$

or in modern terms

$$v_0^2 + 2\int_{r_0}^{r} F dr - \frac{h^2}{r^2} = 0. \tag{5.19}$$

Corollary 1, unlike Corollary 3, does not presuppose that the inverse problem of central forces has been solved by integration. The apsidal distances can be found without the squaring of curves (i.e., without integrations).[14]

Of course, this is still taught in textbooks on rational mechanics. It is interesting, in the context of my exercise in anachronistic history, to remind the reader how we proceed today (De Lange and Pierrus 2010, 216–223). Below I will try to be as complete and exhaustive as possible: the reader will forgive me for my pedantic presentation.

A caveat is in order. While above, I have simplified equations by setting the mass $=$ 1, here I will write m for the mass of the body. Since the problem we are considering is a one-body problem, the mass does not play any significant role.[15] But now I want to obtain the formulas as they are taught nowadays. This shift in notation reminds us that Newton wrote proportions: thus, constants such as m just disappeared from his language. Today, we write equations: all constants are expressed, even though even today there is a habit to set constants such as the velocity of light, or the Planck constant, equal to 1. Further, the physical dimension of our symbols plays an importance that was not understood in Newton's times. Thus, for example, a force F has the dimension of mass times distance divided the square of time. We begin to appreciate the difficulties we encounter when "translating" past mathematics into modern notation. In the case at hand, by such translations, we make the above conceptual differences evaporate.

This said, Eqs. (5.8) and (5.10) are rewritten as:

$$t(r) = \int_{r_0}^{r} \sqrt{\frac{m}{2\left[E - V_e(r)\right]}} dr \tag{5.20}$$

[14] Note that in correspondence of apsidal distances the curves abz and acx have an asymptote. This is not a problem for Newton, who could, of course, square curves with asymptotes, such as the cissoid.

[15] Indeed, in a two-body problem the two-masses, m_1 and m_2, must be taken into consideration because of the third law of motion. Otherwise said, the ratio of the accelerations, a_1 and a_2, is the inverse of the ratio of the masses: $a_1/a_2 = m_2/m_1$.

where E is the energy, and

$$\theta(r) = \theta(r_0) + \frac{h}{\sqrt{2m}} \int_{r_0}^{r} \frac{1}{r^2\sqrt{E - V_e(r)}} dr, \tag{5.21}$$

where m is the mass, the initial position at $t = 0$ is $(r_0, \theta(r_0))$, and $V(r)$ is a spherically symmetric potential

$$V(r) = -\int_{r_0}^{r} F dr, \tag{5.22}$$

so that

$$\mathbf{F} = -\frac{dV}{dr}\hat{\mathbf{r}}, \tag{5.23}$$

and

$$V(r_0) = 0.^{16} \tag{5.24}$$

The energy is

$$E = \frac{1}{2}mv^2 + V(r) = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + V(r) = \frac{1}{2}m\dot{r}^2 + V(r) + \frac{h^2}{2mr^2}, \tag{5.25}$$

and for the chosen potential

$$E = \frac{1}{2}mv_0^2. \tag{5.26}$$

V_e is the so-called 'effective potential'

$$V_e(r) = V(r) + \frac{h^2}{2mr^2}. \tag{5.27}$$

We view the motion of the mass point in a non-inertial frame that rotates in such a way that the motion of the mass point is purely radial. In this frame there is an effective force

$$\mathbf{F_e} = -\frac{dV_e(r)}{dr}\hat{\mathbf{r}} = -\frac{dV(r)}{dr}\hat{\mathbf{r}} + \frac{h^2}{mr^3}\hat{\mathbf{r}}, \tag{5.28}$$

which consists of a physical force, experienced in an inertial frame, plus a non-inertial (centrifugal) force. If the motion is bounded, the mass point oscillates

[16] Note the sign ambiguity in Eq. (5.6). Here we choose the positive root.

between turning points r_1 and r_2 that are roots of the equation:

$$V_e(r) = E, \tag{5.29}$$

so that $\dot{r} = 0$, since

$$\frac{1}{2}m\dot{r}^2 + V_e(r) = E. \tag{5.30}$$

This equation "is just the expression for the energy of a particle of mass m moving in one dimension (along $\hat{\mathbf{r}}$) in an 'effective potential' $V_e(r)$. Evidently, we are now viewing the motion from a frame that rotates in such a way that the angular position of the particle is fixed and the motion is purely radial. Such a frame is clearly non-inertial" (De Lange and Pierrus 2010, p. 218).

We rewrite (5.29) as:

$$V(r) + \frac{h^2}{2mr^2} = -\int_{r_0}^{r} F dr + \frac{h^2}{2mr^2} = \frac{1}{2}mv_0^2, \tag{5.31}$$

which is Eq. (5.19).

The familiar procedure I have reviewed above is the modern equivalent of Newton's method in Corollary 1 for finding "expeditiously " the "greatest and least heights of bodies (that is the apsides of their orbits)." The "rule" for finding the turning points we teach today to our students is the following: since the integrands in Eqs. (5.20) and (5.21) are singular at points in which $E - V_e(r) = 0$, then the motion is possible only in the intervals in which $E \geq V_e(r)$. As already noted, the turning points are found by calculating the roots of Eq. (5.29).

It is interesting to compare our techniques to study the motion of the mass point to Newton's. While we think geometrically in terms of the graph representing the variation of the effective potential V_e in function of distance r, Newton's attention is focused on the motion of the particle. From Newton's perspective, the turning points occur "when the straight lines IK and KN are equal." Instead, we search for the intersection points between the graph of the effective potential and the horizontal line representing the constant energy.

Notwithstanding this notable difference, when we fetch pen and paper in order to calculate the apsidal distances r_1 and r_2, we write very much the same equation as Newton. Before expressing skepticism at such anachronism, the reader is invited to apply Eq. (5.29) to, say, an inverse-square force. The result will be very much the same equation obtained by starting, in accordance with philological accuracy, from the text of the *Principia*, that is from equation $ABFD = Z^2$. Such a translation into calculus or into algebra, I should add, is not so easily achievable for every demonstrative step in the *Principia*.

5.3.2 Corollary 2

We now turn at last to Corollary 2, the main focus of this paper. It reads as follows

> Corollary 2. The angle KIN, in which the orbit anywhere cuts the line IC, is also expeditiously found from the given height of the body, namely, by taking its sine to the radius as KN to IK, that is, as Z to the square root of the area $ABFD$ (Newton 1999, p. 531).

Corollary 2, like Corollary 1, does not presuppose that the problem of central forces has been solved via quadratures. It implies, like Corollary 1, that at the apsidal distances the angle $\angle KIN = \pi/2$, or $KN/IK = 1$.

The content of this corollary is richer though and extremely interesting. Newton shows how to determine the slope of the tangent to the sought orbit.

Indeed, because of formula (5.13), the sine of $\angle KIN$, that is KN/IK, is given by $Z/\sqrt{ABFD}$, or

$$\sin \angle KIN = \frac{KN}{IK} \propto \frac{Z}{\sqrt{ABFD}}. \tag{5.32}$$

What Newton understands here can be expressed in anachronistic terms by saying that he associates a slope field with the 'differential equation' (5.13).[17] The curves which are solutions of the equation must have a tangent at each point whose direction agrees with that of the slope field at that point.

It is important to note that if $\dot{r}$ changes sign, the orbit can have intersection points in which there are two slopes: one with radial component $\dot{r}$ and orthogonal component $r\dot{\theta}$ and another with radial component $-\dot{r}$ and orthogonal component $r\dot{\theta}$ (see Fig. 5.5). Thus, and this should be stressed, Newton's procedure does not define a slope field unambiguously, given the sign-ambiguity in Eq. (5.6).

In the very special case illustrated in Fig. 5.6, for the chosen initial conditions (energy and angular momentum) the particle moves so that $\dot{r}$ does not change sign. Thus, if we draw a circle with centre C, we can unambiguously associate velocity vectors to points on the circumference. All these vectors have the same inclination to the radius given by the formula for the slope field (5.32), and a modulus that we can calculate via Props. 39 and 40 (because of energy conservation, the speed $|v|$ is the same at equal distances from the centre).

It is interesting to note that Corollary 2 paves the way (modulo the above-mentioned sign-ambiguity) to the geometrical construction of a slope field. This concept, in a completely different context, plays a role in the methods by Euler, Heun, and Runge-Kutta, who—it is almost certain—were unaware of the well-

---

[17] I have translated Newton's formula as a differential equation in Sect. 5.2.4.2.

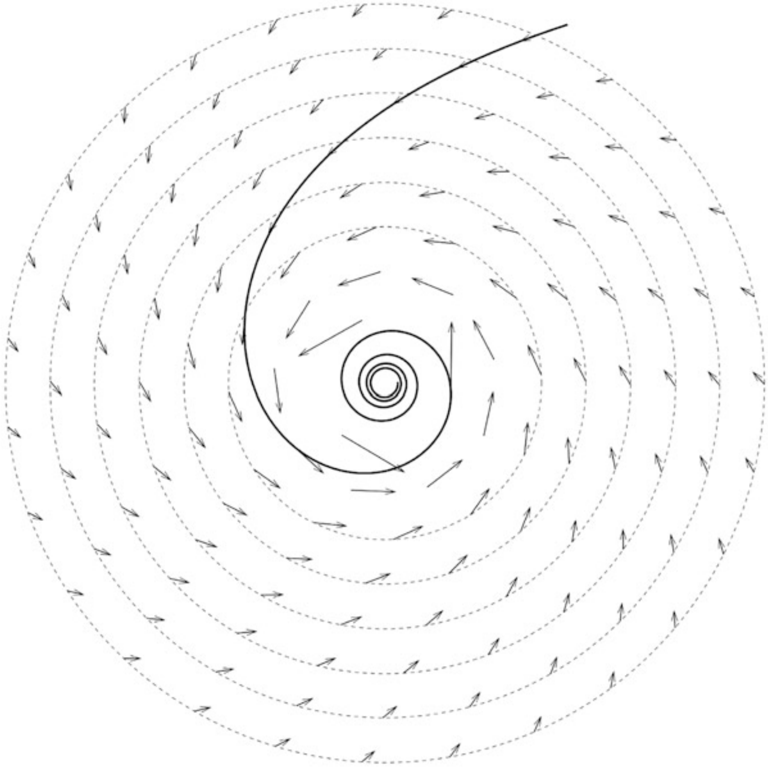


Fig. 5.6 Velocity vector field calculated for an inverse-cube force. For the initial conditions chosen, no apsidal distances r_1 and r_2 exist, since $\dot{r}$ does not change sign. The spiral orbit is graphically constructed so that its tangent at each point has its inclination defined by (5.32). In this case the orbit is a spiral with equation $1/r = \sinh k\theta$. As Newton shows in Corollary 3, Prop. 41, two bounded orbits are also possible (see Fig. 5.4, curves (a) and (c))

hidden Newtonian construction.[18] However, it would be incorrect to suggest that the result stated in Corollary 2 is 'equivalent' to the Runge-Kutta methods.[19]

[18] These methods were first applied to the solution of first-order differential equations. In the case considered in Prop. 41, the problem is to solve the Newtonian equation of motion, a second-order differential equation $F = ma$, applied to one-body central force motion. Runge and Kutta reduced second-order differential equations, which often arise in dynamical problems, to first order systems. For E. J. Nyström's extension of the Runge-Kutta methods to second-order differential equations, see Butcher and Wanner (1966, 120).

[19] As Prof. Siu A.Chin made me aware of (email dated May 23, 2022) one usually uses numerical algorithms to solve Eq. (5.1) directly in Cartesian coordinates. One does not solve equations in polar coordinates, such as equations (5.6) and (5.7), for a variety of reasons. The first is that time is then just a parameter, completely under control. The resulting trajectory is then a simple parametric set of coordinates in t. The second reason is that when solving (5.1) directly in Cartesian coordinates, it is easy to derive symplectic algorithms by sequential updates. One is

Since the inverse problem of central forces is integrable in terms of the techniques of quadrature available to Newton and his contemporaries only for a few cases (corresponding to an elastic, an inverse-square, and an inverse-cube force), it is plausible that Newton conceived Corollary 2 as a way to tackle the problem not via quadratures (as in Corollary 3) but via a graphical approximation technique. The extant evidence suggests that Newton was able to solve via quadratures the inverse problem only for the inverse-cube case we reviewed above. Corollary 2 is cited in only one place in the third edition (1726) of the the *Principia*, namely in Proposition 56, Book 1.[20] Proposition 56 offers a solution to the problem of finding the "curve" traced by a body under any gravity-like centripetal force constrained to a surface of revolution the axis of which extends to the force centre. Since in this case an exact quadrature was not available, Newton referred to Corollary 2, clearly gesturing to the fact that a graphical solution was possible by a construction. The reference to Corollary 2 in Prop. 56 as a way for determining the sought orbit (the "curve APp") supports my reading of this corollary as providing a technique for the graphical construction of orbits.[21]

5.4 Summing Up

Let us briefly recapitulate the main achievements of the Propositions and Corollaries we have analyzed.

In Propositions 39–41, Book 1, of the *Principia*, Newton reduced to quadratures the problem of determining the orbit of a body fired from a given position with a given velocity and accelerated by a given central and isotropic force. He stated the problem of central forces very clearly and for this fact alone these propositions are a notable achievement.

In Corollary 1 Newton was able to evaluate the apsidal distances (when the motion is bounded).

In Corollary 2 he determined the slope of the tangent of the orbit at any distance from the force's centre, anticipating the concept of a slope field.This is a very fruitful concept that is still important today.[22]

Finally, in Corollary 3, Newton squared curves abz and acx (equations (5.11) and (5.12)) for the case of an inverse-cube force and for an initial velocity orthogonal to

then guaranteed the conservation of all Poincaré invariants, including that of phase space volume (Liouville Theorem). Phase volume is not conserved if only the above-mentioned polar coordinate equations are solved in sequence.

[20] Newton (1726, pp. 159–160). In the first and the second edition, reference was made only to Proposition 41. I thank George Smith for his kind suggestion on this issue.

[21] Newton writes: "And accordingly (by consulting Prop. 41 with its corol. 2) the way of determining the curve APp is readily apparent (Newton 1999, 360).

[22] The introduction of the concept of phase space, of course, provided an even richer conceptual framework for representing the solutions of differential equations.

the radius vector, thus solving the problem of central force motion for this particular case.

So far so good. Such a summary helps, at least mnemonically. But, from the historian's point of view, the brief outline provided in this Section hides many interpretative problems, to which we now turn.

5.5 On Presentism, or How to Deconstruct This Paper

In this paper I have freely used concepts (such as slope field, differential equation, potential) and results (such as conservation of mechanical energy, conservation of angular momentum) unavailable to Newton. Basically, I have written an anachronistic piece of work! I have been dabbling with the problem of anachronism for several years, very much helped by a number of friends, including the dedicatee of this volume (Guicciardini 2021).

The reason why I have so freely used anachronisms in this paper is my playful intention to offer Jeremy a divertissement. But there is, perhaps, a more serious reason behind my little hazardous exercise: it is my conviction that anachronism in writing history of mathematics is not per se a sin that should be eschewed at all costs.[23]

In a way, the use of concepts that belong to our present is unavoidable: we are twenty-first-century historians, inevitably situated in our present. To paraphrase Bruno Latour's turn of phrase, we have never been early-moderns! (Latour 1991). In this paper I have pushed the use, and perhaps the abuse, of anachronism to a limit, through a practice that I propose to call 'presentism'. Most notably, as the reader has certainly noticed, I have proposed modern renderings in the form of differential equations of Newton's "quadratures."

These anachronistic renderings are enlightening and deceptive at the same time. They help us familiarize ourselves with the "curves" that must be "squared" in order to solve the inverse problem Newton refers to in Proposition 41. Thanks to these renderings, we can grasp what problem Newton was trying to solve. But they deceptively project our concepts of function and differential equation onto Newtonian mathematics. The presentist translations that I have proposed bestow a generality on Newton's Proposition 41 and its Corollaries that is absent from Newton's conceptual framework. Newton did not conceive his formula (5.13) as an instance of a first-order differential equation. The formula (5.13) was a geometric proportionality that Newton applied to the specific case of motion he was considering. And even less did he have a notion of vector field. On the other hand, there are certain features of Newton's mathematical procedures in Props. 39-41 that are still part of our mathematical toolbox. In this paper I have highlighted some of them: when studying central force motion, we state the initial value problem in a

[23] On anachronism as a 'sin', see Syrjämäki (2011).

way that is equivalent to Newton's, when calculating the apsidal distances we make use of the very same equation, etc.

In order to write an acceptable piece of history, the interpretative anachronistic move explored in this paper—and consisting in a process of assimilation—should be followed by an attempt to situate Newton's text in its proper context. As historians, we have to grasp the distance of Newton's text from our present. That said, one should appreciate that presentist interpretations are not pointless: they offer the historian a chance to appreciate the otherness of past mathematics. We learn something by 'distancing' our historical narrative from presentist interpretations. We can achieve such 'distancing' in three ways.

First, we can set ourselves the task of underscoring the many reasons why Corollary 2 is not mathematically equivalent to the concepts and methods in use today, as I have already partly done above. By highlighting the mathematical differences between Newton's graphical construction and the notion of slope field or the Runge-Kutta methods (see footnotes 18 and 19), we learn something about the mathematics of the *Principia*.

Second, we can study how Newton himself translated Proposition 41 into the language of his methods of quadratures. Indeed, we are lucky to have manuscripts in Newton's own hand dating from the early 1690s in which he explains via algebraic symbols how to "square" the curves leading to the solution of the inverse-problem for inverse-cube forces. This is what I have tried to do in Guicciardini (2016).

Third, we can study the reception of Newton's Propositions 39–41 in the works of competent contemporary readers. Indeed, the mathematical methods deployed by Newton in the *Principia*, and these propositions in particular, were the object of intense debate in the context of the Newton-Leibniz controversy. Newton's detractors claimed that the methods of the *Principia* were old and geometrical, and therefore did not fare well when compared with Leibniz's innovative and symbolical differential equations. Many in Britain, such as Nicolas Fatio de Duillier, David Gregory, and John Keill, disagreed. Studying this debate is a way of reading Newton's text through the eyes of his contemporaries. Rather than siding with one faction, the historian leaves the historical actors to speak for themselves, so that the *Principia* may acquire different meanings via the different, sometimes diverging, interpretations of its readers. This is what I have tried to do in Guicciardini (1999).

The first approach is 'conceptualist'. It consists in a comparison of the concepts and methods deployed in different time periods (and cultures). The (problematic) assumption is that our own mathematical language can be used as a target language in which this comparison can be implemented.

The second approach is 'intentionalist'. It reads an author's text through paratexts by the author himself, in an attempt to capture his intentions (or perhaps, less problematically, his 'tasks and criteria of quality control', as Henk Bos would put it) (Bos 2004). The merit of this approach is that it allows us to situate the text in the author's conceptual framework. In the case at hand, we can grasp how Newton himself understood and performed the quadratures of Prop. 41.

The third approach is 'receptionist'. It reads an author's text through the meaning attributed to it by its readers. The merit of this approach is that it allows us to follow the line of development that emerged in the process of reception, interpretation, and transformation of the text.[24] In the case at hand, we can understand how mathematicians such as Johann Bernoulli, Pierre Varignon, and Leonhard Euler rewrote Prop. 41 in terms of differential equations, a way of writing that presents certain similarities, but also differences, with respect to our presentist rendering. The receptionist approach is, in a way, a form of anachronism. Bernoulli's version of Prop. 41 cannot be attributed to Newton. This attribution would project upon Newton a mathematical conception that was not his own.[25]

It is interesting to distinguish 'receptionism' from 'presentism'. With the receptionist approach we identify aspects of the author's text that have been received and transformed by its readers. We can document this process of reception by studying letters, exchanges of manuscripts, readers' marginalia, explicit quotations, etc. Conversely, with the presentist approach we identify aspects of the author's text that appear to anticipate discoveries which occurred in a much later period, in the absence of any provable continuity in terms of transmission. In the case at hand, we have identified Corollary 2 as an anticipation of the notion of a slope field associated to a differential equation. While with the receptionist approach we write a history of how the text was read, received, and transformed, with the presentist approach we employ the "dangerous" category of precursor, chastized by Helène Metzger (Chimisso and Jardine 2021). In presentist anachronistic history a documentable account of transmission is absent. In the case at hand, it is clear that those who developed the Runge-Kutta method had no direct contact with Newton's text. In a way, while with the receptionist approach we are still probing the author's intentions, with the presentist approach the author— his conceptual world (the author's "other present")—is lost.

Yet, even a presentist approach has its merits. Reading a past mathematical text via the use of concepts and results that were not available to the author can sometimes shed light on its meaning, a meaning that appears to be significant for us today. Perhaps, the most valuable advantage of this approach is that it brings the text to life, making it meaningful for present-day mathematicians: it makes the text available for use in our culture. Mathematical texts display not only a high degree of stability, accounting for the many possibilities of meaningful anachronistic readings. They invite a performative reading. The authors of mathematical texts have an intended reader in mind who is invited to receive the text in a performative way. The reader is asked to appropriate the text by applying, and extending, the methods proposed by the author to new problems, thus producing new texts. This active, performative approach to the mathematical text is very much part of what reading, and appreciating, mathematics consists in Blåsjö and Hogendijk (2018)

[24] On the notion of 'filiation' in Hélène Metzger's work see Chimisso (2019).

[25] See the Chapter on Bernoulli in Guicciardini (1999, pp. 195–249). It is worth stressing that my presentation of Prop. 41 in that chapter is aimed at illustrating how Bernoulli read Newton's treatment of central force in terms of the Leibnizian calculus.

Presentism, however, for many historians of mathematics cannot be a satisfactory aim. One might contend, in agreement with Paolo Rossi, that the historian's ethos consists in capturing the distance between the past and the present.

> The authors are near and far. At one and the same time, they are paradoxically both near and far [...]. All that is needed is to choose aptly among well-known texts. In this case, it is not hard to find passages that present themselves as curiously modern, so that they could be easily integrated into ongoing debates. And then again, it is easy to set them alongside texts by the same authors that awaken the feeling of an unbridgeable distance or a nearness that is merely apparent.[26]

Any anachronism, any 'familiarization' with the text, must be followed by 'foregnization'. If we write the history of mathematics for the purpose of assimilating past texts into our conceptual and linguistic present, we do not honour the task historians should always set themselves, namely: to capture the otherness of past cultures. Yet, as I have hinted above, even presentism, the most extreme form of anachronism, can be a useful hermeneutic move for the historian. In the case at hand, even our shamelessly anachronistic rendering of Corollary 2 in terms of slope fields graphically produced with the help of a computer can offer the historian a chance for re-locating the text into its proper context. How can we do so in this case?

The answer is simple, but still interesting since it opens up a yet unexplored research field. Early-modern historians are aware that mathematical practitioners in Newton's milieu were conversant with graphical methods, which they applied to the construction of ballistic trajectories, as well as to planetary and cometary orbits. An astronomer such as Jeremiah Horrocks constructed the orbit of the Moon via a deferent-plus-epicycle model that Newton adapted in his *Theory of the Moon* (1702). A polymath such as Robert Hooke resorted to a graphical construction in order to approximate the orbit traced by a body under the action of an elastic force. A writer on ballistics such as Robert Anderson used similar construction tools to trace on paper the shape of the path of a projectile fired by a cannon. All these methods were well-known to Newton.[27] It would be extremely interesting to situate Newton's graphical techniques for tracing orbits in the mathematical culture of the astronomers, writers on ballistics, and mechanicians who devised mechanical generations of curves. Thus, what we have here is another interpretative approach to Newton's Props. 39–41, which I would call 'contextual'. In this case, the meaning of the text is accessible after a study of the author's cultural environment.

[26] "Gli autori sono vicine e lontani. Contemporaneamente e paradossalmente vicini e lontani [...] Basta scegliere bene entro testi ben conosciuti. Allora non è difficile trovare pagine che appaiono singolarmente moderne, tali da essere facilmente integrabili nel dibattito contemporaneo. Ed è poi facile contrapporre ad esse testi dello stesso autore che danno il senso di una irrimediabile distanza o di una vicinanza che era solo apparente." (Rossi 1999, p. 28)

[27] The role of graphical solutions in Newton's mathematical work has recently been underscored by Nauenberg in Nauenberg (2018) and, if I understand well, is presently researched by Jed Z. Buchwald.

The moral that we can draw from my little exercise in presentist interpretation is that anachronism, in its various forms, can be a step in a process of engagement with past texts, a process consisting of a series of, sometimes distinct, hermeneutic moves. Reading a text historically is an iterative process in which we start by assimilating the text to our linguistic space (familiarization). This first move (presentism) is followed by a process of distancing (foregnization), in which we resort to a palette of tools that include a technical comparison of present-day concepts and methods with those of the past (conceptualism), the attempt to re-read the text in light of the author's tasks (intentionalism), through the eyes of its contemporary readers (receptionisim), or by situating it in a broader cultural narrative (contextualism). It is possible to write the history of mathematics if we control these moves through an awareness of how they can both enlighten and deceive us. The example we must follow is that provided by Jeremy's work.

Acknowledgments I wish to thank George E. Smith (Tufts University) for an inspiring exchange of emails on this paper and for drawing my attention to the importance of the first two corollaries of Proposition 41, Book 1, and Antonio Giorgilli (Dipartimento di Matematica, Università degli Studi di Milano) for many insightful comments and for providing Figs. 5.4, 5.5, and 5.6. Last but not least, I should thank Siu A. Chin (Texas A & M University) for his competent and eye-opening comments. This research was funded by the Department of Philosophy "Piero Martinetti" of the University of Milan under the Project "Departments of Excellence 2023-2027" awarded by the Ministry of University and Research (MUR).

References

Blåsjö, Viktor, and Jan P. Hogendijk. 2018. On Translating Mathematics. *Isis* 109: 774–781.

Bos, Henk J. M. 2004. Philosophical Challenges from History of Mathematics. In *New Trends in the History and Philosophy of Mathematics*, ed. Tinne Hoff Kjeldsen, Stig Arthur Pedersen, and Lise Mariane Sonne-Hansen, 51–66. Odense: University Press of Southern Denmark.

Brackenridge, J. B. 1995. *The Key to Newton's Dynamics: The Kepler Problem and the Principia*. Los Angeles: Univ. of California Press.

Butcher, John C., and G. Wanner. 1966. Runge-Kutta Methods: Some Historical Note. *Applied Numerical Mathematics* 22: 113–151.

Chimisso, Cristina. 2019. *Hélène Metzger: Historian and Historiographer of the Sciences*. Abingdon, Oxon/New York, NY: Routledge.

Chimisso, Cristina, and Nicholas Jardine. 2021. Hélène Metzger on Precursors: A Historian and Philosopher of Science Confronts Her Evil Demon. *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 11 (2): 331–353.

Chin, Siu A. 2022. Modern Light on Ancient Feud: Robert Hooke and Newton's Graphical Method. *Historia Mathematica*, online 10 March 2022.

De Lange, O. L., and John Pierrus. 2010. *Solved Problems in Classical Mechanics: Analytical and Numerical Solutions with Comments*. Oxford : Oxford University Press.

Guicciardini, Niccolò. 1999. *Reading the Principia: The Debate on Newton's Mathematical Methods for Natural Philosophy from 1687 to 1736*. Cambridge: Cambridge University Press.

Guicciardini, Niccolò. 2016. Lost in Translation? Reading Newton on Inverse-cube Orbits. *Archive for History of Exact Sciences* 70: 205–241.

Guicciardini, Niccolò, ed. 2021. *Anachronisms in the History of Mathematics: Essays on the Historical Interpretation of Mathematical Texts*. Cambridge: Cambridge University Press.

Latour, Bruno. 1991. *Nous n'avons jamais été modernes: Essai d'anthropologie symétrique*. Paris: Editions La Découverte.

Nauenberg, Michael. 1994. Newton's Early Computational Method for Dynamics. *Archive for History of Exact Sciences* 46: 221–252.

Nauenberg, Michael. 2018. Newton's Graphical Method for Central Force Orbits. *American Journal of Physics* 86: 765–771.

Newton, Isaac. 1726. *Philosophiae Naturalis Principia Mathematica*. London: W. & J. Innys.

Newton, Isaac. 1999. *The Principia: Mathematical Principles of Natural Philosophy … Preceded by a Guide to Newton's Principia by I. Bernard Cohen*. Translated by I. Bernard Cohen and Anne Whitman, assisted by Julia Budenz. Berkeley: University of California Press.

Rossi, Paolo. 1999. *Un Altro Presente: Saggi sulla Storia della Filosofia*. Bologna: Il Mulino.

Syrjämäki, Sami. 2011. *Sins of a Historian: Perspectives on the Problem of Anachronism*. Tampere: Tampere University Press.

Part II
Practices of Mathematics

Chapter 6
On Felix Klein's Early Geometrical Works, 1869–1872

David E. Rowe

Abstract Felix Klein's formative years constitute a famous chapter in the history of mathematics, especially familiar because Klein himself wrote about it often. Later writers have often highlighted his collaboration with Sophus Lie and the ideas that led to Klein's "Erlangen Program." Klein re-packaged his early work when he edited his collected works, a project that engaged his attention from 1919 to 1923. By unpacking its first volume, we can begin to appreciate that transformations groups formed a relatively small part of Klein's early geometrical work, whereas a great deal of it was devoted to important new results in line geometry.

6.1 Introduction

The title of this paper refers to the earliest productive period in the long career of Felix Klein (1849–1925). Much of what I have to say here about that fertile period, however, concerns Klein's personal reflections when he looked back on that time some fifty years later. His retrospective account of his early work is reflected in the first of the three volumes of his collected works (Klein 1921–1923), a project that engaged Klein's attention from 1919 to 1923. Klein not only directed the editing of his collected works, he also adopted a novel structure for these three volumes in order to highlight the main themes and subjects of his scientific activity. Thus, for Volume I, he chose a collection of papers, presented chronologically, that dealt with his research in three fields: (1) line geometry, (2) foundations of geometry, and (3) works related to his "Erlangen Program" from 1872. That particular publication has long been the one most strongly associated with Klein's name, but it also brings to mind the Norwegian Sophus Lie (1842–1899), his friend and collaborator from that period (Yaglom 1988). A major motivation behind what I have written here is to underscore how this famous relationship in the history of mathematics has been

D. E. Rowe (✉)
Mathematics Institute, Mainz University, Mainz, Germany
e-mail: rowe@mathematik.uni-mainz.de

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_6

understood far too narrowly. Klein and Lie were both deeply influenced by Julius Plücker's new theory of line geometry, and the account below is an attempt to show the central importance of that first topic in Klein's collected works. By now there is a wealth of secondary literature on the third topic, which has also led to occasional debates, but my intention is not so much to engage with earlier studies as to offer a different perspective that suggests why other factors need to be considered in order to reach a better understanding of the relationship between Klein and Lie.[1]

To begin very broadly, we should recognize that at the time they met Klein and Lie were already both extremely ambitious young men. Not only did they want to make a deep impression on their contemporaries, they also became increasingly concerned with their respective legacies within the history of mathematics as they grew older. A mathematician's reputation can easily change over time, of course, and in Lie's case it began to rise in the 1880s, by which time he was in his forties. By the 1890s, his ambition and pride were as great as ever, but his creative powers were receding. His theory of transformation groups was by then widely celebrated, particularly in France. Supported by Henri Poincaré, among others, Lie became a corresponding member of the French Academy in 1892 (Klein was elected five years later).

When Klein left Leipzig in 1886 to assume his final professorship in Göttingen, he managed to have Lie appointed as his successor, a move rival mathematicians in Berlin rightly saw as part of Klein's strategy of building alliances that undermined the influence of the Berlin network (Biermann 1988). Lie's situation as a foreigner in Leipzig proved difficult from the beginning, however, and he gradually came to resent the way some in Klein's circle treated him as an important figure in the latter's entourage. Lie was more than six years older than Klein, whose support he had long enjoyed, but he never felt the slightest dependence on him intellectually. Indeed, he knew that Klein had never been able to catch up with his new ideas after 1872, when they essentially went their separate ways. Klein's letters to Lie from these years make this abundantly clear, and if we keep in mind that Klein was only twenty years old when they met, it is easy to imagine why Lie thought of him then as a "bright kid."

Twenty years later, both men were well past their prime mathematically, though their ambitions were great as ever. The fact that Klein began pressing Lie to join him in writing a retrospective account of their work together, culminating with a republication of Klein's "Erlangen Programm," only made Lie suspicious and angry. When he learned that Klein had destroyed all of Lie's early correspondence—he, in fact, had at some point in 1877 burned all the letters in his possession—the Norwegian cut off communication with his former friend, preparing the way for his public attack on Klein in 1893. In discussing some of these events from the early 1890s in section 2 below, I take up more or less where Lizhen Ji (Ji and Papadopoulos 2015, 1–58) and Jeremy Gray (Gray 2015) left off in their respective survey articles.

[1] For those with an interest in connections between the ideas in the Erlangen Program and mathematical physics, see the essays in Ji and Papadopoulos (2015).

Klein's name and fame had much to do with his talents as a teacher, editor, and organizer. After his alliance with Lie disintegrated, he developed a new one with David Hilbert, who became his junior colleague in Göttingen and eventually inherited his role as editor-in-chief of *Mathematische Annalen*. Meanwhile, in Leipzig, Lie continued to cultivate his ties with the younger generation of French mathematicians, but even more importantly he gained the support of the young geometer, Georg Scheffers, a talented writer. During the early 1890s, Scheffers wrote two textbooks based on lecture courses taught by Lie, Lie and Scheffers (1891, 1893). Especially important for understanding Lie's early works, however, is their third book (Lie and Scheffers 1896), which Lie understood as elucidating his ideas from the period 1869–1872. Klein's role in this account was reduced to a few special ideas and results, whereas Lie's preface placed his achievements within a grandiose historical framework stretching from Archimedes to his own present days. Sophus Lie no doubt felt he had a deep understanding of the history of mathematics, but he lacked the kind of patience required to read the works of other mathematicians carefully. His impressive creative powers turned nearly everything he wrote about past achievements into a mirror reflecting his own singular accomplishments. In this respect, he probably had no peer as a mathematician who wrote in praise of his own work.

In recent studies of famous figures in the annals of science, historians have taken note of extra-scientific factors that at times enhance a person's visibility or fame. One of the more striking examples came about in November 1919, when the British scientific community announced new experimental evidence confirming Einstein's general theory of relativity, an event that made him famous overnight (Rowe 2012). The fact that Great Britain and Germany had only recently signed the Versailles Treaty ending the First World War added greatly to the sensation this created. Recognizing this, Einstein paid due honor to Newton's immortal theory of gravitation, while the liberal press expounded on how relativity provided a deeper understanding of the universe within a new and yet evolving European political order. Another example is Mario Biagioli's influential study of Galileo's career, which emphasized how in crossing disciplinary boundaries from mathematics to natural philosophy Galileo enhanced his own reputation through a process of "self-fashioning" (Biagioli 1993). Though he lacked any standing as a professional astronomer, his telescopic discoveries called fundamental assumptions of Arisotelian cosmology into question. Galileo's controversial defense of Copernicanism ultimately brought him into conflict with the Catholic Church, a personal defeat that later redounded to his lasting fame.

At this illustrates, fame and controversy are sometimes linked; they have often played a role in contested claims to intellectual rights as well. Competition can foster personal conflicts that occasionally lead to major priority disputes over a brilliant discovery or an ingenious theory.[2] The most notorious such contest in the

[2] In my view, historians should avoid the temptation to adjudicate past contests—such as the "race" between Einstein and Hilbert for generally covariant gravitational field equations that was taken

history of mathematics pitted Leibniz against Newton, a battle that led to a famous public dispute over who deserved credit for being the "first inventor" of the calculus. In his desire to deny Leibniz even the title of "second inventor," Newton used his authority as president of the Royal Society to argue that his German rival had stolen the calculus from him (Hall 1980). Indeed, Newton amassed a mountain of evidence that (in his eyes) proved to the world that Leibniz had been given access to Newton's early (long unpublished) manuscripts, which supposedly contained all he would have needed to pretend he had invented the "Leibnizian calculus." Newton took the latter to be nothing more than a new notation for his own theory of fluxions. One could say about this, in summary, that Newton won this battle, but Leibniz won the war, since the Bernoullis and Euler developed Leibnizian methods much further, whereas the British mathematicians had no comparable legacy to build on. Newton's *Principia* made him exceedingly famous, but the mathematical methods he used could hardly serve as a useful model for modern analysis.[3]

It might seem far fetched to compare the famous controversy between Newton and Leibniz with the rift that developed between Klein and Lie. The former case, after all, was all action at a distance: Newton and Leibniz never even met one another personally. Still, in both cases, severe differences of opinion arose over the respective merits of what the two parties has achieved in their youth. As Niccolò Guicciardini has emphasized, Newton and Leibniz did not even share a common opinion about the methodological importance of the differential and integral calculus (Guicciardini 2009, 381–384). In Lie's case, he was intent on warding off work he saw as encroaching on his own proper mathematical terrain, the theory of transformation groups. As time went on, he kept reiterating how he had staked out these ideas in the mid-1870s, if not earlier, and he adamantly claimed them as his own.

His falling out with Klein was, in fact, only the last in a series of (mainly private) disputes that began much earlier. The most serious of these conflicts took place during the late 1880s when he informed Klein that Friedrich Engel, Lie's assistant at the time, had acted unfaithfully in corresponding with Wilhelm Killing, who published a series of groundbreaking papers on Lie algebras in *Mathematische Annalen*. As editor-in-chief of the *Annalen*, Klein had accepted Killing's new results

up in the late 1990s—nor should they make lightly considered retrospective value judgments about the status of past work.

[3] For an enlightening study of Newton's mathematical views and methods, see Guicciardini (2009). Until well into the twentieth century, mathematicians and historians of mathematics often took sides in the Newton vs. Leibniz story, and much of this literature reflects strongly nationalistic inclinations toward hero worship. A more objective attitude gradually emerged after the Second World War, however, when a great deal of relevant source material came to light. Even so, the Newton scholar A.R. Hall expressed great dismay over some of the opinions expressed by Joseph Ehrenfried Hofmann in Hofmann (1974), the standard account of Leibniz's intellectual journey leading to the calculus (Hall 1980, 65–67). Today, one can easily study Newton's early works in D.T. Whiteside's monumental editions of *The Mathematical Papers of Isaac Newton* in eight volumes (Whiteside 1967–1981), whereas the editors of the Leibniz edition in Hanover continue to turn out new volumes of his mathematical manuscripts.

with delight, especially because they highlighted important new developments in group theory (Hawkins 2000, 165). These events from 1887 onward preceded Lie's breakdown in 1889, though Arild Stubhaug draws no connection between them in his biography of Lie (Stubhaug 2002).[4]

A traditional problem with biographical studies of mathematicians is their tendency to engage in hagiography, often tinged with nationalist sentiments. Stubhaug's book is a glaring example, but Renate Tobies's far more scholarly biography of Felix Klein (Tobies 2021) cannot claim to be free from hero worship either. Her study covers a truly impressive array of topics and favorite causes in Klein's incredibly active life, but without taking any note of the extent to which his "self-fashioning" played a role in this. Unlike Poincaré and Hilbert, the two younger mathematicians who overshadowed him during his lifetime, Klein went to great lengths to ensure that the mathematical world remembered his accomplishments. Throughout the war years, he offered a series of lectures on the mathematics of the preceding century, beginning with the work of Carl Friedrich Gauss. Soon after Einstein's general theory of relativity caught Hilbert's attention, though, Klein broke off these lectures. In 1917, he began a new series devoted to mathematical developments connected with relativity theory, giving special attention to group invariants as sketched in his "Erlangen Program." Once the war ended, though, he dropped all his plans to publish these lectures in order to focus on preparing the three volumes of his collected works (Klein 1921–1923).

Historians have cited Klein's commentaries on numerous occasions, and they can only be grateful when famous mathematicians take the trouble to reflect on their past work and the circumstances that led to it. Mathematicians, on the other hand, rarely hope to find inspiration for their research by going through another's collected works, unless of course they want to hunt down some particular paper that interests them. Probably most experts on Riemann's hypothesis concerning the zeroes of the zeta-function, today one of the Clay Millennium Problems, have tried to read the original paper from 1859, which is only 9 pages long.

A more typical example, though, would be the papers gathered in the seven volumes of Sophus Lie's collected works, which appeared many years after his death in 1899. Even during his lifetime, relatively few mathematicians found Lie's papers readable, and after his death the number who studied them must have been vanishingly small. When the first two volumes, Lie (1934, 1935), finally appeared, the modern theory of Lie groups and Lie algebras had begun to assume a central place in the corpus of mathematics. It would be hard to overstate the gulf separating Lie's earliest geometrical ideas from modern Lie theory. These two volumes contain many of the papers that Lie wrote—or in some cases only drafted, leaving to Klein the task of preparing the final version—during the period from 1869 to 1872, when both worked together closely. Friedrich Engel, the leading expert on Lie's mature

[4] Without naming sources or dates, Stubhaug's study cites passages from several letters from Lie to Klein, in which he vented his anger over Killing's work and especially Engel's role in corresponding with Killing; several of the letters cited can be found in Rowe (1988).

work and the editor of the five other volumes of Lie's collected works, found these early papers so impenetrable that he asked the Danish geometer Poul Heegaard to help him edit them.

Felix Klein's formative years constitute an at once famous and familiar chapter in the history of mathematics. This has much to do with the fact that Klein himself wrote and talked about his early work on numerous occasions, most notably while preparing his collected works (Klein 1921–1923). That project began in 1919, just after the collapse of the German monarchy. As an admirer of Kaiser Wilhelm II, Klein was deeply disheartened by his nation's military defeat and the ensuing political chaos. The opportunity to relive his creative exploits as a young man surely had a therapeutic effect, especially because it gave him the chance to hold special lectures on these topics. These were attended by a small coterie of auditors that included the young Ukrainian Alexander Ostrowski, who assisted Klein in editing works he had written a half-century earlier.

For Volume I, Klein added autobiographical information to help contextualize the events and influences surrounding these works. Some of these papers had already been republished in the 1880s and 1890s, in which cases he had to add yet another layer of footnotes for his new retrospective remarks.[5] This kind of direct engagement with the past, in which a mathematician carefully prepares his own collected works, was quite unparalleled. In Klein's time, when the notion of a self-fashioning scientist had yet to be born, such a project had no precedent at all. Normally, a mathematician took no part in producing compendia of their past work; in most cases, in fact, they were long dead by the time others took this initiative.[6]

As noted above, Klein adopted a tripartite structure for the first volume with: (1) line geometry, (2) foundations of geometry, and (3) works relating to the "Erlangen Program." Since the latter two areas are far better known than the first, most of what follows will concentrate on line geometry, but in the following section I briefly discuss certain little known aspects concerning the last two topics. In discussing line geometry, it should be emphasized that Klein's work in this field had a profound, though by now long since forgotten impact. When Ostrowski signed on as his assistant in 1919, he had no knowledge of line geometry at all, though he soon came to appreciate its beauty.[7]

To a certain extent, my account supplements an older essay (Rowe 1989) and two more recent ones: Rowe (2016, 2019). Those studies all touch on Julius Plücker's approach to line geometry, as set forth in (Plücker 1868) and its posthumously published sequel (Plücker 1869). In the final section, I conclude with some remarks about secondary literature that can serve as particularly useful guides to classical geometry in the spirit of the nineteenth century, including line geometry. Let

[5] Klein signaled the new ones by writing them in square brackets.

[6] Hilbert was still alive and so had the opportunity to read the three volumes of his papers (Hilbert 1932–1935), but he took no part in preparing that edition.

[7] As he told me in an interview conducted in Lugano in September 1984.

me begin, though, with some brief remarks about the other two areas that Klein investigated during his youth.

6.2 "Erlangen Programm" and Non-Euclidean Geometry

When looking back on Felix Klein's career, some writers have focused rather narrowly on his famous "Erlangen Programm" (Klein 1872/1893). This text has rightly been regarded as linked with Klein's early collaboration with Sophus Lie. In fact, under other circumstances they might have even co-authored a work such as this one, just as they had done on other occasions. Had this occurred, it surely would have helped to alleviate some of the bitter feelings that finally ended their friendship in 1893. As it happened, though, Klein wrote the text as an obligatory *Programmschrift* for his inauguration as a full professor in Erlangen. This was a formal requirement, which thus implies that there were many such texts. Klein's successor, Paul Gordan, also composed an "Erlangen Programm", though obviously it was quickly forgotten, unlike Klein's.[8]

Moreover, we should not imagine that those who read Klein's text suddenly saw the light; it took some twenty years before it became widely known (Hawkins 1984). Even when its reception took off in the 1890s, this was still a time when Lie's theory of continuous groups had yet to emerge in its modern form. Back in 1872, the very notion of groups in geometry was largely intuitive, so what Klein had in mind were certain examples of closed families of transformations, most of them well known to contemporary geometers.[9] In short, the informal name "Erlangen Program" reflected the occasion and should not be interpreted as a research program. To the extent the text suggested a new direction for future investigations, this had very little to do with finding or classifying new groups. On the contrary, Klein's principal aim was to promote a broader conceptual understanding of recent geometrical research by highlighting several important traits and ideas. Certainly the notion of transformation groups—understood largely in a heuristic sense—was central to this vision. Although the "Erlangen Programm" became the most famous of all Klein's works, its reputation had far more to do with its guiding ideas rather than the actual content of the text itself. For historians, on the other hand, it provides an interesting snapshot of trends in geometrical research in the early 1870s.

Klein undertook two largely independent efforts to promote the "Erlangen Programm," the first during the early 1890s, whereas the second phase took place in the years leading up to publication of Volume I of Klein (1921–1923). Throughout the 1880s, Klein almost never taught courses in geometry, which reflects the shift

[8] See Konrad Jakobs and Heinrich Utz, Erlangen Programs, *Mathematical Intelligencer* 6(1)(1984): 79.

[9] He had to add a footnote in 1893, when the text of Klein (1872/1893) was reprinted, pointing out that one needed to stipulate that the group contains its inverse transformations.

in his research interests toward complex analysis. After his health collapsed in the midst of a brief, but famous competition with Henri Poincaré, Klein began offering a cycle of courses on elliptic, hyperelliptic, and Abelian functions. This cycle mimicked the courses regularly offered by Karl Weierstrass in Berlin, except that his approach was based on Riemannian ideas or what was often called geometric function theory. By this time, the rivalry between Klein and the aging Berlin school was intense, culminating in 1886 when he left Leipzig for Göttingen and managed to arrange Lie's appointment as his successor. Meanwhile, Klein continued to teach courses on complex function theory in Göttingen. Only toward the very end of the decade did that change, beginning in 1889/90 when he offered a two-semester lecture course on non-Euclidean geometry (Klein 1893).

That course marks the beginning of a crucial four-year period during which Klein lectured extensively on his own early work as well as Lie's. These were unhappy years for both men, and by 1893 the strains in their friendship had reached the breaking point. Klein's situation in Göttingen was hardly what he had hoped, as he found himself constrained by his senior colleague, Hermann Amandus Schwarz. Klein's rival was a disciple of Weierstrass, who was now 75 and in poor health. Adding to the tensions, both men were leading candidates to succeed the elderly analyst. Furthermore, both knew that Schwarz had earlier hoped to succeed Klein in Leipzig, yet another source of acrimony between them. Already concerned about upholding his legacy, Klein offered a year-long lecture course on Riemann surfaces and afterward taught a two-semester course on higher geometry during the academic year 1892/93. The latter course represents the climax of his efforts during this period to elaborate on the themes he had sketched twenty years earlier in his "Erlangen Programm" (Klein 1872/1893).

The original German text of Klein (1872/1893) was barely accessible, so in 1891–1892 Klein began to make plans to republish it with commentary in *Mathematische Annalen*. He hoped to do this as a joint venture with Sophus Lie. Except for Lie (1872), his most important publication from that period, very few of Lie's early papers were known either, since before his arrival in Leipzig in 1886 he had published most of his work in Norway. Klein's promotional plans included showcasing some of his own work alongside several overlooked papers by Lie, thereby putting their earlier collaboration back in the spotlight. Lie no doubt had mixed feelings about this from 1891 onward, but eventually these gave way to intense anger; his outbursts left Klein in a state of shock.[10] In a letter from 2 December 1892 to Adolf Mayer, Lie's Leipzig colleague, Klein wrote: "The Lie who writes us and the Lie who presents himself to us personally are two different people" (Tobies/Rowe 1990, 206).

By this time, Klein had lost out to Schwarz, who was called to Berlin along with the algebraist Georg Frobenius. This seemed to Klein an opportune time to restructure the editorial board of *Mathematische Annalen*, in particular by

[10] For a brief account, see Stubhaug (2002, 386–389), which however fails to engage with many aspects of this conflict.

appointing Lie as an associate editor in order to solidify the Göttingen-Leipzig connection. This plan backfired, however, after Lie insisted on an appointment as one of the principal editors. Klein continued to hope that Lie would reconsider, but the Norwegian wanted to break away completely from what he saw as a web of intrigue within the German mathematical community. In late 1893, he vented his anger in a lengthy preface to the third volume of his monograph on transformation groups. There, he abruptly announced to the mathematical world that "[he] was not a pupil of Klein, ... if anything, the opposite was the case," and later he wrote Mayer how Klein reminded him of "an actress who dazzled everyone in her youth, but who used more and more dubious means in order to attain success on three-rate stages" (Tobies/Rowe 1990, 20).[11] Five years later, Lie returned to Norway, but shortly thereafter he died from pernicious anemia in 1899.

After his official retirement in 1913, Klein made plans to begin a series of lectures on nineteenth-century mathematics. He delivered these talks during the first years of the Great War and planned to end them with an account of the work of Lie and Poincaré. Then came the sudden burst of interest in Einstein's general theory of relativity, which took Hilbert and several other Göttingen mathematicians by storm. For Klein, who had already explored the mathematical underpinnings of Hermann Minkowski's space-time geometry, these fast-breaking developments offered a welcome opportunity to work on the historical and mathematical background of general relativity. By 1919, when he began discussing plans for his collected works, he had found the perfect capstone for contextualizing his "Erlangen Programm," namely his papers on the formal properties of conservation laws in general relativity, work he had taken up with the assistance of Emmy Noether and Hermann Vermeil (Rowe 2022a). Thus, the third section of Volume I, entitled "Zum Erlanger Programm," contains papers Klein wrote over a span of nearly fifty years.

During the period from 1869–1872, though, Klein's reputation had far more to do with his contributions to the second category of his research, which he called foundations of geometry. Here one can be far more precise: in 1871, he published a new approach to non-Euclidean geometry by adapting Arthur Cayley's method for introducing a metric in projective geometry (Gray 2011). Although his original intention was largely technical rather than foundational, he soon became mired in problems that Karl Christian von Staudt had attempted to solve earlier by way of introducing purely projective coordinates in plane geometry.[12] Klein is often credited with having been the first to realize that Staudt's foundational approach depended on Desargues' theorem, an elementary incidence theorem in 3-space that is difficult to prove in a plane geometry without appealing to an embedding in space (otherwise it could even be false, as Hilbert later showed).

[11] Lie was reacting to Klein's interest in launching the *Encyklopädie der mathematischen Wissenschaften*, but he likely also had in mind his recently published *Evanston Colloquium Lectures*.

[12] Detailed accounts of Staudt's work as well as Klein's part in the later developments can be found in two French studies, Nabonnand (2008) and Voelke (2008).

Since Klein, in his youth, had only a meager knowledge of the mathematical literature, he only later came to recognize that this was not an original insight. So in his commentary for the collected works, he indicated that Desargues had been aware of this fact and that Möbius had even emphasized the point (Klein 1921–1923, I:310). He also recalled that:

> [it was] from O. Stolz ... [that] I first heard of non-Euclidean geometry, and I immediately understood that this would have to be quite closely related to Cayley's general projective metric. ... I lectured on the Cayley metric in Weierstrass's final seminar and concluded outright with the question, whether there could be a connection here with non-Euclidean geometry. Weierstrass objected to this, declaring that geometry had to be based on the distance between two points, which meant the straight line needed to be defined as the shortest curve connecting these. (Klein 1921–1923, I:50–51)

This passage has often been cited in the historical literature, but without any critical reflection with regard to the chronology or the bearing these circumstances had on Klein's first publications on non-Euclidean geometry, i.e. Klein (1871b) and the more extensive account in Klein (1871c). Readers of these works, when they first appeared, included Klein's friend Max Noether, with whom he exchanged several interesting letters I will discuss on another occasion. These exchanges and other correspondence from the period provide a vivid picture of Klein's limited knowledge, but also the improvised character of his papers on projective non-Euclidean geometry. As I will show in a subsequent study, a careful reappraisal of the events of that time can shed new light on this important period in the history of geometry.

Noether probably understood the circumstances involved, especially since Klein spent several days with him in Heidelberg soon after August 1871, when he completed his first two papers. Other readers, however, surely gained a faulty impression from these texts, which contain a number of precise references to the literature. Many must have imagined that the author had studied these earlier works carefully, which was not at all the case. In an unpublished manuscript from November 1892 describing his and Lie's work from the years 1870–1872, Klein admitted that his papers on non-Euclidean geometry were based on hearsay information and that this had led him to give incomplete or inaccurate citations of the relevant literature (Klein 1892). Paging through Klein's published work, on the other hand, one finds no traces of any acknowledgment that Otto Stolz had played a decisive role. Even in his lecture course on non-Euclidean geometry from 1889/1890, Klein made no mention of Stolz whatsoever (Klein 1893). This might seem all the more surprising considering that several lectures dealt with the history of the subject, including lengthy remarks on his own publications.

Klein's reticence ca. 1890 stands in sharp contrast to the tone and content of his wartime lectures, which were semi-autobiographical in character. There, he vividly described how Stolz helped him grasp the projective ideas developed by Karl von Staudt, but also to enter into the world of non-Euclidean geometry (Klein 1926, 133, 152). He was even more explicit in his commentary to Volume I, where he wrote:

> I then received a substantial new impulse in the summer semester of 1871 when O. Stolz came to Göttingen. The ideas on the connection between non-Euclidean geometry and

> Cayley's metric, which were already evolving since my time in Berlin, now entered the foreground, and I succeeded in convincing him not only that these were correct, but also that the entire theory could be set out independently by means of v. Staudt's principles. Stolz served all the time not only as my strict critic, but also as my literary support. He had studied Lobachevsky, Joh. Bolyai and v. Staudt carefully, something I could never force myself to do, and he was able to answer all my questions. (Klein 1921–1923, I:51–52)

The fact that Klein's famous papers on non-Euclidean geometry were written without any direct knowledge of the works he cited by Lobachevsky and Bolyai might well have raised a few eyebrows, had this been known at the time. Even more striking, though, was Klein's admission that he only knew about Karl von Staudt's work from conversations with Otto Stolz, who had already left Göttingen by August 1871 when Klein submitted his first note (Klein 1871b) and the longer paper (Klein 1871c). The relevant background events, and in particular Klein's letters to Noether, offer ample opportunity to read those texts anew, as their authoritative ring sounds much weaker when placed against the explanations he gave to his friends.

Furthermore, in view of the intense interest in non-Euclidean geometry at that time,[13] one should not overlook Klein's eagerness to discuss the historical background. In particular, he highlighted the role of Carl Friedrich Gauss, whom he described as the first to realize that Euclid's parallel postulate could not be proved. Why not? Precisely because Gauss recognized that one could erect a "non-Euclidean geometry" in which the sum of the angles in a triangle was less than two right angles (Klein 1921–1923, I:245). Considering how little was known in 1871 about Gauss's work on non-Euclidean geometry, Klein's remarks in this vein appear quite remarkable. Clearly, he had in mind to place his own work in the best possible light, but this also shows how eager Klein was to speak with authority on major topics in the historiography of nineteenth-century mathematics. Let me now turn to the central arena for Klein's early research: line geometry.

6.3 Plücker's *Neue Geometrie*

Klein was still a teenager when he learned about this new approach to the geometry of 3-space at the shoulder of Julius Plücker. In the months before his death in May 1868, Plücker realized he would not be able to complete his monograph on line geometry. He provided Klein with drafts for some parts, entrusting Alfred Clebsch with the publication of Part I (Clebsch 1871). Klein was then charged with the task of completing Part II, which was published the next year in Plücker (1869). This circumstance led to Klein's first contacts with Clebsch, who had only just arrived in Göttingen, having been appointed to Riemann's long vacant chair. The following year, Clebsch and Leipzig's Carl Neumann launched their new journal, *Mathematische Annalen*, which almost from its inception showcased some of the

[13] On the early reception of non-Euclidean geometry in Germany, see Volkert (2013).

works of Klein and his Norwegian friend, Sophus Lie. Both had already established relationships with Clebsch before they met one another in Berlin in the fall of 1869, and as their subsequent collaboration unfolded Klein continually spoke with Clebsch about Lie's work as well as his own. In a word, both were leading protégés of Clebsch, though it was Klein who reaped the immediate benefits from this.

Clebsch had studied under Otto Hesse in Königsberg, an experience that strongly shaped his approach to geometry. Although he was only a decade younger than Plücker, Hesse's analytic style made use of powerful algebraic methods that made the works of earlier writers appear primitive by comparison. Clebsch's reaction to Part I of *Neue Geometrie des Raumes* suggests how he likely felt about all of Plücker's work. Writing to Wilhelm Fiedler on 20 August 1868, he praised the ideas but found the presentation more than wanting.[14]

Although little appreciated in Germany, Julius Plücker enjoyed an excellent reputation in England, especially for his work as an experimental physicist.[15] In 1855 he became a corresponding member of the London Royal Society, which awarded him its Copley Medal in 1866. During the 1850s, Plücker undertook pioneering research on electrical discharges in rarefied gases. Much of the time he was assisted by technicians, including Heinrich Geissler, famous for his invention of the glass tubes that bear his name. This work was closely followed by Michael Faraday, another physicist who thought more in terms of pictures than mathematical formulas.

By the mid-1860s, thus shortly before Klein arrived in Bonn, this phase in Plücker's career was over. He returned to the study of geometry, having been "encouraged by the friendly interest expressed by English geometricians [presumably Cayley, Hirst and Sylvester]" (quoted from Barrow-Green (2021, 44)). Still, there was an indirect connection between his geometrical and physical studies, as one can easily see from his paper "On a New Geometry of Space" (Plücker 1865), in which he pointed to the relevance of line geometry for optics and mechanics. In both of these fields, in fact, as well as in geometry, Plücker attempted to describe complex, never-before-seen spatial phenomena (Clebsch 1871). Felix Klein later recalled how his mentor told him Faraday was the one who urged him to build geometrical models (Klein 1921–1923, II:7). It seems quite unlikely, though, that this was in reference to so-called complex surfaces, which Plücker unveiled as the centerpiece of his new line geometry only a short time before Faraday's death.

As an old-fashioned analytic geometer, Plücker filled his publications with complicated calculations, but from these he extracted new geometrical insights that can be understood quite easily. Since the lines in 3-space form a 4-manifold, one needs at least four parameters to describe them. An elegant system for doing this uses Plückerian coordinates, which assigns six homogeneous coordinates to each line ℓ. One can obtain these either in the form p_{ij}, by taking any two points on

[14] "Haben Sie das Plückersche Werk gesehen, welches unter meinen Auspicien in die Welt gegangen ist? Schöne Gedanken, aber welche Darstellung!" (Confalonieri 2019, 73).

[15] On his family background and youth, see Wiescher (2016).

ℓ, or as r_{ij} by choosing any two planes through ℓ. In the first case, two points $P = (x_1, x_2, x_3, x_4)$ and $Q = (y_1, y_2, y_3, y_4)$ on ℓ lead to six determinants, which define $p_{ij} = x_i y_j - y_i x_j$. These homogeneous coordinates then satisfy the identity $P = p_{12}p_{34} + p_{13}p_{42} + p_{14}p_{23} = 0$, which accounts for why the p_{ij} can be used to coordinatize the 4-manifold of lines in 3-space. In his dissertation, Klein introduced a new system of (Kleinian) coordinates x_i, $i = 1, \ldots, 6$, for which the identity $P = 0$ becomes $\sum x_i^2 = 0$.

A good deal of the classical theory can be developed projectively, so for that purpose one can introduce general homogeneous line coordinates in P^5. These then satisfy a quadratic equation, as for example $P = 0$, so that the lines in space correspond to points on a 4-dimensional quadric hypersurface in projective 5-space. In analogy with projective algebraic geometry, one can then study objects given as the loci of lines that satisfy a homogeneous algebraic equation in line coordinates or various subfamilies of lines given by systems of algebraic equations. Thus, a single algebraic equation corresponds to a 3-parameter family of lines, a line complex, whereas two equations determine a congruence of lines. Such congruences had been studied earlier in geometrical optics as ray systems, which one can transform via reflection and refraction. Line congruences typically envelope surfaces, which in optics are so-called caustics. The standard objects in Plücker's line geometry, on the other hand, were first- and second-degree line complexes, which can only be visualized locally, except for degenerate cases.

A simple, yet instructive degenerate case of a quadratic complex K_2 is given by the lines tangent to a nonsingular quadric surface F_2, for example an ellipsoid. Here the local structure is immediately obvious: for a given point $P \notin F_2$, the lines through P tangent to F_2 form a quadratic cone (real or imaginary). This cone, however, collapses to a (double) tangent plane whenever $P \in F_2$. Since this is a projective theory, we can also consider the dual situation in space by taking arbitrary planes rather than points. A typical plane π cuts out a conic, $\pi \cap F_2 = C_2$, which has tangents belonging to K_2. Moving the plane toward its surface, the conics shrink until they eventually collapse into a pencil of lines (counted twice) precisely when π becomes a tangent plane T_P at a point $P \in F_2$. In the case of a general quadratic complex, the line conics degenerate into two pencils of lines in π, but here these fall together at the point of tangency $P \in F_2$, viewed as the envelope of its tangent planes. This special complex K_2 depends, in fact, on only 9 parameters, whereas a general quadratic complex has 19, which suggests why viewing the lines in quadratic complexes can be very difficult.

Let us now briefly consider the latter case of a general K_2. As noted, when the cone of lines passing through a point P collapses it forms two planar pencils of lines that intersect along a *singular line*. It can then be shown that these singular lines form a congruence of the fourth order and class[16] that envelopes the *singularity surface* S_4 of the complex K_2. Thus, taking an arbitrary plane π with a conic of

[16] The notions of order and class refer here respectively to the number of singular lines passing through a generic point and the number that lie in a generic plane.

lines in K_2, four of these will be singular lines. Furthermore, an arbitrary line ℓ will intersect S_4 in four points and four planes in the pencil passing through ℓ will be tangents to the surface. In the special case just considered, the surface S_4 is simply the quadric F_2 counted twice. One finds mention of this example at the very end of Section 2 in Klein's edition of Plücker's *Neue Geometrie*. There he emphasizes that the theory of quadratic complexes represents a generalization of the theory of quadric surfaces in that the latter corresponds to the case where all the complex lines are singular (Plücker 1869, 336).

Plücker laid out only part of the groundwork for this aspect of the theory, while noting that in the generic case S_4 is self-dual. Thus, one can view it either as the locus of singular points or as the envelope of singular planes of K_2. Plücker described S_4 as a surface of the fourth order and class with 16 double points and 16 double planes (Plücker 1869, 307–321). It seems hard to believe that he was fully unaware of Ernst Eduard Kummer's publications from the mid-1860s devoted to these now famous quartics (Kummer 1975, 418–439), but he made no mention of Kummer's work. Klein wrote frankly about Plücker's longstanding frictions with mathematicians and physicists in Berlin (Klein 1926, 120–121), but he apparently never commented about this curious circumstance. Yet, just a year after his mentor's death, he wrote about Kummer surfaces by name in Klein (1870, 67), an indication that geometers had by then come to adopt this terminology. As we will discuss below, their properties were crucial for Klein's radically new approach to quadratic complexes. We will also describe how Klein focused considerable attention on the properties of singular lines in a general quadratic line complex, a topic Plücker only touched on.

Plücker was mainly interested in the various shapes of special types of quartics he called complex surfaces (Plücker 1869, 337–373). These exotic objects reflected the local structure of a quadratic complex K_2 in relation to a fixed line g (in general, $g \notin K_2$). More precisely, Plücker considered the lines in $K_2 \cap K_1(g)$, where $K_1(g)$ is the first-degree complex consisting of all lines that intersect g. The intersection of two algebraic line complexes yields a 2-parameter family of lines, or a line congruence. In this case the congruence is of the second order and class since two lines lie in an arbitrary plane and two pass through a generic point.

Such systems were familiar from earlier work in ray optics, the background for Kummer's work in the 1860s. Thus it was known that a congruence of lines will envelope a caustic surface (*Brennfläche*), and those studied by Kummer lead to surfaces of the fourth order and class but without a double line g. The double line on a complex surface contains four point singularities, which are pinch points where the leaves of the surface join. They also demarcate the boundaries between real and imaginary portions of the surface, as illustrated in Fig. 6.1.

Plücker designed some 27 models showing different types of complex surfaces. These were originally built by one of his former students, Johannes Epkens; later the commercial firm of Eigel in Cologne built sets of these using a heavy metal, zinc or lead. Many of these models can still be seen in collections throughout Germany, such as the one shown in Fig. 6.2. A particularly well-preserved collection can be seen at Tübingen University (Seidl et al. 2018, 177–185).

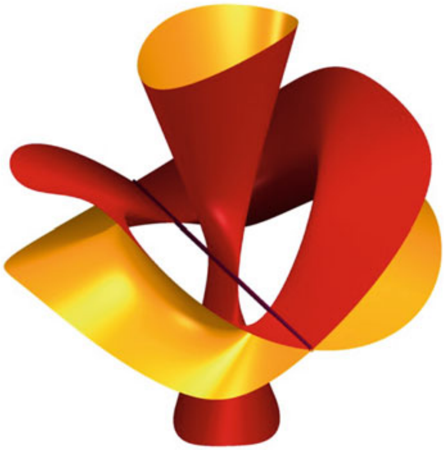


Fig. 6.1 A Plücker complex surface showing four pinch points on its double line. Graphic courtesy of Oliver Labs

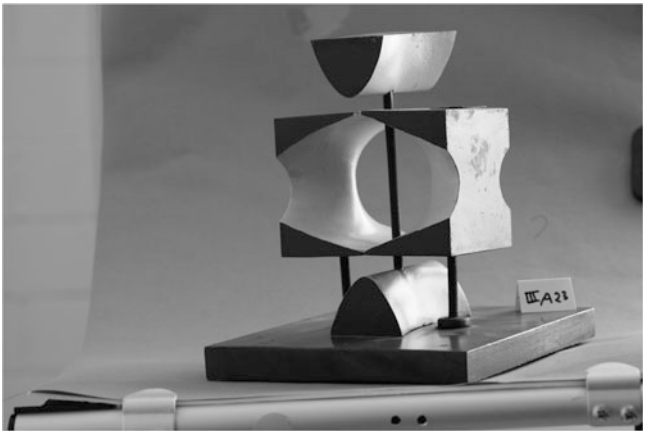


Fig. 6.2 A Plücker complex surface with two real and two imaginary components intersecting along its central axis. From the TU Munich; photo courtesy of Gerd Fischer

Plücker taught exercises in his new line geometry during the summer semester of 1867, when Klein presumably first learned about this new approach to spatial geometry (Tobies 2021, 33). He probably had little or no involvement in designing these models, but he must have studied them carefully during his last semesters in Bonn. During 1919, when Ostrowski was helping him prepare the first volume of his collected works, Klein offered a seminar during which he went through the collection of Plücker models in Göttingen.[17]

[17] A student who attended wrote up a detailed report on these models in the Göttingen collection (see modellsammlung.uni-goettingen.de).

Fig. 6.3 Plücker's Model of a Complex Surface with 8 Real Nodes. Courtesy of the London Mathematical Society

In 1866, Plücker took some models with him to Nottingham, where he spoke about these complex surfaces at a conference. Arthur Cayley and other leading British mathematicians quickly took an interest in them, as they were certainly among the more exotic geometrical objects of their day (Cayley 1871). Plücker sent a smaller set of boxwood models to the London Mathematical Society, one of which is pictured in Fig. 6.3. For most of the models, Plücker chose the line $g \subset E_\infty$, the plane at infinity. The family of planes $E(g)$ that contain the nodal line g are then parallel, and in this case he called the quartic S_4 enveloped by the lines in $K_2 \cap K_1(g)$ an *equatorial surface*. Since $E(g) \cap S_4$ is a family of quartic curves each containing g as a double line, these curves break up into a family of conics C_2 together with g (Rowe 2022b).

Luigi Cremona also took much interest in Plücker's works in geometry from the final years of the latter's life (see their correspondence in Israel (1992, 154–157) and Israel (1994, 73–77)). On 30 December 1868, some seven months after Plücker's death, Klein wrote to Cremona to send him a copy of his recently completely dissertation on the classification of second-degree line complexes. Klein mentioned that Plücker had spoken often of his trip to northern Italy in the fall of 1867 and had told Klein that Cremona was the only geometer who understood him completely.[18]

Six months later, having completed the second-part of Plücker's *Neue Geometrie des Raumes*, Klein wrote again to send Cremona a copy. His letter is highly suggestive for understanding what Klein himself found most interesting about this

[18] Israel (1994, 54); Klein was also eager to learn Cremona's opinion of his work, especially in view of the fact that he had shown that Battaglini's recent study (Battaglini 1868) did not represent the general case of a quadratic line complex.

work, and his comments very likely reflect Plücker's own views as well. He called attention to the final part, which presented a kind of appendix on complex surfaces:

> Apart from the theory of complexes, this section seems to me to be interesting in so far as it treats a manifold family of surfaces in such a way that the different *forms* that occur are made evident. It seems always to me, and this I take to be the sense of the method employed by Plücker, who had models made of the surfaces discussed here—that in the case of geometric problems, it is important not only to express the relationships between the structures treated by means of propositions, but also by an immediate intuition (*Anschauung*) of these structures (Israel 1994, 55).

This was a typically Kleinian pronouncement, though one should not overlook that it was written by a novice, who still imagined himself on the way to becoming a physicist. Klein's broader mathematical interests would gradually come to the surface in his "Erlangen Programm," but what he expresses here reflects his keen interest in concrete geometrical relationships in 3-space. As I will emphasize below, this was one of the most striking features in Klein's early work, an aspect far less apparent in Lie's writings, which are notoriously difficult to read.

In his preface to Plücker (1869, iii–iv), Klein pointed out that Plücker had already discussed complex surfaces in Plücker (1868, 177–225). Moreover, before his death he had also completed the manuscript for the final section of Plücker (1869, 337–373), which deals with the construction and properties of so-called equatorial surfaces, where the double line g lies in the plane at infinity. Klein probably came to realize quickly that all the Plücker quartics are special cases of Kummer surfaces, an idea that led him to consider whether they could be obtained from the latter by a continuous deformation process. Before long, he and his friend Alfred Wenker began building models designed to illustrate the main types. In a more theoretical sense, Klein would continue to address the problem of constructing a quadratic complex, though he shifted the focus to its singularity surface, which in effect was the quartic that arises by dissolving g into two lines each of which joins a pair of singular points.

The contrast between Plücker (1869), which Klein completed in May 1869, and Klein (1870), the paper he submitted just one month later, is truly astonishing. Only a year after Plücker's death, Klein totally transformed the theory of quadratic complexes and in the process forged a deep connection with the theory of Kummer surfaces. Part of his motivation can be traced to the topic of his dissertation, as he explained in his later commentary (Klein 1921–1923, I:3). Here, Clebsch gave him a decisive hint by pointing to a recent study by Giuseppe Battaglini, who studied what later were called harmonic complexes (Battaglini 1868, 241).[19]

In Plückerian coordinates, these take the form $\Omega = \sum k_s p_{ij}^2 = 0,\ i, j = 1, 2, 3, 4$, but Battaglini apparently did not realize that such a harmonic complex Ω depends on 17 parameters rather than 19, the number in a general quadratic complex

[19] In 1870, Ferdinando Aschieri showed that the Battaglini complex could be viewed geometrically as the family of lines that intersect two quadric surfaces harmonically, thus, in four points with cross ratio equal to -1; see Rowe (2016, 246).

(see Hudson (1990, 94–97)). In his dissertation, Klein presented a canonical form for the general case (Klein 1921–1923, I:5–49), but he also showed that one could introduce new coordinates x_i, $i = 1, \ldots, 6$ that made it possible to diagonalize both forms P and Ω so that the two equations for a general complex could be written $\sum k_i x_i^2 = 0$ and $\sum x_i^2 = 0$.

The algebra behind the transition from Plückerian to Kleinian coordinates was straightforward enough, but the underlying geometry behind the latter was subtle and elegant. The six equations $x_i = 0$ represent six fundamental mutually apolar[20] linear complexes, in relation to which the lines in space fall into groups of 32 lines. This follows immediately from the observation that any line $(x_1, x_2, \ldots, x_6)$ must satisfy $\sum x_i^2 = 0$, which means that the same holds for these 32 lines $(\pm x_1, \pm x_2, \ldots, \pm x_6)$. By introducing these coordinates, Klein could view the six fundamental complexes $x_i = 0$ as hyperplanes in P^5. A general linear complex A was then given by $\sum a_i x_i = 0$, which can be taken as a space element with coordinates $(a_1, a_2, a_3, a_4, a_5, a_6)$. The lines $(x_i) \in A$ can thus be expressed by the equation $(x_i) \cdot (a_i) = 0$. A special case arises when the coordinates of A also satisfy $\sum a_i^2 = 0$, which means that A consists of the lines in space that meet a fixed line. The corresponding notion of a special quadratic complex arises when one takes the lines in space that meet a fixed conic, an idea Lie exploited when he began thinking about his line-to-sphere transformation (see below).

As in the usual projective theory for quadratic surfaces, one can derive an equation for the linear tangent complexes to a given quadratic complex simply by calculating its polar form. In the present case, where K_2 is given by $\sum k_i x_i^2 = 0$, this yields $\sum k_i x_i = 0$. Now if $(x_i) \in K_2$ happens to be a singular line, this tangent complex will be special, which means that $(k_i x_i)$ are line coordinates and thus have to satisfy the condition $\sum k_i^2 x_i^2 = 0$. Taking the two lines (x_i), $(k_i x_i)$ together they determine a singular point and plane of K_2, but also a pencil of lines $y_s = k_s x_s - \mu x_s$. Plugging into the equations for (x_s), we obtain

$$\sum (k_s - \mu)^{-1} y_s^2 = 0, \qquad \sum (k_s - \mu)^{-2} y_s^2 = 0.$$

The lines (y_s) and $((k_s - \mu)^{-1} y_s)$ likewise determine a singular point and plane of K_2, which means that by varying the parameter μ in the equation $\sum (k_s - \mu)^{-1} y_s^2 = 0$ Klein obtained a co-singular family of quadratic complexes $K_2(\mu)$, all of which have the same singular points and tangent planes as K_2. (The equation for K_2, $\sum k_i x_i^2 = 0$, corresponds to $\mu = \infty$.) He showed further that the surface S_4 is a self-dual quartic, since for any line ℓ the cross ratio of the four points $\ell \cap S_4$ equals the cross ratio of the four planes through ℓ tangent to S_4.

From these facts, Klein could deduce that S_4 had the same structure of singularities as a Kummer surface. These surfaces have 16 singular points (called nodes) and 16 singular planes (tropes), which form a (16, 6) incidence configuration (Hudson

[20] Klein used the terminology of complexes that lie in involution with one another, but Hudson found this language awkward and I here follow Hudson (1990, 38).

1990, 7–14). This means that 6 tropes pass through each of the 16 nodes and 6 nodes lie in each of the 16 tropes. They lie in special position, since 6 co-planar nodes lie on a conic in the corresponding trope and a conic curve is determined by 5 points in general position. The tangent plane T_P at a generic point $P \in S_4$ contains just one singular line, though there are exceptional points. Since S_4 is of the fourth order and class, a generic line ℓ intersects S_4 in four points P_1, P_2, P_3, P_4 and four planes $\pi_1, \pi_2, \pi_3, \pi_4$ passing through ℓ are tangents to S_4.

When ℓ is a singular line, two points will coincide, say, $P_1 = P_2 = P$. Then ℓ will also lie in a singular plane, say $\pi_1 = \pi_2 = \pi$, and two other tangent planes π_3, π_4, which are spanned by the pencils (P_3, π_3), (P_4, π_4). Moreover, ℓ is the *only* line in the pencil (P, π) that lies in K_2, unless either P_3 or P_4 happens to equal P, in which case ℓ is an inflectional tangent to S_4 and then *every* line in the pencil (P, π) lies in K_2. Klein proved that these exceptional points lie on a curve of 16th degree on S_4 (see below). He also considered the bitangents to S_4, which arise when pairs of points, e.g. $P_1 = P_2 = P$ and $P_3 = P_4 = Q$, fall together. A plane section $S_4 \cap \pi = C_4$ will have 28 bitangents, 16 of which arise as intersection lines between π and the 16 tropes. The remaining 12 form 6 pairs that fall into six different congruences, as was shown in 1866 by Kummer.

As noted above, this is a projective theory,[21] which can be seen by considering a pencil of planes Λ_ℓ that pass through an arbitrary line $\ell \in K_2$, a quadratic complex. Each $\lambda \in \Lambda_\ell$ contains a conic of lines C_2 tangent to ℓ at a point P_λ. As the planes in Λ_ℓ turn they generate a projective correspondence between these λ and the points P_λ. This will be the case for all complex lines ℓ except for those which are singular. In the latter case, the conics in each λ are tangent to ℓ at the singular point $P \in \ell$. This entire theory dualizes, so S_4 is identical as a projective surface, whether viewed as a locus of points or an envelope of planes (Plücker 1869, 315).

Klein's system of coordinates took full advantage of these various symmetries, which were not at all evident following Plücker's approach. Nevertheless, Klein's mentor apparently found many striking properties of quadratic complexes, including their singularity surfaces, quite possibly without having read Kummer's earlier work. Plücker described, for example, how singular lines arise in an arbitrary plane π, which intersects the singular surface in a quartic curve, namely $C_4 = S_4 \cap \pi$. The conic C_2 of complex lines in π then has four points of tangency with C_4, so counting these twice, $C_2 \cap C_4 = 2 \cdot 4 = 8$ points, and their four tangent lines are the singular lines in π.

Klein only presented a counting argument for the degree of the curve of exceptional points $P \in S_4$, namely those where the tangent plane $T_P = \pi$ has an inflectional tangent as singular line (the case discussed above). Taking an arbitrary plane Λ that cuts S_4 in a quartic curve C_4, he reasoned as follows that this plane will contain 16 such points. As noted above, the quartic C_4 and the conic C_2 of complex lines in Λ touch in four points that yield the four singular lines in the chosen plane.

[21] Actually, Plücker freely mixed projective and metric concepts, whereas Klein belonged to a younger generation of geometers who paid careful attention to this distinction.

Klein then noted that there are 16 other lines in Λ that are tangents to both curves at distinct points. This follows from the counterpart to Bezout's theorem for class curves (those arising as envelopes of lines), which states that curves of class m and n will, in general, have $m \cdot n$ common tangents, taking multiplicities into account. Now a curve of degree n in point coordinates will, in general, be of class $n(n-1)$, which means that the conic and quartic are curves of class 2 and 12, respectively, so they have 24 tangents in common. Of these, 8 correspond to the four singular lines, whereas the other 16 are tangents at distinct points. Klein then claims that the 16 points of tangency on C_4 are points where the tangent plane contains a singular inflection line.

To see this, take any one of the 16 tangency points P on the curve $C_4 \subset S_4$. Consider now what happens to the tangent line $\ell_P \subset \Lambda$ when P is held fixed and the plane Λ gradually moves until it coincides with the tangent plane T_P. The line $\ell_P \subset \Lambda$ constrains the movement of the planes to rotate about P, whereas throughout this process the complex conics in each plane continue to touch the corresponding lines starting from a point Q on ℓ_P. In effect, the point Q on the conic slides along the transformed lines until it reaches the point P once the plane coincides with the tangent plane T_P. This straightforward continuity argument then shows how the line ℓ_P passes over to another line in 3-fold contact with S_4 in the plane T_P, thereby becoming an inflexion (or asymptotic) tangent of the singularity surface. Since this argument holds for any of the 16 points in an arbitrary plane, this shows that the asymptotic singular tangents on a Kummer surface lie on a space curve of degree 16.

This was only one of several new results Klein found while working on Plücker (1869, 315) and then compiled in Klein (1870). Yet the significance of this particular finding did not dawn on him until later, as at the time he evidently gave no thought as to how this might be used to investigate the differential geometry of Kummer surfaces. Instead he apparently took this to be an isolated result concerning the singular lines of a quadratic line complex, even though he had come to realize by the summer of 1869 that the entire theory of such complexes could be developed in analogy with families of confocal quadrics (Klein 1870, 79).

In fact, the geometrical ideas behind this approach already appeared in this first paper. There he showed that one could start with a Kummer surface S_4 and a single tangent line ℓ, designated as a singular line in a complex $K_2 \in K(S_4)$, the family of quadratic complexes with the singularity surface S_4. These data then suffice for constructing K_2 uniquely, whereas there will be four complexes in $K(S_4)$ when ℓ is not a tangent to S_4. These observations led Klein to conclude that $K(S_4)$ forms a 1-parameter system of co-singular quadratic complexes in analogy with the theory of confocal quadratic surfaces (Klein 1870, 73). Only after meeting Lie in Berlin did he begin to exploit this approach, however, by using elliptic coordinates as Jacobi had done in studying the differential geometry of confocal quadrics (see the section below). In fact, Klein's motivation to do so came in good part from an independent breakthrough made by Lie during their sojourn in Paris: his recognition that the asymptotic curves on Kummer surfaces were algebraic of degree 16.

Lie's approach to line geometry was closely linked with his interest in differential equations in the tradition of Gaspard Monge, whereas Klein's outlook was shaped by algebra and invariant theory. Klein tried to grapple with Lie's ideas, but he often had difficulty understanding the brilliant Norwegian. Nevertheless, or perhaps even because of their differences in outlook and intellectual maturity, their collaboration proved highly fruitful, even though it lasted only three years. Klein had the opportunity to interact with Lie during the months when the latter's line-to-sphere transformation—or some provisional version of it—was still in a nascent state. Lie only gradually came to recognize the general property that he later in 1871 highlighted in the introduction to his dissertation (Lie 1934, 105–106). Before making this discovery, he spent several weeks groping in the dark.

A glimpse of the various ideas then racing through Lie's mind can be seen from a brief note he sent to the Scientific Society in Christiania on 5 July 1870 (Lie 1934, 86–87). In it he announced seven new results, beginning with his line-to-sphere mapping T_{Lie} for complex projective 3-space. As an important example of a contact transformation, T_{Lie} maps intersecting lines to tangent spheres. Second, it induces a correspondence between surfaces—as envelopes of lines, resp. spheres—taking the asymptotic curves of the first to the lines of curvature of the second. Third, as a special case, Lie could show that the asymptotic curves of Kummer surfaces are of degree 16, which meant the same was true for Plücker's complex surfaces and the Fresnel wave surface. Fourth, in the case of ruled surfaces he could determine their asymptotic surfaces by means of quadrature. Lie's other three findings involved a new class of minimal surfaces, but he only began to publish on this theory in the late 1870s (for the larger context behind all this work, see Lie and Scheffers (1896)).

Klein was not simply an outside observer during this early breakthrough; in fact, he had already in the summer of 1869 identified these 16th-degree curves on Kummer surfaces (see above), but without realizing that they were the asymptotic curves (Klein 1870, 73–74). He was thus already deeply immersed in this realm of possibilities even before he first met Lie. As described in Rowe (2019), he had also undertaken a careful study of a correlative mapping between the lines in a linear complex (thus a 3-manifold) and the points in 3-space (Klein 1869). He learned about this mapping from Max Noether in 1869, when they were together in Göttingen (Klein 1921–1923, I:89), whereas Lie had already found essentially the same mapping on his own, only with a new twist. In Noether's formulation, this $T_{Noe} : K_1 \longrightarrow P^3$ has the property that for any $\ell \in K_1, T_{Noe}(\ell) = p \in P^3$, with one exceptional line ℓ^* for which $T_{Noe}(\ell^*) = C_2$, thus blowing up into a conic. Klein studied the images under T_{Noe} of "configurations" (ruled surfaces) contained within various congruences of lines in K_1, drawing on publications of Cayley, Cremona, and Schwarz. These are mapped to curves in P^3, and if the configuration consists of lines that meet ℓ^*, the image curve will intersect the conic C_2. Klein hoped to gain insights into the types of 4th- and 5th-degree ruled surfaces that can arise in different cases, but this investigation only led to a large number of special results without any apparent general theorem uniting them.

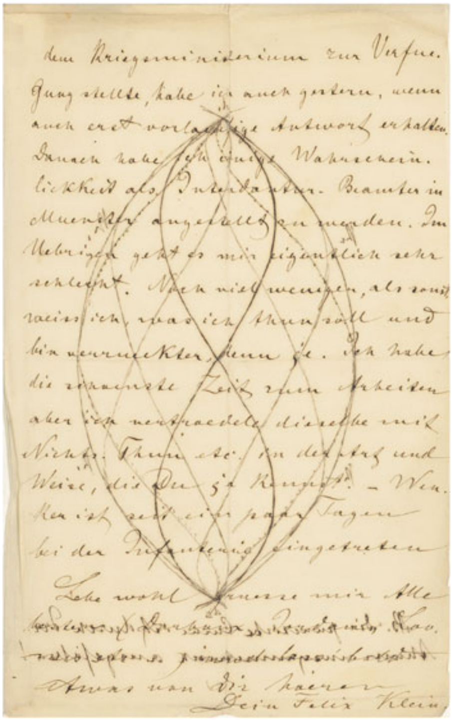
The simple twist that Lie had in mind was to view T_{Noe} as a mapping between two line complexes, $T_{Lie} : K_1 \longrightarrow K_2$, where K_2 is the special quadratic complex consisting of all lines that meet C_2 ($= T_{Noe}(\ell^*)$). Thus, as before $T_{Lie}(\ell) = p$, but now p generates the cone of lines $(p, C_2) \subset K_2$. As noted above, this was only one of many ideas swirling through Lie's mind at the time, but in Paris he began thinking about the special case where C_2 was the imaginary sphere-circle at infinity that lies on all spheres in space. The cones (p, C_2) in this setting can be viewed as spheres of radius zero; they consist of the minimal lines passing through a point p, or "crazy lines" as Lie liked to call them. He eventually came to realize that by viewing an arbitrary sphere S as an envelope of minimal lines, the points $p \in S$ will correspond to lines in a congruence $K(S) \subset K_1$. Such a congruence has two directrices, $g_1, g_2 \in K_1$, with the property that the lines $\ell \in K(S)$ intersect both directrices. Taking $a \in K(S) = \{a | a \cap g_1 \neq \emptyset,\ a \cap g_2 \neq \emptyset\}$, the mapping T_{Lie} is a bijection between the lines in $K(S)$ and the points on the sphere S. Moreover, since the congruence $K(S)$ is completely determined by the directrices g_1, g_2, we can view T_{Lie} as a 2-1 mapping between lines and spheres in space. By July 1870, this became the principal case of interest for Sophus Lie, namely his line-to-sphere transformation.

6.4 Klein and Lie on the Differential Geometry of Kummer Surfaces

Klein later described a key breakthrough Lie made in early July 1870, when both were staying at the same hotel in Paris (Klein 1921–1923, 1: 97). Lie had spent a sleepless night thinking about an induced mapping between two types of quartic surfaces: Kummer surfaces, which arise naturally in line geometry, and quartics with a double conic curve. A special case of the latter had been studied by Gaston Darboux and Théodore Moutard in the context of sphere geometry. When the double conic is the imaginary sphere-circle that lies on all spheres, Darboux showed that the lines of curvature of this quartic are algebraic curves of the 8th degree. Using his mapping, Lie realized these curves must go over into asymptotic curves of degree 16 on the corresponding Kummer surface.

Fifty years later, Klein recalled what Lie had told him that morning. He left the hotel that day to take in the collection of models at the Conservatoire des arts et métiers, all the while thinking about Lie's claim. Since Klein already knew that the singular lines with 3-fold contact on the singularity surface of a quadratic complex enveloped curves of degree 16 (as recounted above), he quickly realized that Lie must have found an important property of these same curves that had escaped him earlier.[22] When he returned to their hotel that afternoon, he saw that Lie had in the meantime gone out, so he left him a message letting him know that he was able to confirm Lie's finding. This unexpected result was surely a major topic in their

[22] Asymptotic curves on surfaces are those along which the tangent plane and osculating plane coincide; see Struik (1961, 96).



dem Kriegsministerium zur Verfügung stellte, habe ich auch gestern, wenn auch erst vorläufige Antwort erhalten. Danach habe ich einige Wahrscheinlichkeit als Intendantur-Beamter in Muenster angestellt zu werden. Im Uebrigen geht es mir eigentlich sehr schlecht. Noch viel weniger, als sonst weiss ich, was ich thun soll und bin unruhiger denn je. Ich habe die genaueste Zeit zum Arbeiten aber ich verbrachte dieselbe mit Nichts. Thun etc. In der Art und Weise, die Du ja kennst. – Wenker ist seit ein paar Tagen bei der Infanterie eingetreten.

Lebe wohl [illegible] mir Alle [illegible]

Etwas von Dir hören

Dein Felix Klein

Fig. 6.4 Klein's sketch of asymptotic curves between two double points on a Kummer surface from his letter to Lie, 29 July 1870 (National Library of Norway, Brevs. 289). CC BY-NCND 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/)

discussions afterward, though their talks were abruptly cut off before they could settle various questions about the properties of these asymptotic curves.

Little more than a week after Lie's breakthrough in early July, Klein had to leave Paris abruptly after the French government declared war on Prussia.[23] He headed back home to Düsseldorf, where he then stayed with his family until called to join an ambulance crew in Bonn. By early September, he and his fellow volunteers were helping care for the wounded at the battle sites near Metz and Sedan, though soon thereafter Klein fell ill and was released from further duty. While still at home, he tried to maintain contact with his Norwegian friend through letters. Unfortunately for Lie, he was carrying some of Klein's letters when French police saw him wandering about near Fontainebleau. They were written, of course, in German, and one contained a drawing that looked suspiciously like lines for troop formations (see Fig. 6.4 and try to imagine knowing that German armies were about to invade France).

[23] The incident that led to this declaration of war bears a striking resemblance to Putin's demand that Ukraine shall never be allowed to join NATO. France insisted that the Catholic branch of the Hollenzollerns, which had declined an offer to assume the throne of Spain, should declare that this decision was valid for all time. Bismarck famously took full advantage of this diplomatic blunder.

The *gendarmes* studied the letters and noticed certain words kept popping up—*Linienkomplex, Kugelkomplex*, etc.—perhaps some kind of code language? Maybe "komplex" meant agent? Lie's attempt to explain that these were mathematical terms only made them more suspicious. They concluded that either this fellow was crazy or he was really dangerous. Either way, they took no chances and locked him up on suspicion of being a German spy, and so Lie spent the next month in a jail cell. Darboux eventually learned of his fate and traveled to Fontainebleau, where he pleaded successfully for his release (Darboux 1899). Klein only learned about this later, as he and Lie were out of touch with one another until late October. Klein spent the months of October and November at home in Düsseldorf on the Rhine recovering from gastric fever.

Lie left France for Italy and wrote to Klein from Turin. After finally reestablishing contact, they began making plans for a meeting. On the final leg of his journey back to Norway in November 1870, Lie stopped to visit Klein at his family's home. During his 10-day stay, Lie discussed future plans with Klein, who was then preparing to habilitate under Clebsch in Göttingen in January. Shortly before Lie left, he and Klein wrote a letter to Kummer, in which they described their various adventures over the last months, including their latest findings with regard to Kummer surfaces.

Prior to his sudden departure from Paris the previous July, Klein had discussed various properties of the asymptotic curves on Kummer surfaces with Lie. As it turned out, however, Klein had been mistaken about some of these. A week or so later, back in Düsseldorf, Klein began to see why certain things he had claimed in Paris about the singularities of these curves were, in fact, wrong. His letter to Lie from 29 July was mainly written in order to alert his friend to these mistakes, which Klein now confidently felt he had corrected. What these mistakes were has little interest, but *how* Klein went about attacking this problem reveals a great deal about his visual approach to mathematical research. Evidently, he brought the models his friend Albert Wenker designed for him in Berlin back to his parent's home. One of these was a Kummer surface with 16 real nodes, precisely the type of model Klein needed in order to follow the course of the asymptotic curves (see Fig. 6.5). What the *gendarmes* perhaps took to be a sketch of "troop movements" in his letter to Lie were, in fact, asymptotic lines bouncing back and forth between two nodes, as Klein briefly explained:

> I came across these things by means of Wenker's model, on which I wanted to draw asymptotic curves. To give you a sort of intuitive idea of how such curves look, I enclose a sketch [Fig. 6.4]. The Kummer surface contains hyperboloid parts, like that drawn; these are bounded by two of the 16 conics ($K1$ and $K2$) and extend from one double point ($d1$) to another ($d2$). Two of the curves are drawn more boldly; these are the two that not only belong to linear complexes but also are curves with four-point contact. They pass through $d1$ and $d2$ readily, whereas the remaining curves have cusps there. This is also evident from the model. At the same time, one sees how $K1$ and $K2$ are true enveloping curves.[24]

[24] Klein to Lie, 29 July 1870, quoted from Rowe (2019, 190).

Fig. 6.5 Copy of the Model of a Kummer Surface designed by Klein in 1871 (Göttingen Collection of Mathematical Models and Instruments, model 95). Mathematical Institute, Georg-August-University, Göttingen

Thus, the drawing in Klein's letter shows some of the asymptotic curves that run between two nodes on a Kummer surface S_4. These are bounded by segments of two conics that pass through these nodes, which determine a hyperbolic region of S_4, one of 48 on the entire surface. Asymptotic curves are only visible in regions of negative curvature as they are otherwise imaginary curves. The two boldly drawn curves are of the 8th order and pass smoothly through the nodes, whereas the 16th-degree curves have cusps at them. This episode has obvious significance in revealing Klein's use of a model to discover or confirm properties of these asymptotic curves, but this was hardly the end of the story.

Soon after Lie had departed, Klein received a friendly reply from Kummer, written on the 26th of November and addressed to both Klein and Lie (Lie 1934, 667). Klein was elated to read Kummer's invitation, offering to submit their new results on the asymptotic curves on Kummer surfaces for publication by the Berlin Academy. This would be the first and last time that Klein's name would appear on a work published by that institution. He set to work right away, sketching a draft in seven points that he sent to Lie, who sent back some remarks that Klein worked into the final manuscript. He thus wrote up this paper for both of them (Klein 1921–1923, 1: 90–97), though in a presentation that underplayed Lie's original accomplishment, which first prompted Klein to study these curves.[25]

Kummer submitted the jointly authored note (Klein and Lie 1870/1884) to the Berlin Academy at its meeting on on December 15 1870. Klein had to compose the final version in great haste the day before, counting on the swift German mail service, which delivered the manuscript to Kummer's hand that morning, so a matter of hours before the academy convened. In a letter to Lie, written on 12 December, he noted that in the final section, which dealt with Lie's original finding, he was able to shorten the argument by invoking a general theorem (Lie 1934, 667). This asserted that when two surfaces touched along a curve so that at each point they have a common asymptotic tangent, then this is an asymptotic curve for both surfaces. Klein mentioned to Lie that he had not found time to prove this, an ironic remark given the circumstances. For over sixty years, this paper was reprinted and cited, but apparently no one ever commented about this claim. Finally, Poul Heegaard, in his commentary on the paper for volume 1 of Lie's collected works, described a simple counterexample that refuted Klein's assertion in the form he had stated it (Lie 1934, 672).[26]

Soon afterward, Klein, who had yet to celebrate his 22nd birthday, began his teaching career as a private lecturer in Göttingen. Clebsch had assured him that he would not have to submit a second dissertation (*Habilitationsschrift*) in order to qualify, but merely deliver a lecture before a small group of professors in the philosophical faculty. Klein's topic, interestingly enough, was a discussion of Plücker's complex surface, which gave him the opportunity to discuss one of the models Wenker had made for him (Tobies 2021, 89).

Only ten days later, Clebsch presented a new paper, Klein's eighth publication, to the Göttingen Scientific Society (Klein 1871a). In this short note, he showed how one can employ Jacobi's so-called elliptic coordinates to parameterize the lines in space relative to a Kummer surface S_4 and its corresponding cosingular family of quadratic complexes. The paper is a noteworthy example of how Klein combined *Anschauung* and heuristic techniques with sharp analytical methods. In spirit, this was much like Plücker's style, but undertaken during an era when more sophisticated tools had become available.

Klein wrote the equations for this family $K(\sigma)$ in the form

$$\sum \frac{x_i^2}{k_i + \sigma} = 0, \quad \sum x_i^2 = 0, \qquad i = 1, \ldots, 6.$$

Recall that a line (x_i) belongs to four complexes corresponding to the four values of σ determined by these equations. Denoting them by (x, y, z, t), Klein called these the elliptic coordinates of the line, which are distinct for a line in general position

[25] This was duly noted by Engel and Heegaard in their notes on this paper; see Lie (1934, 674).

[26] The claim is correct if one assumes the common asymptotic tangents are also tangents to the curve along which the surfaces touch.

relative to S_4. A singular line, on the other hand, together with its associated S_4, determines the parameters uniquely. This then means that two of the coordinates will be equal, say $z = t$; (x, y, z, z) are then tangents to S_4. If one fixes the coordinates x, y, then varying z yields tangent lines through a point $P(x, y) \in S_4$. When three coordinates are equal, the lines are asymptotic tangents; when they are equal in pairs, then they will either lie in one of the 16 tropes or pass through one of the 16 nodes. If all four are equal, then they are either tangents to one of the 16 conics in the tropes or are generators of one of the 16 cones with a node as vertex.

Klein noted further that one can use the coordinates to parameterize the respective complexes and congruences. Thus, fixing a single coordinate, say t, in place of the parameter σ yields a single complex. Fixing two coordinates, say z, t, then produces a congruence, and if $z = t$ this will be a congruence of singular lines tangent to S_4. The six k_i in the equation play a special role in connection with the six fundamental complexes $x_i = 0$. Thus, $z = t = -k_i$ yields one of the six congruences of double tangents to S_4, each of which lies in a $x_i = 0$. As noted, when $y = z = t$ the lines belong to three complexes, so they are generators of a quadric surface that determines an asymptotic curve on S_4. In case $y = z = t = -k_i$, the lines will have four-point contact with S_4, as Klein mentioned in his letter to Lie cited above. Fixing all four coordinates determines 32 associated lines that satisfy the condition $x_i^2 = 0$. Recalling that each plane Π contains four singular lines that are tangent to the quartic curve $\Pi \cap S_4$, the 32 lines that arise when $x = y = z = t$ are pairs of tangents to the 16 conics, which are double curves in the tropes of S_4. Klein then followed Jacobi's analytical argument to derive the differential equation for the asymptotic curves on S_4. This calculation leads to a very simple result, namely if $y = z = t$ are taken as the parameter σ, then the remaining coordinate x remains constant along the asymptotic curve.

Klein had none of these elegant analytic tools at his disposal when he first began trying to visualize how these curves wound around a Kummer surface. The letter he wrote to Lie in July 1870 shows that he got remarkably far by means of qualitative arguments based on general geometrical principles and certainly a good deal of direct experience with models dating back to his student days with Plücker in Bonn. Klein later reproduced nearly the same Fig. 6.4 in the note that he sent Kummer for publication in the Monatsberichte of the Prussian Academy. Afterward, this picture became a standard part of the growing literature on Kummer surfaces; see the discussion in Hudson (1990, 118–121) and Klein's text and commentary in Klein (1921–1923, I: 90–97). Klein had a special fondness for it, too. When five years later he married Anna Hegel, granddaughter of the famous philosopher, he ordered a wedding dress decorated with these arabesque curves (Tobies 2021, 84).

The paper described above, Klein (1871a), was one of several shorter works that Klein decided not to include in the first volume of his collected works, since he published its results as part of a longer paper the next year (Klein 1872). That omission was perhaps unfortunate because of what he wrote at the very end, after deriving the equation $x =$ constant for the envelope of lines that determine the asymptotic curves. Klein's brief remark hints at why he and Lie considered themselves geometers in the tradition of Plücker:

> My friend Lie and I derived this result [in Klein and Lie (1870/1884)] by geometric means. I must add that Lie was the first to recognize the integrability and algebraic nature of the asymptotic curves of Kummer surfaces by means of a method very different from the one set out here and only mentioned in passing in that paper. (Klein 1871a, 49)

6.5 Recommended Reading for Classical Geometry

The mathematical ideas set out in the last two sections fall within the general framework of classical geometry. Much of the vast primary literature from the nineteenth century devoted to that field was soon forgotten once more rigorous language and methods became standard. Another major obstacle for most readers today is that the geometers from that earlier era wrote mainly in German and Italian. The English language only took on prime importance over the course of the twentieth century. Yet, even a well-trained mathematician who can read these older papers in their original languages should not attempt to do so before first gaining a solid understanding of the basic conceptual tools and methods that mathematicians like Klein and Lie took for granted. Luckily, I can suggest some works (in English!) that make it possible to gain a general appreciation for this subject, but which can also suggest why a modern mathematician might be interested to know what line geometry was all about.

A key to understanding nearly all of Felix Klein's mathematics stems from what he and others called *Anschauliche Geometrie*. Although this approach was often attached to Klein's name and legacy, it was also taught with great success by Hilbert in the 1920s. He later enlisted Stefan Cohn-Vossen to prepare the expanded edition of his lectures that they published together as (Hilbert and Cohn-Vossen 1932). Probably at the prompting of Hilbert's former student Max Dehn, who spent his last years at Black Mountain College in the mountains of North Carolina, Dehn's student Peter Nemenyi (not his father Paul, as stated in Wikipedia) published the English translation under the title *Geometry and the Imagination* (Hilbert and Cohn-Vossen 1952/1999). The book is packed with over 300 diagrams, including images of several models, beginning with the simplest examples from line geometry, namely ruled quadric surfaces. Continuity and counting arguments, like those given above, abound, whereas formulas and calculations are held to a minimum. Although the informal presentation is meant to be accessible for those with little background in higher mathematics, its visual dimension has had great appeal for many mathematicians over the last ninety years.

A particularly useful guide to several geometrical topics from the nineteenth century is Dirk Jan Struik's *Lectures on Classical Differential Geometry* (Struik 1961), which was written somewhat in imitation of the informal approach of that era. As Jeremy Gray has often emphasized, this was the century that "invented rigor," especially in analysis. Classical differential geometry eventually became more rigorous, too, but mainly toward 1900, so Struik's book not only imparts a good deal of older geometrical knowledge it does so in the spirit of its time. Line geometry in the tradition of Plücker and Klein does not appear in the book, but many related topics do, some directly relevant to the discussions above.

An excellent way to gain an impression of Klein's work in line geometry is to pick up a copy of *Kummer's Quartic Surface* (Hudson 1990), which is still in print today. This monograph is more technically demanding than the other two books, but the topics it touches on provide an exceedingly clear idea of how a talented writer thought about geometry ca. 1900. The book was first published in 1905, one year after the death of the Cambridge geometer Ronald W.H.T. Hudson, who at age 28 fell while mountaineering in Wales. Much of Hudson's book was based on Klein (1870), Klein's first major paper on line geometry. As described above, that densely written *tour de force* literally rewrote the subject exactly one month after Klein had completed his work on Plücker (1869). Hudson cited Kummer's work, of course, and he noted further how to derive the special case of a Fresnel wave surface. These topics and much more are presented very elegantly in his book, which in some respects can be considered the last word on Kleinian line geometry.

The algebraic geometer Wolf Barth first came upon this classic in 1972 and soon fell in love with it, so he was very pleased when Cambridge University Press decided to reprint it in paperback. Barth's opinions about the book are scattered throughout his description of its contents, which he discusses chapter by chapter (Hudson 1990, xi–xxi). Part of what makes this Foreword so interesting is that Barth often writes as though the reader won't believe his or her eyes. On the one hand, Barth was in awe of Hudson's ability to explain complicated geometry but, on the other, shocked by his clumsy wordiness when it came to something as simple as a group action. The latter problem pops up in the very first chapter, where Hudson describes the configuration of singular points and planes of a Kummer surface (David Mumford dubbed this the Heisenberg group). In Chapter 4, Barth introduces the relevant concepts from line geometry, about which he writes: "the language is old-fashioned, but I have never seen a modern exposition of all this lovely geometry."

In Chapter 5, Hudson unveils the key discovery made by Klein, namely, that a Kummer surface is the common singularity surface for a 1-parameter family of quadratic line complexes. Then, in Chapter 6, he takes up the case of a Plücker complex surface (about which, see Section 6.3 above). In 1990, Barth wrote that he was convinced this book would be useful to many doing active research. Whether or not that actually proved to be the case, it can definitely be said that Hudson's account provides many helpful insights for understanding Klein's work on line geometry.

Acknowledgments Among his many other contributions to the history of mathematics over the course of a distinguished career, Jeremy Gray has also contributed a great deal toward enriching our understanding of major parts of Felix Klein's mathematical work. He gave prominent attention to Klein's papers on algebra and function theory from the late 1870s and early 1880s in *Linear Differential Equations and Group Theory from Riemann to Poincaré* (Gray 2008), which was first published in 1986. For those who read German, I would also recommend Erhard Scholz's history of the emergence of the concept of manifolds (Scholz 1980), as these two studies complement one another very nicely. More recently, Jeremy has made the case for the importance of Klein's contributions to Galois geometry (Gray 2018, 171–188; Gray 2019). Alongside these works, he also dared to write an intellectual biography of Henri Poincaré (Gray 2013), thereby filling one of

the yawning gaps in the literature. Having benefited from these works and many others over the course of my career, it gives me great pleasure to take part in this celebration of Jeremy's 75th birthday.

References

Barrow-Green, June. 2021. "Knowledge Gained by Experience": Olaus Henrici – Engineer, Geometer and Maker of Mathematical Models. *Historia Mathematica* 54: 41–76.

Battaglini, Giuseppe. 1868. Intorno ai sistemi di rette di secondo grado. *Giornale di Matematiche* 6 (1868): 239–283.

Biagioli, Mario. 1993. *Galileo, Courtier: The Practice of Science in the Culture of Absolutism*. Chicago: University of Chicago Press.

Biermann, Kurt-R. 1988. *Die Mathematik und ihre Dozenten an der Berliner Universität, 1810–1933*. Berlin: Akademie Verlag.

Confalonieri, Sara, Peter-Maximilian Schmidt, Klaus Volkert, Hrsg. 2019. *Der Briefwechsel von Wilhelm Fiedler mit Alfred Clebsch, Felix Klein und italienischen Mathematikern*. Siegen: Universitätsbibliothek der Universität Siegen.

Cayley, Arthur. 1871. On Plücker's Models of Certain Quartic Surfaces. *Proceedings of the London Mathematical Society* 3 (1871): 281–285.

Clebsch, Alfred. 1871. Zum Gedächtnis an Julius Plücker. *Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen* Band 16: 1–40.

Cogliati, Alberto, ed. 2019. *Serva di Due Padroni: Saggi di Storia della Matematica in onore di Umberto Bottazzini*. Milano: Egea, 2015.

Darboux, Gaston. 1899. Sophus Lie. *Bulletin of the American Mathematical Society* 5 (7): 367–370.

Gray, Jeremy. 2008. *Linear Differential Equations and Group Theory from Riemann to Poincaré*. 2nd ed. Basel: Birkäuser.

Gray, Jeremy. 2011. *Worlds out of Nothing: A Course on the History of Geometry in the 19th Century*. 2nd ed. Heidelberg: Springer.

Gray, Jeremy. 2013. *Henri Poincaré: A Scientific Biography*. Princeton: Princeton University Press.

Gray, Jeremy. 2015. Klein and the Erlangen Programme. In Ji and Papadopoulos (2015), 59–76.

Gray, Jeremy. 2018. *A History of Abstract Algebra: From Algebraic Equations to Modern Algebra*. Heidelberg: Springer.

Gray, Jeremy. 2019. 19th Century Galois Theory. In Cogliati (2019), 97–127.

Guicciardini, Niccolò. 2009. *Isaac Newton on Mathematical Certainty and Method*. Cambridge: MIT Press.

Hall, A. Rupert. 1980. *Philosophers at War. The Quarrel Between Newton and Leibniz*. Cambridge: Cambridge University Press.

Hawkins, Thomas. 1984. The Erlanger Programm of Felix Klein: Reflections on its Place in the History of Mathematics. *Historia Mathematica* 11: 442–470.

Hawkins, Thomas. 2000. *Emergence of the Theory of Lie Groups. An Essay in the History of Mathematics, 1869–1926*. Heidelberg: Springer.

Hilbert, David. 1932–1935. *Gesammelte Abhandlungen*, 3 vols. Berlin: Springer.

Hilbert, David, and Stephan Cohn-Vossen. 1932. *Anschauliche Geometrie*. Berlin: Springer.

Hilbert, David, and Stephan Cohn-Vossen. 1952/1999. *Geometry and the Imagination*, 2nd ed. Providence: AMS Chelsea Pub.

Hofmann, Joseph E. 1974. *Leibniz in Paris, 1672–1676: His Growth to Mathematical Maturity*. Trans. Adolf Prag and D.T. Whiteside. Cambridge: Cambridge University Press.

Hudson, Ronald W. H. T. 1990. *Kummer's Quartic Surface*. Cambridge: Cambridge University Press, 1905; reprinted in 1990.

Israel, Georgio, et al., eds. 1992. *La corrispondenza di Luigi Cremona (1830–1903)*, volume I, Serie di Quaderni della Rivista di Storia della Scienza, n. 1, Rome, 1992.

Israel, Georgio, et al., eds. 1994. *La corrispondenza di Luigi Cremona (1830–1903), volume II*, Serie di della Rivista di Storia della Scienza; n. 3, Rome, 1994.

Ji, Lizhen, and Athanase Papadopoulos, eds. 2015. *Sophus Lie and Felix Klein: The Erlangen Program and its Impact in Mathematics and Physics*. IRMA Lectures in Mathematics and Theoretical Physics 23. Zürich: European Mathematical Society.

Klein, Felix. 1869. Bewerbungsschrift zur Aufnahme in das Berliner Mathematische Seminar (unveröffentlichtes Manuskript). Klein Nachlass 13A, Niedersächsische Staats- und Universitätsbibliothek Göttingen.

Klein, Felix. 1870. Zur Theorie der Liniencomplexe des ersten und zweiten Grades. *Mathematische Annalen* 2: 198–228. Reprinted in Klein (1921–1923, I: 53–80).

Klein, Felix. 1871a. Zur Theorie der Kummer'schen Fläche und der zugehörigen Linien-Komplexe zweiten Grades. *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen*. Mathematisch-Physikalische Klasse 1871: 44–49.

Klein, Felix. 1871b. Ueber die sogenannte Nicht-euklidische Geometrie, *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen*. Mathematisch-Physikalische Klasse 1871: 419–433.

Klein, Felix. 1871c. Über die sogenannte Nicht-Euklidische Geometrie. *Mathematische Annalen* 4: 573–625. Reprinted in Klein (1921–1923, I: 254–305).

Klein, Felix. 1872. Ueber gewisse in der Liniengeometrie auftretende Differentialgleichungen. *Mathematische Annalen* 5: 278–303. Reprinted in Klein (1921–1923, I: 127–152).

Klein, Felix. 1872/1893. *Vergleichende Betrachtungen über neuere geometrische Forschungen*. Erlangen: A. Deichert. Reprinted in *Mathematische Annalen* 43: 63–100. Reprinted in Klein (1921–1923, 1: 460–496).

Klein, Felix. 1892. Über Lies und meine Arbeiten aus den Jahren 1870–72, Sophus Lie Papers, National Library Oslo.

Klein, Felix. 1893. *Nicht-Euklidischen Geometrie I, Vorlesung gehalten während des Wintersemesters 1889–1890*. Friedrich Schilling, Hrsg., Göttingen.

Klein, Felix. 1921–1923. *Gesammelte Mathematische Abhandlungen*. 3 Bde. Berlin: Julius Springer.

Klein, Felix. 1926. *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, vol. 1. Berlin: Julius Springer.

Klein, Felix, and Sophus Lie. 1870/1884. Ueber die Haupttangentencurven der Kummer'schen Fläche vierten Grades mit 16 Knotenpunkten. *Mathematische Annalen* 23 (1884): 198–228. Reprinted in Klein (1921–1923, I: 90–97).

Kummer, E. E. 1975. *Ernst Eduard Kummer, Collected Papers*, ed. A. Weil, vol. 2. Berlin: Springer.

Lie, Sophus. 1872. Ueber Complexe, insbesondere Linien- und Kugelcomplexe, mit Anwendung auf die Theorie partieller Differential-Gleichungen. *Mathematische Annalen* 5 (1872): 145–256.

Lie, Sophus. 1934. *Gesammelte Abhandlungen*, Bd. 1, Friedrich Engel u. Poul Heegaard, Hrsg. Leipzig: Teubner.

Lie, Sophus. 1935. *Gesammelte Abhandlungen*, Bd. 2, Friedrich Engel u. Poul Heegaard, Hrsg. Leipzig: Teubner.

Lie, Sophus, and Georg Scheffers. 1891. *Vorlesungen über Differentialgleichungen mit bekannten infinitesimalen Transformationen*. Leipzig: Teubner.

Lie, Sophus, and Georg Scheffers. 1893. *Vorlesungen über continuierliche Gruppen mit geometrischen und anderen Anwendungen*. Leipzig: Teubner.

Lie, Sophus, and Georg Scheffers. 1896. *Geometrie der Berührungstransformationen*. Leipzig: Teubner.

Nabonnand, Philippe. 2008. La théorie des Würfe de von Staudt – Une irruption de l'algèbre dans la géométrie pure. *Archive for History of Exact Sciences* 62 (2008): 202–242.

Plücker, Julius. 1865. On a New Geometry of Space. *Philosophical Transactions of the Royal Society of London* 155: 725–791.

Plücker, Julius. 1868. Neue Geometrie des Raumes gegründet auf die Betrachtung der geraden Linie als Raumelement, Erste Abteilung. Leipzig: Teubner.

Plücker, Julius. 1869. Neue Geometrie des Raumes gegründet auf die Betrachtung der geraden Linie als Raumelement, Zweite Abteilung, hrsg. F. Klein. Leipzig: Teubner.

Rowe, David E. 1988. Der Briefwechsel Sophus Lie-Felix Klein, eine Einsicht in ihre persönlichen und wissenschaftlichen Beziehungen. *NTM. Schriftenreihe für Geschichte der Naturwissenschaften, Technik und Medizin* 25 (1): 37–47.

Rowe, David E. 1989. Klein, Lie, and the Geometric Background of the Erlangen Program. In *The History of Modern Mathematics: Ideas and Their Reception*, ed. D. E. Rowe and J. McCleary, vol. 1, 209–273. Boston: Academic Press.

Rowe, David E. 2012. Einstein and Relativity. What Price Fame? *Science in Context* 25 (2): 197–246.

Rowe, David E. 2016. Segre, Klein, and the Theory of Quadratic Line Complexes. In *From Classical to Modern Algebraic Geometry: Corrado Segre's Mastership and Legacy*, ed. G. Casnati, A. Conte, L. Gatto, L., Giacardi, M. Marchisio, and A. Verra, Trends in the History of Science, Trends in the History of Science, 243–263. Basel: Birkhäuser.

Rowe, David E. 2019. Klein, Lie and Their Early Work on Quartic Surfaces. In Cogliati (2019), 171–198.

Rowe, David E. 2022a. Felix Klein and Emmy Noether on Invariant Theory and Variational Principles. In *The Philosophy and Physics of Noether's Theorems*, ed. James Read and Nicholas Teh, 25–51. Cambridge: Cambridge University Press.

Rowe, David E. 2022b. Models from the Nineteenth Century Used for Visualizing Optical Phenomena and Line Geometry. In *Model and Mathematics: From the 19th to the 21st Century*, ed. Michael Friedman and Karin Krauthausen, 175–199. Basel: Springer.

Scholz, Erhard. 1980. *Geschichte des Mannigfaltigkeitsbegriffs von Riemann bis Poincaré*. Boston: Birkhäuser.

Seidl, Ernst, Frank Loose, Edgar Bierende, Hrsg. 2018. *Mathematik mit Modellen*. Tübingen: Schriften des Museums der Universität Tübingen MUT.

Struik, Dirk Jan. 1961. *Lectures on Classical Differential Geometry*, 2nd ed. New York: Dover.

Stubhaug, Arild. 2002. *The Mathematician Sophus Lie*. Heidelberg: Springer-Verlag.

Tobies, Renate. 2021. *Felix Klein: Visions for Mathematics, Applications, and Education*. Cham Switzerland: Birkhäuser.

Tobies, Renate, and David E. Rowe 1990. *Korrespondenz Felix Klein-Adolf Mayer*. Leipzig: Teubner Archiv zur Mathematik.

Voelke, Jean-Daniel. 2008. Le théorème fondamental de la géométrie projective: évolution de sa preuve entre 1847 et 1900. *Archive for History of Exact Sciences* 62 (2008): 243–296.

Volkert, Klaus. 2013. *Das Undenkbare denken. Die Rezeption der nichteuklidischen Geometrie im deutschsprachigen Raum (1860–1900)*. Heidelberg: Springer.

Whiteside, Derek Thomas. 1967–1981. *The Mathematical Papers of Isaac Newton*, 8 vols. Cambridge: Cambridge University Press.

Wiescher, Michael. 2016. Julius Plücker, Familie und Studienjahre. *Sudhoffs Archiv* 100 (1): 52–82.

Yaglom, I. M. 1988. *Felix Klein and Sophus Lie: Evolution of the Idea of Symmetry in the Nineteenth Century*. Trans. Sergei Sossinsky, Boston: Birkhäuser.

Chapter 7
Poincaré and Arithmetic Revisited

Catherine Goldstein

Abstract Henri Poincaré's forays into number theory have often been reduced to his pioneering use of automorphic forms or his contribution to the arithmetic of elliptic curves. We examine here all of his arithmetical papers, in particular the earliest, those devoted to the study of forms. From this apparently marginal standpoint, we will be able to grasp several characteristic features of Poincaré's work at large. We show in particular the coherence of Poincaré's point of view and his mastery of the disciplinary issues of his day. We also come back to his knowledge of the contemporary mathematical literature and to his links with the program of Charles Hermite, in particular for a unity of mathematics built less on the reduction to concepts than on a circulation of methods, tools and inspiration in all fields of mathematics and for a theory of numbers revealed by the domains of the continuous.

Henri Poincaré is not especially renowned as an arithmetician. According to André Weil who devoted an article to the issue,

> Poincaré's writings that touch upon arithmetic occupy an entire volume (volume V) of his Complete Works. One could not deny that they are of unequal value. Some of them have hardly any other interest than to show us how carefully Poincaré studied all of Hermite's work and how he assimilated his methods and results.[. . .] What is particularly striking in this volume is that in it Poincaré shows very little knowledge of German-language works.[1]

[1] Weil (1955, p. 207): " Les écrits de Poincaré qui touchent à l'Arithmétique occupent un volume entier (tome V des *Œuvres*). On ne saurait nier qu'ils sont de valeur inégale. Certains n'ont guère d'autre intérêt que de nous faire voir combien attentivement Poincaré à ses débuts a étudié toute l'œuvre d'Hermite et comme il s'en est assimilé les méthodes et les résultats. [. . .] Ce qui frappe dans le volume de ses *Œuvres* dont il s'agit, c'est surtout qu'il s'y montre fort peu instruit des travaux en langue allemande."

C. Goldstein (✉)
Institut de mathématiques de Jussieu-Paris-Gauche, UMR 7586, CNRS, Sorbonne Université, Université Paris Cité, Paris, France
e-mail: catherine.goldstein@imj-prg.fr

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_7

Nonetheless Weil singled out two major contributions, which may attest to Poincaré's universal and long-lasting genius: Poincaré's use of non-Euclidean geometry in the theory of ternary quadratic forms (Poincaré 1882b, 1886b, 1887), which Weil prefers presenting as the first example of an arithmetically-defined discontinuous group, and Poincaré's celebrated paper on points with rational coordinates on elliptic curves defined over the field of rational numbers (the origin of Weil's own 1928 thesis) (Poincaré 1901). To these two gems, which still survive in the living memory of mathematics, Nicolas Bergeron has more recently added an article on invariants (Poincaré 1905a), published by Poincaré on the occasion of the centenary of Peter Gustav Lejeune Dirichlet, and viewed from the perspective of the modern theory of automorphic forms (Bergeron 2018).

Seen from such points of view which correspond to the disciplinary context of the second half of the twentieth century and beyond, Poincaré's results appear as isolated and scattered, of little significance in illuminating the global characteristics of his work. Here, on the contrary, I would like to restore the synchronic configuration in which these results are inscribed, that of the last third of the nineteenth century. My first and main objective is to show the coherence of Poincaré's interventions in number theory, a coherence that testifies, in a domain that was a priori marginal for him, to several characteristics of his work, in particular his great mastery of the disciplinary issues of his time and his ability to reformulate them in an approach that was specific to him. Such a mastery is rooted in a knowledge of the state of the art in mathematics and thus of previous mathematical literature in one form or another. As Poincaré's relation to his predecessors has often puzzled his commentators, two secondary objectives of the present paper are to assess more clearly Poincaré's awareness of the international literature, in particular of the work of German mathematicians, and to discuss in more detail the components of his interactions with Charles Hermite and with Hermite's mathematics.[2]

7.1 Poincaré's Corpus on Arithmetic

To delineate precisely the corpus to be constructed for our purpose is not as obvious as it may seem. As is well-known, a complete numbered bibliography was provided by Poincaré himself in the classical exercise of his *Notice sur les travaux scientifiques*, preparatory to a candidacy for the French Academy of Sciences.[3] This bibliography was extended by him in 1901, at the request of Gösta Mittag-

[2] In this respect, I am only completing for arithmetic, and confirming, the conclusions of the fine article that Frédéric Brechenmacher devoted to Poincaré's algebraic practices (Brechenmacher 2011).

[3] See in particular Poincaré (1886c). Poincaré was elected in 1887, after having been ranked increasingly higher on the lists of several previous elections, as was the tradition at the time (Gray 2013, p. 162).

Leffler, then reproduced and completed after his death in a special issue of *Acta mathematica*[4] and of course in Volume V of Poincaré's *Œuvres* in 1950.

The bibliography is organized by journal and not by theme. But in the structured presentation of his work that accompanies it, Poincaré gathered 16 articles under the heading "Arithmetic", to which Albert Châtelet, the editor of the Volume V of Poincaré's *Œuvres* in 1950, added several more, most of them published after 1900.[5] However, the Volume V of Poincaré's *Œuvres* is explicitly devoted to both algebra and arithmetic. In the bibliography at the beginning of the volume, Châtelet thus indiscriminately gathered together articles that Poincaré himself had put under the headings "Algèbre" or "Algèbre de l'infini" (Algebra of the infinite) in 1901, as well as later texts belonging (according to Châtelet) to one of these headings ("Arithmétique", "Algèbre", "Algèbre de l'infini"). Then, the whole volume is organized in sixteen sections, the first five being on algebraic, the last eleven on arithmetical themes—this thematic distribution being itself only loosely based on Poincaré's presentation for the earlier texts, without respecting a chronological order. It provides a list of 20 articles devoted to arithmetic (20 among the 491 items of the bibliography published in 1921).

However, the classification of this small set of texts as "arithmetic" was not shared by all mathematicians in Poincaré's era. In the *Jahrbuch über die Fortschritte der Mathematik*, 3 of these 20 articles were not reviewed at all (probably because their journal of publication was not included at the time among the titles taken into account in the *Jahrbuch*). Châtelet had included in his list the part devoted to arithmetic and algebra of a survey on the future of mathematics, presented *in absentia* at the fourth International Congress of Mathematicians, held in Rome, and reproduced several times in a variety of journals (Poincaré 1908a); it was classified as "Philosophy" by the *Jahrbuch*. Moreover, 4 articles among the 20 were classified as algebra; for one of them, this classification is understandable, as the article is the second half of a two-part investigation, one on the algebraic, the other on the arithmetic, theory of forms, and the two parts were simply reviewed together. But, more surprisingly, and despite the word "arithmetic" in their titles, articles on the application of Fuchsian functions to arithmetic and on arithmetical invariants were also classified as algebra or function theory. As for the celebrated paper on the arithmetic of curves (Poincaré 1901), it was classified as analytic geometry![6] Another classification is that of the *Répertoire bibliographique des*

[4] (Poincaré 1921). The number (38) of the volume appears to date it to the year 1915, but, according to Mittag-Leffler, it was finally printed only in 1921 because of the war.

[5] There are small discrepancies in the lists. For instance, the two earliest items, two notes to the *Comptes rendus de l'Académie des sciences* in 1879, are counted as either one or two, depending on the list. Note that there is also a misprint in Châtelet's list, the article 122 should be instead 127 cf. Poincaré (1921, p. 9) and Poincaré (1950, p. 16).

[6] Some of these papers have been reclassified as number theory in the recently-created online database integrating the *Jahrbuch* and zbMATH Open, using a modern classification of all articles. However, the 1901 paper totally escaped the attention of the reviewers, probably because of its original classification and did not receive any modern MSC-headings (September 2022).

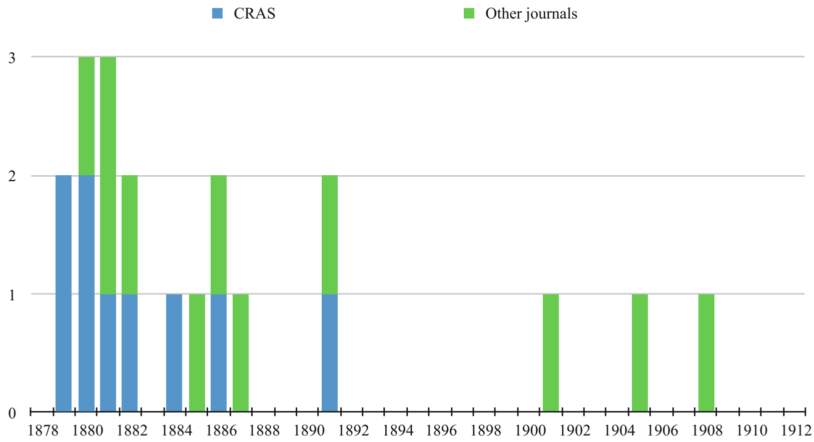


Fig. 7.1 The distribution of Poincaré's articles on arithmetic by year

sciences mathématiques (launched by a committee headed by Poincaré himself!): only 2 articles—on the distribution of prime numbers, published in 1891—were indexed in the section on number theory (section I). The others were either in the section B (which includes linear substitutions, invariants and the algebraic theory of forms) or in the section D (general theory of functions). As for the *Catalogue of Scientific Papers 1800–1900*, issued by a committee of the Royal Society of London in 1908, it also classified most of the relevant part of Poincaré's papers published before 1900 not in Number Theory (rubrics 2800–2920), but under the headings of "Non-Euclidean Geometry" (6410), "Automorphic Functions" (4440) or "General Theory of Quantics" (the English terminology for algebraic forms, 2040).[7]

Such variations are indicative of the uncertain status of number theory circa 1900. They may also hint at the role played by Poincaré in the extension and restructuration of the domain during the twentieth century. How then to select the texts on which to focus for our study? Confronted by the same problem for algebra, Frédéric Brechenmacher took a unique text as his point of departure, from which he unfolded an intricate web of conceptual and disciplinary settings (Brechenmacher 2011). Here, the chronology provides us with a clue; the distribution of our 20 potential candidates for the study of Poincaré's arithmetic (see Fig. 7.1) clearly displays an initial concentrated period, while Poincaré's best known, later papers are rather chronologically isolated.

[7] On the history of mathematical reviewing and the use of these classifications, see Siegmund-Schultze (1993), Nabonnand and Rollet (2002), Goldstein (1999), Goldstein and Schappacher (2007b).

In his 1886 presentation, Poincaré explains that his research on arithmetic concerns exclusively the theory of forms.[8] As explained above, my purpose is not to study Poincaré's influence on the development of number theory, but the coherence of his arithmetical work and its sources. I shall thus first examine in some detail the 15 articles published before 1890, then briefly, for reasons of space, discuss some of their relations to the later articles.[9]

7.2 The Arithmetic Theory of Forms in the Nineteenth Century

Although the work of German number theorists, such as Ernst-Eduard Kummer, Richard Dedekind or David Hilbert, is better known, hundreds of papers on number theory (even according, say, to the classification of the *Jahrbuch*) have been published by French authors and in French journals in the last decades of the nineteenth century, in particular in the *Comptes rendus* of the French Academy. They can be gathered roughly in three main clusters, which can be defined and distinguished by their references, their sources of inspiration, their methods and some of their practices of publication. One of these clusters blossomed in particular thanks to the French Association for the Avancement of Science (launched in 1872), which coordinated teachers, engineers, military and amateurs with a strong interest in mathematics. The two others were mostly restricted to academia and strongly relied on analytic methods as well as complex functions and numbers; the difference between them was mostly thematic and testified to their sources, Hermite and Kronecker's works for one, Dirichlet's and Riemann's on the distribution of prime numbers for the other.[10] Poincaré figured prominently in the second cluster, besides authors like Camille Jordan, Émile Picard, Léon Charve, Georges Humbert or, later, the editor of Volume V of Poincaré's *Œuvres* himself, Albert Châtelet. Their main topic was the arithmetical study of algebraic forms, that is, of homogeneous polynomials in n variables x_i, with coefficients a_{ij} in various sets of numbers (ordinary integers, real numbers, and sometimes algebraic integers, in particular).

Two important aspects should be emphasized. First of all, this cluster was international; we may mention for instance the work of Henry Smith in England, of Eduard Selling or Paul Bachmann in Germany or of Luigi Bianchi in Italy. Then, it largely benefited from the legacy of a research field which blossomed in mid-century and that Norbert Schappacher and myself have called "arithmetic algebraic

[8] Poincaré (1886c, p. 61): "Mes recherches arithmétiques ont exclusivement porté sur la théorie des formes."

[9] This part of Poincaré's work is presented in the ninth chapter of Jeremy Gray's biography of Poincaré (Gray 2013, pp. 466–488), to which my present text can be considered as a footnote.

[10] For the constitution of these clusters and details on each of them, see Goldstein (1994, 1999), Goldstein and Schappacher (2007b).

analysis" (Goldstein and Schappacher 2007a, 24–55); however, after 1870, several parts of this research field became proper disciplines—in the sense of an "object-oriented system of scientific activities" (Guntau and Laitko 1987)—with its own subject matter, its key concepts, its main problems and soon its textbooks and theses. This is particularly the case for the (arithmetical) theory of forms, see for example Smith (1861–1865), Charve (1880), Bachmann (1898).

If the study of sums of squares, in particular, began much earlier, a common source for the arithmetic theory of forms in the nineteenth century was the fifth section of Carl Friedrich Gauss's *Disquisitiones Arithmeticae*, published in 1801. In this section, Gauss studied binary quadratic forms with integer coefficients, that is, expressions $f(x, y) = axx + 2bxy + cyy$, with $a, b, c \in \mathbb{Z}$ and two variables x, y (and launched the study of ternary quadratic forms, with three variables x, y, z). Two such forms f and g are said to be equivalent (or in the same class) if they are the same up to an invertible linear change of variables with integer coefficients: $g(x, y) = f(\alpha x + \beta y, \gamma x + \delta y)$, where α, β, γ, δ are integers and $\alpha\delta - \beta\gamma = \pm 1$.[11] Gauss singled out the determinant $D = b^2 - ac$ of the form f as a key quantity; any two equivalent forms have the same determinant, that is, D is invariant under the linear transformations considered above. Reciprocally, for a given determinant D, Gauss proved that there are only finitely many different classes of equivalent forms. He also defined what would be the two main problems of the theory of forms for the whole of the nineteenth century:

> I. Given any two forms having the same determinant, we want to know whether or not they are equivalent [...] Finally we want to find all the transformations of one form into the other [...].
> II. Given a form, we want to find whether a given number can be represented by it and to determine all the representations (Gauss 1801/1966, §158, p. 113).

While the second problem, for particular forms such as sums of squares, had been in the spotlight previously, the first would take center stage during the remainder of the century. This is indeed the problem that Poincaré (who nonetheless devoted two articles to the second problem) emphasizes at the beginning of his first paper on quadratic forms:

> The main problems relating to quadratic forms can be reduced as one knows to a single one: Recognizing whether two given forms are equivalent, and by what means one can pass from one to the other.[12]

This shift in interest went hand in hand with a view of classification as a central object of research in mathematics as well as in the natural sciences.[13] "Science", Poincaré would write in 1905, "is above all a classification, a manner of bringing

[11] Gauss calls the two forms f and g properly equivalent if $\alpha\delta - \beta\gamma = 1$. I shall not comment here on these different types of equivalences, see Goldstein and Schappacher (2007a, pp. 8–13).

[12] Poincaré (1879a, p. 344): "Les principaux problèmes relatifs aux formes quadratiques se ramènent comme on le sait à un seul : Reconnaître si deux formes données sont équivalentes, et par quel moyen on peut passer de l'une à l'autre."

[13] On this issue, see Knight (1981), Tort (1989), Rey (1994), Lê and Paumier (2016).

together facts which appearances separate, although they are bound together by some natural and hidden kinship. Science, in other words, is a system of relations."[14]

Relations between forms, from the point of view described above, manifest themselves through linear transformations. Their study, including the study of those transforming a form into itself, became an important topic in the nineteenth century, as well as the search for quantities, such as the determinant, that are invariant under such transformations and can be used as characteristics in the classification. Gauss also completed what would be a model for the classifications of forms in the future: to find good, "simple" (in a sense to be explained) representatives of each class of forms (the so-called reduced forms), to study the possible equivalence among the reduced forms and to explain how to transform any form into a reduced form of its class. In the case of the binary quadratic forms with integer coefficients and with a strictly negative determinant (the case of definite forms), for instance, one can choose in each class an essentially unique reduced form such that

$$-a < 2b \leq a < c \qquad \text{or} \qquad 0 \leq 2b \leq a = c. \tag{7.1}$$

There are a finite number of such reduced forms for a given D and every form is equivalent to a reduced form. It is notable that for a reduced form, the coefficient a is then less than $2\sqrt{\dfrac{-D}{3}}$ (a bound which only depends on D, but not on the particular class of the form); since $a = f(1, 0)$ is obviously a value of the reduced form at integer values of (x, y), it is also a value at integers of all the forms in the same class, obtained by a linear change of variables from the reduced form. This remark gave rise to questions about the smallest non-zero value of the numbers represented by a form. For $D > 0$, there is no longer a unique reduced form in each class, but Gauss organized equivalent reduced forms into finite "periods". In this case, there are also infinitely many transformations of a form into itself, such transformations being associated to the solutions of the Pell-Fermat equation $T^2 - DU^2 = 1$.

Equivalent forms obviously represent the same integers, but the reciprocal is not true. Again, in order to solve the second problem, Gauss refined his solution to the first problem—more precisely, to his classification of forms—with new criteria, leading him to the concepts of order and genus, two new characteristics attached to a form. Last, but not least, he defined a relation called the composition of forms: a form F is said to be a compound of the forms g and h if there exists linear functions X and Y of xu, xv, yu, yv such that $F(X, Y) = g(x, y)h(u, v)$ (with extra technical conditions). This construction is of course useful for the representation of integers by forms, as a form composed of two others of the same determinant can represent the product of two integers which are represented respectively by the two forms. But its importance is more subtle; this relation among forms is not a binary operation,

[14] Poincaré (1905b, p. 172): "[Q]u'est-ce que la science? [...] c'est avant tout une classification, une façon de rapprocher des faits que les apparences séparaient, bien qu'ils fussent liés par quelque parenté naturelle et cachée. La science, en d'autres termes, est un système de relations."

that is, it is not possible to define "the" compound of two forms. However, it behaves nicely with respect to equivalence and turns into such a binary operation on classes of forms—that is, Gauss could define "the" composition of two classes—a remarkable idea, as it displayed an operation on less than familiar objects, *classes* of forms, made of infinitely many algebraic expressions. As such it became a model for operation on sets of mathematical objects. In the *Disquisitiones arithmeticae*, however, composition of forms relied upon complicated and extensive algebraic computations; a conundrum for many readers of Gauss who would then try to simplify or redefine it.

Among the numerous works inspired by Gauss, some were decisive in the construction of a discipline around algebraic forms. At the end of the 1830s, Peter Gustav Lejeune-Dirichlet introduced infinite series, built from inverses of quadratic forms with integer coefficients, in order to compute the number of classes of forms for a given determinant. Other analytic means, in particular elliptic functions, were also used by various authors to refine or generalize such computations. A decade later, Hermite began a series of articles devoted to forms; considering first quadratic forms with any number of variables and with real coefficients, he established through a close reading of Gauss's Section V bounds for the values of such forms at integers, which depended only on the determinant and the number of variables (and not on the coefficients of the form). This result, closely linked to the theory of reduction as explained above, led him to a variety of applications, from the approximation of real numbers by rationals to the properties of algebraic numbers or even of complex periodic functions (Goldstein 2007). Hermite also introduced his method of "continuous reduction"; he associated to a given problem a family of positive definite quadratic forms, indexed by *real* parameters, and thus transferred to the initial situation the reduction procedures for this family (continuously, by changing the values of the real parameters, hence the name of the method), in particular through a study of the transformations leading to the reduction.

For instance, if $f(x, y) = ax^2 + 2bxy + cy^2$ is an indefinite quadratic form, with a positive determinant $D = b^2 - ac$, one can write $f(x, y) = (x - \alpha y)(x - \beta y)$ for two real numbers α, β. Hermite thus associated to f the family of definite quadratic forms $f_\Delta(X, Y) = (x - \alpha y)^2 + \Delta(x - \beta y)^2$, with a real positive parameter Δ. For each Δ, there exists a linear transformation such that the transformed form F_Δ of f_Δ is reduced. Applying to f the transformation(s) reducing f_Δ, for each Δ, Hermite showed that there are only a finite number of transformed forms, reproducing themselves periodically; they define the reduced forms associated to f. Remarkably, here, the focus shifted from the forms to the transformations and these transformations became the key elements in the reduction process.

The mid-century witnessed a blossoming of the study of such linear transformations, and it was the nature of their coefficients that defined the domain of research: it was considered to be arithmetic when the coefficients were integers, algebra when they were real or complex general numbers (Brechenmacher 2016). The determinant, as explained, appeared as the simplest instance of invariants of forms—functions built from the coefficients of the forms which certain types of transformations leave unchanged (sometimes up to a well-controlled term). In turn, such invariants played a key role in the classification of forms and invariant theory was seen then as the new and fruitful direction for algebra at large (Fisher 1966; Crilly 1986; Parshall 1989, 2006, 2023).

When Poincaré entered the scene at the end of the 1870s, a whole discipline attached to the arithmetic of forms had thus been established, one of the first in number theory (Goldstein and Schappacher 2007a, p. 54).[15] Besides its subject matter, it included core concepts such as invariants and reduced objects, theorems about the two main problems (equivalence and the representation of integers by forms), a systematization based on the classification of forms, methods of proof based on the study and use of linear transformations such as that of continuous reduction. It constituted a separate subsection of the section on number theory in the recently founded *Jahrbuch über die Fortschritte der Mathematik*. As we will see, Poincaré's memoirs took its place on this map in a quite natural way.

7.3 Lattices as a Framework for Forms

Poincaré's first articles on forms were published at the beginning of his career as a mathematician, in 1879.[16] The two first items on his list of works are these notes at the Academy of Sciences in August and November 1879, parts of a memoir for which Hermite, Joseph Bertrand and Victor Puiseux were designated as reviewers. A version of these results was then expanded into a longer article in *Journal de l'Ecole polytechnique* (Poincaré 1880c). This situation is standard and it is difficult to establish a strict chronology for the subtopics relative to forms that Poincaré handled between 1879 and 1889. Most often, he presented a memoir to the Academy for review, publishing one or two short notes to announce his results, sometimes withdrawing the larger memoir before any referee report and publishing a long version of his results in another journal several years later. Roughly speaking, Poincaré discussed two main situations. First of all, that of quadratic forms, mostly binary and ternary, for which the main results were well-known; for them, Poincaré introduced in particular new geometrical viewpoints, based either on lattices (Poincaré 1879a,b, 1880c) or on non-Euclidean geometry (Poincaré 1882b, 1886b, 1887) and new analytical invariants (Poincaré 1880c, 1882a). Then, that of cubic forms, mostly ternary and quaternary, as Hermite and others had already thoroughly explored the binary cubic case (Poincaré 1880a,b, 1881c, 1882c, 1886a); for these forms, Poincaré aimed at classifying them, establishing in particular relations between his classification and some already known classifications of algebraic curves. Moreover,

[15] It is important to keep in mind that the so-called "algebraic number theory", that is, in fact, the theory of algebraic number fields, has yet not reached that same stage at the time, despite the publication of Dedekind's first papers, nor had analytic number theory.

[16] As a student, Poincaré also published a small contribution to the *Nouvelles annales de mathématiques*, in 1874, see Gray (2013, p. 157).

if some aspect of his research lent itself easily to generalization, for instance, to a greater number of variables, he would discuss this general situation (Poincaré 1881b, 1885). For reasons of space, I will thus not strictly respect chronology, but will discuss each subtopic separately.

The 1879 manuscript and notes on forms were apparently the occasion of a renewal of scientific links with Hermite, who had been his teacher at the Ecole polytechnique a few years earlier.[17] At least until Poincaré's 1887 entry into the Academy of Sciences, which Hermite had supported for several years in a row with laudatory reports and with the promotion of Poincaré among his colleagues, both maintained regular and close relations—and Poincaré would play a decisive role in the organization of Hermite's Jubilee in 1892. Hermite's judgment on Poincaré, entrusted in a letter to Gösta Mittag-Leffler in March 1882, is well-known:

> In confidence, with great fear of being overheard by Madame Hermite, I will tell you that of our three mathematical stars [Paul Appell, Emile Picard and Poincaré], Poincaré seems to me the brightest. And then, he is a charming young man, who, like me, is from Lorraine and who knows my family very well.[18]

This statement is not an isolated one. Despite some opposition, on family as well as on institutional grounds, Hermite insisted that "Poincaré is unquestionably superior to Appell and Picard in terms of both the importance of his discoveries and the number of published works."[19] This on-going support is expressed in several ways: Hermite sent Poincaré's thesis to Mittag-Leffler (Hermite and Mittag-Leffler 1984, I, p. 118), asking him to recommend Poincaré's work to Hugo Gyldén or Karl Weierstrass (Hermite and Mittag-Leffler 1984, I, p. 150); he provided explanations on Poincaré's work to Georges Halphen (Poincaré 1986, p. 158). Reciprocally, he fed Poincaré with mathematical literature, commented on his results and, as we

[17] Hermite and Mittag-Leffler (1984, vol. 5, p. 110) : "Je crois à ce jeune homme, qui a été mon élève à l'Ecole polytechnique en 1875, un véritable génie" (I believe that this young man, who was my student at the Polytechnique in 1875, has a true genius). Hermite wrote to Poincaré on November 22, 1879 that he had not yet seen the August manuscript, but would be delighted to read at the Academy the new note prepared by Poincaré (it would take place on November 24) (Poincaré 1986, p. 164). Let us remind our readers that Poincaré also defended his thesis on differential equations and lacunary series, preceded by a short article on this topic, in August 1879; but Hermite was not a member of the defence committee, composed of Ossian Bonnet, Jean-Claude Bouquet and Gaston Darboux.

[18] Hermite and Mittag-Leffler (1984) : "Tout bas et en confidence, ayant grande crainte d'être entendu de Madame Hermite, je vous dirai que de nos trois étoiles mathématiques [Appell, Picard et Poincaré], Poincaré me semble la plus brillante. Et puis, c'est un charmant jeune homme, qui est lorrain comme moi et qui connaît parfaitement ma famille." Also quoted in Gray (2013, p. 161). Appell was Hermite's nephew by marriage and Picard his son-in-law, and they were thus both supported by Hermite's family members. On the other hand, both had been students at the Ecole normale supérieure, which at that time, under the leadership of Louis Pasteur, was beginning to take over the training of scientific elites, against the influence of the Polytechnique.

[19] Hermite and Mittag-Leffler (1984, vol. 5, p. 214): "Poincaré est incontestablement supérieur à Appell et à Picard sous le double rapport de l'importance des découvertes et du nombre des travaux publiés."

shall see, pushed him to rewrite, develop or explore more deeply and more precisely certain topics.

In Poincaré's first papers on forms, however, Hermite is not mentioned.[20] Poincaré acted here as he often would do in the future; he first re-read or re-established in a specific framework results that were already known, at least partially. In our case, the basic results are those of Gauss's Section V, but the framework is more surprising. In Poincaré's terms:

> The link between Bravais' theory of parallelogrammatic lattices and that of quadratic forms was noticed long ago, but was restricted until now to definite forms. The main objective of this Memoir is to show that nothing is easier than applying the same geometrical representation to indefinite forms. First I had to study the properties of these parallelogrammatic lattices and to sketch out, so to speak, their arithmetic. [...]. The lattices enjoy properties which recall several of the properties of numbers.[21]

The representation of positive forms by lattices had been indeed popularized much earlier among mathematicians through Gauss's review of Ludwig August Seeber's theory of reduction for ternary quadratic forms (Seeber 1824; Gauss 1831). Seeber worked on crystallography and his interest in quadratic forms was primarily linked to the modelization of crystal properties. In his review, Gauss explains that, if on a plane one chooses two coordinate axes making an angle of cosine $\frac{b}{\sqrt{ac}}$, the value at x, y of a positive definite form, $ax^2 + 2bxy + cy^2$, with a, b, c integers, represents the square of the distance to the origin of the point with coordinates $(x\sqrt{a}, y\sqrt{c})$ with respect to these axes. For x, y integers, the form is thus associated to a discrete grid of points, situated at the intersection of two systems of lines which are parallel respectively to each of the two axes and evenly spaced ($\sqrt{a}$ for one system of lines, $\sqrt{c}$ for the other) (see Fig. 7.2); this double system of lines defines a lattice. The plane is thus cut into equal elementary parallelograms (such that no point of the lattice lies inside such a parallelogram); the area of each such parallelogram is $ac - b^2$ (that is, the absolute value of the determinant of the form). Different lattices can be associated to the same distribution of lattice points, for different choices of the systems of lines joining them; the forms associated with these different lattices are then equivalent. In this framework, reduction theory can

[20] The link to Poincaré's other early works, on differential equations, with the topic proposed in 1879 for an Academy Prize has been noted by historians (Gray 2006, 2013, ch. 3). But I have not been able to find an explicit incentive for his work on forms. The theme of the decomposition of a number as a sum of squares was proposed only in 1881. We note that a topic on crystals was also proposed by the Academy for a mathematical prize in 1879 and, given Poincaré's reference to Auguste Bravais's work on crystallography, this might have played a role in his interest in lattices and forms, but I have not found any source to substantiate this.

[21] Poincaré (1880c, pp. 177–178): "Le lien qui existe entre la théorie des réseaux parallélogrammatiques de Bravais et celle des formes quadratiques a été remarqué depuis longtemps, mais on s'est restreint jusqu'ici aux formes définies ; le but principal de ce Mémoire est de faire voir que rien n'est plus facile que d'appliquer la même représentation géométrique aux formes indéfinies. J'ai dû d'abord étudier les propriétés de ces réseaux parallélogrammatiques et en ébaucher pour ainsi dire l'arithmétique [...]. Les réseaux jouissent de propriétés qui rappellent quelques-unes des propriétés des nombres."

also be described in geometrical terms: among the various lattices associated to the same given regular discrete distribution of points, the lattice corresponding to a reduced form in Gauss's sense is the only one for which the fundamental triangle, joining the chosen origin to the nearest points of the lattice, has acute angles; it is also the only one for which the elementary parallelograms have their sides smaller than their diagonals (Gauss 1831; Dirichlet 1850). This geometrical representation of the theory of forms can be extended in the same way to a three-dimensional space and ternary forms $ax^2+by^2+cz^2+2a'yz+2b'xz+2c'xy$, with three variables, the triple system of lines defining a lattice with spacings of $\sqrt{a}$, $\sqrt{b}$, $\sqrt{c}$, respectively and the cosine of the angle between two of the three axes being $\frac{a'}{\sqrt{bc}}$, $\frac{b'}{\sqrt{ac}}$, $\frac{c'}{\sqrt{ab}}$ respectively.

Poincaré, however, did not refer to Seeber or to Gauss's review in his arithmetical work. Besides Gauss's *Disquisitiones arithmeticae* and, for one specific result, Eisenstein, Poincaré mentions only Auguste Bravais in Poincaré (1880c). Bravais, also an engineer from the Polytechnique and a professor at this school (before Poincaré's time) developed his own theory of lattices, first on the plane in a botanical context, then in a three-dimensional setting for crystallography.[22] Bravais's viewpoint was neither arithmetical, nor centered on quadratic forms; it relied on the study of symmetries and of the effect on lattices of various transformations, in particular rotations. But it constituted a common reference among Polytechnicians: Henry Résal, the editor of the *Journal de mathématiques pures et appliquées* following Joseph Liouville, referred for instance to Bravais's in a footnote added to an article by Eduard Selling on the reduction of quadratic forms which Hermite strongly recommended to all his students (Selling 1877).[23] And Camille Jordan explicitly followed Bravais's study of symmetries when he analysed the groups of space motions at the end of the 1860s (Jordan 1868–1869). Poincaré used Bravais not only as a mere designation of objects (the so-called "Bravais lattices"), but also for his proofs of several basic properties on lattices; he would also mention Bravais's work elsewhere, in particular in his lectures on the theory of light (Poincaré 1889, p. 195).

In Poincaré (1880c), Poincaré introduces several notations for plane lattices and their points. First of all, a plane lattice can be defined by four numbers (a, b, c, d), the coordinates of the points of the lattice being given as

$$x = am + bn$$

$$y = cm + dn$$

[22] Bravais (1850, 1851) and the posthumous collection Bravais (1866). Bravais calls a three-dimensional lattice an "assemblage". For a survey of Bravais's approach in crystallography from a group theoretical point of view, see Scholz (1989); for a general presentation of Bravais's work using lattices, see Boucard et al. (2024).

[23] But it is the German version of Selling's paper, published in 1874, and not its French 1877 translation, that Poincaré would mention in Poincaré (1882b).

with integers m, n.[24] Poincaré also designates this lattice (a, b, c, d) by $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. He calls the quantity $ad - bc$ the norm of the lattice (this is the area of an elementary parallelogram of the lattice).

The first part of the 1880 memoir (Poincaré 1880c) is then explicitly devoted to the development of an *arithmetic of lattices*, for the case where a, b, c, d are integers. The objective is to mimic standard concepts of the arithmetic of integers, such as multiples, divisors and primes. A multiple of a lattice, for instance, is simply a sublattice—that is, all its points are contained among the points of the original lattice. Two lattices are then said to be equivalent if each of them is a multiple of the other. It is always possible, up to equivalence, to assume that $d = 0$; then, two lattices such as $(a, b, c, 0)$ and $(a', b', c', 0)$ are equivalent if and only if $c = c'$, $b = b'$, $a \equiv a' \pmod b$.[25]

Using Bravais's results, Poincaré asserts that the norm of a lattice is the limit of the ratio of the area of a circle to the number of lattice points inside the circle when the radius increases indefinitely. This allows him to show that the norm of a lattice is divisible by the norms of its divisors and that the norms of two equivalent lattices are equal. Poincaré then characterizes the smallest common multiple and the largest common divisor of two lattices. He defines a "prime" lattice as a lattice the norm of which is a prime number, and a "second" lattice as one the norm of which is the power of a prime number.[26] He concludes with the theorem that any lattice is the least common multiple of co-prime second lattices.

Poincaré's procedure, and his results, are thus very close to those of Richard Dedekind's theory of ideals, which had been published in French only a few years earlier (Dedekind 1876, 1877). For instance, Dedekind defines a multiple of an ideal in the following way: the ideal $\mathfrak{a}$ is a multiple of the ideal $\mathfrak{b}$ when "all the numbers of the ideal $\mathfrak{a}$ are contained in $\mathfrak{b}$" (Dedekind 1876, p. 287); he then also develops an arithmetic of ideals, in order to generalize and simplify Ernst Kummer's preceding theory of ideal numbers. However, while Poincaré, as we shall see, alluded indeed to "ideal numbers" in his memoirs, he mentioned explicitly neither Kummer nor Dedekind, reinterpreting their theories in his lattice framework.

Following Poincaré, let us then come back to quadratic forms. He represents now a (binary quadratic) form $am^2 + 2bmn + cn^2$ (with $a > 0$) by the lattice

$$\begin{bmatrix} \frac{b}{\sqrt{a}} & \sqrt{a} \\ \sqrt{\frac{b^2-ac}{Da}} & 0 \end{bmatrix}$$

---

[24] The directions of the two systems of lines defining the lattice are then $\begin{pmatrix} a \\ c \end{pmatrix}$ and $\begin{pmatrix} b \\ d \end{pmatrix}$.

[25] As pointed out by Châtelet in the comments of his edition of Poincaré's works, a unique representative of each class of equivalence is obtained if one requires that a, b, c are positive and $0 < a < b$.

[26] In French, the word "premier" means both "prime" and "first", which explains this somewhat strange terminology. We would now prefer "primary" instead of "second".

or by

$$\begin{bmatrix} b & a \\ \sqrt{\frac{b^2-ac}{D}} & 0 \end{bmatrix}$$

and its multiples.[27] We recall that $D = b^2 - ac$, which makes it obvious that the entries in the last expression are integers when the form has integer coefficients. A key point of this expression is to underline its relation with the usual association between a quadratic form and a lattice that we have outlined above. In the usual representation, valid for definite binary quadratic forms (with $D < 0$), the third term would have been $\sqrt{\frac{ac-b^2}{a}}$. As Poincaré explains, his representation allows a similar treatment for definite and indefinite forms (those with $D > 0$). For definite forms, he introduces it as a projection of the usual plane representation on another plane, which makes an angle with the first that depends on D. It means in particular that the plane hosting the lattices depends on the determinant D (or, from our modern perspective, that the representation is proper to one specific quadratic field $\mathbb{Q}(\sqrt{D})$).[28] The whole theory Poincaré then develops is associated to the forms of a given determinant D, up to the square of an integer (which amounts to normalizing the size of the elementary parallelograms). Again, on the plane, the same regular distribution of points may correspond to different lattices, and thus give rise to different quadratic forms, depending on the way the two systems of parallel lines joining these points are chosen.

Let us now explain Poincaré's geometrical definition of the theory of reduction for binary quadratic indefinite forms. Let O, A, B be a fundamental triangle of the lattice associated to an indefinite form, that is, a triangle with three points of the lattice as vertices and such that no other point of the lattice lies inside the triangle. Poincaré then constructs an elementary parallelogram of the lattice $OABC$, OA and OB being sides of this parallelogram.[29] Then OBC and OAC are two other fundamental triangles of the lattice, that Poincaré calls "derived" from the first one OAB. In the same way, OAB is itself derived from two other fundamental triangles OAE and OBD (that Poincaré calls the "primitive" of OAB) (see Fig. 7.2).[30]

[27] There is here of course an abrupt change of notation, as a, b, c do not mean the same thing as before! This type of change is frequent in Poincaré's early papers. Darboux, for instance, complained of precisely this problem when he reviewed Poincaré's thesis in 1878–1879.

[28] This apparent disadvantage could be seen in a positive way as changing the implicit Euclidean metric of the Gauss-Seeber representation into a Lorentzian one.

[29] In the original text, the drawings do not show the lattices, they only sketch the construction of the derived and primitive triangles. For convenience, our representation Fig. 7.2 is thus slightly different from the original, as well as its normalization.

[30] Again, Poincaré's terminology is slightly confusing. The words "derived" and "primitive" (in Latin or German, in particular) appear in the theory of forms developed by Gauss and some of his successors. In their work, a "derived form" is also obtained by applying a certain transformation to the initial form. But "primitive" here describes an intrinsic property of the form, for instance that

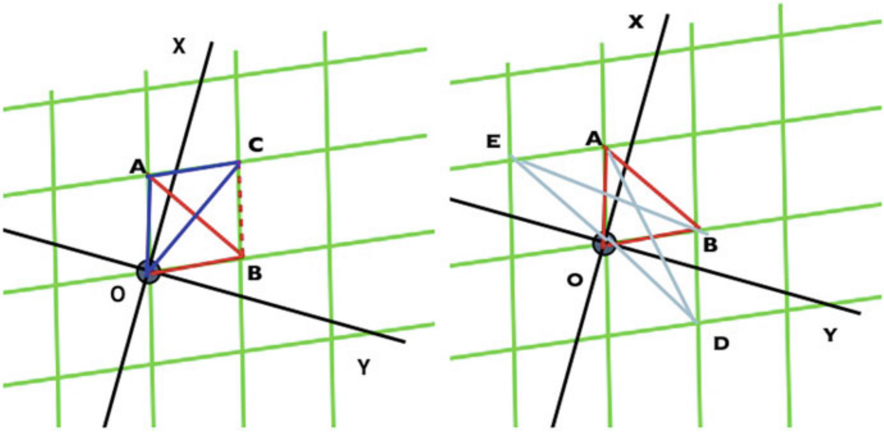


Fig. 7.2 A fundamental triangle OAB of a lattice, its derivatives OAC and OBC (left) and its primitives OEB and OAD (right)

For instance, let $x^2 - 2y^2$ be our initial form: its determinant is $D = 2$, the associated lattice is $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. On the appropriate projection plane, the elementary parallelograms of the lattice are squares (the area being renormalized as 1 in the projection). The points of the lattice are those of coordinates $(x = n, y = m)$, with n and m integers; the corresponding points A and B are respectively (0, 1) and $(1, 0)$, the point C being thus $(1, 1)$. The new elementary parallelogram associated to the derived triangle $0AC$, for example, is thus 0, A, C', C, with $C' = (1, 2)$ (the other derived triangle $0BC$ gives rise to $0BCC"$ with $C" = (2, 1)$). Reciprocally, $0AB$ is one of the two primitive triangles of $0AC$. The new lattice corresponding to the elementary parallelogram 0, A, C', C is thus given by $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, and the new quadratic form of determinant 2 associated to it is $x^2 + 2xy + 2y^2$.

The lines $X : \sqrt{D}x = y$ and $Y : \sqrt{D}x = -y$ are called the asymptotes. Poincaré calls "ambiguous" a fundamental triangle (such as OAB) such that the first asymptote cuts it and the second does not. Exactly one derivative and one primitive of an ambiguous triangle are also ambiguous (in our example, OAC is ambiguous and OBC is not). The procedure of derivation thus allows Poincaré to construct sequences of ambiguous triangles. Relying on a theorem of Bravais, Poincaré is then able to show that the binary quadratic forms associated with the successive triangles of such a sequence are periodically reproduced. As seen in our example Fig. 7.2,

its coefficients a, b, c are co-prime. In this perspective, unlike Poincaré, 'primitive' and 'derived' do not correspond to inverse transformations of each other. Poincaré's terminology seems more akin to that coming from function theory.

a triangle has a common side with its derivatives; thus, among the finite sequence of triangles in a period, subsequences of ambiguous triangles and their successive derivatives share a common side, before another side occurs as the common side of another subsequence. Poincaré shows that the last triangle (or the first) of such a finite subsequence is associated to a reduced binary quadratic form.

Poincaré also displays the correlated geometrical interpretation of the development into continued fractions of $\frac{a\sqrt{D}}{1-b\sqrt{D}}$: the successive reduced fractions provide the coordinates of a vertex of a triangle associated to a reduced form. This interpretation will be used by Poincaré in a later note to embody the rational approximation of a real number α, by using the line $y = \alpha x$ and the Bravais lattice of the points with integral coordinates (Poincaré 1884).[31]

The following section of Poincaré (1880c) reinterprets in the language of lattices the difficult composition of forms introduced by Gauss in the *Disquisitiones arithmeticae*. For this, Poincaré introduces a new multiplication on lattices, which is different from the non-commutative general multiplication on which the arithmetic of lattices discussed above was based. Let $Am + Bn$ and $A_1m_1 + B_1n_1$ be two lattices (with A, B, A_1, B_1 complex quadratic numbers for the same D), the result of this new multiplication is the lattice $AA_1\mu_1 + AB_1\mu_2 + BA_1\mu_3 + BB_1\mu_4$, with $m, n, m_1, n_1, \mu_1, \mu_2, \mu_3, \mu_4$ integers.[32] Poincaré proves that if a form is composed of two others, the corresponding lattice is the (new) product of the lattices associated to the two forms, that is, the composition on forms corresponds to a true operation on lattices, from which he can deduce Gauss's results relative to composition.

At the beginning of the last section of this memoir, Poincaré announces:

> The above considerations make it possible to present in a simple and concrete manner the theory of ideal complex numbers that correspond to quadratic forms of determinant D.[33]

This reference to ideals is very rare in the French landscape at the time (Goldstein 1999). Hermite had displayed an interest in Kummer's work on ideal numbers early on (Hermite 1850), but his research aimed at offering alternative proposals to handle the arithmetic of algebraic numbers rather than promoting the reception of Kummer's (or later, Dedekind's) conceptual enterprise.[34] Poincaré situates himself in a Hermitian vein here, when he proposes (at least for the quadratic case) to substitute plane lattices—themselves embodied in tables of familiar "true" numbers—to an

[31] In this note, he also generalizes this construction to the rational approximation of two real numbers, by using this time the three-dimensional version of Bravais lattices. We recalled earlier the link between the theory of reduction and bounds on the values of the form at integers. Hermite had already applied it, for suitable quadratic forms, to the rational approximation of several real numbers (Hermite 1850).

[32] The infelicitous notation, again, is Poincaré's.

[33] Poincaré (1880c): "Les considérations qui précèdent permettent d'exposer d'une manière simple et concrète la théorie des nombres complexes idéaux qui correspondent aux formes quadratiques de déterminant D."

[34] See Goldstein (2007) and on Hermite's emphasis on a "clear" and "concrete" approach in mathematics, see Goldstein (2011), and below.

"ideal" family constructed by means of divisibility properties, as Kummer did, or with a set-theoretical perspective, as Dedekind did.[35] To a (quadratic) real complex number $\lambda + \mu\sqrt{D}$, Poincaré associates the lattice $\begin{bmatrix} \lambda & \mu\sqrt{D} \\ \mu & \lambda \end{bmatrix}$. The points of the lattice represent the multiple of this number and the (new) product of lattices, as defined for the composition of forms, corresponds to the product of two associated complex numbers.[36] An ideal complex number is then defined by Poincaré as a lattice with some simple conditions (which are of course verified for the lattices associated to a "true" complex number). The main arithmetical properties of lattices are then transferred to these ideal numbers; a prime ideal number, for instance, is one for which the norm (of the defining lattice) is a prime and Poincaré proves in particular that every ideal number, in his sense, can be decomposed in a unique way into ideal prime factors.[37]

Although Poincaré only addresses well-known cases in his first articles on forms, he clearly had hopes regarding the framework of lattices he has deployed, in particular for a generalization to any number of variables. The equivalence of two forms (under linear, invertible, transformations with integer coefficients) correspond to the equality of the two associated lattices, that is, (as in Bravais's memoirs), the two lattices differ from each other only by a rotation around the fixed origin O, of an angle θ. Poincaré computes the transformations of the forms in terms of their coefficients and the angle θ and reciprocally. But an original and, for Poincaré, decisive step is the introduction of new types of invariants. As explained before, for a binary quadratic form $ax^2+2bxy+cy^2$, the quantity b^2-ac is invariant under all linear transformations of the variables of determinant 1, whatever the nature of the coefficients of this transformation; this property had given rise to the search for other invariant algebraic expressions, associated to forms of various degrees and number of variables, during all of the nineteenth century; as the nature of the coefficients does not intervene, it was considered to be the *algebraic* part of the study of forms by most authors, in particular Hermite, Jordan or Poincaré himself (Brechenmacher 2011, 2016). For binary quadratic forms, $b^2 - ac$ was the only algebraic invariant. But, as noted by Poincaré, there exist (many) arithmetical invariants, that is, expressions which are unchanged under a linear, invertible, transformation with *integer* coefficients. For instance, the series $\sum_{-\infty}^{\infty} \frac{1}{(am^2+2bmn+cn^2)^k}$ (where the sum is taken over all the integer couples $(m, n) \neq (0, 0)$) are such arithmetical invariants.

[35] On Kummer's construction of ideal numbers, see Edwards (1977). On the issues at stake with ideal numbers at the time and Dedekind's approach, see Edwards (1980, 1992), Haubrich (1992). Once again, Poincaré seems close to the practical preferences expressed by Dedekind, even if their proposals to solve the problems differ (Ferreiros 2007; Haffner 2014, 2019).

[36] See Edwards (2007) on the link between the composition of forms and a theory of ideal numbers.

[37] As pointed out by Châtelet, Poincaré's viewpoint amounts for complex numbers to the representation by a matrix of the multiplication by this number of the elements of a well-chosen basis of the quadratic field (Poincaré 1950, p. 174, footnote 2). This matrix point of view on ideal theory was adopted by several authors after Poincaré, including Châtelet himself.

Poincaré also considers $\sum_{-\infty}^{\infty} \frac{1}{(\sqrt{a}m+\frac{b+\sqrt{ac-b^2}}{\sqrt{a}}n)^{2k}}$, which a linear transformation with integer coefficients changes by a function of the angle of rotation θ.[38]

He then provides an effective (if not efficient) procedure to decide if two definite binary quadratic forms $ax^2 + 2bxy + cy^2$ and $a'x^2 + 2b'xy + c'y^2$ with the same determinant $D = ac - b^2 = a'c' - b'^2$ are equivalent (Poincaré 1879b, 1882a). Let us consider the (convergent) series

$$\phi_k(q) = \sum \frac{1}{(qm+n)^{2k}},$$

the sum being taken over all integers m, n, except $(0, 0)$. Assuming that the two forms are equivalent, Poincaré expresses the coefficients $\alpha, \beta, \gamma, \delta$ of a linear transformation between the forms as a function of the coefficients of the forms and the real and imaginary parts of

$$\sqrt{\frac{a\phi_1(\frac{b'+i\sqrt{D}}{a'})}{a'\phi_1(\frac{b+i\sqrt{D}}{a})}}.$$

If one computes the values of ϕ_1 with a sufficient approximation, the values of $\alpha, \beta, \gamma, \delta$ can be known up to less than $1/2$ and one thus gets exact values for these integers. It is then sufficient to check if this transformation indeed sends the form $ax^2 + 2bxy + cy^2$ into the form $a'x^2 + 2b'xy + c'y^2$. Other series are introduced for deciding on the equivalence of indefinite forms (Poincaré 1882a). In both cases, Poincaré also shows that the series can be represented by definite integrals (using in particular then recent results on elliptic functions by Appell).

7.4 Representation of Numbers

Poincaré's foray into ideal theory did not go unnoticed. For instance, Arthur Cayley wrote to Poincaré on October 12, 1883:

> I have to thank you very much for the valuable series of memoirs which you have kindly sent me. I see that you have in one of them applied the theory of ideal numbers to the case of binary quadratic forms ; it had occurred to me that a very good illustration of the general theory would thus be obtained and I am very glad to find that the case has been worked out (Poincaré 1986, p. 116).

Nor was this an isolated instance. As explained earlier, a classical question of the theory was the study of the values of forms at integers, and, in particular, the

[38] Analogous series had appeared in Dirichlet's work on the computation of class numbers of binary quadratic forms (as well as in his work on prime numbers in arithmetic progressions).

representation of integers by such values.[39] After a short communication to the Academy of Sciences in 1881 (Poincaré 1881b), Poincaré handled this problem for general binary forms in a memoir in the *Bulletin of the French Mathematical Society*, of which he had been accepted as a member on April 21, 1882 (Poincaré 1885). In this article, he puts ideal theory center-stage, but this time he explicitly adopts Dedekind's terminology, mentioning ideals instead of ideal numbers.[40] Poincaré's point of departure, however, is a type of form that Hermite had singled out in several memoirs (Goldstein 2007):

$$\Psi(x_0, x_1, \cdots, x_{m-1}) = (x_0 + \alpha_1 x_1 + \alpha_1^2 x_2 + \cdots + \alpha_1^{m-1} x_{m-1})$$
$$(x_0+\alpha_2 x_1+\alpha_2^2 x_2+\cdots+\alpha_2^{m-1} x_{m-1}) \cdots (x_0+\alpha_m x_1+\alpha_m^2 x_2+\cdots+\alpha_m^{m-1} x_{m-1}),$$

where $\alpha_1, \cdots, \alpha_m$ are the roots of an algebraic equation. In other terms $\Psi(x_0, x_2, \cdots, x_{m-1})$ is the norm of the complex integer $x_0 + \alpha_1 x_1 + \alpha_1^2 x_2 + \cdots + \alpha_1^{m-1} x_{m-1}$ (as well as of its conjugates).

Such decomposable norm-forms were here used by Poincaré (as had been done earlier by Hermite) as a link between ideal theory and the representation of integers by binary forms. More precisely, let F be an arbitrary binary form

$$F(x, y) = B_m x^m + B_{m-1} x^{m-1} y + \cdots + B_1 x y^{m-1} + B_0 y^m$$

with the B_i integers. The question of the representation of an integer N by F is easily reduced to that of the representation of $B_m^{m-1} N$ by the form

$$(x + \alpha_1 y)(x + \alpha_2 y) \cdots (x + \alpha_m y) = \Psi(x, y, 0, \cdots, 0)$$

(where the α_i are now the roots of the algebraic equation obtained by dehomogenizing the form F). Poincaré was thus led to study in general the representation of an integer, say N', by a form $\Psi(x_0, x_1, \cdots, x_{m-1})$. To do this, he proceeds by studying all the ideals of norm N' in what we would call the ring generated by the α_i. The question is thus to decide if the ideals representing N are principal, that is, if they are composed of the multiples of one complex number, as required for the initial problem.

Poincaré represents the elements of such an ideal (this concept being more or less understood in Dedekind's sense as a module stable under multiplication by any complex number of the type $x_0 + \alpha_1 x_1 + \alpha_1^2 x_2 + \cdots + \alpha_1^{m-1} x_{m-1}$)[41] as $x^{(1)} m_1 +$

---

[39] Poincaré also made use of ideals in 1891 while extending to them some analytical results of Pafnuty Chebyshev on the distribution of prime numbers, but his application was limited to the ring $\mathbb{Z}[i]$ of Gaussian integers, for which all ideals are principal (Poincaré 1891a,b).

[40] Besides Dedekind, this article includes references to Eisenstein and Kummer, and, as we shall see, to Hermite.

[41] By taking only integral coefficients, Poincaré does not obtain in all cases the principal ideal generated by an element in the complete ring of integers, and here, as elsewhere in his writings,

$x^{(2)}m_2 + \cdots + x^{(n)}m_n$, with integers m_i and $(x^{(i)})$ generators of the ideal. The norm of the elements of the ideal then defines a form of the same degree and the same number of variables as Ψ, and it is possible to study its equivalence with the form Ψ by using the Hermitian method of continuous reduction we mentioned earlier (Hermite 1851).

To summarize, Poincaré's procedure is to construct all ideals of norm N, then to examine if they are or are not principal by deciding on the equivalence of two forms. Hermite's technique even provides theoretically the transformation that is needed to express N as the value at integers of the initial binary form.

Most of Poincaré's article is thus devoted to the determination of ideals with a given norm. The generators $(x^{(i)})$ of an ideal are represented as a function of the powers of the α_j by a table of coefficients.[42] The first decisive step is to reduce this table to a triangular form which describes this possible ideal-solution—a step also arising from Hermite's work (Hermite 1851). Then, Poincaré computes successively the conditions required such that a table represents an ideal (in his sense), then an ideal with a prime number as its norm, then an ideal with a power of a prime number as its norm, which finally allows him to exhibit ideals with a given norm N.

To give a flavor of the computations involved, let us illustrate them by the first step, in the case of a reduced 3 by 3 table $\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}$.[43] The three generators are thus here $a, b + d\alpha_1, c + e\alpha_1 + f\alpha_1^2$, with a, b, c, d, e, f integers. If a complex integer $x_0 + x_1\alpha_1 + x_2\alpha_1^2$ is in this module, it should be a linear combination with integral coefficients of the generators, say $pa + q(b + d\alpha_1) + r(c + e\alpha_1 + f\alpha_1^2)$ (with p, q, r integers), thus the coefficient of the term in ${\alpha_1}^2$ should be a multiple of f. Moreover, for this module to be an ideal, the multiplication by α_1 of the generators should again be in the module, that is, $a\alpha_1 = qd\alpha_1$ (the term $rf{\alpha_1}^2$ being 0, one should have here $r = 0$), thus d divides a. In the same way, expressing that $b\alpha_1 + d\alpha_1^2$ is in the module provides the fact that f divides d. A repetition of the same argument with ${\alpha_1}^2$ provides the final conditions:

$$a \equiv d \equiv 0 \mod f \quad b \equiv e \equiv c \equiv 0 \mod f \quad a \equiv 0 \mod d \quad b \equiv 0 \mod d.$$

several assumptions are missing. Châtelet completed them carefully in his comments to the *Œuvres* and I will not discuss them further.

[42] We would now call this table a matrix, but F. Brechenmacher (2011) has convincingly discussed the conceptual nuances of the two terms. As for Poincaré, he spoke of "notation" and later of "tableaux" (tables, or charts).

[43] In his paper, Poincaré uses a representation with 3, 4 or 5 variables, while asserting the generality of his construction. Such a tension appears again in other papers of the same period and has also been analyzed (Brechenmacher 2011).

7.5 Cubic Ternary Forms: Another Geometrical Outlook

As we have seen, the use of lattices could be extended to ternary (quadratic) forms. But Poincaré came back to the basics of the theory of forms when he turned to cubic forms (Poincaré 1881c, 1882c):

> The various problems connected with the theory of binary quadratic forms have long been solved by the notion of reduced forms [...]. To generalize such a useful idea, to find forms playing in the general case the same role as reduced forms do in the case of quadratic forms, such is the problem which naturally arises and which M. Hermite has solved in the most elegant way. M. Hermite has confined himself to the study of definite or indefinite quadratic forms and of forms decomposable into linear factors; but his method can be extended without difficulty to the most general case. I believe that this generalization can lead to interesting results, and this is what determined me to undertake this work. [...] The simplest of all forms, after the quadratic forms and the forms decomposable into linear factors, are the ternary cubic forms. [...] In addition to [their] simplicity, other considerations have influenced my choice. These forms have indeed, from the algebraic point of view, been the object of very interesting and very complete works, and thanks to the close connection between Higher Algebra and Higher Arithmetic, these results have been of great help to me.[44]

The "very interesting and very complete works" mentioned by Poincaré were, according to his references, those of Otto Hesse, Siegfried Aronhold, Jakob Steiner and Alfred Clebsch, which concern invariant theory. Though Poincaré describes them as "algebraic", their relevancy here relies on a geometrical interpretation of the problem; for a ternary form $F(x_1, x_2, x_3)$, the equation $F(x_1, x_2, x_3) = 0$ indeed gives rise to a (projective) plane curve. Following Hermite, Poincaré first studies and classifies the *linear transformations* reproducing the form (that is, leaving the form unchanged when the transformation is applied to its variables), by means of what we now call eigenvalues. Then, he transfers the results to the corresponding plane curves, linking their (algebraic) invariants and their geometrical characteristics to the various categories of transformations. For each associated family of forms,

[44] Poincaré (1881c, pp. 190–191): "Les divers problèmes qui se rattachent à la théorie des formes quadratiques binaires ont été résolus depuis longtemps, grâce à la notion de réduite [...]. Généraliser une idée aussi utile, trouver des formes jouant dans le cas général, le même rôle que les réduites remplissent dans le cas des formes quadratiques, tel est le problème qui se pose naturellement et que M. Hermite a résolu de la façon la plus élégante [...]. M. Hermite s'est borné à l'étude des formes quadratiques définies ou indéfinies et des formes décomposables en facteurs linéaires ; mais sa méthode peut s'étendre sans difficulté au cas le plus général. Je crois que cette généralisation peut conduire à des résultats intéressants ; et c'est ce qui m'a déterminé à entreprendre ce travail. [...] Les plus simples de toutes les formes, après les formes quadratiques et les formes décomposables en facteurs linéaires, sont les formes cubiques ternaires. [...] Outre [leur] simplicité [...] d'autres considérations ont influé sur mon choix. Ces formes ont été en effet, au point de vue algébrique, l'objet de travaux très intéressants et très complets, et grâce au lien étroit qui rapproche l'Algèbre supérieure de l'Arithmétique supérieure, ces résultats m'ont été d'un grand secours."

Poincaré also provides a canonical one, whose equation is considered as particularly simple, and he explicitly computes the invariants.[45]

In the second part of his memoir, Poincaré addresses the properly arithmetical problems of the cubic ternary forms: their equivalence and classification, and the description of the transformations (this time with integer coefficients) which reproduce them (we now call them automorphisms and for sake of simplicity, we will use this terminology freely). To do this, Poincaré, following Hermite and other authors, in particular Selling, uses a (real) transformation sending the form into a canonical one (which is thus algebraically equivalent) and then transfers to the original form the reduction and the automorphisms of the canonical form. As for the quadratic case, there exist several possible definitions of the reduction of a form and/or of the canonical forms; the explicit description of the reduced forms depends on these choices, but their general properties, in particular the finiteness or unicity of the reduced forms in each class of algebraically equivalent forms, do not. For instance, the first family (in Poincaré's terminology) of ternary cubic forms identified by him is made of forms algebraically equivalent to the form $6\alpha xyz + \beta(x^3 + y^3 + z^3)$, with $\alpha \neq 0$,[46] chosen as the canonical form. Poincaré computes its Hessian $\Delta = 6(\beta^3 + 2\alpha^3)xyz - 6\alpha^2\beta(x^3 + y^3 + z^3)$ and the two Aronhold invariants $S = 4\alpha(\alpha^3 - \beta^3)$ and $T = 8\alpha^6 + 20\alpha^3\beta^3 - \beta^6$. The distribution of the nine inflection points on the associated cubic curve shows that a real transformation reproducing the canonical form can only exchange the three lines $x = 0$, $y = 0$, $z = 0$, i.e., that these transformations (except the identity, which is never mentioned by Poincaré) should be given by one of the following five "tables":

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Poincaré then proves that there is in general a unique reduced form arithmetically equivalent to a given form of this family and provides bounds on the coefficients of the reduced forms in terms of the invariants S and T. In the case of cubic ternary forms, this gives a new proof of a recent result of Camille Jordan, stating there are only finitely many classes of forms with integer coefficients algebraically equivalent to a given form (here the chosen canonical form).[47]

A more subtle case arises when the cubic form can be decomposed into several factors. For instance, in the case where the form represents a conic and a line which are not tangent, the canonical forms can be chosen to be $6\alpha xyz + z^3$ or $3\alpha x^2 z + 3\alpha y^2 z + z^3$, whether or not the line and the conic intersect each other.

[45] On this algebraic work, and its relation to both Hermite and Jordan's works, see Brechenmacher (2011). On the history of the classification of algebraic curves, see Lê (2023).

[46] This condition means that the equation is not reducible to a sum of three cubes.

[47] Jordan (1879). Jordan excludes the case of determinant 0.

The first canonical form is reproducible by the family of transformations with one parameter $\begin{bmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\lambda} \end{bmatrix}$. When the double points of the associated curve are imaginary (in particular when the canonical form is $3\alpha x^2 z + 3\alpha y^2 z + z^3$), there are a finite number of reduced forms, thus a finite number of classes. But when the double points are real, several cases occur, whether or not the invariant $4S$ is a fourth-power: there may be a finite number of classes or infinitely many classes, each of them containing a finite number of reduced forms.

The preliminary notes presented to the Academy of Sciences, as well as the longer memoirs on ternary cubic forms, are, as we have seen, quite explicit, giving for each family the concrete equations of canonical forms, their automorphisms, and the possible distribution of the reduced forms in arithmetical classes (and also genus). However, this did not satisfy Hermite completely.

Hermite, who presented Poincaré's note (Poincaré 1880a) in June 1880 to the Academy, wrote to him a few days earlier to fix an appointment in order to discuss his memoir. He also suggested some further readings and concluded:

> Your search for the substitutions that reproduce a given form, and the distinction between the cases where these substitutions are entirely determined or depend on one or two variable parameters, seem to me to be entirely new, and I attach great importance to them. You have seen perfectly well that there is no arithmetical question in the search for the equivalence of cubic forms unless there are an infinite number of algebraic substitutions that change them into themselves. But then we leave the field of cubic forms and the question that you had the merit of posing—an entirely new question and one that I consider very beautiful and very fruitful —is that of the simultaneous reduction, that is to say, by the same linear substitutions, and with integer coefficients, of the system of a ternary quadratic form and of a linear function. [. . .] But you must not be satisfied with having thus opened the way, you must, in reality and in fact, give the means of calculating these reduced forms, and produce numerical applications. Many things can be revealed in this way of which neither you nor anyone else has any idea, so hidden are the properties of numbers and so far beyond any prediction. It is with regard to them that observation plays an absolutely necessary role; you need elements of observation, and these elements you will be the first to have obtained and to have given.[48]

[48] Poincaré (1986, pp. 164–165): "Votre recherche des substitutions qui reproduisent une forme donnée, et la distinction des cas où ces substitutions sont entièrement déterminées ou bien dépendent d'un ou deux paramètres variables, me semblent entièrement nouvelles, et j'y attache une grande importance. Vous avez parfaitement vu qu'il n'y a de question arithmétique, dans la recherche de l'équivalence des formes cubiques qu'autant qu'il existe une infinité de substitutions algébriques qui les changent en elles-mêmes. Mais alors on quitte le champ des formes cubiques et la question que vous avez eu le mérite de poser, question entièrement neuve et que je juge très belle et très féconde, est celle de la réduction simultanée, c'est-à-dire par la même substitution linéaire, et à coefficients entiers, du système d'une forme quadratique ternaire et d'une fonction linéaire. [. . .] Mais il ne faut point vous contenter d'avoir ainsi ouvert la voie, il faut, en réalité et en fait, donner les moyens de calculer ces réduites, et faire des applications numériques. Bien des choses peuvent se révéler ainsi dont ni vous ni personne n'a eu l'idée, tant les propriétés des nombres sont cachées et en dehors de toute prévision. C'est à leur égard que l'observation joue un

Hermite's aphorism about observation is a recurrent one (Goldstein 2011), but Poincaré took the request seriously. "Following M. Hermite's advice", he wrote later, he investigated more deeply the simultaneous reduction of a quadratic and a linear form—which corresponds to one of the more complicated case alluded to above, in which the cubic form can be decomposed and which is more delicate to handle from an arithmetical point of view. Using both complex congruences and Pell-Fermat type equations, Poincaré exhibited for instance the reduced forms associated to the system $x + y + z, x^2 + 4y^2 - z^2 + 2xy + 2xz + 2yz$ or the automorphisms of the system $14x + y + 2z$, $y^2 - 6z^2$ as the powers of the transformation[49]

$$\begin{bmatrix} 1 & 5981360 & 14651280 \\ 0 & 46099201 & 112919520 \\ 0 & 18819920 & 46099201 \end{bmatrix}.$$

Hermite continued to encourage him to make new explicit calculations.

> Your result on the transformations of a system composed of a ternary [quadratic] form and a linear form is excellent, but I confess that I would have preferred that, at the cost of greater difficulty, you had been led to a new algorithm of calculation, instead of reducing the solution to the transformation into themselves of the simple binary forms. It is therefore necessary to persevere in the research which concerns only ternary forms.[50]

He had in fact a model for this, as his student Léon Charve defended a thesis on November 1880, a few months after this letter, on the reduction of ternary quadratic forms with completely effective (and extremely laborious) computations (Charve 1880). Hermite mentioned this thesis several times in his own correspondence in laudatory terms.[51] However, when Poincaré returned to this quadratic case, it was not to have the effect that Hermite wished for.

---

rôle absolument nécessaire : il faut donc des éléments d'observation, et ces éléments vous serez le premier à les avoir obtenus et donnés."

[49] Poincaré (1880b) and the developed article (Poincaré 1886a, pp. 135–142).

[50] Poincaré (1986, p. 168): "Votre résultat sur les transformations semblables d'une système composé d'une forme ternaire [quadratique] et d'une forme linéaire est excellent, mais je vous avoue que j'aurais préféré qu'au prix d'une difficulté plus grande vous eussiez été amené à un nouvel algorithme de calcul, au lieu de ramener la solution à la transformation en elles-mêmes des simples formes binaires. Il faut donc persévérer dans la recherche qui concerne les seules formes ternaires."

[51] For instance, to Thomas Stieltjes (Hermite and Stieltjes 1905, II, p. 12): "La réduction n'est point un procédé facile ni commode et il n'a rien moins fallu que le talent et l'opiniâtreté de M. Charve pour en faire application dans quelques cas particuliers, et cependant il serait si utile et même absolument indispensable de pouvoir faire de nombreuses applications, pour s'éclairer et se diriger, j'ajouterai pour s'inspirer puisqu'il s'agit d'Arithmétique" [Reduction is neither an easy nor a convenient procedure and it took nothing less than the talent and obstinacy of Mr. Charve to apply it in a few specific cases, and yet it would be so useful and even absolutely indispensable to be able to make numerous applications, to enlighten and to direct, and since we are dealing with Arithmetic I will add, to inspire ourselves.].

7.6 Back to Quadratic Forms: Fuchsian Functions and Non-Euclidean Geometry in Arithmetic

Poincaré's discovery of a link between Fuchsian functions and non-Euclidean geometry is well-known, as the mathematician used it to illustrate the art of invention in mathematics (Poincaré 1908b). It is the famous story of the omnibus:

> At the moment I set foot on the step, the idea came to me, without anything in my previous thoughts seeming to have prepared me for it, that the transformations I had used to define the Fuchsian functions are identical to those of non-Euclidean geometry.[52]

The scene probably took place in June 1880,[53] at a time when Poincaré was working on the classification of cubic ternary forms, and more specifically on the case when the form is composed of a linear and a quadratic factor, which led him again to ternary quadratic forms. As is well-known, he was involved simultaneously in the writing of a contribution to the 1880 prize in mathematics of the French Academy of Sciences, on differential equations, and in his own creation, description and classification of specific Fuchsian functions and of their associated transformations.[54]

I would like to underline once more the close relations, at several levels, between these works and the way Poincaré transfers methods, intuitions and objets from one topic to another, as we have already seen for the invariants of quadratic forms or the classification of cubic ones. In his very first note on Fuchsian matters, Poincaré defines a Fuchsian function as a uniform function on the plane which is reproduced by a discontinuous subgroup of the homographic transformations on the unit disk. He then remarks that some of these subgroups are isomorphic to groups of linear transformations with integer coefficients that reproduce an indefinite ternary form with integer coefficients, concluding that this "highlights the intimate links between number theory and the analytical question in hand" (Poincaré 1881a, p. 335). Indeed the omnibus story continues on, displaying influences in both directions:

> On my return to Caen [...] I then began to study questions of Arithmetic without much result and without suspecting that this could have the slightest connection with my previous research. Disgusted with my failure, I went to spend a few days at the seaside and thought of other things. One day, while walking on a cliff, the idea came to me, always with the same characteristics of brevity, suddenness and immediate certainty, that the arithmetic transformations of the ternary indefinite quadratic forms are identical to those of non-Euclidean geometry. Back in Caen, I meditated on this result and drew the consequences:

[52] Poincaré (1908b, p. 363): "Au moment où je mettais le pied sur le marche-pied, l'idée me vint, sans que rien de mes pensées antérieures parut m'y avoir préparé, que les transformations dont j'avais fait usage pour définir les fonctions fuchsiennes sont identiques à celles de la Géométrie non-euclidienne."

[53] See Gray (2013, pp. 216–217).

[54] We will not restate the details of this episode, which has been thoroughly studied, in particular by Jeremy Gray (2000, 2013, ch. 3).

> the example of quadratic forms showed me that there were Fuchsian groups which are different from those corresponding to the hypergeometric series.[55]

Poincaré was here perfectly in line with several of his fellow mathematicians, in particular the contemporary research of Émile Picard on substitutions with 3 variables on the hypersphere or Camille Jordan's study of the groups of motions, which Jordan directly connected to Bravais's works.[56] While the appearance of non-Euclidean geometry in the story is a more spectacular feature, a close reading of all these papers suggests that the explicit writing out of the various transformations used in these different situations was a driving element and a decisive factor in favoring the thematic rapprochements.

As Poincaré explained at the Algiers meeting of the French Association for the Advancement of Science where he presented his new viewpoint in April 1881 (Poincaré 1882b), his point of departure was Hermite's method for ternary quadratic forms (Hermite 1854). In 1854, Hermite had studied the reduction of indefinite ternary quadratic forms—the forms algebraically equivalent to, say, $X^2+Y^2-Z^2$—by a variant of his technique of continuous reduction. To such a form f, Hermite associated a family of *definite* ternary quadratic forms

$$\phi(x, y, z) = f(x, y, z) + 2(\lambda x + \mu y + \nu z)^2,$$

λ, μ, ν real numbers with some suitable conditions. Each ϕ could be reduced by a suitable transformation, by the general theory of definite quadratic forms. Hermite then applied this transformation to the initial f and thus obtained, by varying ϕ, a family of transformed forms, which he considered as the reduced forms of the initial form f. He showed that the coefficients of these reduced forms satisfy certain bounds; in particular, if the coefficients of the initial f are integers, it implies that there are a finite number of such reduced forms. Hermite also studied the automorphisms of the form, proving for instance that what we now call their characteristic equation has solutions of the type $\pm 1, l, \frac{1}{l}$.[57]

Poincaré followed exactly the same path in 1881. He associated to the indefinite form $F(x, y, z) = (ax + by + cz)^2 + (a'x + b'y + c'z)^2 - (a''x + b''y + c''z)^2 = \xi^2 + \eta^2 - \zeta^2$ the definite forms $\xi^2 + \eta^2 - \zeta^2 + 2(\xi_1\xi + \eta_1\eta - \zeta_1\zeta)^2$, with ξ_1, η_1, ζ_1

[55] Poincaré (1908b, p. 363): "De retour à Caen [...] je me mis alors à étudier des questions d'Arithmétique sans grand résultat apparent et sans soupçonner que cela pût avoir le moindre rapport avec mes recherches antérieures. Dégoûté de mon insuccès, j'allai passer quelques jours au bord de la mer et je pensai à tout autre chose. Un jour, en me promenant sur une falaise, l'idée me vient, toujours avec les mêmes caractères de brièveté, de soudaineté et de certitude immédiate, que les transformations arithmétiques des formes quadratiques ternaires indéfinies sont identiques à celles de la Géométrie non-euclidienne. Étant revenu à Caen, je réfléchis sur ce résultat, et j'en tirai les conséquences ; l'exemple des formes quadratiques me montrait qu'il y a des groupes fuchsiens autres que ceux qui correspondent à la série hypergéométrique."

[56] The analogies and differences with Jordan, in particular with respect to the concept of group, are discussed in Brechenmacher (2011).

[57] The description of the automorphisms was completed by several authors c. 1870, in particular by Georg Cantor in his Habilitationschrift, by Paul Bachmann, and by Hermite himself.

satisfying the condition (analogous to that of Hermite) $\xi_1^2 + \eta_1^2 - \zeta_1^2 = -1$. Again, he used the transformations reducing these definite forms, applying them in turn to F to get what he defined as the reduced (forms) for F.

As he noted, however, since $\xi_1^2 + \eta_1^2 - \zeta_1^2 = -1$, the point with coordinates $\frac{\xi_1}{\zeta_1+1}, \frac{\eta_1}{\zeta_1+1}$ is inside the unit disk. To each definite form of the family is then associated such a point, as well as a reduced form. When the parameters ξ_1, η_1, ζ_1 change, the point moves inside the disk. However, the reduced form remains the same so long as the point lies inside a certain region of the disk, then it changes. The transformations providing the reduction can be then studied geometrically, by looking at the corresponding regions delimited inside the disk.[58] To do this, Poincaré used non-Euclidean geometry on the disk, more specifically, a non-Euclidean description of the tessellation of the disk in domains delimited by polygons. This approach, as might be expected, was not to the taste of Hermite who asked Poincaré several times to reformulate his results:

> In renewing my request to you to present your results on the classification of functions $\frac{az+b}{cz+d}$ in order to obtain the elements of the formation of the Fuchsian functions, without resort to the use of non-Euclidean geometry, and after having presented them by the method by which you discovered them, I beg you, Sir, to receive the renewed assurance of my highest esteem for your work and of my most devoted sentiments.[59]

The explicit, detailed, connection with the Fuchsian groups was presented only a few years later, Poincaré choosing at that time another expression of the canonical ternary quadratic form (Poincaré 1886b):

> An indefinite ternary quadratic form may always be written ... in the following way:
>
> $$F(x, y, z) = Y^2 - XZ,$$
>
> where
>
> $$X = ax + by + cz, \quad Y = a'x + b'y + c'z, \quad Z = a''x + b''y + c''z,$$
>
> a, b, c being arbitrary real numbers.

Let now $\alpha, \beta, \gamma, \delta$ be four real numbers such that $\alpha\delta - \beta\gamma = 1$. Poincaré introduces the transformations:

$$X' = \alpha^2 X \qquad + 2\alpha\gamma Y \quad + \gamma^2 Z$$

[58] In other words, as Châtelet explains in a footnote, Poincaré studies the fundamental domain of the automorphisms, seen as homographic transformations.

[59] For instance, Poincaré (1986, p. 174): "En vous renouvelant la prière de présenter sans recourir à l'emploi de la géométrie non euclidienne, après les avoir exposés par la méthode qui vous les a fait découvrir, vos résultats sur la classification des fonctions $\frac{az+b}{cz+d}$ afin de posséder les éléments de la formation des fonctions fuchsiennes, je vous prie, Monsieur de recevoir la nouvelle assurance ma plus haute estime pour vos travaux et de mes sentiments bien dévoués." Hermite was not hostile to all geometrical arguments; he did not complain about Poincaré's lattices or Hermann Minkowski's geometry of numbers (also based on lattices). But for him, non-Euclidean geometry was not helpful in representing analytical facts, see Goldstein (2011).

$$\begin{aligned} Y' &= \alpha\beta X &+ (\alpha\delta + \beta\gamma)Y &\quad + \gamma\delta Z \\ Z' &= \beta^2 X &+ 2\beta\delta Y &\quad + \delta^2 Z \end{aligned}$$

If $X = ax' + by' + cz'$, $\quad Y = a'x' + b'y' + c'z'$, $\quad Z = a''x' + b''y' + c''z'$, it is then easy to check that the transformation changing x, y, z into x', y', z' leaves F invariant. If the coefficients of F and the α, β, γ, δ are integers, these transformations form a discontinuous group and the associated substitutions $z \to \frac{\alpha z+\beta}{\gamma z+\delta}$ form a Fuchsian group. In a longer memoir, Poincaré emphasizes the particular properties of these arithmetically-defined Fuchsian groups, in particular the algebraic relations satisfied by the associated Fuchsian functions, analogous to those already known for elliptic and modular functions (Poincaré 1887).

7.7 Classification Again

In the wake of his research on the classification of forms, Poincaré also devoted two short notes to a generalization of the tools Gauss had introduced for his own refined classification of binary quadratic forms, "order" and "genus", Poincaré (1882d).[60] In the *Disquisitiones arithmeticae*, two classes of binary quadratic forms, represented for example by the forms $ax^2 + 2bxy + cy^2$ and $a'x^2 + 2b'xy + c'y^2$, are said to belong to the same order if the g.c.d. of (a, b, c) is equal to the g.c.d. of (a', b', c') and the g.c.d. of $(a, 2b, c)$ is equal to the g.c.d. of $(a', 2b', c')$. Following a proposal of Eisenstein for ternary forms, for forms of higher degree or with more variables, Poincaré imposes equality conditions not only on the coefficients of the forms, but also on those of some of their invariants and covariants. For instance, for the binary cubic form $f = ax^3 + 3bx^2y + 3cxy^2 + dy^3$, the Hessian $6(ac - b^2)x^2 + 6(ad - bc)xy + 6(bd - c^2)y^2$ should be taken into account; the order is determined by four quantities, the g.c.d of a, b, c, d, the g. c. d. of $a, 3b, 3c, d$, the g.c.d. of $ac - b^2, ad - bc, bd - c^2$ and the g.c.d. of $2(ac - b^2), ad - bc, 2(bd - c^2)$.

As for the genus, its definition in the *Disquisitiones arithmeticae* relied on the following fact: for a binary quadratic form $ax^2 + 2bxy + cy^2$, such that g.c.d.(a, b, c)=1, and a prime factor p of its determinant $ac - b^2$, two integers which are represented by the form are both quadratic residues modulo p or both non-quadratic residues. The various cases ("characters") for the various p then define the genus of the form (in fact, of its whole class). Poincaré defined the equivalence of two forms f and f' according to a modulus m when there exists a linear transformation T with integer coefficients and determinant $\equiv 1 \pmod{m}$ such that $f \circ T \equiv f' \pmod{m}$. Two (algebraically equivalent) forms are then in the same

[60] Several authors, in particular Eisenstein and Dirichlet, had contributed to simplifications, reformulations and partial extensions of these notions, which also appear in other domains of mathematics, see Lê (2023).

genus if they are equivalent according to all moduli. Again, in these notes, Poincaré gives only very general statements, without proofs, and illustrates his definitions with a few numerical examples. He computes in particular the distribution of binary cubic forms according to moduli 2, 3 and 5, but he does not seem to have gone further in the 1880s in this attempt of a classification of higher-degree forms. In particular, he does not appear to have then seen Eisenstein's suggestion, developed in particular by Smith in the 1860s, and then much later by Minkowski, of defining the genus by means of transformations with *rational* coefficients (Smith 1861–1865; Dickson 1919).

7.8 Poincaré's Arithmetic Revisited

We have tried to show that Poincaré's arithmetical work is highly coherent as soon as one restores the collective program in which it is embedded, i.e. the disciplinary configuration of the theory of forms in the last third of the nineteenth century. With its own questions, concepts and resources, it largely guided Poincaré's objectives, in the perspective of the classification of forms: to find well-chosen, preferably effective, invariants; to identify adequate representatives of classes (and of other levels of classification), such as the canonical forms of the algebraic classification and the reduced forms of the arithmetic one; to explain the operations allowing the transformation of each form into its representatives; and to study the automorphisms of a form. Conversely, we have seen that the recourse to ideal numbers did not indicate a change of discipline (for instance, as David Hilbert would define it in his own presentation of the theory of algebraic numbers), but an attempt to integrate (or even disintegrate …) these concepts (whose importance Poincaré clearly perceived) in the disciplinary framework of forms.

We have seen this program at work in all his early research, including the famous papers linking quadratic forms, Fuchsian functions and non-Euclidean geometry. What shaped these papers was also reflected in his later, more famous, work. His 1905 article, on the centenary of Dirichlet's birth, is thus in many respects, a microcosm of the larger mathematical world we have just presented. Its results, as their recent commentator Nicolas Bergeron describes them, may seem disparate, the only clear arithmetical application being a new proof of a well-known formula of Dirichlet on the number of classes of forms. But, as before, essential and varied analytical tools (in particular those linked to automorphic functions) were mobilized to search for invariants of linear and quadratic forms (Bergeron 2018).

As for Poincaré's celebrated memoir of 1901, today's readers see it as one of the main origins of so-called Diophantine geometry and focus on the way Poincaré, with the help of the parametrization of cubic curves by elliptic functions, defined more or less adequately the rank of (the group of points with rational coordinates on) an elliptic curve.[61] However, our study provides another context for this article. In 1880, Poincaré had employed a geometric interpretation of ternary forms, considering them as defining equations of plane (algebraic) projective curves. It allowed him to link the classifications of curves (by their invariants, singular points and other geometric characteristics) with those of forms (by linear transformations). In 1901, Poincaré proposed a new classification of algebraic curves directly inspired by the theory of forms, one based on birational transformations between curves. The link to the mode of classification of forms is explicit, Poincaré refers directly to the *Disquisitiones arithmeticae* as providing the principles to classify conics according to his own program, principles that he applies to higher-degree curves (Goldstein and Schappacher 2007b, pp. 95–96).[62] And it is indeed not to the structure of points on a cubic, but to the study of birational transformations between curves defined by equations of different degrees that the major part of Poincaré's memoir is devoted, as well as to the reduction of any algebraic curve to a curve defined by equations of the lowest possible degree—coherent with the guidelines of the theory of forms.

In this respect, the presentation of arithmetic in Poincaré's lecture on the future of mathematics is illuminating :

> Among the words which have had the happiest influence, I would mention "group" and "invariant". [...] Progress in arithmetic has been slower than that in algebra and analysis and it is easy to understand the reasons. The feeling for continuity is a precious guide which the arithmetician lacks [...] [He] must therefore take analogies with algebra for his guide [...] The theory of forms, and in particular that of quadratic forms, is intimately bound to the theory of ideals. One of the earliest to take form among arithmetic theories, it arose with the successful introduction of unity through the use of linear transformation groups. These transformations have allowed a classification with its consequent introduction of order.[63]

This coherence also manifests itself on a more subterranean level, that of practices. As the word has been widely used recently in the philosophy of mathematics, let me specify that I use it here in a rather informal way, to designate a concrete way to carry on an activity (as opposed to official rules or principles, and to theory). They have to do with "real individuals, their actions and their material conditions

[61] See Weil (1955), Gray (2013, pp. 486–488). The parametrization was already widely used, see Schappacher (1991), Lê (2018). Norbert Schappacher (1991) has discussed the problems raised by Poincaré's definition of rank. On the geometrical viewpoint of Poincaré 1901 paper, see Schneider (2000).

[62] We do not know if Minkowski's use of transformations with rational coefficients to define the genus of forms then played a role in Poincaré's conception. Nor does Poincaré mention contemporary work on birational geometry, even that connecting it with Diophantine equations.

[63] Poincaré (1908a, p. 175, pp. 179–180): "Les progrès de l'Arithmétique ont été plus lents que ceux de l'Algèbre et de l'Analyse, et il est aisé de comprendre pourquoi. Le sentiment de la continuité est un guide précieux qui fait défaut à l'arithméticien. [...] L'arithméticien doit donc prendre pour guide les analogies avec l'Algèbre. [...] La théorie des formes, et en particulier celle des formes quadratiques, est intimement liée à celle des idéaux. Si parmi les théories arithmétiques elle a été l'une des premières à prendre figure, c'est quand on est parvenu à y introduire l'unité par la considération des groupes de transformations linéaires. Ces transformations ont permis la classification et par conséquent l'introduction de l'ordre."

of life"[64] In mathematics, practices can thus be attached to the pervasive use of a certain tool or technique, or to a way of reading the articles of other mathematicians or of publishing one's own work or of exchanging mathematical information. They can also be detected through an epistemic privilege attached to specific features, like effectivity or proofs, or a recurrent representational device, be it diagrams or lattices.

Poincaré's way of practising mathematics in his early arithmetical work displays a striking mixture of a particularly vague mode of writing and of an impressive mobilisation of ideas and techniques from several branches of mathematics. The second point has been obvious on several occasions here, with the recourse to several kinds of geometry or to a large variety of analytical tools. On the first point, let us note, for example, that necessary hypotheses are often missing—the examples include the irreducibility of an algebraic equation under scrutiny, the non-vanishing of certain expressions like the determinant or the fact that some of his ideals are not defined over the ring of integers, but only over a subring. As seen in the section on quadratic forms and lattices, Poincaré may also use the same symbols to designate different things in the same article[65]. Browsing through his articles gives the impression of flying over a vast textual landscape (in the vernacular) with the occasional example or calculation serving as anchor points. More than the precise statements and detailed proofs we are now used to, computations of examples are the warrants of the solidity of Poincaré's whole construction.[66]

It is also quite tricky to identify Poincaré's sources of inspiration—he often quotes some predecessors in a general way at the beginning of his text, very rarely for a specific result inside the text (we have nonetheless seen some references to Bravais or Eisenstein). As those close to him sometimes explained after his death, Poincaré was particularly gifted for roughly grasping ideas or problems and

[64] Marx and Engels (1846/1969, p. 20): "Es sind die wirklichen Individuen, ihre Aktion und ihre materiellen Lebensbedingungen." Or, as Michel Foucault writes (Foucault 1982/2001, p. 1039): "L'on tient plus aux manières de voir, de dire, de faire et de penser qu'à ce qu'on voit, qu'à ce qu'on pense, qu'à ce qu'on dit" ("We care more about the ways of seeing, saying, doing and thinking than about what we see, what we think, what we say"). On this issue, see among many others, Bourdieu (1994), Lepetit (1995), Chateauraynaud and Cohen (2016)."

[65] As mentioned earlier, Gaston Darboux wrote to Poincaré in 1878 about his thesis: "I still believe that we will make a good thesis out of it, but it seems essential to me to recast the writing and to correct all the errors of calculation or the changes of notation which make it almost unreadable." ["Je persiste à croire que nous en ferons une bonne thèse, mais il me parait indispensable de fondre la rédaction et de corriger toutes les erreurs de calcul ou les changements de notation qui la rendent presque illisible."] (Poincaré 1986, p. 132).

[66] Poincaré is almost describing his own practice when, advising Mittag-Leffler on the translation of Georg Cantor's memoirs on set theory, he writes (Hermite and Mittag-Leffler 1984, p. 278): "To make it accessible, it would be necessary to give a few specific examples after each definition and then put the definitions at the beginning instead of at the end" ("Il faudrait pour la rendre accessible donner quelques exemples précis à la suite de chaque définition et puis mettre les définitions au commencement au lieu de les mettre à la fin.")

then integrating them into his own framework.[67] Moreover, several correspondents pointed out to Poincaré that such and such a result had already appeared in one of the sources that he mentioned. However, he was also one of the rare authors (French or not) to mention Dedekind's theory of ideals in the 1880s and he quoted and relied on numerous German authors; we have mentioned Gauss of course, but also Eisenstein, Dirichlet, Selling, Hesse, Steiner.... Weil's assertion in this respect seems a little misleading—or perhaps a little anachronistic, in that he seems to be referring to Poincaré's neglect of what will be considered in the interwar period only as "the royal road" to a structuralist point of view.[68]

An obvious source, however, is Hermite, whose influence operates at several levels, in addition to the direct interactions we have already mentioned. Like Poincaré in 1880, Hermite had revisited the classical results of the classification of forms of the *Disquisitiones arithmeticae*, in the light of his procedure of continuous reduction (Hermite 1851). At another level, the emphasis on (linear) transformations is one of the characteristics of Hermite's work during his whole career. His use of "tableaux" (our matrices) to work out transformations is pervasive, as it is in Poincaré's work.[69]. The reduction of transformations, in particulier, is carried out on these "tableaux", playing a key role for both mathematicians. Poincaré also took from Hermite the idea that decomposable forms constitute a fruitful entry into the study of algebraic numbers. Some specific constructions were directly borrowed from Hermite's articles: for example, Poincaré followed and generalized the approach Hermite had introduced to factorize into complex factors prime numbers congruent to 1 modulo 5 or 7 (Hermite 1850; Goldstein 2007).

Two other instructive shared features deserve to be highlighted. First, the importance of reduced forms in their scheme of work. Reduced forms are particular representatives of classes (sets of forms connected by suitable linear transformations). The later, structural, viewpoint would privilege classes, which are intrinsic. Poincaré, like Hermite, was perfectly aware that several (rather arbitrary) choices were possible for the reduced forms; indeed, he modified his choice, for example, in the course of his research on ternary quadratic forms. This freedom of choice, however, like that of the "tableaux" (whose writing depends on a choice of generators), favors calculation. Hermite is quite explicit about his predilection (Goldstein 2011) and Poincaré-, who nevertheless built, as we have shown, an arithmetic of lattices, followed him on this point. This can be seen in particular in what Poincaré called a classification: his are not based on classes per se but

[67] The contrast with Châtelet's painstaking corrections and complements of Poincaré's memoirs is in this respect quite striking.

[68] Despite Hermite's protests, the rumor that Poincaré did not know or mention German sources spread through Mittag-Leffler, in particular among French students in Germany at the time of the rivalry between Poincaré and Klein around automorphic functions; see for instance (Hermite and Mittag-Leffler 1984, I, pp. 129, 251).

[69] This has already been underlined on the basis of algebraic works in the same period, in particular Poincaré's 1884 paper on complex numbers (Brechenmacher 2011).

on the construction of specific, and in principle, calculable characteristics such as invariants.

Another point that brought the two mathematicians together was their vision of a larger research field that would merge arithmetic, algebra and analysis, and exclude the disciplinary purity which was at the time defended by many mathematicians, such as Edouard Lucas or Leopold Kronecker. On the contrary, the use of continuous tools in arithmetic was favored and praised by Hermite as well as by Poincaré. We have emphasized this direction here, but, reciprocally, the search for automorphisms, for instance, was exported into the study of Fuchsian or Abelian functions, as well as differential equations.[70] Poincaré's famous sentence—"The only natural object of mathematical thought is the integer"—might thus lead to a misinterpretation if read in isolation. It is in fact a mere concession to the defenders of a pure number theory, stripped of its analytical tools, a concession immediately corrected into a promotion of a unified field of mathematics.[71]

> The only natural object of mathematical thought is the integer. [... We] have devoted almost all our time and energy to the study of the continuous. Who will regret it? Who will believe that this time and these efforts have been wasted? Analysis unfolds for us infinite perspectives that arithmetic does not suspect, it shows us at a glance a grandiose whole, the order of which is simple and symmetrical; on the contrary, in the theory of numbers, where the unforeseen reigns, the view is, so to speak, blocked at every turn. [...] Let us be grateful to the continuum which, if everything comes out of the whole number, was alone capable of bringing out so much. Need I remind you, moreover, that M. Hermite drew a surprising advantage from the introduction of continuous variables into the theory of numbers? Thus, the proper domain of the whole number is itself invaded, and this invasion has restored order where disorder reigned.[72]

In Poincaré, however, the continuous is not restricted to the theory of functions. It extends to geometric representations or even geometric techniques, themselves borrowed from several branches of mathematics, from Bravais's theory of polyhedra

---

[70] For other examples, see Brechenmacher (2008, 2011), de Saint-Gervais (2011).

[71] Hermite repeated on several occasions that the theory of numbers is only an anticipation of the theory of elliptic functions. For the importance he attached to the use of analytical tools, see Goldstein (2007, 2011). It should be noted that Hermite himself, a priori unaware of advances in geometry, sometimes used elementary geometric representations.

[72] Poincaré (1897): "Le seul objet naturel de la pensée mathématique, c'est le nombre entier. [... N]ous avons consacré à l'étude du continu presque tout notre temps et toutes nos forces. Qui le regrettera? Qui croira que ce temps et ces forces ont été perdus? L'analyse nous déroule des perspectives infinies que l'arithmétique ne soupçonne pas; elle nous montre d'un coup d'œil un ensemble grandiose, dont l'ordonnance est simple et symétrique; au contraire, dans la théorie des nombres, où règne l'imprévu, la vue est pour ainsi dire arrêtée à chaque pas. [... S]oyons reconnaissants au continu qui, si tout sort du nombre entier, était seul capable d'en faire tant sortir. Ai-je besoin, d'ailleurs, de rappeler que M. Hermite a tiré un parti surprenant de l'introduction des variables continues dans la théorie des nombres? Ainsi, le domaine propre du nombre entier est envahi lui-même, et cette invasion a rétabli l'ordre là où régnait le désordre."

and lattices to that of projective curves or to non-Euclidean geometry.[73] This justifies his well-known reputation as one of the last universalist mathematicians. But what is striking when reading his early work on arithmetic is his professionalism (all the more paradoxical for us who are now used to a very different writing style); his mastery of both the disciplinary issues and the tools available, his ability to intervene effectively in order to fill in all the gaps in a program, rarely explained in detail, but whose reconstruction allows us to see that Poincaré had identified its stakes and components perfectly.[74]

The historian Gil Bartholeyns suggests that:

> The evolution of the object of history during the twentieth century can be described as the change from the extraordinary (the particular, the unique) to the ordinary (the collective, the structural, the trivial). In place of the exceptional individuals, the chefs-d'oeuvre, the memorable events, [the historians] have preferred the forgotten, the unpretentious documents, the repetitive and shared dimensions of existence.[75]

As I have tried to show here, the forgotten and the repetitive may also draw a path to a better understanding of the exceptional.

Acknowledgments This text has benefited greatly from careful review and helpful suggestions for which I am deeply grateful. For them I would like to warmly thank Jinze Du, Lizhen Ji, François Lê, Juan Li, Jim Ritter, Norbert Schappacher, Erhard Scholz and Jiwei Zhao.

References

Bachmann, Paul. 1898. *Die Arithmetik der quadratischen Formen*. Leipzig: Teubner.

Bartholeyns, Gil. 2010. Le paradoxe de l'ordinaire et l'anthropologie historique. *L'Atelier du Centre de recherches historiques* 6: http://journals.openedition.org/acrh/1928.

Bergeron, Nicolas. 2018. Les "invariants arithmétiques" de Poincaré. *Graduate Journal of Mathematics* 3 (1): 1–14.

Boucard, Jenny, Catherine Goldstein, and Valéry Malécot. 2024. Dispositions discrètes: transferts entre botanique et cristallographie dans les travaux d'Auguste Bravais. In *Arranger, disposer,*

[73] Châtelet also underlines Poincaré's hope, again like Hermite, in the specific approach of Hermann Minkowski with his geometry of numbers (Poincaré 1908a). Minkowski's approach would be regularly integrated into the arithmetical work developed in France before the First World War, in particular by Châtelet himself (Gauthier 2009, 2011). On the importance of geometry in Poincaré's arithmetical works, see Schneider (2000).

[74] This situation can be compared to that described for Albert Einstein's famous articles of 1905, also often perceived as isolated, but whose coherence can be restored (Rynasiewicz and Renn 2005).

[75] Bartholeyns (2010): "L'évolution de l'objet de l'histoire au XXe siècle peut être décrite comme le passage de l'extraordinaire (le particulier, l'unique) à l'ordinaire (le collectif, le structurel, le banal). Aux individus exceptionnels, aux chefs-d'œuvre et aux événements mémorables, on a donné préférence aux oubliés, aux documents sans prétention, aux dimensions répétitives et partagées de l'existence."

combiner: Théories de l'ordre dans les sciences, les arts d'ornement et la philosophie (1770-1910), ed. Jenny Boucard and Christophe Eckes, pp. 161–208. Paris: Hermann.
Bourdieu, Pierre. 1994. *Raisons pratiques: sur la théorie de l'action*. Paris: Seuil.
Bravais, Auguste. 1850. Mémoire sur les systèmes formés par des points distribués régulièrement sur un plan ou dans l'espace. *Journal de l'École polytechnique* 19: 1–128.
Bravais, Auguste. 1851. Études cristallographiques. *Journal de l'École polytechnique* 20: 101–278.
Bravais, Auguste. 1866. *Études cristallographiques*, ed. Léonce Élie de Beaumont. Paris: Gauthier-Villars.
Brechenmacher, Frédéric. 2008. La controverse de 1874 entre Camille Jordan et Leopold Kronecker. *Revue d'histoire des mathématiques* 2 (13): 187–257.
Brechenmacher, Frédéric. 2011. Autour de pratiques algébriques de Poincaré : héritages de la réduction de Jordan. Preprint, https://hal.archives-ouvertes.fr/hal-00630959.
Brechenmacher, Frédéric. 2016. Algebraic generality vs arithmetic generality in the controversy between C. Jordan and L. Kronecker (1874). In *The Oxford Handbook of Generality in Mathematics and the Sciences*, ed. Karine Chemla, Renaud Chorlay, and David Rabouin, 433–467. Oxford: Oxford University Press.
Charve Léon. 1880. De la réduction des formes quadratiques ternaires et de leur application aux irrationnelles du 3e degré. *Annales scientifiques de l'École normale supérieure* 2e s. 9: 3–156.
Chateauraynaud, Francis, and Yves Cohen (dir.). 2016. *Histoires pragmatiques*. Paris: Editions EHESS.
Crilly, Tony. 1986. The Rise of Cayley's Invariant Theory (1841–1862). *Historia Mathematica* 15: 241–254.
de Saint-Gervais, Henri Paul. 2011. *Uniformisation des surfaces de Riemann : Retour sur un théorème centenaire*. Lyon: ENS Éditions.
Dedekind, Richard. 1876. Sur la théorie des nombres entiers algébriques. *Bulletin des sciences mathématiques* 11: 278–288, 310.
Dedekind, Richard. 1877. Sur la théorie des nombres entiers algébriques. *Bulletin des sciences mathématiques et astronomiques* 2e s. 1: 17-41, 69–92, 144–164, 207–248.
Dickson, Leonard Eugene. 1919. *History of the Theory of numbers,* vol. III*: Quadratic and Higher Forms*. Washington: Carnegie Institute. Repr. New York: Chelsea, 1952.
Dirichlet, Peter Gustav Lejeune-. 1850. Über die Reduction der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen. *Journal für die reine und angewandte Mathematik* 40: 209–227.
Edwards, Harold. 1977. *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*. Berlin: Springer.
Edwards, Harold. 1980. The Genesis of Ideal Theory. *Archive for the History of Exact Sciences* 23: 321–378.
Edwards, Harold. 1992. Mathematical Ideas, Ideals and Ideology. *The Mathematical Intelligencer* 14: 6–19.
Edwards, Harold. 2007. Composition of Binary Quadratic Forms and the Foundations of Mathematics. In *The Shaping of Arithmetic after C. F. Gauss's Disquisitiones Arithmeticae*, ed. Catherine Goldstein, Norbert Schappacher, and Joachim Schwermer, 129–144. Heidelberg, Berlin: Springer.
Ferreiros, Jose. 2007. The Rise of Pure Mathematics as Arithmetic with Gauss. In *The Shaping of Arithmetic after C. F. Gauss's Disquisitiones Arithmeticae*, ed. Catherine Goldstein, Norbert Schappacher, and Joachim Schwermer, 235–268. Heidelberg, Berlin: Springer.
Fisher, Charles. 1966. The Death of a Mathematical Theory. *Archive for the History of Exact Sciences* 3: 137–159.
Foucault, Michel. 1982/2001. *Dits et Ecrits*. Paris : Gallimard.
Gauss, Carl Friedrich. 1801. *Disquisitiones Arithmeticae*. Leipzig: Fleischer. English transl. by Arthur Clarke, New Haven: Yale University Press, 1966.
Gauss, Carl Friedrich. 1831. Recension der Untersuchungen über die Eigenschaften der positiven ternären quadratischen Formen von Ludwig August Seeber. *Göttingische gelehrte Anzeigen* 2: 1065-1077.

Gauthier, Sébastien. 2009. La géométrie dans la géométrie des nombres : histoire de discipline ou histoire de pratiques à partir des exemples de Minkowski, Mordell et Davenport. *Revue d'histoire des mathématiques* 15 (2): 183–230.

Gauthier, Sébastien. 2011. Justifier l'utilisation de la géométrie en théorie des nombres : des exemples chez C.F. Gauss et H. Minkowski. *Justifier en mathématiques*, ed. Domminique Flament and Philippe Nabonnand: 103–128. Paris : Éditions de la Maison des sciences de l'homme.

Goldstein, Catherine. 1994. La théorie des nombres dans les *Notes aux Comptes rendus de l'Académie des sciences* (1870–1914): un premier examen. *Rivista di Storia della scienza*, II, 2, 2:137–160.

Goldstein, Catherine. 1999. Sur la question des méthodes quantitatives en histoire des mathématiques: le cas de la théorie des nombres en France (1870–1914). *Acta historiae rerum naturalium nec non technicarum* New series 3-28: 187–214.

Goldstein, Catherine. 2007. The Hermitian Form of Reading the *Disquisitiones*. In *The Shaping of Arithmetic after C. F. Gauss's Disquisitiones Arithmeticae*, ed. Catherine Goldstein, Norbert Schappacher, and Joachim Schwermer, 377–410. Heidelberg, Berlin: Springer.

Goldstein, Catherine. 2011. Un arithméticien contre l'arithmétisation : les principes de Charles Hermite. *Justifier en mathématiques*, ed. Dominique Flament and Philippe Nabonnand, 129–165. Paris: MSH.

Goldstein, Catherine, and Norbert Schappacher. 2007a. A Book in Search of A Discipline (1801–1860). In *The Shaping of Arithmetic after C. F. Gauss's Disquisitiones Arithmeticae*, ed. Catherine Goldstein, Norbert Schappacher, and Joachim Schwermer, 3–65. Heidelberg, Berlin: Springer.

Goldstein, Catherine and Norbert Schappacher. 2007b. Several Disciplines and a Book (1860–1900). In *The Shaping of Arithmetic after C. F. Gauss's Disquisitiones Arithmeticae*, ed. Catherine Goldstein, Norbert Schappacher, and Joachim Schwermer, 67–103. Heidelberg, Berlin: Springer.

Gray, Jeremy. 2000. *Linear differential equations and Group theory from Riemann to Poincaré*. 2nd ed. Boston: Birkhäuser.

Gray, Jeremy. 2006. A History of Prizes in Mathematics. *The Millennium Prize Problems*, ed. James Carlson, Arthur Jaffe, and Andrew Wiles, 3–30. Cambridge: CMI/AMS.

Gray, Jeremy. 2013. *Henri Poincaré: A Scientific Biography*. Princeton: Princeton University Press.

Guntau, Martin, and Hubert Laitko. 1987. *Der Ursprung der modernen Wissenschaften: Studien zur Entstehung wissenschaftlicher Disziplinen*. Berlin: Akademie Verlag.

Haffner, Emmylou. 2014. D'un point de vue rigoureux et parfaitement général : pratique des mathématiques rigoureuses chez Richard Dedekind. *Philosophia Scientiae* 18: 131–156.

Haffner, Emmylou. 2019. From modules to lattices: Insight into the genesis of Dedekind's Dualgruppen. *British Journal for the History of Mathematics* 34 (1): 23–42.

Haubrich, Ralf. 1992. *Zur Entstehung der algebraischen Zahlentheorie Richard Dededkinds*. Dissertation. Göttingen: Georg-August-Universität.

Hermite, Charles. 1850. Lettres à M. Jacobi sur différents objets de la théorie des nombres. *Journal für die reine und angewandte Mathematik* 40: 261–315.

Hermite, Charles. 1851. Sur l'introduction des variables continues dans la théorie des nombres. *Journal für die reine und angewandte Mathematik* 41: 191–216.

Hermite, Charles. 1854. Sur la théorie des formes quadratiques ternaires indéfinies. *Journal für die reine und angewandte Mathematik* 47: 307–312.

Hermite, Charles, and Gösta Mittag-Leffler. 1984. Lettres à Gösta Mittag-Leffler, [ed. P. Dugac]. *Cahiers du séminaire d'histoire des mathématiques* (1 part. 1874–1883) 5: 49–285; (2nd part.1884–1891): 6 (1985), 69-217; (3rd part. 1891–1900): 10 (1989), 1–82.

Hermite, Charles, and Thomas Stieltjes. 1905. *Correspondance*, ed. Benjamin Baillaud and Henri Bourget. Paris: Gauthier-Villars.

Jordan, Camille. 1868–1869. Sur les groupes de mouvements. *Annali di matematica* s. 2, 2: 167–215, 322–345.

Jordan, Camille. 1879. Sur l'équivalence des formes algébriques. *Comptes rendus hebdomadaires des séances de l'Académie des sciences* 88: 906–908.

Knight, David. 1981. *Ordering the World: A History of Classifying Man*. London: Burnett Books.

Lê, François. 2018. Le paramétrage elliptique des courbes cubiques par Alfred Clebsch. *Revue d'histoire des mathématiques* 24(1): 1–39.

Lê, François. 2023. Des taxons et des nombres : quelques remarques sur les ordres, classes et genres des courbes algébriques. *Revue d'histoire des sciences*, à paraître.

Lê, François, and Anne-Sandrine Paumier. 2016. De la science comme classification à la classification comme pratique scientifique. *Cahiers François Viète*, s. III, 1: 9–34.

Lepetit, Bernard (dir.). 1995. *Les formes de l'expérience*. Paris: Albin Michel.

Marx, Karl, and Friedrich Engels. 1846/1969. Die Deutsche Ideologie. *Werke*, vol. 3. Berlin: Dietz Verlag.

Nabonnand, Philippe, and Laurent Rollet. 2002. Une bibliographie mathématique idéale? Le Répertoire bibliographique des sciences mathématiques. *Gazette des mathématiciens* 92: 11–26.

Parshall, Karen. 1989. Toward a History of Nineteenth-Century Invariant Theory. In *The History of Modern Mathematics*, ed. David E. Rowe and John Mc-Cleary, vol. 1, 157–206. Boston: Academic Press

Parshall, Karen. 2006. The British development of the theory of invariants (1841–1895). *The Bulletin of the British Society for History of Mathematics* 21: 186–199.

Parshall, Karen. 2023. A Convergence of Paths: Arthur Cayley, Charles Hermite, James Joseph Sylvester, and the Early Development of a Theory of Invariants. *Revue d'histoire des mathématiques*, to appear.

Poincaré, Henri. 1879a. Sur quelques propriétés des formes quadratiques. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences* 89: 344–346.

Poincaré, Henri. 1879b. Sur les formes quadratiques. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences* 89: 897–899.

Poincaré, Henri. 1880a. Sur les formes cubiques ternaires. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences* 90: 1336–1339.

Poincaré, Henri. 1880b. Sur la réduction simultanée d'une forme quadratique et d'une forme linéaire. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences* 91: 844–846.

Poincaré, Henri. 1880c. Sur un mode nouveau de représentation géométrique des formes quadratiques définies ou indéfinies. *Journal de l'École polytechnique* 47: 177–245.

Poincaré, Henri. 1881a. Sur les fonctions fuchsiennes. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences* 92: 333–335.

Poincaré, Henri. 1881b. Sur la représentation des nombres par des formes. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences* 92: 777–779.

Poincaré, Henri. 1881c. Sur les formes cubiques ternaires et quaternaires. Première partie. *Journal de l'Ecole polytechnique* 50: 190–253.

Poincaré, Henri. 1882a. Sur les invariants arithmétiques. In *Comptes rendus de la 10e session - Alger 1881*, 109–117. Paris: Association française pour l'avancement des sciences.

Poincaré, Henri. 1882b. Sur les applications de la Géométrie non-euclidienne à la théorie des formes quadratiques. In *Comptes rendus de la 10e session - Alger 1881*, 132–138. Paris: Association française pour l'avancement des sciences.

Poincaré, Henri. 1882c. Sur les formes cubiques ternaires et quaternaires. Seconde partie. *Journal de l'Ecole polytechnique* 51: 45–91.

Poincaré, Henri. 1882d. Sur une extension de la notion arithmétique de genre. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences* 94: 67–69, 124–127.

Poincaré, Henri. 1884. Sur une généralisation des fractions continues. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences* 99: 1014–1016.

Poincaré, Henri. 1885. Sur la représentation des nombres par les formes. *Bulletin de la Société Mathématique de France* 13: 162–194.

Poincaré, Henri. 1886a. Mémoire sur la réduction simultanée d'une forme quadratique et d'une forme linéaire. *Journal de l'Ecole polytechnique* 56: 79–142.

Poincaré, Henri. 1886b. Sur les fonctions fuchsiennes et les formes quadratiques ternaires indéfinies. *Comptes rendus des séances hebdomadaires de l'Académie des sciences* 102: 735–737.

Poincaré, Henri. 1886c. *Notice sur les travaux scientifiques*. Paris: Gauthier-Villars.

Poincaré, Henri. 1887. Les fonctions fuchsiennes et l'arithmétique. *Journal de mathématiques pures et appliquées* 4e s., 3: 405–464.

Poincaré, Henri. 1889. *Leçons sur la théorie mathématique de la lumière*. Paris : Georges Carré.

Poincaré, Henri. 1891a. Sur la distribution des nombres premiers. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences* 113: 819.

Poincaré, Henri. 1891b. Extension aux nombres premiers complexes des théorèmes de M. Tchebicheff. *Journal de mathématiques pures et appliquées* 4e s., 8: 25–68.

Poincaré, Henri. 1897. Les rapports de l'analyse pure et de la physique mathématiques. *Acta Mathematica* 21: 331–341.

Poincaré, Henri. 1901. Sur les propriétés arithmétiques courbes algébriques. *Journal de mathématiques pures et appliquées* 5e s. 7 (2): 161–233.

Poincaré, Henri. 1905a. Sur les invariants arithmétiques. *Journal für die reine und angewandte Mathematik* 129 (2): 89–150.

Poincaré, Henri. 1905b. La Valeur de la science. Paris: Flammarion. English transl. by George Bruce Halsted. New York and Garrison: The Science Press, 1913.

Poincaré, Henri. 1908a. L'Avenir des mathématiques. *Bulletin des sciences mathématiques* 2e s. 32: 168–190; Eng. transl. The future of mathematics, *The Monist* 20 (1910): 76–92.

Poincaré, Henri. 1908b. L'invention mathématique. *L'Enseignement Mathématique* 10, 357–371.

Poincaré, Henri. 1921. Analyse des travaux scientifiques. *Acta Mathematica* 38: 3–135.

Poincaré, Henri. 1950. *Œuvres*, Tome V, ed. Albert Châtelet. Paris: Gauthier-Villars.

Poincaré, Henri. 1986. La correspondance d'Henri Poincaré avec des mathématiciens (A-J), [ed. P. Dugac]. *Cahiers du séminaire d'histoire des mathématiques* 7: 59–219.

Rey, Roselyne. 1994. La Classification des sciences (1750–1850). *Revue de synthèse* ser. IV, 1–2: 5–12.

Rynasiewicz, Robert, and Jürgen Renn. 2005. The Turning Point for Einstein's Annus mirabilis. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 37 (1): 5–35.

Schappacher, Norbert. 1991. Développement de la loi de groupe sur une cubique. *Séminaire de théorie des nombres de Paris 1988/1989*: 159–184. Basel: Birkhäuser.

Schneider, Martina. 2000. *Poincarés Beitrag zum Studium rationaler Punkte auf elliptischen Kurven: Eine Verknüpfung von Funktionentheorie, Geometrie und Zahlentheorie*. Unpublished Diplomarbeit, Fachbereich Mathematik, Bergische Universität Gesamthochschule Wuppertal.

Scholz, Erhard. 1989. *Symmetrie, Gruppe, Dualität*. Basel: Birkhäuser.

Seeber, Ludwig August. 1824. Versuch zur Erklärung des inneren Baues der festen Körper. *Annalen der Physik* 16 (3): 229–248, 349–372.

Selling, Eduard. 1877. Des formes quadratiques binaires et ternaires. *Journal de mathématiques pures et appliquées* 3e s. 3: 21–60.

Siegmund-Schultze, Reinhard. 1993. Mathematische Berichterstattung in Hitlerdeutschland: der Niedergang des *Jahrbuchs über die Fortschritte der Mathematik*. Göttingen: Vandenhoeck & Ruprecht.

Smith, Henry John Stephen. 1861–1865. Report on the Theory of Numbers, part III-V. *Report of the British Association for the Advancement of Science* 1861: 292–340; 1862: 503–526; 1865: 768–786.

Tort, Patrick. 1989. *La Raison classificatoire*. Paris: Aubier.

Weil, André. 1955. Poincaré and arithmetic. In *Livre du centenaire de la naissance de Henri Poincaré*, 206–212. Paris: Gauthier-Villars.

Chapter 8
Simplifying a Proof of Transcendence for e: A Letter Exchange Between Adolf Hurwitz, David Hilbert and Paul Gordan

Nicola M. R. Oswald

Abstract This article is inspired by a letter exchange between Adolf Hurwitz (1859–1919), David Hilbert (1862–1943) and Paul Gordan (1837–1912) concerning their proofs of the transcendence for Euler's number e. This correspondence took place in the period from 1892 to 1894 and accompanies the process of developing mathematical conclusions. We analyze the evolution of the three proof variations, in particular with a focus on ideas of simplifying the lines of argumentation. This is contrasted by a look on the later reception of the proofs.

8.1 A Brief Introduction

A number is called algebraic if it is the root of a not identically vanishing polynomial with rational coefficients; otherwise the number is called transcendental. The first to prove the transcendence of some number was Joseph Liouville (1809–1882) in 1844; his approach was constructive but does not apply to any given significant number.[1] The first who proved that $e = \exp(1) = \sum_{m\geq 0} 1/m!$ is transcendental was Charles Hermite (1822–1901) in 1873.[2] In 1882, The not unrelated case of the transcendence of π was solved by Ferdinand Lindemann (1852–1939); his result also gave a negative answer to the classical problem of squaring the circle by ruler and compass only.

In a standard reference for number theory, one can read that the original proofs of Hermite and Lindemann "*were afterwards modified and simplified by Hilbert, Hurwitz, and other writers*" (see Hardy and Wright 1959, p. 177). In fact, David Hilbert (1862–1943), Adolf Hurwitz (1859–1919), and Paul Gordan (1837–1912)

[1] See Liouville (1844).

[2] See Hermite (1873).

N. M. R. Oswald (✉)
University of Wuppertal, Wuppertal, Germany
e-mail: oswald@uni-wuppertal.de

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_8

published in 1893 successively proofs of transcendence for Euler's number e. That these were directly inspired by each other is shown by an accompanying correspondence between the three mathematicians. We consider the proofs as well as the exchange of ideas in the following.

Although there is certainly much to say about the influential mathematicians Hilbert and Gordan and their relationship to each other,[3] the focus here is on their common correspondence partner Adolf Hurwitz. In this sense we begin with some introductory biographical notes about him, particularly with regard to his work on the transcendence of e.

8.2 Historical Context

Throughout his life, the mathematician Adolf Hurwitz maintained an active network of correspondence partners in contemporary mathematics. Both, his scientific estate and his remaining collection of correspondence with several hundred letters with about forty correspondents, testify his silent yet constant involvement in the mathematical community. There seem to have been several reasons for his continuous and extensive letter exchanges: From the very beginning of his higher education Adolf Hurwitz was fortunate that his mathematical talent was recognized by influential and excellent mathematicians. During his school years he received private lessons from Hermann Caesar Hannibal Schubert (1848–1911),[4] later he became doctoral student of Felix Klein (1849–1925) in Munich and Leipzig, and was supported by the Berlin mathematicians Karl Weierstrass (1815–1897) and Leopold Kronecker (1823–1891). In particular, Weierstrass must have had a large influence on Hurwitz's occupation with analytic issues and encouraged him finding further fields of research.[5] By the age of 23, Adolf Hurwitz accomplished his habilitation in Göttingen where again he was promoted by influential mathematicians such as Hermann Amandus Schwarz (1843–1921) and Moritz Abraham Stern (1807–1894).

Hurwitz's early interest in transcendence is best reflected in his third mathematical diary (Hurwitz 1919, No. 3), which he started in January 1883 and where one can find a detailed reflection on the celebrated theorem of Lindemann–Weierstrass. In his resulting article (Hurwitz 1883), he investigated generalizations by studying arithmetical properties of transcendental functions satisfying a homogeneous linear

[3] For example, both were presidents of the German Mathematical Society (from 1894 and 1900) and both worked and discussed on invariant theory, see a.o. Weyl (1944, p. 622). Gordan was "sometimes referred to as the King of invariant theory" (cf. Gray 2018, p. 265). For the story on Hilbert's finite basis theorem (from 1888), his encounter with Gordan and the latter one's assessment of it, we refer to Jeremy Gray (Gray 2018, Ch. 25).

[4] At age 17 Hurwitz published his first result (on Chasles's theorem); it was a joint paper with Schubert.

[5] For more details about periods in Hurwitz's professional and personal life we refer to Oswald and Steuding (2014).

differential equation. His later wife Ida Samuel-Hurwitz (1864–1951) mentioned in a biographical essay that his work on transcendence was at least one of the reasons why Lindemann became aware of Adolf Hurwitz and initiated the invitation of Hurwitz to an extraordinary professorship at the Albertus University at Königsberg in 1884.[6] In the eastern province of Prussia his good luck to be surrounded by talented mathematicians unexpectedly continued. It is well known that his students David Hilbert and Hermann Minkowski (1864–1909) became successful mathematicians and the interaction of the three went on for all of their lives.[7] Hurwitz was indeed not only an interested and helpful mathematical colleague of both but also became a friend. This is revealed by the correspondence of the three and brought to the point in Hilbert's obituary[8] for his former teacher:

> Here [in Königsberg] I was, at that time still a student, soon asked for scientific exchange and had the luck by being together with Hurwitz to get to know in the easiest and most interesting way the directions of thinking of the at time opposite, however, each other excellently complementing schools, the geometrical school of Klein and the algebraic-analytical school of Berlin. [...] On numerous, sometimes day by day undertaken walks at the time for eight years we have browsed through probably all corners of mathematical knowledge, and Hurwitz with his as well wide and multifaceted as also established and well-ordered knowledge was always our leader.[9] (Hilbert 1921, p. 162)[10]

David Rowe describes the relationship of Hilbert and Hurwitz in Königsberg with the following words: "Adolf Hurwitz was at the height of his powers and he opened up whole new mathematical vistas to Hilbert who looked up to him with admiration mixed with a tinge of envy" (Rowe 2007, p. 25). Certainly it is difficult to assess Hilbert's nature or degree of admiration, however, the subsequent letter exchange may provide further insights into their relationship.

When Adolf Hurwitz moved to Switzerland in 1892 as full professor at the *Polytechnikum* in Zurich (since 1911 named *Eidgenössische Technische Hochschule*, ETH) the exchange with his colleagues became more important. Only with Hilbert one can find around two hundred exchanged letters in about thirty years. From the

[6] See Samuel-Hurwitz (1984).

[7] An approach on Hilbert and Hurwitz's student-teacher-relationship developing to a collegial level can be found in Oswald (2014). The collected letters from Minkowski to Hilbert are published in Minkowski (1973).

[8] Taken from his commemorative speech 'Adolf Hurwitz' published in Hilbert (1902).

[9] "Hier [in Königsberg] wurde ich, damals noch Student, bald von Hurwitz zu wissenschaftlichem Verkehr herangezogen und hatte das Glück, durch das Zusammensein mit ihm in der mühelosesten und interessantesten Art die Gedankenrichtungen der beiden sich damals gegenüberstehenden und doch einander so vortrefflich ergänzenden Schulen, der geometrischen Schule von Klein und der algebraisch-analytischen Berliner Schule kennenzulernen. [...] Auf zahllosen, zeitweise Tag für Tag unternommenen Spaziergängen haben wir damals während acht Jahren wohl alle Winkel mathematischen Wissens durchstöbert, und Hurwitz mit seinen ebenso ausgedehnten und vielseitigen wie festbegründeten und wohlgeordneten Kenntnissen war uns dabei immer der Führer."

[10] **Remark.** Here and in the sequel all German and French texts were freely translated by the author to the best of her knowledge.

first two years in Zurich originates the subsequently discussed correspondence on simplifications of the proofs of transcendence for e.

Throughout his life, Hurwitz remained an active mathematician in research and teaching. He stayed with his wife and their three children in Zurich until the end of his life in 1919.

8.3 Three Variations of a Proof of Transcendence for *e*

In this section we focus on the mathematical methods and ingredients of the different approaches taken by David Hilbert, Adolf Hurwitz and Paul Gordan. Therefore, we give and compare the three proofs in detail, starting with some notes on their mathematical and historical background.

The first proof of transcendence for e was given by Hermite (1873). His pathbreaking approach relies on an analogy between classical diophantine approximation (that is approximation of real numbers by rationals) and approximating analytic functions of one variable by rational functions. The central idea is to approximate the exponential function $\exp(x)$ by a rational function in x. For his explicit solution of this problem by using what is now known as Padé approximants[11] for the exponential function and a detailed discussion of his transcendence proof for e we refer to Mahler (1976) (in its appendix). In the proofs we are going to analyze below this approximation is realized by the function $F(x)$.

In the late nineteenth century, several mathematicians tried to simplify Hermite's original, rather lengthy and technical proof as well as the related one for π, found by Lindemann (1882b,a). Besides Karl Weierstrass' important contribution on generalizations of their methods,[12] there are Andrei Andreevich Markoff (1856–1922),[13] as well as Oswald Venske (1867–1939) and Thomas Jan Stieltjes (1856–1894)[14] to be mentioned. The latter two contributions are more relevant since they both focus on e and struggle for a simplification of Hermite's reasoning. Besides those publications from 1890 there is another paper due to Victor Jamet[15] from 1891 which is very close to Stieltjes' proof and contains the footnote "On pourra consulter aussi sur une simplification de la methode de M. Hermite une Note de M. Stieltjes (*Comptes Rendus*, 1890)." Then, in 1893, three *simplifications* were published, the first by David (Hilbert 1893), followed by Adolf (Hurwitz 1893a), and, finally, Paul (Gordan 1893d). Proofs of transcendence were definitely a hot topic at that time.

[11] Named after Henri Padé (1863–1953) who received his doctorate under Hermite with a systematic study of these approximations in 1892.

[12] See Weierstrass (1885).

[13] See Markoff (1883).

[14] See Venske (1890) and Stieltjes (1890).

[15] See Jamet (1891).

Stieltjes' reasoning relies on an identity due to Hermite, resp. the following variation thereof

$$\int_0^c e^{-xy} f(x)\,\mathrm{d}x = F(0) - e^{-cy} F(c), \tag{8.1}$$

where $F(x) = \sum_{m\geq 0} f^{(m)}(x)/y^{m+1}$ and, as usual, $f^{(m)}$ denotes the mth derivative of $f = f^{(0)}$ which is here a suitable polynomial. This simple case of Hermite's identity can easily be proved by partial integration and induction. In fact, Hilbert's later proof does not differ too much from Stieltjes' proof.

Therefore, we begin with a sketch of Hilbert's proof and afterwards point out the little differences to Stieltjes' proof. His reasoning is indirect (as well as all the other proofs), so we assume that e satisfies an algebraic equation,[16]

$$0 = \sum_{0\leq j\leq n} a_j e^j \qquad \text{with} \quad a_j \in \mathbb{Z},\ a_0 \neq 0; \tag{8.2}$$

in fact, if a_0 would vanish, we could divide the equation by e in order to get an algebraic equation for e of smaller degree. Next we define a polynomial

$$f(x) = x^{\mu} \prod_{1\leq c\leq n} (x-c)^{\mu+1},$$

where μ is a positive integer to be chosen later, and multiply the corresponding integral $\int_0^{\infty} e^{-x} f(x)\,\mathrm{d}x$ with the algebraic equation for e. This leads to

$$0 = \sum_{0\leq j\leq n} a_j e^j \left\{ \int_0^j + \int_j^{\infty} \right\} e^{-x} f(x)\,\mathrm{d}x;$$

we observe that Hermite's integral (8.1) with parameter $y = 1$ appears side by side with an infinite integral. We may rewrite the latter equation as

$$0 = A + B,$$

where

$$A = \sum_{1\leq j\leq n} a_j e^j \int_0^j e^{-x} f(x)\,\mathrm{d}x$$

[16] Please notice that we use abbreviatory notation including $\sum$, $\prod$ which was less common at the end of the nineteenth century. Moreover, we have adjusted the choice of letters in the different proofs to stress similarities.

and

$$B = a_0 \int_0^\infty e^{-x} f(x)\,\mathrm{d}x + \sum_{1 \le j \le n} a_j e^j \int_j^\infty e^{-x} f(x)\,\mathrm{d}x;$$

here we have split the sum according to certain arithmetical properties we shall investigate now. For this purpose we use the elementary formula

$$\int_0^\infty x^u e^{-x}\,\mathrm{d}x = u! \tag{8.3}$$

which is derived by partial integration and induction (similar to Hermite's identity (8.1) above). Hence, the expression $a_0 \int_0^\infty e^{-x} f(x)\,\mathrm{d}x$, appearing in B, is an integer multiple of $\mu!$ but not divisible by $\mu + 1$ provided that μ is chosen as a large integer multiple of $a_0 n!$. More precisely, expanding the polynomial $\prod_{1 \le c \le n} (x - c)^{\mu+1}$, this integral equals

$$\begin{aligned} &\int_0^\infty e^{-x} x^\mu \prod_{1 \le c \le n} (x - c)^{\mu+1}\,\mathrm{d}x \\ &= \int_0^\infty e^{-x} x^\mu\,\mathrm{d}x \cdot \prod_{1 \le c \le n} (-c)^{\mu+1} + \sum_{k \ge \mu+1} b_k \int_0^\infty e^{-x} x^k\,\mathrm{d}x \end{aligned} \tag{8.4}$$

with some integer coefficients b_k, the sum being finite. In order to show that also the other summands in B are integers we substitute $x = \omega + j$ for $j = 1, 2, \ldots, n$, which yields

$$\begin{aligned} e^j \int_j^\infty e^{-x} f(x)\,\mathrm{d}x &= \int_0^\infty (\omega + j)^\mu \prod_{1 \le c \le n} (\omega + j - c)^{\mu+1} e^{-\omega}\,\mathrm{d}\omega \\ &= \int_0^\infty \sum_{k \ge \mu+1} b_k \omega^k e^{-\omega}\,\mathrm{d}\omega \end{aligned} \tag{8.5}$$

by the binomial theorem with another set of integer coefficients b_k. Again in view of (8.3) each of these expressions in B is an integer multiple of $(\mu + 1)!$, and therefore B is an integer divisible by $\mu!$ such that

$$B/\mu! \equiv \pm a_0 (n!)^{\mu+1} \not\equiv 0 \bmod (\mu + 1),$$

as follows from (8.4).

Now denote by K the maximum of $x\prod_{1\le c\le n}|x-c|$ and by k the maximum of $e^{-x}\prod_{1\le c\le n}|x-c|$, both taken for values x from the closed interval $[0,n]$. Then,

$$|A| \le \sum_{1\le j\le n}\left|a_j e^j \int_0^j e^{-x} f(x)\,\mathrm{d}x\right| < \kappa K^{\mu},$$

where $\kappa = \sum_{1\le j\le n}|a_j e^j|jk$. Since $K^{\mu}/\mu!$ are the summands in the convergent exponential series for $\exp(K)$, we may choose μ in addition to satisfy $\kappa\frac{K^{\mu}}{\mu!} < 1$. Then $B/\mu!$ is an integer not divisible by $\mu+1$, hence $B/\mu!$ does not vanish. Furthermore, $|A/\mu!| < 1$, thus $A/\mu! + B/\mu! \neq 0$, giving the desired contradiction.

Comparing Hilbert's proof with the one by Stieltjes, one observes that Hermite's identity appears in disguise: taking $y = 1$ in (8.1), we find the corresponding terms $F(0) - e^{-c}F(c)$ as the building blocks of B, and so the parameter y is superfluous in Hilbert's proof. Moreover, Hilbert omitted a lengthy and technical discussion of a certain integral expression involving a piecewise constant function similar to $A + B$ by using μ as a parameter and congruence calculus for his arithmetical argument. Another difference is within the polynomial f. Where Stieltjes assumes arbitrary integral roots, Hilbert specifies them to be the integers $c = 0, 1, 2, \ldots, n$. As a matter of fact, a closeness to Stieltjes' reasoning is obvious though Hilbert's presentation is a little shorter, definitely more elegant and more to the point. This is well documented by Hurwitz in his mathematical diary (see Figs. 8.1 and 8.2) and in his correspondence.

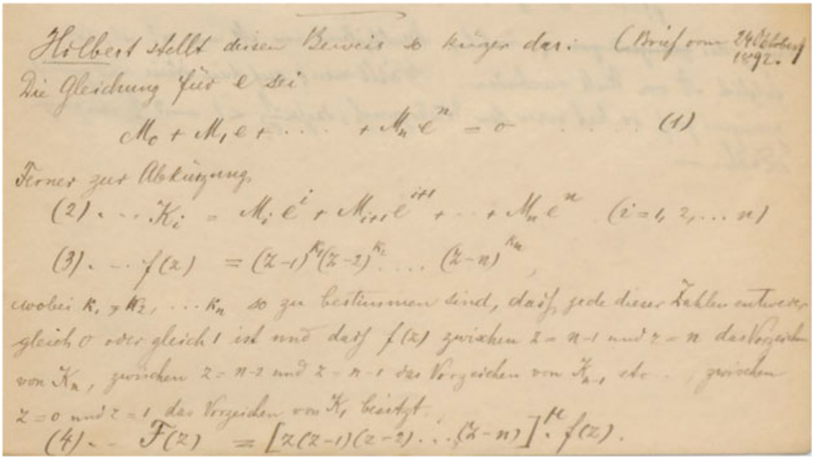
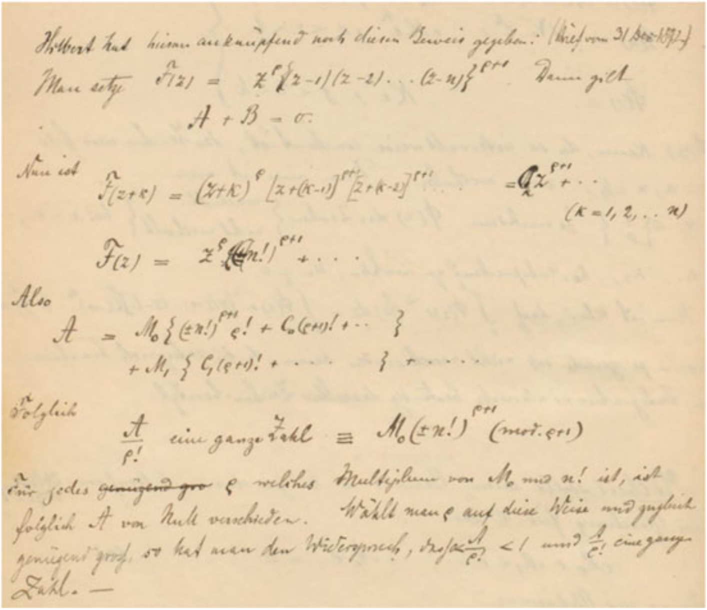

Hilbert stellt diesen Beweis so kürzer dar: (Brief vom 24 Oktober 1892.)
Die Gleichung für e sei
$M_0 + M_1 e + \cdots + M_n e^n = 0$ (1)
Ferner zur Abkürzung
(2) ... $K_i = M_i e^i + M_{i+1} e^{i+1} + \cdots + M_n e^n$ $(i = 1, 2, \ldots n)$
(3) ... $f(z) = (z-1)^{k_1}(z-2)^{k_2}\ldots(z-n)^{k_n}$
wobei $k_1, k_2, \ldots k_n$ so zu bestimmen sind, daß jede dieser Zahlen entweder gleich 0 oder gleich 1 ist und daß $f(z)$ zwischen $z = n-1$ und $z = n$ das Vorzeichen von K_n, zwischen $z = n-2$ und $z = n-1$ das Vorzeichen von K_{n-1}, etc. zwischen $z = 0$ und $z = 1$ das Vorzeichen von K_1 besitzt.
(4) ... $F(z) = [z(z-1)(z-2)\ldots(z-n)]^{\mu} \cdot f(z)$.

Fig. 8.1 An excerpt from Hurwitz's diary, notebook no. 10 (Hurwitz 1919, No. 10, p. 153), referring to a letter from Hilbert bearing the date 24 October 1892. After a brief outline of Stieltjes' proof Hurwitz starts his sketch of Hilbert's reasoning with the words "Hilbert presents [Stieltjes'] proof in a shorter way as follows." (All diaries are published online on the e-manuscripta.ch platform)



Hilbert hat hieran anknüpfend noch diesen Beweis gegeben: (Brief vom 31 Dec 1892)

Man setze $F(z) = z^\rho\{(z-1)(z-2)\ldots(z-n)\}^{\rho+1}$ Dann gilt

$$A + B = 0.$$

Nun ist

$$F(z+k) = (z+k)^\rho\,[z+(k-1)]^{\rho+1}\,[z+k-2]^{\rho+1}\ldots = Q z^{\rho+1} + \ldots \quad (k=1,2,\ldots n)$$

$$F(z) = z^\rho (\pm n!)^{\rho+1} + \ldots$$

Also

$$A = M_0\{(\pm n!)^{\rho+1}\rho! + C_0(\rho+1)! + \cdots\} + M_1\{C_1(\rho+1)! + \cdots\} + \cdots$$

Folglich

$$\frac{A}{\rho!} \text{ eine ganze Zahl } \equiv M_0(\pm n!)^{\rho+1} \pmod{\rho+1}$$

Für jedes ~~genügend gro~~ ρ welches Multiplum von M_0 und $n!$ ist, ist folglich A von Null verschieden. Wählt man ρ auf diese Weise und zugleich genügend groß, so hat man den Widerspruch, daß $\frac{A}{\rho!} < 1$ und $\frac{A}{\rho!}$ eine ganze Zahl. —

Fig. 8.2 Another excerpt from Hurwitz's diary, notebook no. 10, on the next page (Hurwitz 1919, No. 10, p. 154). Hurwitz refers to another letter from Hilbert bearing the date 31 December 1892. He writes "Following up on this Hilbert has also given this proof" and continues with a brief sketch of Hilbert's reasoning. This notebook also contains an early draft of Hurwitz's proof of the transcendence of e, p. 175

Although Hilbert went beyond Stieltjes in proving the transcendence of π by almost similar means, it is noteworthy that he did not refer to Stieltjes in his article at all. Actually, Lindemann is the only mathematician that is mentioned in Hilbert's article. David Rowe wrote:[17]

> From one of his letters to Hurwitz, we learn that Hilbert got the initial idea for this new proof by reading a paper on the transcendence of e published by Th.J. Stieltjes in 1890. Stieltjes' paper was hardly longer than Hilbert's, and had been published in the widely read *Comptes Rendus* of the French Academy. Remarkably, Hilbert made no mention of this in his publication; in fact, his notice contains no references to the literature on these problems whatsoever! Was this in the name of simplicity, or perhaps just a sign of supreme intellectual arrogance?

One could take the ambitious Hilbert in shelter by stressing his age and experience as well as the standards of citation at that time (see also Fig. 8.3). Similarly, Stieltjes did not refer to Venske, Hurwitz mentioned only Hilbert and Gordan (although he wrote reviews for the Jahrbuch of both, Venske's and Stieltjes' article, namely for the *Jahresberichte* `JFM 22.0435.03` and `JFM 22.0437.01`, the first one

[17] Rowe (2007), p. 229.

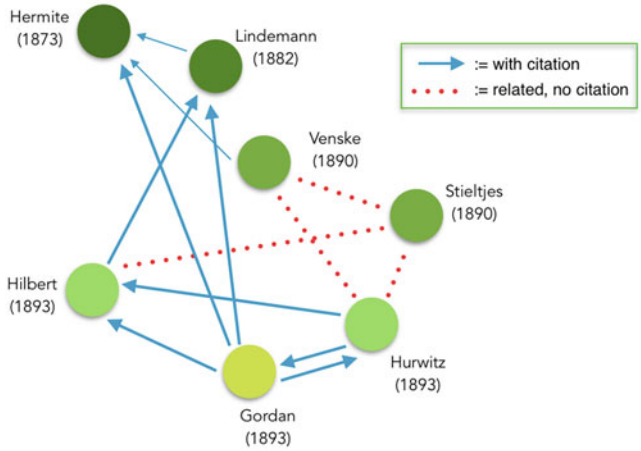


Fig. 8.3 Illustration of the citations in publications

containing several typos), and Gordan cited Hermite, Lindemann, Hurwitz, and Hilbert (see Fig. 8.3). However, as already pointed out by David (Rowe 2007), Felix Klein had the idea to publish the three different proofs of Hilbert, Hurwitz and Gordan in one volume of the well distributed and renowned *Mathematische Annalen* of which he was the managing editor. Different to the *Göttinger Nachrichten*[18], where Hilbert and Hurwitz had published their variations, a reader of the *Annalen* would expect proper references for a hot topic result as these transcendence proofs, however Hilbert and Hurwitz were not much interested in a further publication besides the *Göttinger Nachrichten* (see Rowe 2007, p. 230, for further details). Gordan's proof appeared in slightly different form and in French, see Gordan (1893c).[19]

We continue with a sketch of Hurwitz's proof. For $F(x) = \sum_{m\geq 0} f^{(m)}(x)$, differentiation with respect to x shows

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{-x}F(x)\right) = -e^{-x}f(x), \tag{1'}$$

which is the differential analogue of Stieltjes' variant (8.1) of Hermite's integral identity (and its proof relies on the product rule which is the counterpart of partial integration that has been used to derive (8.1)). Applying the mean-value theorem from differential calculus, implies the existence of a real number ξ in between 0 and x such that

$$e^{-x}F(x) - F(0) = x\left(-e^{-\xi}f(\xi)\right).$$

[18] *Nachrichten von der königlichen Gesellschaft der Wissenschaften zu Göttingen.*

[19] Besides there were a few translations of these proofs, e.g., Hilbert's article in Polish in the Polish journal *Prace Matematyczno-Fizyczne*.

Now, for $x = c$ ranging from 1 up to n, we write the intermediate value ξ as δc with some $\delta \in (0, 1)$ (depending on c) and get

$$F(c) - e^c F(0) = -ce^{c(1-\delta)} f(\delta c) =: \epsilon_c, \tag{8.6}$$

say. For a large prime number p, we consider the polynomial

$$f(x) = \frac{1}{(p-1)!} x^{p-1} \prod_{1 \le c \le n} (c - x)^p, \tag{8.7}$$

which differs only slightly from Hilbert's polynomial. Then, all the expressions ϵ_c get arbitrarily small with increasing p. Next we expand $f(x)$ around $x = c$ for $c = 1, 2, \dots, n$ into a power series by applying the binomial theorem, namely

$$f(c + h) = \frac{h^p}{(p-1)!} \left(b_0 + \sum_{k \ge 1} b_k h^k \right),$$

where the sum is finite and the b_k's are integers depending on c. Comparing with the Taylor series expansion

$$f(c + h) = \sum_{m \ge 0} \frac{f^{(m)}(c)}{m!} h^m,$$

it follows that $f^{(m)}(c)$ vanishes for $m < p$ and

$$f^{(m)}(c) = \frac{m!}{(p-1)!} b_{m-p} \equiv 0 \bmod p$$

for all $m \ge p$. For the root $x = 0$ of f, however, the situation is slightly different: here we observe by the same reasoning that $f^{(m)}(0)$ is vanishing for $m < p - 1$, is equal to $(n!)^p$ for $m = p - 1$, and is an integer multiple of p whenever $m \ge p$. Thus, the numbers $F(c)$ for $c = 1, \dots, n$ are all integer multiples of p if $p > n$, and $F(0)$ is an integer too but not divisible by p.

Now, assuming an equation of the form (8.2) for e, it follows that

$$a_0 F(0) + \sum_{1 \le c \le n} a_c F(c) = \sum_{1 \le c \le n} a_c \epsilon_c.$$

Since the left hand side is an integer, the right hand side is an integer too, however, it gets as small as we please by increasing p, so the quantity on the left vanishes. This vanishing contradicts that the integer $a_0 F(0) + \sum_{1 \le c \le n} a_c F(c)$ is not divisible by p when p is chosen to be sufficiently large.

Both proofs so far rely essentially on the differential equation $\exp' = \exp$ for the exponential function (or, equivalently, the integral equation for exp). This kind of reproduction property appears in the proof of Hermite's identity (8.1) as well as in its differential counterpart (1'). The additional analytic ingredients can also not be called advanced. The estimate of the quantities ϵ_c in Hurwitz's proof seems formally easier than bounding the integrals in Hilbert's reasoning, though it includes the mean-value theorem from differential calculus and the infinitude of prime numbers although Hurwitz (and we as well) did not explicitly mention the latter simple fact at all.

Hurwitz (1893a) contains a footnote:

> By the way, Mr. Hilbert, as I learned from him recently, has already occasionally given hints in a lecture how one can avoid the integrals (and at the same time the differentiation) in his proof by replacing the integrals by limit values.[20]

This idea had been realized by the following third proof published in 1893, found by Paul Gordan, professor in Erlangen and a former colleague of Felix Klein. Jeremy Gray characterizes Gordan as "a master at manipulating long algebraic expressions" (Gray 2018, p. 265).

This ability together with his urge for explicitness can also be found in Gordan's proof of transcendence. His reasoning relies only on the convergence of the exponential series,

$$e^x = \sum_{i \geq 0} \frac{x^i}{i!},$$

and a subtle notation, namely denoting $r!$ as h^r, where r is any non-negative integer. This allows to write

$$h^r e^x = r! \left\{ \sum_{0 \leq i \leq r} + \sum_{i > r} \right\} \frac{x^i}{i!} = (x+h)^r + x^r u_r, \tag{1''}$$

where

$$\sum_{0 \leq i \leq r} \frac{r!}{i!} x^i = \sum_{0 \leq i \leq r} \binom{r}{i} x^i h^{r-i} = (x+h)^r, \tag{8.8}$$

(since $\binom{r}{i} h^{r-i} = \frac{r!}{i!(r-i)!} h^{r-i} = \frac{r!}{i!}$ and the last equality follows *symbolically* from the binomial theorem), and

$$u_r := \sum_{i > r} \frac{r!}{i!} x^{i-r} = \sum_{k \geq 1} \frac{x^k}{(r+1) \cdot \ldots \cdot (r+k)}.$$

[20] "Herr Hilbert hat übrigens, wie ich von ihm neuerdings erfahre, auch schon gelegentlich eines Vortrages Andeutungen gegeben, wie man die Integrale (und zugleich das Differenziren) bei seinem Beweise vermeiden kann, indem man die Integrale durch Grenzwerthe ersetzt."

In view of the convergence of the exponential series, for $x \geq 0$,

$$u_r \leq \sum_{k \geq 1} \frac{x^k}{k!} < e^x;$$

hence $u_r = q_r e^x$ with some real number q_r from the unit interval.

Now, for arbitrary numbers $b_0, \dots, b_d$, we deduce from (1”) that

$$\sum_{0 \leq r \leq d} b_r h^r e^x = \sum_{0 \leq r \leq d} b_r (x + h)^r + \sum_{0 \leq r \leq d} b_r q_r e^x x^r,$$

resp.

$$e^x f(h) = f(x + h) + e^x g(x), \tag{8.9}$$

where

$$f(\mathrm{x}) := \sum_{0 \leq r \leq d} b_r \mathrm{x}^r \quad \text{and} \quad g(\mathrm{x}) := \sum_{0 \leq r \leq d} b_r q_r \mathrm{x}^r.$$

Next we assume that e is algebraic, i.e. an equation of the form (8.2) holds.

Then (8.9) with $x = j$ and $j = 0, 1, \dots, n$ implies

$$\begin{aligned} 0 &= \sum_{0 \leq j \leq n} a_j e^j \cdot f(h) \\ &= \sum_{0 \leq j \leq n} a_j f(j + h) + \sum_{0 \leq j \leq n} a_j g(j) e^j = F + G, \end{aligned}$$

say.

Now we choose f the same way as Hurwitz did, that is (8.7). For $j = 1, 2, \dots, n$, writing $p! = h^p$, we observe that

$$\begin{aligned} f(j + h) &= \frac{(j + h)^{p-1}}{(p - 1)!} \prod_{1 \leq c \leq n} (c - j - h)^p \\ &= -(j + h)^{p-1} \frac{h^p}{(p - 1)!} \prod_{\substack{1 \leq c \leq n \\ c \neq j}} (c - j - h)^p \\ &= -p(j + h)^{p-1} \prod_{\substack{1 \leq c \leq n \\ c \neq j}} (c - j - h)^p. \end{aligned}$$

Since by (8.8) every factor is an integer, it follows that $f(j+h)$ is an integer divisble by p. For $j=0$, however, we find similarly

$$f(h) = \frac{h^{p-1}}{(p-1)!} \prod_{1\le c\le n} (c-h)^p = \prod_{1\le c\le n} \left(- \sum_{0\le i\le p} \frac{p!}{i!}(-c)^i \right),$$

which is an integer not divisible by the prime p. Hence, F is a non-zero integer.

Since with increasing p the coefficients b_k of f tend to zero (thanks to the factor $\frac{1}{(p-1)!}$), it follows that G gets as small as we please, contradicting $F+G=0$.

Probably, Gordan came to his symbolical proof by introducing formal differentiation in Hurwitz's reasoning. In fact, this can be seen in (8.8), namely,

$$\begin{aligned} f(x+h) &= \sum_{0\le r\le d} b_r (x+h)^r = \sum_{0\le r\le d} b_r \sum_{0\le i\le r} \frac{r!}{i!} x^i \\ &= \sum_{0\le m\le d} \sum_{m\le r\le d} b_r \frac{r!}{(r-m)!} x^{r-m} = \sum_{m\ge 0} f^{(m)}(x), \end{aligned} \tag{8.10}$$

which is equal to $F(x)$ in other proofs. Later Heinrich Weber (1842–1913) reinvented in his important monograph (Weber 1899) derivatives without any comment; his intention might have been the counter-intuitive notation Gordan used for circumventing derivatives. Whereas Hilbert and Hurwitz use the differential equation for the exponential function, the convergence of the used exponential series is indeed the only property Gordan needed for his proof.

8.4 Hessenberg's Analysis

Concerning Gordan's symbolic proof, Gerhard Hessenberg (1874–1925)[21] wrote:

> [...] so it is no miracle that G o r d a n 's symbolism raises commonly some gentle shudder, not least because the imposition to think of h^μ as $\mu!$ collides with the habit not to think at all while calculating something.[22]

In 1912, Hessenberg analyzed the various proofs of transcendence for e and π in detail, in some places a little polemic. The only proofs he missed in his careful study are those by Venske and Stieltjes. The first one relies on a certain non-vanishing determinant and is therefore closer to Hermite's original reasoning

[21] Professor at Bonn, Breslau and Tübingen; Hessenberg is known for his work in geometry and set theory.

[22] "[...] so nimmt es kein Wunder, daß die Gordansche Symbolik gemeinhin ein gelindes Gruseln zu erwecken pflegt, nicht zuletzt auch darum, weil die Zumutung, unter h^μ sich $\mu!$ zu denken, mit der Gewohnheit kollidiert, beim Rechnen überhaupt nichts zu denken." (Hessenberg 1912, p. 50)

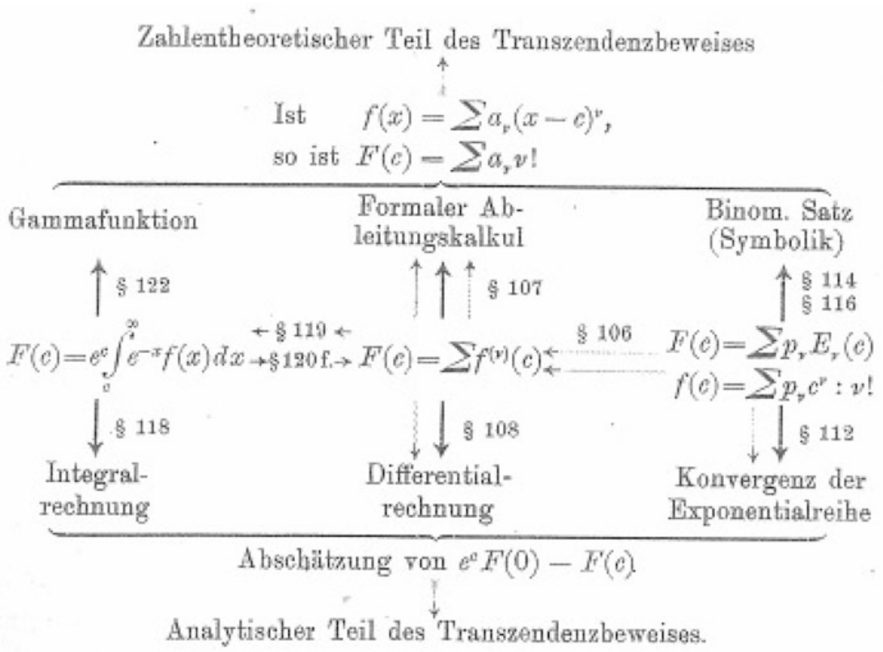


Fig. 8.4 A diagram from Hessenberg (1912, p. 53) illustrating the proof essentials of Hilbert (left), Hurwitz (middle) and Gordan (right). For details Hessenberg refers to various paragraphs. We have adjusted our presentation of the three proofs in question with respect to Hessenberg's analysis. Here the Gamma-function is mentioned in the upper left with respect to Formula (8.3) in Hilbert's proof since $\Gamma(u+1) = \int_0^\infty x^u e^{-x}\,\mathrm{d}x$

whereas Stieltjes' proof is quite similar to Hilbert's reasoning (as already pointed out above). According to Hessenberg, the proofs can be dissected into three parts (see Fig. 8.4).

First of all, there is an arithmetical part linking $f(x) = \sum a_\nu (x-c)^\nu$ with the values $F(c) = \sum a_\nu \nu!$, where the coefficients are all integers. In all three proofs it is shown, by slightly different means, that the numbers $F(c)$ are integers satisfying certain divisibility properties with respect to the multiplicities of the roots of f. It might be worth to notice that Hermite's line of argument for this purpose differs significantly by a rather lengthy calculation of a certain determinant.

The middle part is the most interesting since this is essentially the only part where the proofs differ. Hilbert's reasoning relies on integration (and is therefore building on Hermite's original proof and his identity) by using (8.1) implicitly with $y = 1$, resp.

$$F(c) = e^c \int_c^\infty e^{-x} f(x)\,\mathrm{d}x,$$

and explicit calculations (by partial integration, resp. the gamma-function; see (8.3), (8.4) and (8.5)).

Hurwitz replaced integration by differentiation and the corresponding formula is indeed simpler. More precisely, he observed $\frac{\mathrm{d}}{\mathrm{d}x}\Big(e^{-x}F(x)\Big) = -e^{-x}f(x)$ and applied the mean value theorem in order to obtain (8.6), i.e.,

$$F(c) - e^c F(0) = -ce^{-\delta c} f(\delta c)$$

for some intermediate value δc from $[0, c]$. Maybe this formula shows best that the polynomial F works as an approximation for the exponential function. The link to the arithmetical part follows from the Taylor expansion $F(c) = \sum f^{(\nu)}(c)$.

Finally, Gordan argued symbolically without using any integral or differential calculus; even formal differentiation is replaced by using the binomial theorem. Whereas Hilbert and Hurwitz worked with the differential equation for the exponential function in addition with some fundamental results from integral and differential calculus, respectively, the only ingredient needed for Gordan's argument is the convergence of the exponential series linking both functions F and f. Hessenberg rewrote this shortly by the formulae

$$F(c) = \sum p_\nu E_\nu(c), \quad f(c) = \sum p_\nu c^\nu/\nu!;$$

for the sake of brevity we do not explain the numbers p_ν and functions E_ν appearing here though the reader may guess by comparing these expressions with $F(c) = \sum_{m\geq 0} f(c) = f(c+h)$ according to (8.10).

The final third part of Hessenberg's dissection of the transcendence proofs is analytic and consists of an estimate of the quantities $e^c F(0) - F(c)$. Also here differences appear according to the use of integration, differentiation or just the convergent exponential series.

Hence, and here we follow Hessenberg's judgement, Hilbert's proof can be considered purely analytical, Hurwitz's reasoning is mixed, and Gordan argues purely formal. Moreover, Hessenberg considered also further proofs (e.g. those by Weber, Mertens, Vahlen, and Schottky), however, those are too close to the above discussed ones to be interesting in our framework.

8.5 Letter Exchanges

In the following we turn to the letter exchange that accompanied the publication of the proofs. These letters give the opportunity to look over the shoulders of mathematicians in their research (although this can only be a brief look, quoting only excerpts of the correspondence). Correspondents were, of course, the three authors Hurwitz, Hilbert and Gordan, supplemented by references to Felix Klein and Charles Hermite. In particular, we consider six related letters from Hurwitz to

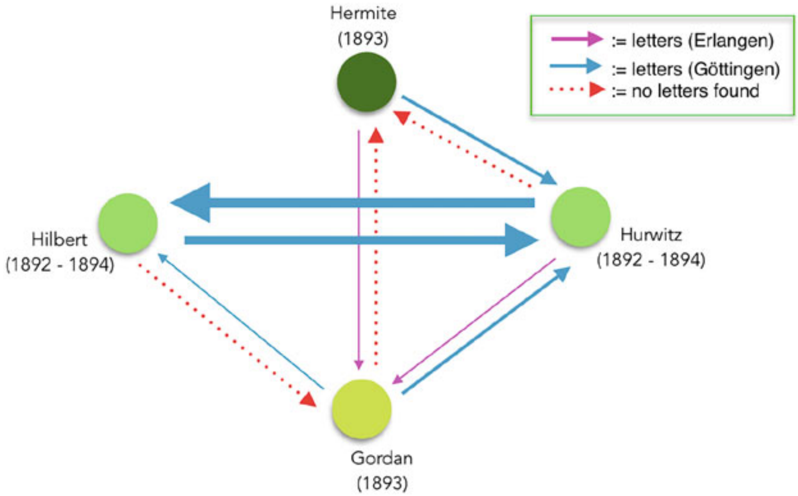


Fig. 8.5 Illustration of the letters that were exchanged; the thickness of the arrows is related to the number of letters

Hilbert as well as five letters from Hilbert to Hurwitz, furthermore, one letter from Gordan to Hilbert and two letters to Hurwitz, one letter from Hurwitz to Gordan. Besides we give quotes from two letters from Hermite to Hurwitz and one letter to Gordan (see Fig. 8.5).

8.5.1 Hilbert and Hurwitz

"I draw your attention to a very nice proof of the transcendence for e by Stieltjes (Comptes R. 1890)."[23] (6.9.1892, Adolf Hurwitz to David Hilbert (Hurwitz 1893b, Br. 7))

As mentioned above, Adolf Hurwitz was reviewer of the article by Thomas (Stieltjes 1890) in the annual *Jahresberichte*. Hilbert's answer to the brief reference to Stieltjes' proof from October 24, 1892, can be regarded as the impetus of the three proof variations for the transcendence of e discussed above. Herein the Königsberg mathematician informed Hurwitz about a new discovery which he had investigated while preparing his lecture on integral calculus. Hilbert stated that "Stieltjes' proof of the transcendence of e can be presented [...], avoiding Hermite's integrals" (Hilbert 1894, Br. 240).

[23] "Ich mache Sie auf einen sehr schönen Beweis für die Transzendenz von e von Stieltjes (Comptes R. 1890) aufmerksam."

Although the diary entry quoted above (Fig. 8.1) proves that Hurwitz was interested in Hilbert's idea, it took some weeks before he commented his proof on 4 December 1892 with a short remark:

> Your simplification of Stieltjes' proof is very nice; I directly wrote it down into my "book of numbers". One should try to also simplify the proof of the transcendence of π.[24] (Hurwitz 1893b, Br. 8)

After a further simplification by Hilbert, communicated in a letter of 31 December 1892, Hurwitz's own activity was aroused. Here, Hilbert gave a proof avoiding even "Stieltjes' point"[25] (Hilbert 1894, Br. 243). He chose the function $F(z) = z^{\sigma}\{(z-1)(z-2)...(z-n)\}^{\sigma+1}$, an explicit simplification, and showed that for sufficiently large σ the terms with improper integrals (see B in Section 2) contribute to an integer not equal to zero.

> You can see that altogether herewith the poof receives another proof line, in which the inequality to 0 is not shown by the sums of integrals [A] but by the integer number [B]. The use of this conclusion leads also to a simplification for the transcendence of π, which does not seem to be irrelevant to me.[26] (Hilbert 1894, Br. 243)

One may assume that Hurwitz shared his opinion: already some days later, despite of New Year's Day and holidays, he answered on 10 January 1893 with not only another new proof by his own, but, furthermore, with a certain idea:

> Your scientific note concerning the number e was, as you may imagine, very interesting for me. [...] I could not stay calm and discovered a further simplification [...] in a way that one can now give the proof in one of the first lessons of a lecture on differential calculus. (Hurwitz 1893b, Br. 9)

As outlined in Sect. 8.3, Hurwitz's proof is based on differential calculus and a combination of divisibility and estimates. In the further course of the letter, Hurwitz makes the following suggestion:

> Did you already edit your proof? If yes, please write me if you have already sent it to Klein. Then I would let the above mentioned further simplification be printed in a short note subsequent to your work. I would prefer if we would choose the Göttinger Nachrichten. [...] So please answer quickly, also if only with a postcard. It is clear that the punch line can also

24 "Ihre Vereinfachung des Stietjes'schen Beweises ist sehr schön; ich habe sie mir gleich in mein "Zahlenbuch" eingetragen. Man sollte doch versuchen auch den Beweis der Transcendenz von π ähnlich zu vereinfachen."

25 "Stieltjesche Pointe"

26 "Sie sehen, dass hiermit der Beweis auch überhaupt eine andere Schlussfolgerung enthält, indem nicht von der Integralsumme II sondern von der ganzen Zahl I das Verschiedensein von 0 gezeigt wird. Die Benutzung dieses Schlusses gibt auch dem Beweise für die Transcendenz von π eine Vereinfachung, welche mir nicht unerheblich erscheint."

> be applied to π, however, I have not really thought through this yet.[27] (Hurwitz 1893b, Br. 9)

But Hilbert did not agree with his former teacher Hurwitz. Within only three days, he answered:

> In fact, I have already worked out my proof for e and π in Christmas holidays, therein - in particular in the part concerning π - some advantageous and simplifying points arouse, such that the whole thing will be on 4–5 printed pages and my presentation is not even short. Of course, in your proof the integral is avoided; however, if the representation of the proof becomes shorter and clearer, is not yet obvious to me. [In a further of my ideas] the opportunity is given to reduce all to a simple considering of limits and the summation of geometric series. Of course, this thing has to be carefully studied yet. [...] However, it is in my opinion, that the proof with the help of integrals will always be the clearest and most developable one. For p or in my notation $\rho + 1$ I did not choose a prime number, since it is easier to define a number which is divisible by c than a prime number of the needed value. Furthermore, I have given how big ρ is to be chosen and I also put a lot of care on the editing. Later on I please you to give me your opinion also concerning incidental parts. It is strange that I also have chosen the Göttinger Nachrichten; Klein wrote me directly, that he would present the note in this week. However, by no means this does affect your intention to show your proof in the next meeting [...] [28] (Hilbert 1894, Br. 244)

This letter includes a number of interesting details. What had happened? In short time of ten days (including New Year's Day) Hilbert had completed his proof line and he had sent his paper to Felix Klein, member of the Göttingen Academy of Sciences and therewith able to communicate articles for the *Göttinger Nachrichten*.

[27] "Ihre wissenschaftliche Mittheilung die Zahl e betreffend hat mich, wie Sie sich denken können, sehr interessiert. [...] Mich hat die Sache nicht ruhen lassen und ich habe eine weitere Vereinfachung entdeckt, [so] daß man jetzt den Beweis in den ersten Stunden einer Vorlesung über Differentialrechnung bringen kann. [...] Haben Sie Ihren Beweis schon redigiert? Wenn ja so schreiben Sie mir bitte, ob Sie ihn schon an Klein geschickt haben. Ich würde dann die vorstehende weitere Vereinfachung in einer kurzen Mittheilung hinter Ihrer Arbeit abdrucken lassen. Am liebsten wäre es mir, wenn wir die Göttinger Nachrichten wählten. [...] Also antworten Sie mir bitte rasch, wenn auch nur per Karte. Daß sich Ihre Pointe auch auf π anwenden lässt, ist klar; ich habe das aber noch nicht ganz durchdacht."

[28] "Meinen Beweis für e und π habe ich in der That bereits in den Weihnachtsferien ausgearbeitet, es hat sich dabei - zumal bei dem über π handelnden Theile - noch mancherlei Vorteilhaftes und Vereinfachendes ergeben, so dass die ganze Sache jetzt auf 4–5 Druckseiten herausgehen wird und dabei ist meine Darstellung durchaus nicht knapp. Bei Ihrem Beweis wird freilich das Integral vermieden; ob aber die Darstellung des Beweises kürzer und übersichtlicher wird, ist mir doch noch nicht ganz einleuchtend. [Bei einer meiner weiteren Idee] ist die Möglichkeit gegeben alles auf eine einfache Grenzbetrachtungen und die Summation geometrischer Reihen zurückzuführen. Die Sache müsste natürlich noch genau durchgedacht werden [...]. Doch ist es meine Überzeugung, dass der Beweis mit Hilfe des Integrals immer der übersichtlichste und entwicklungsfähigste bleiben wird. Für p oder in meiner Bezeichnung für $\rho+1$ habe ich absichtlich nicht eine Primzahl gewählt, weil es doch einfacher ist, eine durch c teilbare Zahl als eine Primzahl von der nötigen Grösse zu bestimmen. Auch habe ich angegeben wie gross ρ zu wählen ist und auch auf die Redaktion recht viel Sorgfalt verwandt. Ich bitte nachher sehr um Ihr Urteil auch im Nebensächlichen. Kurios ist es, dass auch ich gerade die Göttinger Nachrichten gewählt habe; Klein schrieb mir umgehend, dass er die Note bereits in dieser Woche vorlegen würde. Aber dies schadet ja durchaus nicht Ihrer Absicht, Ihren Beweis in der folgenden Sitzung vorzulegen [...]"

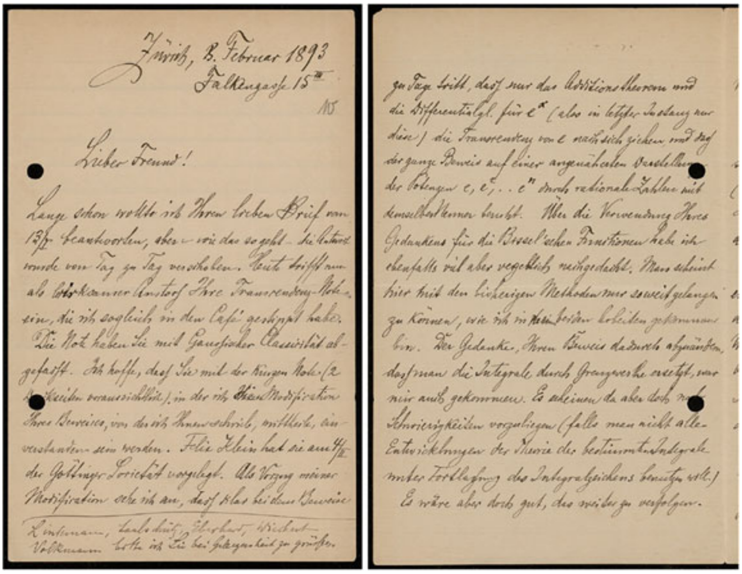


Fig. 8.6 First two pages of the letter from Adolf Hurwitz to David Hilbert, from February 8 or 13, 1893 (Hurwitz 1893b, Br. 10)

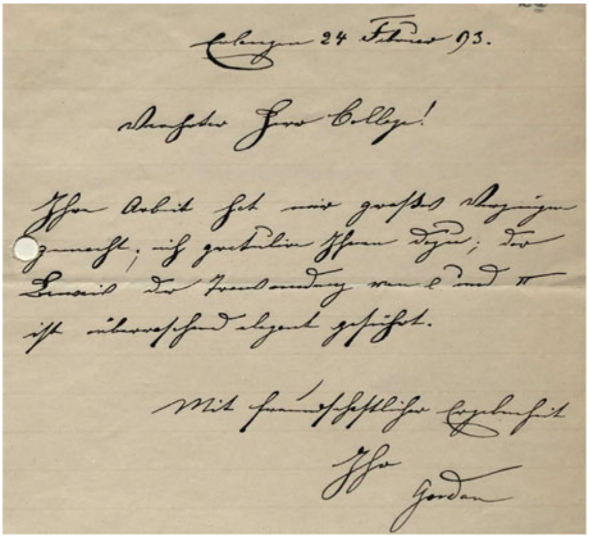
Hilbert is not willing to supplement his publication with Hurwitz's proof and his assessment that his approach will always "be the clearest and most developable" is quite explicit.

When Hurwitz answered one month later (see Fig. 8.6), he is defending his proof:

> It's been a long time since I wanted to answer your kind letter from 13/I, but - how it goes - the answer was shifted from day to day. Today now your note on transcendence, which I have directly dipped into a coffee, appears as guiding hint. You have written the note with Gaussian classicity. I hope that you agree with the note (probably 2 printed pages) in which I wrote down the modification of your proof. Felix Klein showed it 4/II to the Göttingen Society. As advantage of my modification I consider that the proof shows that only the theorem of additivity and the differential equation for e^x (so in last instance only this) imply the transcendence of e, and that the proof relies on an approximate representation of the powers $e, e^2, \ldots e^n$ by rational numbers with the same denominators. [...] I also had the idea of modifying your proof by replacing the integrals by limits. However it seems that there are some difficulties. [...] Perhaps one is lead to a proof which only uses the definition
>
> $$e^z = (1 + \frac{z}{n})^n (n = \infty).$$
>
> [29](Hurwitz 1893b, Br. 10)

---

[29] "Lange schon wollte ich Ihren lieben Brief von 13/I beantworten, aber - wie das so geht - die Antwort wurde von Tag zu Tag verschoben. Heute trifft nun als lenksamer Anstoß Ihre Transzendenz-Note ein, die ich sogleich in einen Café gestippt habe. Die Note haben Sie mit Gaußscher Classizität abgefasst. Ich hoffe, daß Sie mit der kurzen Note (2 Druckseiten voraus-


Erlangen 24 Februar 93.

Verehrter Herr College!

Ihre Arbeit hat mir grosses Vergnügen gemacht; ich gratulire Ihnen dazu; der Beweis der Transcendenz von e und π ist überraschend elegant geführt.

Mit freundschaftlicher Ergebenheit

Ihr Gordan

Fig. 8.7 Letter from Paul Gordan to David Hilbert from February 24, 1893: "Your work has given me great pleasure; I congratulate you on it; your proof of the transcendence of e and π is surprisingly elegantly conducted"

Hurwitz thus insisted on the "advantage of his modification". At the same time using terms like "dipping into coffee" and "Gaussian classicity" he expressed joviality towards his former student and successor at the Königsberg University. But when he wrote about his own accepted, shorter (only two pages) paper and finally mentioned that he had had similar ideas as Hilbert and, moreover, even ideas for further simplifications, Hurwitz nearly seems to hit the ball back. Astonishingly, it was taken up by Hilbert who declared in a letter from March 8, 1893: 'Gordan wrote me a special acknowledgement for my proof of transcendence."[30] (Hilbert 1894, Br. 245) (which indeed had been done in a letter (Gordan 1893a) from 24 February 1893, see Fig. 8.7).

sichtlich) in der ich die Modifikation Ihres Beweises, [...] einverstanden sein werden. Felix Klein hat sie am 4/II der Göttinger Sozietät vorgelegt. Als Vorzug meiner Modifikation sehe ich an, daß klar bei dem Beweis zu Tage tritt, daß nur das Additionstheorem und die Differentialgleichung für e^x (also in letzter Instanz nur diese) die Transzendenz von e nach sich ziehen, und daß der Beweis auf einer angenäherte Darstellung der Potenzen $e, e^2, ...e^n$ durch rationale Zahlen mit demselben Nenner beruht. [...] Der Gedanke, Ihren Beweis dadurch abzuändern, daß man die Integrale durch Grenzwerte ersetzt war mir auch gekommen. Es scheinen aber noch Schwierigkeiten vorzuliegen [...]. Vielleicht wird man auf einen Beweis geführt, der nur die Definition $e^z = (1 + \frac{z}{n})^n (n = \infty)$ benutzt."

[30] "Gordan hat mir extra eine Anerkennung für meinen Transcendenzbeweis geschrieben."

Some weeks later, the slightly elder informed the younger about a recognition of his proof from the initiator of the proofs of transcendence. In a letter from 8 April 1893 he wrote in a surprisingly comparative manner:

> Hermite wrote me in his kind way a letter about my e-proof which is obviously much more convenient to him than yours. He asked me to be allowed to show it to the academy and so it will probably be printed in the next Comptes Rendus [...].[31] (Hurwitz 1893b, Br. 11)

Indeed, Hermite expressed that after Stieltjes and Hilbert, Adolf Hurwitz had probably said "le dernier mot"[32] to the important question of transcendence of e.[33] Explicitly he mentioned:

> Studying the fine work of M. David Hilbert I was a bit embarrassed about the essential point at which he established that the integer number [...] is necessarily different to zero. But everything is of extreme clarity in your analysis and what seems to me the foremost remarkable and completely new is to be led to an impossibility regarding integer numbers, when before we encountered another type of contradiction, an integer equal to a quantity decreasing to an inferior.[34] (Hermite 1893a, Br. 218)

Although Hermite declared a certain preference for Hurwitz's proof line, a second letter from 19 April 1893 shows that he considered Hilbert's work to be mentioned as preparatory:

> I made it my duty to fulfill your intention to mention that [your result] was already published in the Nachrichten der Gesellschaften der Wissenschaften zu Göttingen and at the same time I cited the work of M. David Hilbert in the same collection.[35] (Hermite 1893a, Br. 217)

Hurwitz's proof was published on not more than one and a half printed pages in number 116 of the *Comptes Rendus*[36] which covered the period of January to June 1893.[37]

[31] "Hermite hat mir in seiner liebenswürdigen Weise einen Brief über meinen e-Beweis geschrieben, der ihm offenbar viel bequemer liegt als der Ihrige. Er hat mich gebeten, denselben der Akademie vorlegen zu dürfen und so wird er voraussichtlich [...] in der nächten Comptes Rendus abgedruckt."

[32] "the last word".

[33] See (Hermite 1893a, Br. 218).

[34] "J'avais été un peu embarassé en étudiant le beau travail de M. David Hilbert sur ce point essentiel òu il établit que le nombre entier désigné par [...] est nécessairement différent de zéro. Mais tout est d'une extrême clarté dans votre analyse, et ce qui me semble on ne peut plus remarquable et entièrement nouveau, c'est d'être amené à une impossibilité relative à des nombres entiers, lorsque précédemment on aboutissait à cette autre nature de contradiction, à un entier égal à une quantité inférieurement décroissante."

[35] "Je me suis fait un devoir de remplir votre intention en mentionnant qu'elle a été publiée précédemment dans les Nachrichten der Gesellschaften der Wissenschaften zu Göttingen, et j'ai cité pareillement le travail de M.David Hilbert, dans le même recueil."

[36] *Comptes rendus hebdomadaires des séances de l'Académie des sciences*

[37] See Hurwitz (1893d).

8.5.2 Gordan and Hurwitz

The above briefly indicated involvement of Paul Gordan goes in fact much deeper. The Erlangen mathematician was not only curious about Hurwitz's and Hilbert's work, he had foremost his own ideas of simplifying the proof itself. Only a short time after his appreciative words to Hilbert, Gordan wrote a letter to Hurwitz on 25 April 1893:

> I read your work on the transcendence of the number e with highest interest, however, I believe that some of your conclusions can still be simplified: [...] I beg for your opinion on this.[38] (Gordan 1893b, Br. 176)

What follows is the very elegant proof (see Sect. 8.3) based only on the exponential series concluded with the final remark: "Please tell me whether I am right!". And Adolf Hurwitz approved, answering three days later, on 28 April,

> I have no doubt that your new method only relying on the series expansion of e^x for the proof of the vanishing of $F(k) - e^k F(0)$ for $p = \infty$ is completely in order [...] such that one now only has to operate with the simplest tools of analysis. Mister Hermite had asked me if he could show my proof to the French Academy (of course I agreed to this request with many thanks at that time.) At the same time he meant that now the last word about e had been spoken. I see now that Mister Hermite was in a gentle error at this point and I congratulate you that the last word was reserved to you.[39] (Hurwitz 1893c, Br. 58)

Indeed, Hermite agreed. In a letter from 6 May 1894 he answered to Gordan's submission: "Your new proof which you honored me to communicate to me of the transcendence of the number e seems to me of foremost worth of interest and will arise the attention of all geometers."[40] (Hermite 1893b, Br. 71)

Returning to the correspondence of Hurwitz and Hilbert, it seems that none of them could give up having "the last word". In a letter from May 1, 1893, Hurwitz passed on the information to Hilbert about Gordan's new proof only relying on the series expansion of e^x and refers directly to his remark: "The letter was written quite typically for Gordan. He closed with the words: "Please tell me whether I

---

[38] "Ihre Arbeit über die Transzendenz der Zahl e habe ich mit großem Interesse gelesen, doch glaube ich, daß einige Ihrer Schlüße sich noch vereinfachen lassen. Ich bitte um Ihre Ansicht darüber: [...] Ich bitte mir zu schreiben, ob ich Recht habe!"

[39] "[...] so zweifle ich doch nicht daran, daß Ihre neue nur auf die Reihenentwicklung von e^x sich gründende Methode zum Beweis des Verschwindens von $F(k) - e^k F(0)$ für $p = \infty$ vollständig in Ordnung ist [...] insofern man nun nur mehr mit den einfachsten Hülfsmitteln der Analysis zu operieren braucht. Herr Hermite hatte mich gebeten, daß er meinen Beweis der franz. Akademie vorlegen dürfe (eine Aufforderung der ich natürlich damals mit herzlichem Danke entsprochen habe.) Zugleich meinte er, daß nun mehr über e das letzte Wort gesprochen sein. Ich sehen nun, daß Herr Hermite sich in diesem Punkte in einem liegenswürdigen Irrthume befunden hat und ich beglückwünsche Sie, daß Ihnen das letzte Wort vorbehalten blieb."

[40] "La nouvelle démonstration que vous m'avez fait l'honneur de me communiquer de la transcendance du nombre e me semble on ne peut [plus] digne d'intérêt et méritera l'attention de tous les géomètres."

am right!" "[41] Half a year later (and after a couple of letters without touching the subject of proofs of transcendence), on 6 January 1894, in a letter concerning his visit to Berlin the younger added: "I met Weierstrass in good health and he said that he read my e and π-note with pleasure."[42] (Hilbert 1894, Br. 150) Hurwitz replied on January 30, 1894:

> Your reports from Berlin were of course very interesting to me. So Weierstrass is again completely mobile![43] (Hurwitz 1893b, Br. 17)

8.6 Further Reception

In view of the various proofs of the three and their ambitions, so explicitly marked in the correspondence, regarding simplicity and capacity for further development of the argumentation, we distinguish in the following their reception in teaching and research (until the 1970s).

We have already mentioned Felix Klein and his role in publishing the three proofs on consecutive pages in an issue of the *Mathematische Annalen*. One may be curious about his judgement as to which of the given proofs is the most "simple". In his famous lectures at the Evanston Colloquium , "delivered from Aug. 28 to Sept. 9, 1893, before members of the Congress of Mathematics held in connection with the World's Fair in Chicago" (a forerunner of the later International Congresses of Mathematicians), Klein chose (in Chapter VII) Hilbert's "very simple proof".[44] In addition, he commented

> Immediately after Hurwitz published a proof for the transcendency of e based on still more elementary principles; and finally, Gordan gave a further simplification. [...] The problem has thus been reduced to such simple terms that the proofs for the transcendency of e and π should henceforth be introduced into university teaching everywhere.

In the probably first textbook presenting proofs of transcendence for e and π, namely the *Lehrbuch der Algebra* (Weber 1899, volume 2, §226) of Heinrich Weber, the author praised the proofs of Hilbert, Hurwitz and Gordan as conducted "*with very elementary means and in the simplest way*"[45] (see page 829) and his proof relies mainly on Hurwitz's reasoning. The later textbook of Edmund Landau (1877–1938) presented a similar proof although without mentioning the *simplifiers* explicitly but just the names of Hermite and Lindemann (see Landau 1927, pp.

---

[41] "Der Brief war recht Gordan'sch abgefasst. Er schloß: "Ich bitte Sie mir zu schreiben, ob ich Recht habe! [...]" "

[42] Weierstraß traf ich wohlauf und er sagte, dass er meine e und π-Note mit Vergnügen gelesen habe."

[43] "Ihre Berichte aus Berlin haben mich natrürlich sehr interessiert. Also Weierstraß ist wieder ganz mobil!"

[44] See Klein (1894).

[45] "mit ganz elementaren Mitteln und auf die einfachste Weise"

90–93). In his textbook, Oskar Perron (1880–1975) reproduced Hurwitz's proof, however he writes (see Perron 1921, on page 174)

> The first proof of this theorem is by H e r m i t e. The proof has been significantly simplified later by several authors, namely by H i l b e r t and W e b e r, keeping H e r m i t e 's basic idea, so that it can be made quite elementary today.[46]

The standard reference for number theory for decades, written by Godfrey Hardy (1877–1947) and Edward Wright (1906–2005), gives as well Hurwitz's proof; as already mentioned in the introduction, the authors remarked that the original proofs of Hermite and Lindemann "*were afterwards modified and simplified by Hilbert, Hurwitz, and other writers*".[47] In the North-American community, the booklet of Ivan Niven (1915–1999) was one of the first sources including proof of transcendences and there also Hurwitz's proof is reproduced.[48] The only exception in early sources with a broad readership is the textbook of Alexander Ostrowski (1893–1986) who gave the proof of his mentor Hilbert (see Ostrowski 1954, pp. 85–88).

If we take a look at specialized literature and research in particular, a somewhat different picture emerges. In his important book *Transcendental Numbers*,[49] relying on notes from a lecture given at Princeton in 1946, Carl Ludwig Siegel (1896–1981) presents the transcendence results of Hermite, Lindemann, and the more general result of Weierstrass; his reasoning relies on Hermite's original approach. Siegel wrote at the end of his proofs (on page 30):

> It should be mentioned that the preceding proofs of the transcendency of e and π [...] are not the simplest to be found in literature. Our proofs are related to the original work of Hermite; however, our procedure in constructing the approximation forms is somewhat more algebraic, and this has been decisive for the generalization which we shall investigate in the next chapter.

Siegel refers to his work on what is now called E-functions (from 1929; see also Shidlovskii (1989)) and later he discusses path-breaking results of his pupil Theodor Schneider (1911–1988), who proved Hilbert's seventh problem on the transcendence of α^β with algebraic α, β (such that $\alpha \neq 0, 1$ and β is irrational) in 1934 independently to Aleksandar Gelfond (1906–1968); both reasonings rely on an analytic treatment of complex-valued functions.

That a transcendence problem found its way into Hilbert's famous list of 23 problems for the twentieth century is, in retrospect, no wonder. But it is remarkable to realize at this point to what extent Hilbert developed in the short time from "Hurwitz's pupil" in 1892 to one of the few universal mathematicians in 1900,

---

[46] "Der erste Beweis dieses Satzes stammt von H e r m i t e. Der Beweis ist später unter Beibehaltung des H e r m i t eschen Grundgedankens von mehreren Autoren, namentlich von H i l b e r t und W e b e r, bedeutend vereinfacht worden, so daß er sich heute ganz elementar gestalten läßt."

[47] See Hardy and Wright (1959).

[48] See Niven (1963).

[49] See Siegel (1949).

parallel to the "changing of the generations in the German mathematical world" (cf. Gray 2000, p. 33). It is an interesting side note that in the context of his seventh problem (Hilbert 1902) referred to Hurwitz (1883)[50] indicating the latter one's early interest in questions around transcendency.

Besides the Gelfond-Schneider theorem another important transcendency result of that time was Schneider's treatment of abelian functions (in 1937). What we observe here is indeed a return to the roots! In his monograph (Schneider 1957, p. 47), Schneider makes his perspective pretty clear:

> The fact that after the above-mentioned proofs of HERMITE on the transcendence of e and of LINDEMANN on the transcendence of π numerous further proofs on the same subject have been published, especially in the nineties of the last century, shows that these proofs, although perfect, appeared to be either unsatisfactory with respect to transparency or capable of improvement with respect to the means used. Special emphasis was put on the most possible elimination of analytical means and thus a situation arose which caused G. Hessenberg in 1911 to write his book on the "Transcendence of e and π" in order to work out clearly the basic ideas of the proofs [...]. Just by elementarization these basic ideas were veiled and the resulting generalizable approaches of HERMITE and LINDEMANN were so narrowed and tailored to the exponential function and its inverse function, that also for this reason further results were not obtained.[51]

Schneider's book was published in the renowned Springer series *Die Grundlagen der Mathematischen Wissenschaften in Einzeldarstellungen* (as volume LXXXI) and became a standard reference for some period. Concerning the issue of simplification, Baker wrote in (Baker 1975, p. 3) two decades later

> The work of Hermite and Lindemann was simplified by Weierstrass in 1885, and further simplified by Hilbert, Hurwitz and Gordan in 1893. We proceed now to demonstrate the transcendence of e and π in a style suggested by these later writers.

After giving the proofs, hoewever, he added (on page 8)

> The above proofs are simplified versions of original arguments of Hermite and Lindemann and their motivation may seem obscure; indeed there is no explanation *a priori* for the introduction of the functions I and f. A deeper insight can best be obtained by studying the basic memoir of Hermite where, in modified form, the functions first occured, but it may be said that they relate to generalizations, concerning simultaneous approximation, of the convergents in the continued fraction expansion of e^x.

---

[50] The 1883 article had been accompanied by a second part in 1888.

[51] "Daß nach den genannten Beweisen von HERMITE zur Transzendenz von e und von LINDEMANN zur Transzendenz von π zahlreiche weitere Beweise über den gleichen Gegenstand, vor allem in den neunziger Jahren des vorigen Jahrhunderts veröffentlich worden sind, zeigt, daß diese Beweise, obzwar einwandfrei, so doch entweder in bezug auf Durchsichtigkeit unbefriedigend oder in bezug auf die verwendeten Hilfsmittel verbesserungsfähig erschienen waren. Man legte gerade auf möglichste Ausschaltung analytischer Hilfsmittel ein besonderes Gewicht und es entstand so eine Situation, die G. HESSENBERG im Jahre 1911 veranlaßte, sein Buch zur "Transzendenz von e und π" zu schreiben, um die Grundgedanken der Beweise deutlich herauszuarbeiten [...]. Gerade durch die Elementarisierung waren diese Grundgedanken verschleiert und die daraus verallgemeinerungsfähigen Ansätze von HERMITE und LINDEMANN so sehr eingeengt und auf die Exponentialfunktion und deren Umkehrfunktion zugeschnitten, daß auch deshalb weitere Ergbenisse nicht erzielt wurden."

In his lecture notes, Kurt Mahler (1903–1988) discussed the various proofs arising in the 1890s in a final chapter in detail. He begun with the words (see Mahler 1976, p. 213)

> We shall here collect a number of such proofs by different mathematicians and explain their relations. Since Siegel's fundamental paper of 1929 (Chapters 4–9), entirely new and very powerful methods have been introduced into the theory of transcendental numbers. This has the danger that the easier and often very ingenious classical proofs may be forgotten.

Concerning the flexibilities, he later wrote (on page 243)

> A. Hurwitz (1893) satisfies the divisibility conditions in a seemingly simpler way by taking his parameter equal to a sufficiently large prime. This is a very convenient choice, but it imposes unnecessary restrictions on the parameter without actually simplifying the proof. The disadvantage of Hurwitz's choice becomes particularly evident if one wants to derive a measure of transcendence, say for e.

But of course, this was not the intention of Hurwitz and others who struggled for a *simple* proof; the measure of transcendence is a concept of the twentieth century.

We conclude with a brief evaluation of this compilation of receptions. If the attribute *simple* or *elementary* is intended to mean that a certain argument can be accomplished with as few additional aids as possible, then, of course, Gordan's proof is the simplest. Comparing the proofs of Hilbert and Hurwitz, Hurwitz's reasoning has to be regarded more simple since Hilbert's argumentation needs in addition to Hurwitz's use of the differential equation $\exp' = \exp$ also the fundamental theorem of differentiation and integration as well as partial integration (or basic knowledge of the Gamma-function); on the contrary, Hurwitz applies the mean-value theorem and the infinitude of primes, two results that can be considered more fundamental. This ranking may as well explain the choices of Hurwitz's proof for general textbooks. If the focus is on research (e.g., the generalization to E-functions), however, then simplification is not of major interest but issues like flexibility or options for generalizations are relevant. In this sense, the generalization by Weierstrass (1885) had proven to be the most fruitful for the advancement of the transcendency proofs of the nineteenth century.

Acknowledgments The author is grateful to the anonymous referees for their valuable comments and suggestions.

References

Baker, A. 1975. *Transcendental Number Theory*. Cambridge: Cambridge University Press.

Gordan, P. 1893a. *Briefe und Postkarten von Paul Gordan an David Hilbert*. Signatur: Cod. Ms. D. Hilbert 116: Nachlass David Hilbert, Niedersächsische Staats- und Universitätsbibliothek Göttingen.

Gordan, P. 1893b. *Briefe von Paul Gordan an Adolf Hurwitz*. Signatur: Cod. Ms. Math.-Arch. 76: Mathematiker-Archiv, Niedersächsische Staats- und Universitätsbibliothek Göttingen.

Gordan, P. 1893c. Sur la transcendance du nombre e. *Comptes Rendus de l'Académie des Sciences (Paris)* 116: 140–141.

Gordan, P. 1893d. Transzendenz von e und π. *Mathematische Annalen* 43: 222–224.
Gray, J. 2000. *The Hilbert Challenge*. Oxford: Oxford University Press.
Gray, J. 2018. *A History of Abstract Algebra*. Berlin: Springer.
Hardy, G., and E. Wright. 1959. *The Theory of Numbers*. 4th ed. Oxford: Clarendon Press.
Hermite, C. 1873. Sur la fonction exponentielle. *Comptes Rendus de l'Académie des Sciences (Paris)* 77: 18–24, 74–79, 226–233, 285–293.
Hermite, C. 1893a. *Briefe von Charles Hermite an Adolf Hurwitz*. Signatur: Cod. Ms. Math.-Arch. 76: Mathematiker-Archiv, Niedersächsische Staats- und Universitätsbibliothek Göttingen.
Hermite, C. 1893b. *Briefe von Charles Hermite an Paul Gordan*. Signatur: Ms. 2148/13: Universitätsbibliothek Erlangen-Nürnberg.
Hessenberg, G. 1912. *Transzendenz von e und π*. Leipzig: Teubner.
Hilbert, D. 1892–1894. *Briefe von David Hilbert an Adolf Hurwitz*. Signatur: Cod. Ms. Math.-Arch. 76: Mathematiker-Archiv, Niedersächsische Staats- und Universitätsbibliothek Göttingen.
Hilbert, D. 1893. Über die Transcendenz der Zahlen e und π. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*: 113–116.
Hilbert, D. 1902. Mathematical Problems. *Bulletin of the American Mathematical Society* 8: 437–479.
Hilbert, D. 1921. Adolf Hurwitz. *Mathematische Annalen* 83: 161–168.
Hurwitz, A. 1883. Ueber arithmetische Eigenschaften gewisser transcendenter Functionen. *Mathematische Annalen* 22: 211–229.
Hurwitz, A. 1884–1919. *Die Mathematischen Tagebücher und der übrige handschriftliche Nachlass von Adolf Hurwitz*. Hs 582:1–30: University Archives ETH Zurich Library.
Hurwitz, A. 1893a. Beweis der Transzendenz der Zahl e. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*: 153–155.
Hurwitz, A. 1893b. *Briefe von Adolf Hurwitz an David Hilbert*. Signatur: Cod. Ms. D. Hilbert 160: Nachlass David Hilbert, Niedersächsische Staats- und Universitätsbibliothek Göttingen.
Hurwitz, A. 1893c. *Briefe von Adolf Hurwitz an Paul Gordan*. Signatur: Ms. 2148/13: Universitätsbibliothek Erlangen-Nürnberg.
Hurwitz, A. 1893d. Démonstration de la transcendance du nombre e. *Comptes Rendus de l'Académie des Sciences (Paris)* 116: 788–789.
Jamet, V. 1891. Sur le nombre e. *Nouvelles Annales de Mathématiques* 10: 215–218.
Klein, F. 1894. *The Evanston Colloquium*. New York: Macmillan.
Landau, E. 1927. *Vorlesungen über Zahlentheorie*. Leipzig: 3. Verlag von S. Hirzel.
Lindemann, F. 1882a. Über die Ludolphsche Zahl. *Sitzungsberichte der Preussischen Akademie der Wissenschaften*: 679–682.
Lindemann, F. 1882b. Über die Zahl π. *Mathematische Annalen* 20: 213–225.
Liouville, J. 1844. Nouvelle démonstration d'un thèorème sur les irrationnelles algébriques. *Comptes Rendus de l'Académie des Sciences (Paris)*: 910–911.
Mahler, K. 1976. *Lectures on Transcendental Numbers*. Number 546. Lecture Notes in Mathematics. Berlin: Springer.
Markoff, A. 1883. *Proof of the Transcendence of the Numbers e and π*. St. Petersburg (in Russian).
Minkowski, H. 1973. *Briefe an David Hilbert*, ed. L. Ruedenberg and H. Zassenhaus. Berlin: Springer.
Niven, I. 1963. *Irrational Numbers*. Hoboken: John Wiley & Sons.
Ostrowski, A. 1954. *Vorlesungen über Differential- und Integralrechnung*, vol. 3. Boston: Birkhäuser.
Oswald, N. 2014. David Hilbert, ein Schüler von Adolf Hurwitz? *Siegener Beiträge zur Geschichte und Philosophie der Mathematik* 4: 11–30.
Oswald, N., and J. Steuding. 2014. Complex Continued Fractions: Early Work of the Brothers Adolf and Julius Hurwitz. *Archive for History of Exact Sciences* 68: 499–528.
Perron, O. 1921. *Irrationalzahlen*. Göschens Lehrbücherei. Berlin und Leipzig: Walter de Gruyter.
Rowe, D. 2007. Felix Klein, Adolf Hurwitz, and the "Jewish Question" in German Academia. *Mathematical Intelligencer* 28: 18–30.

Samuel-Hurwitz, I. 1984. *Erinnerungen an die Familie Hurwitz, mit Biographie ihres Gatten Adolph Hurwitz, Prof. f. höhere Mathematik an der ETH*. Signatur: Hs 583a:2: University Archives ETH Zurich Library.

Schneider, T. 1957. *Einführung in die transzendenten Zahlen*. Berlin: Springer.

Shidlovskii, A. 1989. *Transcendental Numbers*. Berlin: Walter de Gruyter.

Siegel, C. 1949. *Transcendental Numbers*. Princeton: Princeton University Press.

Stieltjes, T. 1890. Sur la fonction exponentielle. *Comptes Rendus de l'Académie des Sciences (Paris)* 110: 267–270.

Venske, O. 1890. Über eine Abänderung des ersten HERMITEschen Beweises für die Transzendenz der Zahl e. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*: 335–338.

Weber, H. 1899. *Lehrbuch der Algebra*, 2nd ed., vol. 2. Braunschweig: F. Vieweg.

Weierstrass, K. 1885. Zu Lindemann's Abhandlung. "Über die Ludolph'sche Zahl". *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, Berlin*: 1067–1085.

Weyl, H. 1944. David Hilbert and His Mathematical Work. *Bulletin of the American Mathematical Society* 50: 612–654.

Chapter 9
Current and Classical Notions of Function in Real Analysis

Colin McLarty

Abstract Not only historians of mathematics but also working analysts know how seventeenth through nineteenth century mathematicians advanced from vaguer notions to the set theoretic idea of function. The celebrated *Princeton Lectures in Analysis* of Elias Stein and Rami Shakarchi are shaped around that history, and review it at some length in both volumes 3 and 4. Stein and Shakarchi take the set theoretic notion as their official definition of function, but the reason they review that history twice is explicitly to contrast it with the modern forms of two other classical notions of function that they use informally. Terence Tao studied with Stein in the 1990s and emphasizes an even wider view of the function concept. All these authors show both how and why these generalizations of the set theoretic notion of function suit the purposes of classical and current analysis.

This note on the notion of function is part of a project on intertwined classicism and innovation in analysis. This project benefitted early from Jeremy Gray's advice on the Navier-Stokes equation and then from his book *Change and Variations: A History of Differential Equations to 1900* (2021) which arrived at the start of the writing this note. His work and his advice have been decisive for my sense of historical method and my concrete historical projects.

Not only historians of mathematics but also working analysts know how seventeenth through nineteenth century mathematicians advanced from vaguer notions to the set theoretic idea of function. Take Elias Stein and Rami Shakarchi's celebrated *Princeton Lectures in Analysis*. Volume 3 on real analysis opens with the history of the set theoretic function concept (Stein and Shakarchi 2005, pp. xvff.). Volume 4 on functional analysis goes over it again (Stein and Shakarchi 2011, pp. 98ff.).

However, the purpose of the *Princeton Lectures* review of that history is to show how the set theoretic definition is not now and has never been the most general notion of function in analysis. Those lectures review the history twice expressly

C. McLarty (✉)
Case Western Reserve University, Cleveland, OH, USA
e-mail: colin.mclarty@cwru.edu

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_9

in order to contrast it with the modern forms of two other classical notions of function that they will use. While they take the set theoretic definition as official, the real analysis volume contrasts the set theoretic idea with *measurable functions*. The lectures informally, but constantly, treat measurable functions as real-valued functions on the real vector spaces $\mathbb{R}^d$ which violate the set theoretic definition of function by having no determinate value at any single point $x \in \mathbb{R}^d$ (Stein and Shakarchi 2005, pp. 69, 157). See Sect. 9.3 below. The functional analysis volume contrasts the history with *distributions* which it treats as *generalized functions* notably including the Dirac delta function $\delta(x)$. These often violate the set theoretic definition of function by allowing no possible value, finite or infinite, at crucial points, as Sect. 9.2.3 below explains in the case of $\delta(x)$. Terence Tao studied with Elias Stein in the 1990s, and has emphasized an even wider view of the function concept as we will see in Sects. 9.2.2–9.2.4.

There is no controversy here. This is not about foundations of mathematics. Nor is there active debate over calling measurable functions or distributions "functions." Stein, Shakarchi, and Tao, like many analysts, often call them that and this does not change the mathematics at all. Exactly the same theorems are proved by analysts who take care to avoid such usage as by analysts who consider the usage informal, convenient, indispensable in practice, or actually correct. The topic is historically interesting, though, because the variant notions of function in analysis retain usages older than the set theoretic definition.

9.1 The Modern Definition of Function Circa 1930

Édouard Goursat's influential *Cours d'analyse mathématique* looked back one hundred years to say:

> The modern definition of the word *function* is due to Cauchy and to Riemann. One says *y is a function of x when to each value of x corresponds a value of y*. One indicates this dependence by the equality $y = f(x)$. (Goursat 1933, p. 13)

Numerous historical sources show the practical meaning of that definition was much more widely and clearly understood by Goursat's time than in Cauchy's or even Riemann's. Notably, over that span of time, Cantor, Dedekind and others created set theory, and one major motivation was to clarify this definition of function. Working standards of explicit rigor are higher today than Goursat's standards in 1930. Yet to this day the definition of *function* in analysis textbooks is little changed verbatim from what Goursat credits to Cauchy and Riemann.

Take this example from Tao's undergraduate textbook *Analysis I*:

> A function $f : X \to Y$ from one set X to another set Y is an operation which assigns to each element (or "input") x in X, a single element (or "output") $f(x)$ in Y. (Tao 2014, p. 49)

To put it in a fuller context, though, Tao calls this the notion of function specific to *set theory*. Section 9.2.1 returns to this. He stresses that some central notions of function in other fields of mathematics, notably including advanced analysis, do not fit this definition. Section 9.2.2 goes into that. Most of Tao's examples originated in the mid- to late-twentieth century, but his examples from analysis go back before Fourier in 1820.

Real analysis is an unusually classicist branch of mathematics. A time-travel metaphor can explain my meaning: Leonhard Euler would meet striking new ideas if he could read Fefferman (2000) describing the Millennium Prize problem on the Navier-Stokes equations. But after a brief introduction to vector calculus notation, Euler could follow the outlines of the new ideas right there in Fefferman's essay. Number theory for example is not like that. Euler would need substantial preparation in Galois theory and Riemann surfaces before getting any idea of the Frobenius endomorphisms and modularity in Bombieri (2000) describing the Millennium Prize for the Riemann Hypothesis. Euler's ideas were formative for both problems, as Fefferman and Bombieri each describe. The other Millennium Prize problems in algebraic topology, algebraic geometry, computation theory, and mathematical physics are all farther from what Euler knew.

Gray calls Euler prescient on the prospect for fluid dynamics, and quotes a passage of Euler illuminating the classicism of real analysis. After successful work on equilibrium in fluids, Euler wrote:

> I now propose to deal with the motion of fluids in the same way.... a much more difficult undertaking.... Nevertheless I hope to arrive at an equally successful conclusion, so that, if difficulties remain, they will pertain not to Mechanics but purely to Analysis....

Euler was confident his tools, including his idealizations such as perfect continuity and incompressibility of the fluid, were the right mathematics for fluid dynamics. He knew this mathematics would be extremely difficult to use. But he was sure this Analysis would give "the general principles on which the entire science of fluid motion is based" (Gray 2021, p. 44).

History to date bears out Euler's confidence. His own *Euler equations* are basic to the slightly later *Navier-Stokes equations* (Fefferman 2008). Both are obviously physically unrealistic in many ways, and famously difficult to use. Applications virtually always require very coarse simplification of the equations and often rely on ad hoc corrections. But these equations express basic fluid motion so well that today they are more securely central to theoretical and practical fluid dynamics than ever: "they are nowadays regarded as the universal foundation of fluid mechanics" (Galdi 2011, p. 3). Pierre Gilles Lemarié-Rieusset's book *The Navier-Stokes Problem in the twenty-first century* (2015) presents two centuries of deep technical innovations, and current wide open problems, focusing on the pure mathematics. That book would offer Euler many new methods (especially using topological vector spaces) and correspondingly many new theorems. But I urge, without attempting a rigorous argument, he would recognize the sense of it all as his own idea of analysis.

9.2 Functions as a Theme

Terence Tao, known among other things for his 2006 Fields Medal in analysis, and quality exposition at many levels of mathematics, discusses the notion of function in two contexts. In the research context he describes an array of function-like devices which are not set theoretically functions but have been used as functions in analysis since at least Joseph Fourier in the 1820s. New kinds are still being created today especially for use with non-linear differential equations. Tao urges thinking of these notions as forming a spectrum of *smoother* and *rougher* functions (Tao 2008a,b).

In the teaching context, Tao's undergraduate textbook *Analysis I,II* (Tao 2014, 2016b) uses an original axiomatic set theory giving functions more independent reality than they have in Zermelo-Fraenkel set theory and less than in categorical set theory. These axioms are not presented as a rival to any other foundations of mathematics as a whole. Tao offers these as helping teach a unified grasp of analysis.[1]

These topics are best approached in the opposite order to their publication.

9.2.1 Set Theory in Analysis I, II

Through several years teaching an honors class in real analysis, Tao produced his *Analysis I,II*. The topics and theorems are classical. The exposition conveys Tao's distinct vision of the path through it all. For this, Tao gives his own axiomatic set theory. These axioms aim to do exactly what the preface says the whole book aims to do. That is, make logical rigor an aid rather than an obstacle to substantial insight for beginners in real analysis.

Some students asked why they were spending time on apparent trivia, but:

> when [the students] did persevere and obtain a rigorous proof of an intuitive fact, it solidified the link in their minds between the abstract manipulations of formal mathematics and their informal intuition of mathematics (and of the real world), often in a very satisfying way. (Tao 2014, p. xv)

Few analysis textbooks include an explicit axiomatic set theory. Fewer still include an original one. One famous example is John L. Kelley's *General Topology* (1955). The most influential aspect of Kelley's book was the clear, concise organization of the definitions and theorems of point set topology. But Kelley also created a version of what is now called Morse–Kelley set theory, to accommodate large, proper-class sized, categories of topological spaces and maps.

While sharing Kelley's interest in rigor from the ground up, Tao has quite different concrete motives for his axioms than Kelley had. Tao offers the honors

[1] Tao posted more foundational discussions on line, partly collected in Tao (2013). Here I discuss only the set theory in Tao (2014) and subsequent editions of that book.

analysis students a conceptual alternative to the formal reduction of functions to sets in Zermelo-Fraenkel set theory. In fact his 2014 edition merely said functions might not themselves be sets. But two years later he wrote:

> Strictly speaking, functions are not sets, and sets are not functions; it does not make sense to ask whether an object x is an element of a function f, and it does not make sense to apply a set A to an input x to create an output $A(x)$. (Tao 2016a, p. 51)

Rather, Tao defines functions in set theory as objects related to sets in the following way. For any property $P(x, y)$:

> **Definition 3.3.1** (Functions). ... [if] for every $x \in X$, there is exactly one $y \in Y$ for which $P(x, y)$ is true.... [then] define the function $f : X \to Y$ *defined by* P *on the domain* X *and range* Y to be the object which, given any input $x \in X$, assigns an output $f(x) \in Y$, defined to be the unique object $f(x)$ for which $P(x, f(x))$ is true. (Tao 2014, p. 51)

That same page defines equality of functions, $f = g$ as holding if and only if $f(x) = g(x)$ holds for all $x \in X$. And so a function f is fully described by a certain set called its *graph*:

$$\{\langle x, f(x)\rangle \mid x \in X\} \subset X \times Y.$$

But Tao is explicit the function is not a set. It is distinct from its graph.

To emphasize the centrality of functions, Tao states his power set axiom in terms of function sets:

> **Axiom 3.11** (Power set axiom). Let X and Y be sets. Then there exists a set, denoted Y^X, which consists of all the functions from X to Y. (Tao 2014, p. 59)

All this emphasis on functions suits how Tao wants the students in honors analysis to see set theory. And yet he notes other parts of mathematics use *morphisms* the way set theory uses functions, and indeed morphisms can be functions in the set theoretic sense, but are not always (Tao 2014, p. 49).

9.2.2 The Best Ways of Describing Functions

Tao (2008a) says in set theory the fundamental operation on a function $f : X \to Y$ is evaluation at an element $x \in X$ to give a value $f(x) \in Y$:

> However, there are some fields of mathematics where this may not be the best way of describing functions. In geometry, for instance, the fundamental property of a function is not necessarily how it acts on points, but rather how it *pushes forward or pulls back* objects that are more complicated than points.... Similarly, in analysis, a function need not necessarily be defined by how it acts on points, but may instead be defined by how it acts on other objects [Tao mentions sets, and test functions] the former leads to the notion of a *measure*, the latter to the notion of a *distribution*. (Tao 2008a, p. 184f.)

Tao cites other sections of the *Princeton Companion to Mathematics* for geometric examples. These more or less arose around the 1930s and Tao describes the conception of them that became prominent in and after the 1950s (McLarty 2007). The examples he gives from analysis were also extensively re-organized through the mid- and late twentieth century, and Tao describes how they are still developing today, but these ideas arose much earlier.

9.2.3 Analysis: Fourier and the Dirac Delta Function

In 1822 Joseph Fourier's *Analytical Theory of Heat* (2009) drew on a century of incisive work by Euler and other great mathematicians, to present utterly confident solutions to differential equations from physics, by calculating integrals that many of Fourier's contemporaries found dubious or nonsensical. Fourier's ideas had a compelling coherence and produced very valuable purported results, yet were too swift to entirely trust. In fact they cannot be justified in the sweeping generality that Fourier claimed. Efforts to clarify these methods became a major force driving the next century and more of rigorization of analysis. See numerous references to Fourier in the index of Kline (1972). Stein and Shakarchi (2003, p. 23) give a précis of this history, and shape the four volumes of their *Princeton Lectures in Analysis* largely around key steps in the long, rigourous development of Fourier's ideas.

Fourier's boldest single calculation was his claim that "any function" g is expressible in the following form (Kline 1972, p. 680). We lightly adapt the notation:

$$g(y) \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} g(x)\int_{-\infty}^{\infty} \cos q(x-y)\, dq\, dx. \tag{9.1}$$

Comparing Eq. (9.1) with Eq. (9.3) below shows that according to this calculation by Fourier, for $y = 0$, the inside trigonometric integral defines the Dirac delta function:

$$\delta(x) \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} \cos(q\cdot x)\, dq. \tag{9.2}$$

Some mathematicians objected that this and related integrals are ill-defined and actually impossible. But Fourier did not care. He felt he knew how to calculate with them. See discussion by Lützen (1982, p. 113).

Today the Dirac delta $\delta(x)$ is usually defined by simply stipulating that, for all functions $g:\mathbb{R}\rightarrow\mathbb{R}$,

$$\int_{-\infty}^{\infty} g(x)\cdot\delta(x)\, dx \;=\; g(0). \tag{9.3}$$

It is widely described as a function from $\mathbb{R}$ to $\mathbb{R}$ with $\delta(x) = 0$ whenever $x \neq 0$, and $\delta(0)$ so high that the area under the graph is 1.

The inspired core calculus textbook (Strang 2015, p. 23) gives a concise practical account of $\delta(x)$ in that way, and then chapters full of examples. Strang immediately notes a problem: no function in the set theoretic sense meets Eq. (9.3) for all g, or has any such graph. But a textbook on the essential ideas of differential equations, largely for engineering students, is not the place to worry about that. Strang uses $\delta(x)$ throughout his book to teach correct intuitions and calculations with differential equations including Fourier and Laplace transforms. Strang warns the reader that $\delta(x)$ is "by no means an ordinary function." By the same token, Eq. (9.3) is no ordinary integral.

Someone might try to make $\delta(x)$ an ordinary function by adjoining infinities $-\infty, \infty$ at each end of the real line $\mathbb{R}$. This extended real line is widely used in analysis. It works well for many purposes. But defining $\delta(0) = \infty$ will not work. Equation (9.3) requires that doubling $\delta(x)$ gives a different function. Doubling ∞ on the extended real line just gives back ∞. Indeed, no kind of trick can make $\delta(x)$ an ordinary function if the integral in Eq. (9.3) is to be much like either a Riemann integral or a Lebesgue integral. Famously, both the Riemann and the Lebesgue integrals keep the same value when the integrand is changed at any single point. Changing $\delta(x)$ at the single point $x = 0$ by setting $\delta(0) = 0$ would give the constant 0 function, utterly violating Eq. (9.3).

Set theory can formalize $\delta(x)$ as a part of an operator $\int_{\mathbb{R}} \_ \cdot \delta(x)\, dx$ taking each suitable function $g(x)$ to the number $\int_{-\infty}^{\infty} g(x) \cdot \delta(x)\, dx$. Equation (9.3) does exactly that. Compare Stein and Shakarchi (2011, pp. 100f.). This equation defines $\delta(x)$ by its action on test functions g, just as Tao meant in the quote above saying a function in the wider sense could be defined by how it acts on test functions. Set theory can also formalize $\delta(x)$ in other ways we will not go into. The point for this essay is, none of these ways make $\delta(x)$ itself a function in the set theoretic sense. None defines $\delta(x)$ by its values at points $x \in \mathbb{R}$ or even allows it to have a value at the decisive point $x = 0$.[2]

The whole point of the Dirac delta $\delta(x)$ is its behavior at $x = 0$. That behavior contradicts any possible value $\delta(0)$.

[2] Key steps in the history of distributions: Leray (1934) developed a version to solve the 3 dimensional Navier-Stokes equation (Lemarié-Rieusset 2015, Ch. 12). And in 1944 Laurent Schwartz gave them a rigorous foundation using topological vector spaces. Barany (2018, p. 263) documents how early adopters of Schwartz's distributions took them variously as "a banal trick for applied calculations, a difficult intervention in the recent theory of topological vector spaces, a profound realignment of established methods, a radical departure from familiar concepts, and many things in between." Cartier (2021) gives a Bourbaki insider's perspective: "this was the great talent of Schwartz: to give a simple idea that works."

9.2.4 Smoother and Rougher Functions in Analysis

> In analysis, it is helpful to think of the various notions of a function as forming a spectrum, with very "smooth" classes of functions at one end and very "rough" ones at the other. (Tao 2008a, p. 185)

The exposition of *function spaces* and *harmonic analysis* in Tao (2008b,c) shows how concretely helpful this is. Tao not only shows the various notions of functions can be ordered as more-or-less a spectrum, but illustrates how his concept of roughness guides intuition about what can and cannot be done with functions at different points along the spectrum. Compare Sect. 9.3 below.

Tao uses "smooth" itself for the functions traditionally called smooth: those which have derivatives of all finite orders. Smooth functions which furthermore are the limits of their Taylor series, Tao calls "very smooth." Continuous functions which lack some derivatives are less smooth, and discontinuous are rougher yet. The rough end of Tao's spectrum includes function-like things that are not set-theoretically functions: Tao names Borel measures, Distributions, and hyperfunctions. We will take the Dirac delta $\delta(x)$ as a paradigm of distributions. For the general definition of distributions see Tao (2008a), Stein and Shakarchi (2011, Ch. 3), and philosophic discussion in McLarty (2023).

The smoother a function is, the more things you can be sure you can do with it. Smooth, or even just continuous, functions can be added to each other, or multiplied by each other. Distributions in general cannot by multiplied by each other, and cannot even be multiplied by arbitrary smooth functions. Smooth, or even just continuous, functions can be evaluated at every point of their domain, while $\delta(x)$ cannot be evaluated at 0. On the other hand, the rougher classes are more inclusive, and in that sense rougher functions may be easier to find.

A sequence of smooth functions may converge to a rougher limit. Intuitively, a series of smooth curves with tighter and tighter bends in the middle of each may have a limit with a sharp kink there. A sequence of smooth approximate solutions to some differential equation may converge to a rough solution. By this means in the 1930s Jean Leray proved the 3-dimensional Navier-Stokes equation has distribution solutions for all initial conditions—while a still-open Clay Millennium problem asks whether it has smooth solutions (Fefferman 2000).

Distributions and hyperfunctions often serve as *weak* or *generalized* solutions to differential equations when solutions by set theoretic functions are impossible—or at least currently unknown. They are often lumped together as *generalized functions* because that is how analysts think about them. In the opening quote for this section Tao calls them "various notions of a function" and his cited articles show how well this viewpoint works for a number of topics.

9.3 Measuring Size in Two Contrasting Examples

These examples follow the organizing theme of Tao's (2008b), namely that different measures of the "size" of a function will work for functions at different points along the smoothness spectrum. They illustrate why rougher functions may not have well-defined values at single points. And the mathematics of continuous functions versus integrable functions on an interval $[-1, 1]$ goes back to Fourier's representation of heat distribution along a bar (Gray 2021).

First consider real-valued functions defined and continuous on the closed interval $[-1, 1] \subset \mathbb{R}$. The closed interval is compact so the absolute value of such a function actually achieves some finite maximum value $|f(x)|$. So one useful measure of the size of these functions is their maximum absolute values. Call this the *max-norm*.

$$\|f\|_{max} = Max\,\{|f(x)| \in \mathbb{R} \mid x \in [-1, 1]\}.$$

A trivial fact is central to the use of this norm:

$$\|f\|_{max} = 0 \text{ if and only if } f \text{ is the constant function } 0. \tag{9.4}$$

Functional analysis makes great use of the *normed vector space* $C^0([-1, 1])$. That is an infinite dimensional vector space where the points are these functions. Addition of these vectors $f + g$ and scalar multiplication $c \cdot f$ are defined in the natural way:

$$(f + g)(x) = f(x) + g(x) \qquad (c \cdot f)(x) = c \cdot (f(x)).$$

And $\|f\|_{max}$ is a well defined vector norm. Notably, the only function with $\|f\|_{max} = 0$ is the constant function 0.

A function f may have very small values, even value 0, across most of the interval $[-1, 1]$ and yet be "large" in the max-norm. It only needs to have a large value somewhere.

Continuous functions sit towards the smooth end of Tao's spectrum. Intuitively: the graph of a continuous function may have kinks but is unbroken. Analysts often need rougher functions, with graphs broken at many points, to say the least. Classical analysts, notably including Fourier, did this over 200 years ago.

For the second, rougher example, take *integrable* real-valued functions on $[-1, 1]$. That is, possibly discontinuous real-valued functions f which have a well-defined integral over $[-1, 1]$.[3]

[3] For our purposes think of either the improper Riemann or the Lebesgue integral.

Consider the discontinuous integrable function defined for all $x \in [-1, 1]$ by

$$f(x) = \begin{cases} 0, & \text{if x=0;} \\ |x|^{\frac{-1}{2}}, & \text{if x } \neq 0. \end{cases}$$

The value $f(x)$ rises towards infinity as x approaches 0 from either side though it leaps back to 0 at the single point $x = 0$. The function attains no maximal value on the interval $[-1, 1]$. Yet it is integrable because it rises and falls so narrowly just around 0 that the area under the graph is finite. Specifically that area is 4.

So the max-norm is not defined for all integrable functions. But integration itself defines a norm, called the *1-norm*:

$$\|f\|_1 = \int_{-1}^{1} |f(x)|\, dx.$$

The analogue for $\|f\|_1$ of Eq. (9.4) on $\|f\|_{max}$ says:

$$\|f\|_1 = 0 \text{ if and only if } f = 0 \text{ almost everywhere.} \tag{9.5}$$

We will not define *almost everywhere* beyond this germane but vastly understated sufficient condition: Take a function g on $[-1, 1]$ and change its value at any one point in $[-1, 1]$. Then the new function is equal to the original g almost everywhere.

Every continuous function f is integrable on the interval $[-1, 1]$. If f has generally small values but is very large near just a few points, then it will be very large in the max-norm and small in terms of $\|f\|_1$. Conversely if the value of f is not much larger at any point than at any other, then it may have $\|f\|_{max} < \|f\|_1$. There is just one necessary relation between the two size measures. Because the interval $[-1, 1]$ has length 2:

$$\|f\|_1 \leq 2\|f\|_{max}.$$

Stein and Shakarchi (2005, p. 69) use a well chosen parenthesis to say the size measure $\|f\|_1$ "gives a (somewhat imprecise) definition" of the function space $L^1([-1, 1])$. Here $L^1([-1, 1])$ is a normed vector space of integrable functions just as $C^0([-1, 1])$ is a normed vector space of continuous functions. The imprecision is that a norm must take value 0 only for the 0 vector, and not also for many others near to it. The solution universally adopted by analysts is not to change the definition of the norm. It is to define the points of $L^1([-1, 1])$ as not actually set theoretic functions, but equivalence classes of set theoretic functions, where functions f and g are equivalent in $L^1([-1, 1])$ if $f(x) = g(x)$ almost everywhere in the interval $[-1, 1]$.

Analysts commonly refer to elements of $L^1([-1, 1])$ (and numerous related spaces) as functions. But a well known key fact is: a given $f \in L^1([-1, 1])$ has no definite value $f(x)$ at any point x. Changing the value of f just at one point x

does not change the equivalence class of f, and so does not change f as an element of $L^1([-1, 1])$. Analysts never lose sight of that, even when they refer to elements of $L^1([-1, 1])$ as functions.

The normed vector spaces $C^0([-1, 1])$ and $L^1([-1, 1])$ typify the very many different *function spaces* constitutive of functional analysis.

9.4 Conclusion

> At present the concept of function is not as finally crystalized and undeniably established as it seemed to be at one time at the end of the 19th century. It is no exaggeration to say that at present the function concept is still evolving and that the controversy about the vibrating string continues,[4] except for the obvious fact that the scientific circumstances, the personalities involved, and the terminology are different. (Shenitzer and Luzin 1998, p. 66)

That quote is from a translation by Abe Shenitzer of an article by Nikolai Luzin in the 1930s. The evolution is still going on. But today it would be exaggerated to call the evolution "controversial." New concepts are constantly being developed, and are often at least informally called "functions." Few people if any oppose that usage. Nor does anyone object that set theorists have too narrow a definition of "function," for the purposes of set theory. There is just a difference in practice over what can be called a function outside of set theory.

To put this in our context, Stein and Shakarchi (2005, p. 69) on real analysis officially keeps the set theoretic definition of "function," while making explicit the "practice we have already adopted not to distinguish two functions that agree almost everywhere." They constantly use this practice through the cited volume and the next one on functional analysis (Stein and Shakarchi 2011). They are explicit enough to preserve rigor, but they normally say "function f" where a set theorist would rather they say "class of functions equal to f almost everywhere." This kind of equivalence is just so natural for the topics they cover. These volumes give a masterful, elegant introduction on how and why to study in detail what a function $f : \mathbb{R}^d \to \mathbb{R}$ does "almost everywhere" while never supposing it has any specific value $f(x)$ at any single point $x \in \mathbb{R}^d$.

Tao writes more explicitly about functions that are not set theoretically functions. He says: "Formally a function space is a NORMED SPACE X the elements of which are functions (with some fixed domain and range)" (2008b, p. 210). His examples $C^0[-1, 1]$ and $C^1[-1, 1]$ do use the set theoretic idea of function. The Lebesgue spaces $L^p[-1, 1]$ and Sobolev spaces $W^{k,p}[-1, 1]$ which he also describes go outside that set theoretic sense to follow his idea of the smoothness spectrum of notions of function. Tao also gets the set theory right. But he does it with an expertly light touch. By *not* treating the set theoretic notion of function as the only notion

[4] For the history of the vibrating string see Gray (2021, Ch. 3) and many other references.

of function, set theory can aid rather than obscure the analytic insights that got measurable functions and generalized functions called "functions" in the first place.

References

Barany, M. 2018. Integration by Parts: Wordplay, Abuses of Language, and Modern Mathematical Theory on the Move. *Historical Studies in the Natural Sciences* 48: 259–299.

Bombieri, E. 2000. *The Riemann Hypothesis*. Cambridge: Clay Mathematical Institute.

Cartier, P. 2021. Il a tué l'analyse fonctionelle. In *Lectures grothendieckiennes*, ed. F. Jaëck, 27–46. Paris: Spartacus IDH, Societé Mathématique de France. English translation forthcoming from the same publisher.

Fefferman, C. 2000. *Existence and Smoothness of the Navier Stokes Equation*. Cambridge: Clay Mathematical Institute.

Fefferman, C. 2008. The Euler and Navier-Stokes Equations. In *Princeton Companion to Mathematics*, ed. T. Gowers, J. Barrow-Green, and I. Leader, 193–196. Princeton: Princeton University Press.

Fourier, J. B. J. 2009. *Théorie Analytique de la Chaleur*. Cambridge Library Collection. Cambridge: Cambridge University Press.

Galdi, G. 2011. *An Introduction to the Mathematical Theory of the Navier-Stokes Equations: Steady-State Problems*, 2nd ed. New York: Springer-Verlag.

Goursat, E. 1933. *Cours d'analyse mathematique*. Paris: Gauthier-Villars.

Gray, J. 2021. *Change and Variations: A History of Differential Equations to 1900*. Berlin: Springer.

Kelley, J. 1955. *General Topology*. New York: Van Nostrand.

Kline, M. 1972. *Mathematical Thought from Ancient to Modern Times*. Oxford: Oxford University Press.

Lemarié-Rieusset, P. 2015. *The Navier-Stokes Problem in the 21st Century*. Milton Park: Taylor & Francis.

Leray, J. 1934. Sur le mouvement d'un liquide visqueux emplissant l'espace. *Acta Mathematica* 63: 193–248.

Lützen, J. 1982. *Prehistory of the Theory of Distributions*, vol. 7. *Studies in the History of Mathematics and the Physical Sciences*. New York: Springer-Verlag.

McLarty, C. 2007. The Rising Sea: Grothendieck on Simplicity and Generality I. In *Episodes in the History of Recent Algebra*, ed. J. Gray and K. Parshall, 301–26. Providence: American Mathematical Society.

McLarty, C. 2023. Fluid Mechanics for Philosophers, or Which Solutions Do You Want for Navier-Stokes? In *Physical Laws and the Limits of Explanation – What the Equations Don't Say*, ed. L. Patton and E. Curiel. Berlin: Springer-Verlag.

Shenitzer, A., and N. Luzin. 1998. Function: Part I. *American Mathematical Monthly 105* (1): 59–67.

Stein, E., and R. Shakarchi. 2003. *Fourier Analysis: An Introduction*, vol. 1. *Princeton Lectures in Analysis*. Princeton: Princeton University Press.

Stein, E., and R. Shakarchi. 2005. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, vol. 3. *Princeton Lectures in Analysis*. Princeton: Princeton University Press.

Stein, E., and R. Shakarchi. 2011. *Functional Analysis: Introduction to Further Topics in Analysis*, vol. 4. *Princeton Lectures in Analysis*. Princeton: Princeton University Press.

Strang, G. 2015. *Differential Equations and Linear Algebra*. Wellesley: Wellesley-Cambridge Press.

Tao, T. 2008a. Distributions. In *Princeton Companion to Mathematics*, ed. T. Gowers, J. Barrow-Green, and I. Leader, 184–187. Princeton: Princeton University Press.

Tao, T. 2008b. Function Spaces. In *Princeton Companion to Mathematics*, ed. T. Gowers, J. Barrow-Green, and I. Leader, 210–213. Princeton: Princeton University Press.

Tao, T. 2008c. Harmonic Analysis. In *Princeton Companion to Mathematics*, ed. T. Gowers, J. Barrow-Green, and I. Leader, 448–455. Princeton: Princeton University Press.

Tao, T. 2013. *Compactness and Contradiction*. Providence: American Mathematical Society.

Tao, T. 2014. *Analysis I*. New Delhi: Hindustan Book Agency.

Tao, T. 2016a. *Analysis I*. New Delhi: Hindustan Book Agency.

Tao, T. 2016b. *Analysis II*. New Delhi: Hindustan Book Agency.

Chapter 10
"No Mother Has Ever Produced an Intuitive Mathematician": The Question of Mathematical Heritability at the End of the Nineteenth Century

Jemma Lorenat

Abstract In January 1893, Christine Ladd Franklin published an article on "Intuition and Reason" in *The Monist*—a new philosophy of science periodical based in Chicago. On the question of heritability, Ladd Franklin was absolute: "No mother has ever produced an intuitive mathematician." This contribution examines why Ladd Franklin wrote "Intuition and Reason" and how her expertise in psychology, mathematics, and logic informed her arguments. In particular, Ladd Franklin turned to the history of mathematics as a source for counterexamples against the "ancient opinion" that coupled women with intuition and men with reason.

Is mathematical intuition heritable?

In 1893 Felix Klein pondered a variation of this question in his *Evanston Colloquium Lectures*. Speaking "On the Mathematical Character of Space-Intuition, and the Relation of Pure Mathematics to the Applied Sciences" Klein infamously speculated:

> Finally, it must be said that the degree of exactness of the intuition of space may be different in different individuals, perhaps even in different races. It would seem as if a strong naïve space-intuition were an attribute pre-eminently of the Teutonic race, while the critical, purely logical sense is more fully developed in the Latin and Hebrew races. A full investigation of this subject, somewhat on the lines suggested by Francis Galton in his researches on heredity, might be interesting. (Klein 1894, 42)

It is well-known how Klein's suggestion wormed its way into the rhetoric of Nazis in Germany, serving as historical justification for racial types of mathematicians. In " 'Jewish Mathematics' at Göttingen in the Era of Felix Klein," David Rowe examines Klein's suggestion and concludes that "men of integrity like Klein, Weierstrass, and Sommerfeld were incapable of freeing themselves from the conventional racial thinking of their day is certainly suggestive of how pervasive these prejudices must

J. Lorenat (✉)
Mathematics, Pitzer College, Claremont, CA, USA
e-mail: Jemma_Lorenat@pitzer.edu

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_10

have been" (Rowe 1986, 443).[1] That is, Klein's speculations on racial difference and heritable intuition in mathematics may not have seemed out of place to many of his German colleagues.

In the United States, too, Klein's talk was well-received. Through in-person attendance, transcripts published in 1894, and then republished by the American Mathematical Society in 1911, *The Evanston Colloquium Lectures on Mathematics* achieved a widening circulation. Among the unregistered attendees at the Colloquium was Christine Ladd Franklin. In January that same year, she had published an article on "Intuition and Reason" in *The Monist*—a new philosophy of science periodical based in Chicago (Ladd Franklin 1893). On the question of heritability, Ladd Franklin was absolute: "No mother has ever produced an intuitive mathematician."

She was not engaged in a debate about racial abilities, but a similarly structured discourse on sex. These were separate, though overlapping, discussions. For instance, Klein's suggestion of racial difference does not seem to have been matched by a belief in sex difference, as will be elaborated below. However, each of these stereotype-driven arguments extrapolated from underrepresentation to conclude that certain groups were biologically limited in their mathematical abilities.

Speculations on "woman's intuition" have a long and well-documented history.[2] To explain the underrepresentation of women in the arts and sciences, late-nineteenth-century writers pointed at biologically-circumscribed intuition as inhibiting logical thinking, creative originality, and genius. Against these claims, the "European women's movement of the 1880s–1920s" developed an encyclopedic "strategy of emphasizing the achievements of exceptional women" (Schiebinger 1989, 4). The results were wielded by feminists and antifeminists alike. At a conference in France in 1908, mathematician Maurice d'Ocagne found that "among women in science, the majority were women mathematicians" and justified this finding, in part because "mathematics relied first on a form of divination, of intuition before the phases of rigorous reasoning." Even so, d'Ocagne concluded that "most women were able to understand but not to invent mathematics" (Boucard 2020, 208). More generally, for every woman with documented scientific success there were so many more men. The encyclopedic strategy withered under such strictly quantitative comparisons.

[1] On Klein's use of "intuition" as well as the relationship between mathematical intuition, modernism, politics, and culture see Mehrtens (1990) and Gray (2008).

[2] An overview on the intersections between history of science and research on gender can be found in Kohlstedt and Longino (1997). Of particular relevance to the application of scientific arguments to women's social position and mental capacity see Fausto-Sterling (1985). On the "influence of evolutionary theory on the psychology of women" see Shields (1975). Note that here I write "sex" rather than "gender" to reflect the language of the late-nineteenth century. For the nineteenth-century authors considered here, sex was understood as a binary category.

Nineteenth-century women in mathematics both challenged and succumbed to the conclusions drawn from historical precedent. In her book *Femininity, Mathematics and Science, 1880–1914*, historian Claire G. Jones considers the gendered inheritance of mathematical genius at Cambridge and Göttingen University through the writings and experiences of mathematics student Grace Chisholm Young (Jones 2009). Jones demonstrates how Chisholm Young eventually internalized the perceived masculinity of mathematical research at both of these institutions. After her marriage to William Young, "surrounded by prescriptions that served to limit female ambition and opportunity, overwhelmed by romantic notions of genius, and practising a pure mathematics that privileged the male intellect, she decided to transfer her mathematical ambitions onto her husband" (65). Most of their subsequent mathematical research was published under his name alone.

In the 1890s both Chisholm Young and Ladd Franklin contributed to the admission of foreign women to mathematics lectures and seminars at the University of Göttingen. However, while the former participated in the mathematical community as an exception, such was not the aim of Ladd Franklin. Rather, she uncovered underlying reasons for sex prejudice and exclusion. In what follows, I will examine why Ladd Franklin wrote "Intuition and Reason" and how her expertise in psychology, mathematics, and logic informed her arguments against the "ancient opinion" that coupled women with intuition and men with reason.

One late nineteenth-century variant of this "ancient opinion" took shape in the writings of the novelist and biologist Grant Allen, the subject of Sect. 10.2. Allen's theory pervaded his fiction and nonfiction, and his paper "Woman's Intuition" served as a catalyst for Ladd Franklin. Allen showed little concern for positive evidence, still his work enjoyed popular acclaim.[3]

Ladd Franklin endeavored to communicate with a similar audience. Her path to publication will be considered in Sect. 10.3. Beginning with "Intuition and Reason"—the subject of Sect. 10.4—her rhetorical strategy can be divided into two complementary parts: against intuition as an innate faculty and in favor of women who reason. First, Ladd Franklin pulled from recent experimental studies and her experience as a mathematics teacher to dismantle the belief in inherited intuition. To show that women could reason, she might have followed other late-nineteenth-century feminists, who had emphasized "the achievements of exceptional women." Indeed, as will be shown in Sect. 10.5, Ladd Franklin looked to the history of mathematics. But rather than exceptions that left the rules intact, she identified counterexamples that demanded a different logical conclusion.

[3] Allen was part of a broader current. A biography on the Higher Education of Women compiled in 1905 by the Association of Collegiate Alumnae lists around eighty texts over three decades under the heading "II. Mental and Physical Status. Includes Questions of Curriculum and Physical Training" including Allen's text, though not Ladd Franklin's rebuttal (of the Association of Collegiate Alumnae 1905).

10.1 On the Difficulty of Talking About Intuition

What Klein intended by "spatial intuition" is clearly not identical to Ladd Franklin's juxtaposition of reason and intuition.[4]

First, there was the problem of translation. Klein was German, but he delivered the Evanston lectures in English before an audience consisting mostly of American mathematicians. According to the editor's introduction, Klein "carefully revised" the manuscript and proofs before publication (Klein 1894). Thus Klein can be credited with choosing "intuition." But this was not the only possible translations of the German *Anschauung*. Notably, in Mellen Woodman Haskell's English translation of Klein's 1872 "Recent Researches in Geometry" published in the *Bulletin of the New York Mathematical Society* in 1893, he translated Anschauung and its derivatives as perception, interpretation, or view. Haskell's one use of "intuitively" is a translation of "begrifflich." Still Klein described this translation as "absolutely literal" and so seems to have accepted these alternates as appropriate (Klein 1893). Secondly, in his *Evanston Colloquium Lectures*, Klein engaged with multiple and somewhat contradictory meanings of intuition, initially distinguishing between naïve and refined intuition, and then clarifying that "*the naïve intuition is not exact, while the refined intuition is not properly intuition at all, but arises through the logical development from axioms considered as perfectly exact*" (Klein 1894, 42). The same author could invoke "intuition" across multiple meanings and any single use of "intuition" was not guaranteed to be understood across languages, individuals, or texts.[5]

Despite these inconsistent connotations, the concept of intuition persisted across debates on race, sex, and mathematical capacity.

10.2 Grant Allen and 'Woman's Intuition'

Grant Allen was born in Canada and lived in England, but "Woman's Intuition" appeared in the American monthly magazine, *The Forum*, published in New York

[4] In her 2011 study of Charles Hermite, Catherine Goldstein she singled out how "the word Anschauung in Felix Klein's entourage, [enjoyed] the role of a banner, rallying mathematicians, types of explanations, methods. Employed together or no, rather rarely in association with other terms ("simple and general", a little more often "precise"), these words recurrently come to oppose the pair "rigorous" and "complicated"[...]" (Goldstein 1994, 145).

["[. . .] Elles jouent, tout comments dans l'entourage de Felix Klein le mot *Anschauung*, le rôle d'une bannière, reliant des mathématiciens, des types d'explication, des méthodes. Employés ensemble ou non, assez rarement en association avec d'autres termes ("simple et général", un peu plus souvent "précis"), ces mots viennent s'opposer de manière récurrente au couple "rigoureux" et "compliqué" [. . .]"]

[5] In this volume, Nicolas Michel's chapter "In Search of Absolute: On H. G. Zeuthen's geometrical holism" defines Zeuthen's intuition as "an ability to perceive at once a connected whole where symboblic cognition and rigid computational rules present one's mind with disconnected particulars," which Zeuthen in turn based on the work of French philosopher Henri Bergson.

(Allen 1890). *The Forum* covered a wide range of political topics and opinions. Allen's article is preceded by a piece speculating on how to improve battleships and succeeded by a call for prohibition.

Allen's paper is at once a celebration of woman's intuition and a warning that such intuition might be in peril. He attributed current differences between men and women to an evolutionary narrative beginning with a hypothetical past of hunting tribes. In this whole history, "man has specialized himself on logical intelligence and practical handicraft; woman has specialized herself upon the emotions and intuitions, the home and the family."' Accordingly, at present, while women had maintained "the common endowment of all animals possessing nervous systems at all" men had lost the gift "through the gradual evolution, training, and discipline of his logical faculties" (337).

Despite this model of past growth and change over time, Allen was pessimistic that women could learn to reason effectively. Those who have tried only became "feeble, second-rate copies of men." He lamented the "mannish women of our age" created by the "women's cause" in the form of "lady lecturers and anti-feminine old maids" (333). Not only did such women ruin themselves, they also diminished the capacity for human (male) genius. As Allen explained, "in all genius, however virile" there could be observed "a certain undercurrent of the best feminine characteristics." A masculine woman was an aberration, but "the man of genius is comprehensively human." Allen attributed this rare faculty in some men "to the imaginative faculty and to the intuitive faculty that they derive from their mothers." He recognized this "oft-repeated fallacy" as grounded in truth and matching his own observations (340).

These observations were both personal and historical. Allen recalled a conversation he had with "the greatest living mathematician" who remarked "that Laplace, in summarizing a mathematical argument, often wrote, 'Hence it obviously follows that $x = f \times ab^2 + y$,' or whatever it might be; when he, the great living mathematician, could see the truth of the inference only after working out a page or two of elaborate calculations" (335). Allen speculated that "Laplace's mind cleared at a bound the 'obvious' intervening steps" by employing "what we call intuition." It was the same intuition "coupled of course with high masculine qualities—knowledge, application, logical power, hard work—that gives us the masterpieces of the world's progress" including "steam engines and locomotives, telegraphs and telephones, Hamlets and Richard Feverls, Newton's 'Principia" and Spencer's 'First Principles."' So while most men possessed little intuition, those who inherited it from their "most purely womanly" mothers were among the great minds of the British and European past.

Allen implored society to not "deliberately educate out the intuitive faculty in woman" and thereby impoverish humanity (340). He remained hopeful that the "celibate lady lecturer will die unrepresented" while "the woman with grace, tact, high emotional endowments, pure womanly gifts, will hand down her exquisite and charming qualities to other women, her likes, after her."

Allen's tone is jocular and his peculiar theory of inheritance might be dismissed as hyperbolic if not for its pervasiveness in his other essays and fiction. In the latter, the reader can witness both the inexplicable ways of woman's intuition and the harm caused to the sex by too much reasoning, such as in the case of a mathematical education. Allen was neighbors with Arthur Conan Doyle and successfully dabbled in detective fiction. But whereas Sherlock Holmes is the epitome of logical reasoning, Allen invented heroines who solved mysteries without being able to explain how.

What's Bred in the Bone was published serially between 1890 and 1891 and won a 1000-pound prize from the editor of "Tit-Bits" who then published the book in its entirety (Allen 1891). The book opens when Elma enters a train car and knows immediately that her companion is an artist. Allen offered an explanation to the presumably male reader:

> Now, you and I, to be sure, most proverbially courteous and intelligent reader, might never have guessed at first sight, from the young man's outer aspect, the nature of his occupation. The gross and clumsy male intellect, which works in accordance with the stupid laws of inductive logic, has a queer habit of requiring something or other, in the way of definite evidence, before it commits itself offhand to the distinct conclusion. But Elma Clifford was a woman; and therefore she knew a more excellent way. *Her* habit was, rather to look once fairly and squarely in the face, and then, with the unerring intuition of her sex, to make up her mind about them firmly, at once and for ever. That's one of the many glorious advantages of being born a woman. You don't need to learn in order to know. You know instinctively. And yet our girls want to go to Girton, and train themselves up to be senior wranglers!
>
> Elma Clifford, however, had *not* been to Girton, so, as she stumbled into her place, she snatched one hurried look at Cyril Waring's face, and knew at a glance he was a landscape painter. (2)

As the story continues, one learns that Elma inherited her "most extraordinary intuition" from a long line of female descendants traced back to "some kind of Oriental gipsy" (117). In her formidable ability Elma is at once superior and inept. When asked for her reasons, Elma has none and shows "a true woman's contempt for anything so unimportant as mere positive evidence" (357). Elma embodies the womanly gifts Allen so admires, and ends her adventures appropriately in matrimony.

The mathematical training at Girton College reappears as a threat to intuition in *Miss Cayley's Adventures*. The story begins shortly after the heroine "had taken high mathematical honors at Cambridge" (Allen 1899, 59).[6] Miss Cayley somehow escaped the worst of her education, and is only criticized once for being "extremely Girtony," that is, "unnaturally and unfemininely reasonable" (138). By contrast, her former classmate Elsie intends to teach higher mathematics in high school and is "lacking in feminine intuition." Cayley reflects that she should "be sorry if I had

[6] The name choice here is striking to a historian of mathematics, but without further evidence it may be only a coincidence that the Girtonite shared a surname with the recently deceased Sadleirian professor.

allowed the higher mathematics to kill out in me the most distinctively womanly faculty" (208).

For a casual reader, Allen's asides may pass unnoticed in the lively story and adorable characters. Once aware of his theory, however, they take on a dull pedantry. Allen described women as "mere passive transmitters of these male acquisitions"—in this case, his heroines operate as shells for his theory of biological intuition (Allen 1888, 263). Yet it is only this fictional evidence along with vague references to his observations that support Allen's claims. He touted his expertise "as a biologist" but did not engage in actual research on women and intuition.

10.3 Ladd Franklin to 1892

Allen's "Woman's Intuition" featured in the May, 1890 issue of *The Forum*. By this time, Ladd Franklin had already begun circulating her own findings on the relationship between women and intuition. Without any mention of her title or education, she plumbed her transdisciplinary professional training to gather precise and well-documented evidence.

Ladd Franklin's own biography contradicts Allen's stereotypes. She grew up in the Northeastern United States and was among the first to take advantage of the college education made available with the founding of Vassar College in 1861. She learned of the college in 1863 and enrolled a few years later. Following her own biographical account, Ladd Franklin had initially been drawn to the study of physics and "would have devoted herself [to it] after graduation had it not been for the impossibility, in those days, in the case of women, of obtaining access to laboratory facilities. She therefore took up as the next best thing mathematics, which can be carried on without any apparatus."[7] She graduated in 1869 then taught mathematics at private schools while continuing her studies and offering solutions to posed mathematics problems in the *Educational Times*.

Shortly after Johns Hopkins University opened in 1876, Ladd Franklin wrote to the head of the mathematics department, James Joseph Sylvester, inquiring "Will you kindly tell me whether the Johns Hopkins University will refuse to permit [listening to mathematical lectures] on account of my sex?".[8] Based on her publications in mathematics, she was admitted as an "invisible student"—able to attend lectures, but without any official record of enrollment. Even after winning the stipend of a fellowship for three successive years, "to avoid making a precedent, her name was carefully separated from the list of fellows" (Cameron 1928). Working with logician Charles Peirce, Ladd Franklin completed a dissertation on "The Algebra of Logic," which was published in *Studies in Logic by Members of the Johns*

[7] Ladd Franklin wrote her own biography, published in Cameron (1928).

[8] The letter is quoted in Parshall (2006, 255).

Hopkins University in 1883. But she was denied the degree from Johns Hopkins until 1926.[9]

In 1882 Christine Ladd married mathematics lecturer Fabian Franklin. Ten years later, the Franklin family (including their eight-year-old daughter Margaret) traveled to Germany for Fabian's sabbatical year. There Ladd Franklin met Klein, David Hilbert, and Adolf Horowitz, but turned her research to psychology, in particular the study of color-vision. She began in G. E. Müller's lab in Göttingen and later spent time in Berlin with Arthur König. She first publicly communicated the findings from her experimental research at the International Congress of Psychology in London during that sabbatical year in Europe.[10]

In this same period, Ladd Franklin served an integral role in the organization of the American Association of Collegiate Alumnae, which had been founded in 1882. For the next decade, Ladd Franklin advocated to create a European Fellowship that would allow women to pursue graduate research opportunities abroad. In particular, she hoped that this funding would further serve to open German Universities, like the University of Göttingen, where women had been excluded.[11]

Thus Ladd Franklin engaged with the question of inherited intuition as an academic and as an activist. In writing "Intuition and Reason" she aimed for a broad, international audience. At the beginning of 1890, she sent the manuscript to G. Croom Robertson, then editor of *Mind*. Ladd Franklin had already contributed as a reviewer of technical texts in logic and Robertson replied encouragingly to her submission expressing "much interest" and thinking "the argument both sound and sensible." However, he advised, "*Mind* seems hardly the right channel for it to public attention." In particular, Robertson worried that "the Leipsic [sic] experiments" though "full of interest and instruction to "the general reader" would be less effective in *Mind*." Robertson suggested offering the paper to "the worldlies" and promised to send it to *The Nineteenth Century*, a monthly review published in London (Robertson 1890a). That October, Robertson received a positive response from editor James Thomas Knowles, who "will be very glad to accept it" for *The Nineteenth Century* with the proviso that "Mrs Ladd Franklin [. . .] leave the date of publication entirely in his discretion." Robertson advised Ladd Franklin to accept with the hope that it would appear in December. He also suggested a "less severe

[9] Her early mathematical studies are investigated in Green and LaDuke (2009), Fenster and Parshall (1994), and Parshall and Rowe (1994). In the history of logic, David W. Agler and Deniz Durmuş provide a philosophically-oriented approach to Ladd Franklin's feminism and concern with the historical record (Agler and Durmuş 2013).

[10] In the history of psychology, Laurel Furumoto has situated Ladd Franklin's contributions to color theory and lifelong efforts to secure credit, despite formidable obstacles, in Furumoto (1992, 1994).

[11] The importance of the Association of Collegiate Alumnae for women in science can be seen in Rossiter (1982), particularly chapter 2. More recently, Scott Spillman has documented Ladd Franklin's contributions and negotiations to create opportunities within the Association of Collegiate Alumnae in Spillman (2012).

title" something like "The Female Mind," "The Mind of Woman," or "The Male and Female Mind" (Robertson 1890b).[12]

The paper sat unpublished until January 1892 when Knowles' secretary wrote to Ladd Franklin explaining the former's illness, repeating the conditions of acceptance, "but as he can see no chance of his being able to publish it for a considerable time," ultimately returning the unpublished paper (Knowles 1892).

With her manuscript in hand, Ladd Franklin sent it off again, writing to William James to see where he would recommend publication. James was an excellent resource for finding venues of popular psychology. He expressed delight in anticipation of the manuscript, initially suggesting *Popular Science Monthly*. Further, he approved of higher education for women, affirming that "of course we are going to have women at Harvard soon—Göttingen mustn't be allowed to get ahead here." He also complimented Ladd Franklin as among the "mathematically minded geniuses" thus confirming that women could achieve such an epithet (James 1892). Yet, James may have agreed with Allen on certain points. His *Principles of Psychology*, which Ladd Franklin had read closely, suggests a more conventional expectation for women's reasoning abilities (James 1890).

In their correspondence, Ladd Franklin critiqued the last chapter of *Principles of Psychology* on "Necessary Truths and the Effects of Experience" that treated formal logic, mathematical propositions, arithmetic, and geometry. James had hoped for her "approval as a logician + mathematician" and attributed her response to "The unfathomable ways of woman!" (James 1892). Indeed, this exclamation betrayed a sex distinction between intuition and reason that also informed his chapter on "Reasoning."

As compared to the masculine brain, "the feminine method of direct intuition" performed "admirably and rapidly" within its limits but "can vainly hope to cope" with any "new and complex matter" (James 1890, 368). But here James was more descriptive than prescriptive. While James aligned with Allen in his assessment of women at present, James also recognized the importance of situation and environment. He criticized "the paucity of empirical evidence" in support of the hypothesis "that what was acquired habit in the ancestor may become congenital tendency in the offspring."

Such a stance was incompatible with Allen's theory of how men had lost their intuition through evolution and training. James determined that the "technical differentia of reasoning" was the "ability to deal with NOVEL data." Thus reasoning was antithetical to inheritance. And since mathematics was based on reasoning, mathematical ability would also seem to be beyond the bounds of congenital tendency.[13] Even if James viewed Ladd Franklin as an exception, she clearly

[12] G. Croom Robertson was also a committee member of the National Society for Women's Suffrage in Britain and may have wanted a more substantial audience for Ladd Franklin's work for political as well as literary reasons.

[13] In "Indebted to No One': Grounding and Gendering the Self-Made Mathematician" Ellen Abrams describes how biographies of American mathematicians used evidence of "self-making" to demonstrate masculinity. The relative unimportance of inheritance in these narratives aligns with

recognized him as a reader who would entertain her argument and recognize its value to a broader audience.

But Ladd Franklin's paper was not published in *Popular Science Monthly*. James may have later recommended *The Monist*, where "Intuition and Reason" would eventually land. Suitably, *The Monist* included a wide range of discussions on psychology, logic, and heredity. Further, the mutual friend of James and Ladd Franklin, Charles Peirce, had already written two articles for the new periodical and was much admired by editor Paul Carus.

While fittingly interdisciplinary, *The Monist* had a more modest audience than that of the other venues considered. After reading Ladd Franklin's article several years later, Kate Holladay Claghorn—a fellow leader of the Association of Collegiate Alumnae—wrote on "how much I enjoyed your article" but regretted "that it was not published in the Nineteenth Century, because the results are so wholesome, and the matter put so clearly and interestingly that it would have done the 'greater public' a lot of good" (Claghorn 1899). Some public was better than no public. Ladd Franklin's perseverance attests to her conviction in the wholesome results of "Intuition and Reason."

10.4 "Intuition and Reason"

Ladd Franklin identified intuition as "a word of double meaning." First, "it covers those actions which we go through with by instinct, or inherited experience ingrained from the beginning in our nervous structure, and those which we perform automatically" (Ladd Franklin 1893, 212).[14] This inherited intuition concerned actions necessary for survival. As an example, Ladd Franklin asked her reader whether they "know that a certain feeling of strain in the muscles which move the eyes is a sign of a certain distance of an object looked at, and a different feeling of strain, a sign of different distance." She explained that the "common man *knows* that one object is near and the other far" even though he was "not *conscious* even of the feeling of strain." Further, though "the physiological psychologist" was aware of "the unconscious syllogism by which he *must* reach his conclusion." He still could not "by any possibility, make it cease to be instinctive." This automatic, instinctive intuition was present in all humans, regardless of sex. For this paper, Ladd Franklin focused her attention on the second meaning of intuition— covering actions that "by individual experience become so familiar that [intuition] can act as a guide without

James' theory to some extent, but also suggests a difference in the value of accomplished ancestors between the British and American contexts (Abrams 2020).

[14] Though she did not provide these credentials in her article, Ladd Franklin was very familiar with biological theories of evolution and the necessary timescale for evolutionary processes. In the 1890s she was in the middle of research that would lead to her "idea of the evolution of a colour molecule"—that the sensation of red and green colors "constitute a real evolution out a more primitive yellow sense" (Ladd Franklin 1929, 183).

the aid of conscious reflection." This is the intuition that had been positioned as in opposition to reason.

On the question of the heritability of mathematical intuition, Ladd Franklin's offered multiple grounds for opposition. She first turned to new results from psychological experiments to demonstrate that intuition is acquired through experience. From these empirical findings, Ladd Franklin then attacked the dichotomy that positioned men as reasoning and women as intuitive. Instead of possessing different faculties, she explained that men and women have different life circumstances. So, contra Allen's "apotheosis of the uneducated woman," it was impossible that "mothers should occasionally transmit their powers of intuition to favored sons" (215). Finally, she found his claim particularly preposterous with respect to mathematics. Ladd Franklin saw no role for intuition in the discipline. Balanced between scientific studies and personal experience, "Intuition and Reason" is redundant in its variety of evidence.

Drawing on recent psychological studies, Ladd Franklin showed that the purported dichotomy between intuition and reason was illusory: "reason is merely intuition in its formative stage" (216). In particular, Wilhelm Wundt and the students in his Leipzig laboratory had performed tests in which subjects were instructed to press a button when a bell sounded. One group would purposely focus on the tap of the bell and then "decide consciously what to do in response," for the other group "the process is unconscious." In comparing reaction times, the "first is nearly twice as long as the second" suggesting a qualitative difference. Further, each subject could "teach himself to give either reaction time at his pleasure" by training his consciousness on the bell or not. Ladd Franklin praised this experiment as catching "automatism in the very act of formation." In this simple matter, the unconscious observers relied on intuition to perform more efficiently.

Ladd Franklin compared Wundt's findings with other tests. Following a suggestion by William James (and discussed briefly in their correspondence) she described "a good experiment that some one who has eyes that he is not afraid of injuring" could try at home (217). Despite this clear injunctive not to perform such an experiment, she then detailed how one could put a finger under their eyelid in order to push their eye ball for "several hours a day." At first this process would produce the illusion of moving objects, but over time one might be able to "force conscious reason to do her work and to make him *see* that the objects are not moving." Through this process, an effect that occurred intuitively could be gradually eliminated through purposeful effort. Together these experiments revealed that intuition could be learned and unlearned at will. This kind of intuition was consequently not inherent nor inheritable.

Consequently there was no justification "that men's minds and women's minds have a different way of working." Given the recent advances in "the psychology of the working of the human mind" Ladd Franklin characterized such a view as "old-fashioned" and "out-of-date." Instead, Ladd Franklin looked to environmental factors and determined "that the circumstances of women's lives have hitherto been such as to make their interests lie somewhat more exclusively in those regions in which conduct is intuitive than in those in which it is long thought out" (212).

Not everything could be made intuitive. While Allen had posited that the genius of mathematics was grounded in intuition, Ladd Franklin asserted that no one "who knew anything about the higher mathematics for a moment" would "suppose that when a great mathematician leaves out intermediate steps in a printed book, he had jumped at his conclusions by instinct." Rather "with his thorough knowledge of this particular subject, the intermediate steps have seemed to him too easy to set down." Ladd Franklin thus explained why "the greatest living mathematician" (according to Allen) might have found some difficulty with Laplace: "If his book is hard to read, it is simply because he has assumed a greater amount of learning in his readers than they are in possession of." What applied to the higher mathematics also applied to school mathematics. Consequently, geometry "is a branch of learning which is entirely built up out of abstract reason, pure and undefiled" and "no one, whether man or woman, can pass from one proposition in geometry to another by a process which is in any sense unconscious, though one person may be obliged to give a much more strained attention to what he is doing than another" (217).[15]

Because mathematics was the epitome of reason, the mathematics classroom served as a perfect situation in which to judge such ability. With recent educational opportunities in the past few decades, women had done quite well. Ladd Franklin reported that in the United States geometry is studied in high schools by "*three times* as many girls as boys [. . .] it cannot be said, therefore, (as is said of girls who go to college) that the girls who go to the high school are a selected lot; they are the very bone and fibre of the women who make up the country."[16] Ladd Franklin hypothesized that "if women could not reason, we ought to hear a great hue and cry from the teachers of the geometry classes about the difficulty of teaching that subject to girls, and the girls ought to lament and moan over the impossibility of getting safely through with their demonstrations." Yet, in her years of teaching she had "never met with a teacher of geometry who thought his boys did better than his girls,—I have met with several who thought the reverse."[17] Similarly, the students appeared unharmed in the process, as she observed with some sarcasm. "Day after day an army of girls goes smiling into the class-room and comes smiling out, utterly unaware that an unnatural wrench has been given to their delicate minds, and that they are rapidly transformed into monstrous products of over-reason" (218). Ladd

[15] For an opposing viewpoint on the role of the unconscious in mathematics, also published in *The Monist*, see Poincaré (1910). Ladd Franklin's use of "intuition" here is certainly narrower than that of Poincaré in what was translated as "Mathematical Creation."

[16] Ladd Franklin repeated this statistic on several occasions. For instance, in an earlier book review she had applied the finding to prophecy that "if geometry is as good a specific against bad reasoning as is commonly supposed, logicalness will soon become a feminine instead of a masculine characteristic" (Ladd Franklin 1886).

[17] Compare Ladd Franklin's classroom observations to the responses recorded by Florian Cajori in *The Teaching and History of Mathematics in the United States* (Cajori 1890). Participants from schools, colleges, and universities were asked which sex had greater aptitude for mathematics (pp. 316–319). The responses are varied, but the majority favor males.

Franklin's cheerful army of girls contrasted sharply with Allen's lonely, feeble "lady lecturers."

Ladd Franklin corroborated these population statistics with her own data. She had "kept a record for many years of errors committed by boys and by girls, and I have not been able to detect any difference in their character" with the exception that boys have stronger "intuition about stretched strings and lines on balls." She found no correlation between poor intuition and skilled reasoning.

This lack of causation worked in both directions. If women had historically demonstrated greater intuition "in social matters" there was no danger these would be lost "because she has made herself familiar with the speculations of philosophers, and can turn to them for guidance in the intricate questions of conduct which the complexities of modern life give rise to" (219). Women could reason and all would benefit if they should be able to do so through access to higher education.

This call to action drives at Ladd Franklin's overarching programmatic aim to remove women's barriers to educational and professional institutions. Embedded in this lifelong project, her precise claim that mathematical intuition is not heritable supports this practical end. There is no reason to preserve women's intuition so that their sons might become mathematical geniuses. Instead, if the social goal is to nurture mathematical genius, then the solution lies in increased access to learning and practicing higher mathematics. Ladd Franklin had presented precisely this suggestion in an earlier anonymous review of George Bruce Halsted's *The Elements of Geometry*, where she proposed organizing "some method for picking out the clever girls from among those who cannot afford to go to college" and providing them with scholarships (Ladd Franklin 1886).

"Intuition and Reason" showed women had no biological impediment to mathematical reasoning and exhibited national trends that girls were learning mathematics alongside boys and outnumbering them. It is also a very personal article. Who else but Ladd Franklin could report authoritatively on both cutting-edge German psychology and high school mathematics in the United States? But the paper is not about Ladd Franklin's life story, nor did she invoke biographies of other women in mathematics. This latter absence is made more acute in comparison with her subsequent writings in the same vein. Ladd Franklin continued to push against restricting women to domestic intuition, but increasingly turned to history for evidence.

10.5 History as Evidence

By the 1890s, history was a common source of evidence for determining the role of inheritance. Francis Galton's *Hereditary Genius* is based on research from "a large amount of carefully selected biographical data" including 65 scientific men "who have achieved an enduring reputation, or who are otherwise well known in the present generation" (Galton 1869, 193). Though Galton did not directly study women for this book, they entered his research as mothers. His treatment of "men of science" found "the maternal influence to be unusually strong" with "fully eight cases out of the forty-three" in which "the mother was the abler of the two parents."[18] Galton attributed this finding to early childhood education from an "able mother" at home:

> Happy are they whose mothers did not intensify their naturally slavish dispositions in childhood, by the frequent use of phrases such as, "Do not ask questions about this of that, for it is wrong to doubt;" but who showed them, by practice and teaching, that inquiry may be absolutely free without being irreverent, that reverence for truth is the parent of free inquiry, and that indifference or insincerity in the search after truth is one of the most degrading of sins. It is clear that a child brought up under the influences I have described is far more likely to succeed as a scientific man than one who was reared under the curb of dogmatic authority. (195)

Ladd Franklin may have been aware of Galton's work when she noted that "Hume and James Mill are two men who are supposed to owe much to their mothers" (Ladd Franklin 1893, 215). But while biographical studies could show that educated women might bear scientific sons, the same approach had potential to demonstrate that women themselves could be capable practitioners.

Rather than amassing large numbers, Ladd Franklin positioned the few women in the history of mathematics as counterexamples. This rhetorical strategy is most explicit in her popular article on "Sophie Germain: an unknown mathematician" (Ladd Franklin 1894). After introducing Germain, Ladd Franklin explained the logical implications of Germain's accomplishments.

> As proof that women may be pure mathematicians, Mrs. Somerville has had, outside of Italy and Russia, to stand alone. This is unfortunate, for the detractors of her sex have maintained that her work, though exceedingly profound, was not remarkable for originality. That charge cannot be brought against Sophie Germain. She showed great boldness in attacking a physical question which was at that time entirely outside the range of mathematical treatment, and the more complicated cases of which have not yet submitted themselves to analysis. (946)

To prove existence, Germain alone sufficed—a recent woman in pure mathematics and one whose contributions had been judged by her contemporaries as original.

Ladd Franklin extrapolated from Germain's early biography to a "general law that women's learning must be got by heroic measures, if at all" (947). Such heroism was exhibited in three cases. First, there was Germain "absorbed in her studies in a room so cold that the ink was frozen in the inkstand." Meanwhile "Mrs. Somerville, at that very same time, in her little village in Scotland, was obliged to wrap herself up in blankets to pursue her studies before breakfast." Finally,

[18] Notably, his list of men of science is also much less *English* than his other samples, including Leibniz, Buffon, Condorcet, Cuvier, D'Alembert, Ampère, Arago, and Jussieu. Recalling Klein's suggestion to conduct a study "on the lines suggested by Francis Galton in his researches on heredity" it is worth adding that Galton wrote that he "should have especially liked to investigate the biographies of Italians and Jews, both of whom appear to be rich in families of high intellectual breeds. Germany and America are also full of interest. It is a little less so with respect to France, where the Revolution and the guillotine made sad havoc among progeny of her abler races." (4)

and most tragically, "Ellen Watson, the highly gifted young woman, Clifford's pupil, who died at the Cape of Good Hope at an early age, did all her studying before breakfast, because she was required to spend the day-time in teaching her younger brothers and sisters."[19] By fitting these three women as manifestation of a general law, Ladd Franklin transcended anecdotal lists. Even if such invocations were only metaphorical, they elevated the discourse beyond armchair speculations and fictional narratives.

Thus, Ladd Franklin was similarly impatient when another biologist, George Romanes, advanced the theory—without "confirmation by facts"—that "women of unusual mental powers" were always unhappy and "too exceedingly obnoxious" (Ladd Franklin 1896, 315).[20] Ladd Franklin responded with "The Higher Education of Women" published in *Century*. As Romanes offered no reasons for his conclusions, Ladd Franklin sardonically suggested that "the strong intuitive powers of his sex can perceive [it] to be true at a glance." She corrected Romanes' claim for *all* women by citing the "social success" of Mrs. Somerville; the "remarkable fascination" and charms of Mme. Kovalévsky; the "wide circle of friends, who all spoke with enthusiasm of the charm and grace" of Sophie Germain; and the "love and reverence" bestowed upon Maria Mitchell. They were not only women mathematicians, but *feminine* women mathematicians.

In her commitment to sound reasoning and feminism, Ladd Franklin was accompanied by her husband, Fabian, who was working as a journalist and editor by the late 1890s. Just as Ladd Franklin's training in logic can be identified in the language and structure of her articles, so too Fabian Franklin exhibited a mathematical framework in "The Intellectual Powers of Woman" (Franklin 1898). But while his mathematical research had been in algebra, analysis, and number theory, to situate the history of women in mathematics he invoked comparative statistics.

Franklin's article for *The North American Review* is a response to a "spirited discussion by Mrs. G. G. Buckler" that had appeared a few months before. Buckler had determined that women had not and could not attain "the highest rank in science, literature, or art" (40). After reminding his readers of "the hindrances to woman's intellectual achievements," Franklin showed "their bearing upon that matter of numbers, which, while it is the vital element of the whole question, is so strangely ignored by the supporters of the view maintained in the article under discussion."

Buckler had found "few individual instances of female achievement in science" and interpreted this small sample as demonstrating "the rule that women as discoverers are inferior to men." However, Franklin recalled the difference in

[19] According to her obituary in 1881 Ellen Watson "was a favourite pupil of the late Professor W. K. Clifford, at the University College, and earned off the highest scholarship for mathematics in a mixed class of men and women." Watson qualified to continue her studies at London University, but had to delay on account of poor health (Anonymous 1881).

[20] Romanes found the biological inferiority of women "almost painfully obvious" resting on his belief that "in many department of intellectual work the field has been open, and equally open, to both sexes" without producing any women of note (Romanes 1887).

denominators. He reminded the reader “that the whole number of women who acquired the elements of the infinitesimal calculus, in the two centuries from its creation by Newton and Leibnitz, to the opening of Vassar College in 1865, was probably less than the number of mathematical honor men the single University of Cambridge turns out in a single year.” Over a hundred years “ten thousand men or so” had graduated, but only “two, or at most three, have achieved high rank as discoverers in pure mathematics” (47). If this proportion was accurate, then there was no reason to expect any women to earn such a level of rare attainment at present.

For a more appropriate comparative population, Franklin turned to his national context. Even if no woman had “achieved the very highest distinction” the same could be said “with equal truth of Americans, and with vastly greater emphasis of the inhabitants of almost any of our great States, say Pennsylvania; yet no one thinks of inferring from this that Americans or Pennsylvanians are utterly barred by inherent defect from ever attaining the highest intellectual glory” (48). Such a conclusion held no “logical weight.” As women gradually gained access to mathematical studies, they had been judged by experts such as Professor Klein, who “assures us that the women who have attended the mathematical courses at Göttingen ‘have constantly shown themselves from every point as able as their male competitors.’”[21] In particular, Franklin warned Americans to “not talk glibly of women’s power in scientific discovery being essentially inferior to men’s, until such time as some American mathematician receives as high recognition as that bestowed by the French Academy on the work of Sonia Kovalewski, the judgement being pronounced without knowledge of the writer’s sex” (53). For a nation that ranked itself according to European metrics of success, a statistical comparison of past greatness could be an effective weapon.

There were very few women in mathematics during the 1890s, but a few was enough to demonstrate what was possible. In these review journals and magazines, Fabian and Christine Ladd Franklin endeavored to be part of a national conversation on women and education. Their arguments exhibit fluency in logical reasoning and a commitment to concrete evidence. Were they convincing? Allen, Romanes, and Buckler did not publish counterattacks, but they also did not alter their respective positions. The Franklins succeeded in dismantling their opponents posture of scientific rigor. What remained were prejudiced beliefs immune to logic.

10.6 Conclusion

If they could have clarified their disparate meanings of intuition, Ladd Franklin may have agreed with Klein’s suggestion at the Evanston Colloquium. Her remarkable advocacy for women should not be read as demanding equal opportunity for

[21] Klein’s work in admitting foreign women to mathematics lectures at the University of Göttingen is elaborated in Parshall and Rowe (1994, pp. 240–244). On the significance of Klein and Göttingen for women studying mathematics, see Tobies (2020).

all people. First, Ladd Franklin clearly did not believe that mathematical talent was uniformly distributed. Instead, she proposed that for most children abstract reasoning was repulsive if not impossible. For the majority, it was not "worth while to expend the higher education." Only a few should "dwell long among the geometrical concepts, should become throughly imbued with the bare and rigid form of reasoning, and should have the results as familiar as his mother-tongue" (Ladd Franklin 1886).[22] Secondly, Ladd Franklin's claims for "women" might be more narrowly read as for white women. At the beginning of "Intuition and Reason" she claimed that "we all—men, women, and negroes alike—act from intuition" (Ladd Franklin 1893). While included in "we" the addition of a distinct third category certainly suggests a view of racial difference. Ladd Franklin otherwise did not discuss race in "Intuition and Reason" nor am I aware of any of her other writings on racial issues. Still, the detail is worth noticing as a caution to not interpret her calls to end discrimination against some women as anything more inclusive.

Even in this more limited scope, Ladd Franklin's message and method in "Intuition and Reason" continue to resonate. Much has changed since 1893, but beliefs around biological intuition and gendered mathematical ability persist. To take one recent example, a 2014 statistical survey by Gerd Gigerenzer, Mirta Gaelic, and Rocio Garcia-Retamero on "Stereotypes about men's and women's intuitions" in Germany and Spain found that while "the majority of Spaniards believe that men and women have equally good intuitions for scientific discoveries, [...] only one third of Germans think the same" which correlates with the percentage of female researchers in the natural sciences and engineering and technology between the two nations (Gigerenzer et al. 2014, 69).[23] This finding exemplifies a potential vicious cycle. The researchers suggest that "negative self-perception among German females might decrease their interest in science—and thus contribute to the actual difference between professional scientists in the two countries."

More optimistically, the potential impact of Ladd Franklin's argument in the context of the late-nineteenth century can be glimpsed in a letter among her papers—the kind of letter any teacher would hope to receive from a past student (Cary 1896). Many years after attending Miss Randolph's School, Pearl Buckner Cary recalled "that delightful geometry class where you taught us our first lessons in logic and where I first began to realize that I possessed that inestimable treasure called 'reason,'—from the cultivation of which I have since derived the purest intellectual delight." Cary further expressed thanks "for that *fine* article in the last 'Century'—it is *so inspiring* and so like you that I was carried back to the old school days and

[22] Ladd Franklin's elitism in this passage resonate with the political writings of Karl Pearson, a feminist and a eugenicist. On the relationship between these identities for Pearson, see MacKenzie (1981), especially pp. 84–87.

[23] In the higher education sector in 2006 the percent of female researchers in natural sciences was 24% in Germany and 39% in Spain, and in engineering & technology was 16% in Germany and 35% in Spain. In the government sector the percent of female researchers in natural science was 28% in Germany and 42% in Spain, and in engineering & technology was 20% in Germany and 39% in Spain (67).

seized with a great longing to tell you what a beautiful influence you have been in my life."

What Ladd Franklin had achieved for Cary in the classroom, she hoped to emulate for a wider audience in her popular articles. At a personal and professional level, Ladd Franklin chipped away at barriers for women. In her struggle for a better future, she showed that a single case could negate a false conclusion.

References

Abrams, E. 2020. 'Indebted to No One': Grounding and Gendering the Self-Made Mathematician. *Historical Studies in the Natural Sciences* 50 (3): 217–247.

Agler, D. W., and D. Durmuş. 2013. Christine Ladd-Franklin: Pragmatist Feminist. *Transactions of the Charles S. Peirce Society* 49 (3): 299–321.

Allen, G. 1888. Woman's Place in Nature. *The Forum* 7: 258–263.

Allen, G. 1890. Woman's Intuition. *The Forum* 9: 333–340.

Allen, G. 1891. *What's Bred in the Bone*. London: Tit-Bits.

Allen, G. 1899. *Miss Cayley's Adventures*. London: G. P. Putnam's Sons.

Anonymous. 1881. Obituaries. In *The Englishwoman's Review of Social and Industrial Questions*, ed. J. H. Murray and M. Stark, 9–38. Cham: Routledge Library Editions.

Boucard, J. 2020. Arithmetic and Memorial Practices by and around Sophie Germain in the 19th Century. In *Against All Odds: Women's Ways to Mathematical Research Since 1800*, ed. E. Kaufholz-Soldat and N. M. R. Oswald, 185–230. Cham: Springer.

Cajori, F. 1890. *The Teaching and History of Mathematics in the United States*. Washington: Bureau of Education, Government Printing Office.

Cameron, M. W., ed. 1928. *Biographical Cyclopedia of American Women*, vol. 3. New York: Halvord Publishing Co.

Cary, P. B. 1896. Letter to Mrs. Ladd Franklin. December 27 1896. The Ladd-Franklin Papers. Box 3. Letters A to F. Butler Library. Columbia University.

Claghorn, K. H. 1899. Letter to Mrs. Franklin. March 26, 1899. The Ladd-Franklin Papers. Box 3. Letters A to F. Butler Library. Columbia University.

Fausto-Sterling, A. 1985. *Myths of Gender: Biological Theories About Women and Men*. New York: Basic Books, Inc.

Fenster, D., and K. H. Parshall. 1994. Women in the American Mathematical Research Community: 1891 – 1906. In *The History of Modern Mathematics. 3 : Images, Ideas, and Communities*, ed. D. E. Rowe, J. McCleary, and E. Knobloch. London: Academic Press.

Franklin, F. 1898. The Intellectual Powers of Woman. *The North American Review* 166 (494): 40–53.

Furumoto, L. 1992. Joining Separate Spheres: Christine Ladd-Franklin, Woman-Scientist (1847–1930). *American Psychologist* 47 (2): 175–182.

Furumoto, L. 1994. Christine Ladd-Franklin's Color Theory: Strategy for Claiming Scientific Authority? *Annals of the New York Academy of Science* 727 (1): 91–100.

Galton, F. 1869. *Hereditary Genius: An Inquiry Into Its Laws and Consequences*. London: Macmillan.

Gigerenzer, G., M. Galesic, and R. Garcia-Retamero. 2014. Stereotypes About Men's and Women's Intuitions: A Study of Two Nations. *Journal of Cross-Cultural Psychology* 45 (1): 62–81.

Goldstein, C. 1994. Un arithméticien contre l'arithmétisation : les principes de Charles Hermite. In *Justifier en mathématiques*, ed. D. Flament and P. Nabonnand, 129–165. Paris: MSH.

Gray, J. 2008. *Plato's Ghost : The Modernist Transformation of Mathematics*. Princeton: Princeton University Press.

Green, J., and J. LaDuke. 2009. *Pioneering Women in American Mathematics: The Pre-1940 PhD's*. Providence: American Mathematical Society.
James, W. 1890. *The Principles of Psychology*. London: Macmillan.
James, W. 1892. Letter to Mrs. C. L. Franklin. March 3, 1892. The Ladd-Franklin Papers. Box 4. Letters G to M. Butler Library. Columbia University.
Jones, C. G. 2009. *Femininity, Mathematics, and Science c. 1880–1914*. New York: Palgrave Macmillan.
Klein, F. 1893. A Comparative Review of Recent Researches in Geometry. *Bulletin of the American Mathematical Society* 10 (2): 215–249.
Klein, F. 1894. *The Evanston Colloquium. Lectures on Mathematics Delivered from August 28 to September 9, 1893 at Northwestern University. Reported by Alexander Ziwet*. New York: Macmillan & Co.
Knowles, J. 1892. Letter to Mrs. Ladd Franklin. 6 January 1892. The Ladd-Franklin Papers. Box 4. Letters G to M. Butler Library. Columbia University.
Kohlstedt, S. G., and H. Longino. 1997. The Women, Gender, and Science Question: What do Research on Women in Science and Research on Gender and Science Have to Do with Each Other? *Osiris* 12: 3–15.
Ladd Franklin, C. 1886. The Study of Geometry. *Science. Supplement* 7 (152): 15–16.
Ladd Franklin, C. 1893. Intuition and Reason. *The Monist 3*: 211–219.
Ladd Franklin, C. 1894. Sophie Germain. An Unknown Mathematician. *Century Illustrated Magazine* 48: 946–948.
Ladd Franklin, C. 1896. The Higher Education of Women. *The Century* 53 (2): 315–316.
Ladd Franklin, C. 1929. *Colour and Colour Theories*. New York: Harcourt, Brace & Company.
MacKenzie, D. A. 1981. *Statistics in Britain, 1865–1930: The Social Construction of Scientific Knowledge*. Edinburgh: Edinburgh University Press.
Mehrtens, H. 1990. *Moderne Sprache, Mathematik : eine Geschichte des Streits um die Grundlagen der Disziplin und des Subjekts formaler Systeme*. Frankfurt: Suhrkamp.
of the Association of Collegiate Alumnae, A. C., ed. 1905. *Contributions toward a Bibliography of the Higher Education of Women. Supplement No. I.* Boston: The Trustees of the Public Library.
Parshall, K. H. 2006. *James Joseph Sylvester: Jewish mathematician in a Victorian World*. Baltimore: Johns Hopkins University Press.
Parshall, K. H., and D. Rowe. 1994. *The Emergence of the American Mathematical Research Community, 1876–1900: J. J. Sylvester, Felix Klein, and E. H. Moore*. Providence: American Mathematical Society.
Poincaré, H. 1910. Mathematical Creation. *The Monist* 20: 321–335.
Robertson, G. C. 1890a. Letter to Mrs. Ladd Franklin. 1 January 1890. The Ladd-Franklin Papers. Box 5. Letters N to S. Butler Library. Columbia University.
Robertson, G. C. 1890b. Letter to Mrs. Ladd Franklin. 25 October 1890. The Ladd-Franklin Papers. Box 5. Letters N to S. Butler Library. Columbia University.
Romanes, G. 1887. Mental Differences of Men and Women. *Popular Science Monthly* 31: 383–401.
Rossiter, M. 1982. *Women Scientists in America: Struggles and Strategies to 1940*. Baltimore: Johns Hopkins Press.
Rowe, D. E. 1986. 'Jewish Mathematics' at Göttingen in the era of Felix Klein. *Isis* 77: 422–449.
Schiebinger, L. 1989. *The Mind Has No Sex? Women in the Origins of Modern Science*. Cambridge: Harvard University Press.
Shields, S. 1975. Functionalism, Darwinism, and the Psychology of Women. *American Psychologist* 30 (7): 739–754.
Spillman, S. 2012. Institutional Limits: Christine Ladd-Franklin, Fellowships, and American Women's Academic Careers, 1880–1920. *History of Education Quarterly* 52 (2): 196–221.
Tobies, R. 2020. Internationality: Women in Felix Klein's Courses at the University of Göttingen (1893–1920). In *Against All Odds: Women's Ways to Mathematical Research Since 1800*, ed. E. Kaufholz-Soldat and N. M. R. Oswald, 9–38. Cham: Springer.

Chapter 11
Learning from the Masters (and Some of Their Pupils)

John Stillwell

Abstract Historians are trained to read original sources, and mathematicians in general are also advised to "study the masters." In practice, this is difficult to do, because some masters (and some languages) are easier to read than others. Nevertheless, it can fruitful to try, as I hope to show from my own experiences of reading and learning.

11.1 Introduction

I sometimes wonder, and maybe others do too, how I ended up among the historians of mathematics, where I am by no means a professional. Some people think I'm a logician, some think I'm a topologist, some think I'm a translator, and all these things are partly true. But today I am mainly an expositor with historical leanings, and I would like to explain how this came about. Books had a lot to do with it, and I like to think that books may nourish future mathematicians, and historians of mathematics, as much as they nourished me.

We have all heard the story of Abel, who advised that "to make progress in mathematics one should study the masters and not the pupils," and many of us have read the article *Read the Masters!* of Edwards (1981). And, of course, historians are *always* reading the masters, so we don't need to be convinced that it is a good idea. My story is about reading the masters partly by accident, first to learn things that were not taught at the schools I attended, and later finding that it was a good way to learn.

No doubt, other historians have had similar moments of revelation, but I hope that mine are of some interest because of their peculiar origins in Australia, with unexpected twists and turns through the work of Turing, Post, Dehn, Nielsen, Poincaré,

J. Stillwell (✉)
University of San Francisco, San Francisco, CA, USA
e-mail: stillwell@usfca.edu

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_11

Dedekind, and Dirichlet, and chance encounters with influential mathematicians. It will be seen later that Jeremy Gray makes an appearance at a crucial moment.

I give hearty thanks to David Rowe, whose reading of an early version of this article resulted in many improvements. I also benefited from comments by Chinese readers, and I hope that my experience reading the classics chimes with the Chinese cultural tradition, which has been based on the classics for millennia.

11.2 From Melbourne High School to University

When I started high school in 1956 I was not very interested in mathematics—nor in history, which was in fact my worst subject. This began to change in the final 2 years, when I chanced upon books on the history of mathematice in the school library, such as Florian Cajori's *A History of Mathematics* (Cajori 1919). (And, yes, I also loved E. T. Bell's *Men of Mathematics*, as did many of us before we knew better.) I was attracted by the mysterious formulas discovered by the early exponents of calculus, perhaps all the more so because I had no idea why they were true. Calculus was taught at school, but not the infinite series, infinite products, and continued fractions that fascinated me.

Instead I turned to some books on mathematics that were popular at the time. The first mathematics books I bought for myself were the books by W. W. Sawyer: *Mathematician's Delight* and *Prelude to Mathematics* (Sawyer 1943, 1955), which are still in print as Dover books. Others that I read in high school or in the first years at university were the Newman (1956) anthology *The World of Mathematics*, which includes digestible excerpts from many great mathematicians, and the slightly less digestible *A Course of Pure Mathematics* of Hardy (1958). Reading the latter book was undoubtedly my first experience of what could be classed as "learning from the masters, not their pupils." I did not fully appreciate it at the time, but it gradually sank in, with the help of two Cambridge-educated professors at the University of Melbourne, T. M. Cherry and E. R. Love.

Other important influences in my undergraduate years were two German-speaking mathematicians, Felix Behrend and Karl Moppert. Behrend was a refugee mathematician who taught a rigorous analysis course in our first year, including Dedekind cuts. I found the cut concept extremely irritating, but in a good way. As a naive beginner I felt sure there must be a better way to approach the real numbers, so I tried very hard to find it—only to conclude, eventually, that Dedekind cuts *were* the best way. Karl Moppert, in contrast to Behrend, was intuitive in the extreme. He viewed himself as a descendent of Felix Klein because one of his teachers had been Ostrowski, who was a student of Klein.

From both Behrend and Moppert I learned the importance of German mathematics, and they whet my appetite for later study of Dedekind and Klein. Klein came first, because he was more visual, and his very engaging *Elementary Mathematics from an Advanced Standpoint* was available in English as a cheap Dover paperback. I met Dedekind briefly through the English version of his little book on cuts

(Dedekind 1872), but as an undergraduate I wasn't ready for his works on algebraic integers and ideal theory.

11.3 University of Melbourne: 1960–1965

In the early 1960s, when I entered university, a lot of important mathematical literature was available only in German, French, and Russian. Hence part of the requirements for an advanced degree in science was a reading knowledge of at least two of these languages. For my master's degree I opted for German, which I had studied in high school, and Russian.

I discovered that translating from German was both enjoyable and enlightening, so much so that I went way beyond the requirements of the Science German course. Around 1963 I somehow found the time to translate most of Kamke's *Mengenlehre* (Kamke 1947), which I found in a foreign-language bookstore in Melbourne. I did not know that the book had already been translated into English. Kamke was not a giant of set theory, like Cantor or Hausdorff, but it was his readable little book that sparked my interest in set theory—the first of several "off syllabus" subjects I studied on my own.[1] Others were the related subjects of logic and computability theory.

Here I read my first masters: Alan Turing and Emil Post. Turing's paper on computable numbers (Turing 1936), was recommended to me by my honors superviser, Bruce Craven, who was kind enough to indulge my interest in logic, though it was not his field. I did not know at the time that Turing's concept of computation was what convinced Gödel (and hence logicians in general) that the intuitive concept of computability could be formalized. Turing certainly convinced me; I actually tried simulating a Turing machine on an old typewriter belonging to my mother, and I quickly rediscovered some elementary theorems about Turing machines, such as their ability to compute using only two different symbols, and without erasing.

Post's paper on recursively enumerable sets (Post 1944), I discovered more or less by chance, after a long search for an account of Gödel's incompleteness theorem that I could understand. Gödel's own paper was impossible, I thought, and the popularization by Nagel and Newman (1958) didn't help me much either. Quine (1951) was another washout, because it persisted in following Gödel's approach through a "self-referential" sentence, which I did not like. After these false starts, Post's idea of viewing incompleteness as a corollary of unsolvability, obtainable easily by a diagonal argument, was a revelation. It immediately struck me as the "right" approach to Gödel's theorem. I have since been on a crusade to make

[1] I have a habit of going off-syllabus, not only because syllabi have gaps but also because I prefer to learn at my own pace. I am generally not quick enough to grasp ideas in lectures or conversations, so learning from books has always been more important to me than to most mathematicians.

Post's name better known, describing his work in various books and in the article (Stillwell 2004). Post and Turing both died in 1954, but their posthumous careers have followed very different paths—Turing receiving the recognition he deserves, but Post . . . I'm afraid not.

Reading Post led to my master's thesis on recursively enumerable sets, which proved to be my ticket to graduate school at MIT, in 1965.

But this was not the only inspiration I received from Post. His 1947 paper on the word problem for semigroups also haunted me for the rest of my career. I was then very ignorant about algebra, and barely knew the difference between groups and semigroups. But semigroups were about to enjoy their day in the sun, with mid-1960s discoveries of connections between semigroups and automata. And Melbourne was about to become a hotspot for semigroup theory, with the foundation of Melbourne's second university, Monash, in 1961, its mathematics department headed by one of the world's leading semigroup theorists, Gordon Preston.

Gordon attended some talks I gave on the word problem, and other topics in logic, in 1965. (I did not know at this time that Gordon had actually been a colleague of Turing, among the cryptographers at Bletchley Park during WWII.) He informally offered me a job at Monash whenever I should complete my Ph.D. at MIT, and that is precisely what happened. That was the hiring process in the 1960s!

11.4 MIT: 1965–1970

The Science Russian lessons from Melbourne University bore fruit when I met similar requirements in graduate school. In early summer 1966 my advisor, Hartley Rogers, handed me a copy of the Russian edition of Mal'cev's *Algorithms and Recursive Functions* (Mal'cev 1970), as practice for MIT's Science Russian exam. He was surprised when I came back with a complete translation at the end of summer, and immediately wrote a potential publisher on my behalf. It turned out (not for the last time) that someone already had a contract to translate the book; however my translation was used in a small way to polish up the English edition. This was a bonus for me, since I still viewed translation as a way to learn and had not expected to get any credit or see my name in print. I did, however, learn the lesson that it was not hard to get translations published.

But other areas of mathematics, such as algebra and topology, remained a closed book to me. I received no training in these fields at Melbourne, where the emphasis was on classical analysis, but I thought I would like topology because knots and surfaces looked interesting. I had even given a lecture on intuitive topology when I was a freshman. So it was a terrible shock to run into the *Foundations of Algebraic Topology* of Eilenberg and Steenrod (1952), the assigned text in Daniel Kan's algebraic topology course at MIT. Far from being about surfaces or knots, this book did not have a single picture, other than commutative diagrams. Knowing what I know now, I am sure it is utterly the wrong way for beginners to approach topology, but all I could do then was drop out of the course (on Kan's advice) after a couple of weeks.

That was the only course I ever dropped out of, and I felt guilty about it. Even though Eilenberg and Steenrod wasn't my idea of topology, I thought that I *ought* to know something about ... whatever it was they were doing. I didn't realize that, from a historical point of view, Eilenberg and Steenrod fits better into the history of *category theory* than the history of topology. I didn't know that Kan was a major contributor to category theory, with his discovery of adjoint functors in Kan (1958), and that category theory was probably the hidden agenda of his algebraic topology course. How little we really know when we are graduate students!

In contrast to the fog surrounding algebra and topology in my graduate education, logic, set theory, and computability theory seemed dramatically clear. Not only could I study from current masters in the subject, such as Hilary Putnam and Tony Martin, virtually all the big names visited MIT and gave seminar talks. For example, we heard Robert Solovay explain his model of set theory in which every set of reals is Lebesgue measurable, years before he got around to writing it up for the *Annals*. About the only no-show was Paul Cohen, who by that time had settled the big independence questions of set theory and was taking aim at something completely different: the Riemann Hypothesis. We did, however, use his book Cohen (1966) in the graduate set theory course. It was then hot off the press, and a rare example of a research monograph that makes a good textbook (possibly because Cohen was not a professional set theorist, but an outsider with remarkable insight).

11.5 1970s: Group Theory and Topology

My training in logic, computability, and set theory at MIT was of lasting value, but in the early 1970s these fields became discouragingly complex, and I started looking for a bridge to somewhere else in mathematics. I found it in the computational aspects of topology and group theory, where the immediate learning challenge was to understand the word problem for *groups*, rather than semigroups. Seeking the origin of this problem led me to one of the most inspiring papers I ever read: Max Dehn's *Ueber unendliche diskontinuierliche Gruppen* (Dehn 1911). The richness of this paper—with its glorious blend of algebra, topology, and hyperbolic geometry—led to many things, including a collection of my translations of Dehn's papers, in Dehn (1987). I spent much of the 1970s in a frenzy of translation, because I found it best suited my style of learning.

The pace of translation, being slower than reading, gives just the right amount of time (for me) to absorb one idea before going on to the next. I found that the language didn't matter much, as long as I knew the basic words and very basic grammar, because it didn't take long to learn the vocabulary of a particular field of mathematics. So, I translated not only from German (where I was competent) but also from French (where I got by on the analogy with Spanish) and even Danish (where some simple words are weird and I didn't see the difference between singular and plural for a while). The other criteria were that the subject had to be potentially understandable—so that I could guess a correct translation for mathematical reasons when my language knowledge failed—and well-written.

Luckily for me, the authors I was most interested in wrote clearly and simply, so my pace of translation was equal to my pace of understanding. The authors in question were Dehn, Poincaré, and the duo Seifert and Threlfall.[2] I came to Poincaré through the series of topology papers he wrote under the title *Analysis Situs*, which translations were later collected in Poincaré (2010). When revising the translations, 30 years after I first made them, I was surprised to find that they no longer seemed quite so easy to follow. I think I was by then more sensitive to Poincaré's sloppy arguments, gaps, and occasional mistakes. When I was a beginner, I didn't notice them. The Zen concept of "beginner's mind" perhaps applies here: Poincaré wrote, and I first read, as a beginner.

In the 1970s I also translated Seifert and Threlfall's *Lehrbuch der Topologie* (Seifert and Threlfall 1934). This was a wonderful learning experience for me, and it also greatly pleased Seifert when I wrote to him in 1977, suggesting publication. He replied (14 October 1977):

> Ich freue mich, dass zu der russischen, chinesischen und spanisch Übersetzungen nun auch eine englische erscheinung soll. [I am delighted that an English translation will now appear in addition to those in Russian, Chinese, and Spanish.]

But, unbeknownst to either of us, another English translation had already been completed and accepted for publication (Seifert and Threlfall 1980).

Partly out of exasperation at losing the translation both Seifert and I had hoped for, I wrote my own book, *Classical Topology and Combinatorial Group Theory* (Stillwell 1980), based on my knowledge of the history of these subjects. While writing, I sent some of the group theory chapters to Dehn's eminent student, Wilhelm Magnus, who was very encouraging. It was also fortunate, I think, that I sent the finished manuscript to Paul Halmos at Springer, since it evidently struck a chord with him. Halmos was instrumental in getting the book published in Springer's GTM series, and led to my publishing several subsequent books at Springer.

The aim of the book was to revive the visual style of topology that prevailed from Dehn to Seifert-Threlfall, and I spent a month of a sabbatical at MIT just making the diagrams for the book. Not trusting the mail, I carried the diagrams personally (by Greyhound bus) to my publisher Springer in New York. While there, I had the pleasure of visiting Wilhelm Magnus at his home in New Rochelle. He encouraged me to continue translating Dehn's work and gave me some copies of unpublished Dehn manuscripts. These finally saw the light of day in Dehn (1987). In *Classical Topology* I also took the opportunity to make a connection with computability theory, which was fully realized in the second edition of the book in 1993, with the first textbook proof of the result of Markov (1958) that the homeomorphism

---

[2] Another translation I made around this time was of Reidemeister's *Kombinatorische Topologie*. I LaTeXed this a few years ago and posted it on the arXiv: arXiv:1402.3906.

problem for manifolds is unsolvable. (Reading Markov's paper, incidentally, was the last time I used my knowledge of Russian.)

The book also served, in a way, to exorcise the demons of algebraic topology that had haunted me since 1965. However, I didn't stop translating topology and combinatorial group theory: in 1980 I also published translations of Zieschang's *Flächen und ebene diskontinuierliche Gruppen* and (my first translation of a modern master) Serre's *Arbres, Amalgames*, SL_2 (English title simply *Trees*). The latter book starts off very simply with an introduction to graph theory, then proceeds very briskly to free groups and amalgamated free products. I got a little lost in the last chapter of *Trees*, but Serre helped me out. His bare-bones style seemed a bit austere to me then but now I can't fault it.

11.6 1980s: Geometry

My translations took a more geometric turn in the 1980s, foreshadowed by the hyperbolic geometry in Dehn's 1911 paper. One of the unpublished Dehn papers given to me by Magnus contained an early version of what is now called the Dehn-Nielsen theorem, relating mappings of a surface to the automorphisms of its fundamental group. This led me to the full version of the theorem proved by Dehn's student, Jakob Nielsen—like Wilhelm Magnus, a pupil who was himself a master[3]—and I couldn't stop there. Nielsen's use of hyperbolic geometry in topology and group theory was so brilliant that I eventually translated almost all of his papers, some of them in Danish. These papers were later collected in a two-volume edition of Nielsen's works (Nielsen 1986a,b), thanks to the generous support of the Danish Mathematical Society.

My fleeting competence in Danish had one more product: the English translation Topsøe (1990) of a book on the Poisson distribution. This book came at the request of the author who, while perfectly competent in English, wanted a native speaker to translate it. This experience taught me something I should have realized earlier: it is more important to be competent in English than in the language being translated. And, indeed, translating into English is good practice for *writing* in English. It is, in effect, "writing with trainer wheels," because the ideas are already in the right order, and it remains only to express them well in English.

In the early 1980s I became interested in classical algebraic geometry and complex function theory. In 1978 I had already translated Riemann (1851), the paper that introduces Riemann surfaces, but I was not satisfied with my work. Like many, I found Riemann's writing hard to follow. The right approach for me turned out to

[3] Another important Dehn student was Ruth Moufang, who made some spectacular discoveries on connections between the octonions and projective geometry. About 20 years ago I discovered some notes she wrote on this subject in the Frankfurt University mathematics library, and I later posted a translation of them on the arXiv: arXiv:2012.05809.

be the masterly *Ebene algebraische Kurven* of Brieskorn and Knörrer (1981). I was so excited by the long historical introduction in this book that I translated its whole 900 pages—the longest book I ever translated. It taught me a lot about the history of algebraic geometry, and also gave me a new insight into topology, where Riemann surfaces arise from algebraic curves and knots arise from their singularities.

Filling in more details in the history of topology and geometry, I came back to Poincaré, making a translation of his main papers on Fuchsian functions, including his approach to hyperbolic geometry in two and three dimensions (Poincaré 1985). These papers are foundational not only in non-Euclidean geometry, but also in the "geometrization" of topology, which was enjoying a revival at this time thanks to the work of Thurston. In the process, the work of Dehn and Nielsen also enjoyed a revival, when it was noticed that Thurston had actually rediscovered some of Nielsen's results on the topology of surfaces. The confluence of Dehn, Nielsen, and Thurston took place around the time of a visit by Thurston to University of Melbourne in the late 1970s, so I was well aware of it, and it informed my writing of *Classical Topology and Combinatorial Group Theory*. Thurston's geometrization was so influential that the latter part of my title was eventually rendered obsolete: "combinatorial group theory" is now "geometric group theory."

In the 1980s I had the opportunity to teach history of mathematics for the first time. I felt ready, because at this stage I had already learned geometry, topology, and a bit of algebra, from the masters. Also, in the mid-1980s there was growing anticipation of a proof of Fermat's last theorem, thanks to results of Serre, Frey, and Ribet that reduced it to a problem about elliptic curves. The decisive moment, for me, occurred when I heard Ken Ribet give a talk on his contribution at Cambridge University in 1987. With Fermat in mind, I studied the number theory and geometry around elliptic functions and elliptic curves, and included its history in my course. This is where Jeremy Gray comes into my story. I sent him some of my lecture notes and, with his encouragement, I started writing *Mathematics and Its History* (Stillwell 1989). The first edition came out in 1989, and there have been two more editions since. It turned out to be my most popular book. (Of course, it was only in the second edition (2002) that I could report that Fermat's last theorem had been proved.)

11.7 1990s: Number Theory and Geometry

In the 1990s I followed the number theory thread into the writings of my next master, Richard Dedekind, and to his master, Dirichlet. First I made the translation of Dedekind's 1877 *Theory of Algebraic Integers*, which came out as Dedekind (1996). Strangely, the original version of Dedekind's monograph seems to be lost, so I worked from the French translation (Dedekind 1877). Dedekind writes so well I suspect that could be part of the reason his approach (ideals versus divisors) is more popular than that of Kronecker. I found Kronecker's writings impenetrable, with one exception—his paper Kronecker (1870) that gives the structure theorem for finite Abelian groups. I shoehorned Kronecker's proof of this theorem into my *Classical Topology and Combinatorial Group Theory*.

As for Dirichlet, I worked from the first version of the work that later became known as Dirichlet-Dedekind, consisting of Dirichlet's lectures on number theory and Dedekind's supplements. Most of the supplements remained the same after the first version (Lejeune-Dirichlet 1871), but the supplement on algebraic integers and ideal theory grew until it dominated the book in the final version (Lejeune-Dirichlet 1894). I omitted this supplement because I felt that Dedekind (1996) was a more approachable treatment of the subject, though some number theorists have been disappointed by my decision. Dirichlet's book was the last I translated by pen onto paper. I translated the last 34 pages of the book in a single day, which is my all-time record. I had the luxury of writing by hand then because my son Michael had the time to type it up in LaTeX. After that, like most of us, I had to do my own LaTeX, which is a lot slower but nevertheless satisfying.

The important paper Dedekind and Weber (1882) on the arithmetic theory of algebraic functions that Dedekind co-authored with Weber was, for me, his most difficult. It took me a couple of attempts to translate it, finally getting it published in 2012. Maybe the reason is that Dedekind left most of the writing to Weber, who was perhaps a less gifted expositor; for whatever reason, their joint paper is not as readable as Dedekind's other writings.

Also in this decade I collected my translations of Beltrami, Klein, and Poincaré on hyperbolic geometry in Stillwell (1996), *Sources of Hyperbolic Geometry*, and wrote my own books *Geometry of Surfaces*, *Elements of Algebra* and *Numbers and Geometry*. By this time I was beginning to think it was time to do less translation, and to write up what I had learned. There is no shortage of topics that I think can be better explained with the help of some historical insight.

11.8 2000s

The books I wrote in the 1990s established a pattern I continued in the 2000s: books on undergraduate topics written from a historical perspective. I did revisit the history of topology in a survey article on Poincaré and the history of 3-manifolds (Stillwell 2012), and I returned somewhat to my roots in logic and set theory with books on infinity, reverse mathematics, and the history of proof.

The quotation from Abel that prompted the title of this article—"study the masters, not the pupils"—was of course intended to explain his productivity as a research mathematician, which I cannot claim to have achieved. But it is good advice for everyone and, as long as the masters do not have the time to write expository books, it seems to be up to us, their pupils, to do the job. Otherwise, few will know what the masters have done.

* * * * * * * * * * * * *

The motto of Monash University, where I spent more than half of my career, is "Ancora imparo," meaning "I am still learning." I very much agree with this saying.

I have never written a book or article on how to learn, or teach, or read, or write, or translate mathematics, and I probably never will. That is because *I still don't know*, exactly, how to do it. But I hope to improve.

References

Brieskorn, E., and H. Knörrer. 1981. *Ebene algebraische Kurven*. Basel: Birkhäuser Verlag. English translation *Plane Algebraic Curves*, by John Stillwell, Birkhäuser-Verlag, 1986.

Cajori, F. 1919. *A History of Mathematics*, 2nd ed. New York: Macmillan.

Cohen, P. J. 1966. *Set Theory and the Continuum Hypothesis*. New York: W. A. Benjamin, Inc.

Dedekind, R. 1872. *Stetigkeit und irrationale Zahlen*. Braunschweig: Vieweg und Sohn. English translation in: *Essays on the Theory of Numbers*, Dover, New York, 1963.

Dedekind, R. 1877. Sur la théorie des nombres entiers algébriques. *Darboux Bulletin (2)* 1: 17–41.

Dedekind, R. 1996. *Theory of Algebraic Integers*. Cambridge Mathematical Library. Cambridge: Cambridge University Press. Translated from the 1877 French original and with an introduction by John Stillwell.

Dedekind, R., and H. Weber. 1882. Theorie der algebraischen Functionen einer Veränderlichen. *Journal für die reine und angewandte Mathematik* 92: 181–290. English translation *Theory of Algebraic Functions of One Variable*, by John Stillwell, American Mathematical Society, 2012.

Dehn, M. 1911. Über unendliche diskontinuierliche Gruppen. *Mathematische Annalen* 71 (1): 116–144.

Dehn, M. 1987. *Papers on Group Theory and Topology*. New York: Springer-Verlag. Translated from the German and with introductions and an appendix by John Stillwell, With an appendix by Otto Schreier.

Edwards, H. M. 1981. Read the masters! In *Mathematics Tomorrow*, 105–110. New York: Springer.

Eilenberg, S., and N. Steenrod. 1952. *Foundations of Algebraic Topology*. Princeton: Princeton University Press.

Hardy, G. H. 1958. *A Course of Pure Mathematics*, 10th ed. New York: Cambridge University Press.

Kamke, E. 1947. *Mengenlehre*. Berlin: Walter de Gruyter & Co.

Kan, D. M. 1958. Adjoint functors. *Transactions of the American Mathematical Society* 87: 294–329.

Kronecker, L. 1870. Auseinandersetzung einiger Eigenschaften der Klassenzahl idealer complexer Zahlen. *Monatsbericht der Königlich-Preussischen Akademie der Wissenschaften zu Berlin*, 881–889.

Lejeune-Dirichlet, P. G. 1871. *Vorlesungen über Zahlentheorie, herausgegeben von R. Dedekind. Zweite Auflage.* Braunschweig. Vieweg. English translation *Lectures on Number Theory*, by John Stillwell, American Mathematical Society, 1999.

Lejeune-Dirichlet, P. G. 1894. *Vorlesungen über Zahlentheorie. Hrsg. und mit Zusätzen versehen von R. Dedekind. 4. umgearb. u. verm. Aufl.* Braunschweig. F. Vieweg u. Sohn. XVII + 657 S. 8°.

Mal'cev, A. I. 1970. *Algorithms and Recursive Functions*. Groningen: Wolters-Noordhoff Publishing. Translated from the first Russian edition by Leo F. Boron, with the collaboration of Luis E. Sanchis, John Stillwell and Kiyoshi Iséki.

Markov, A. 1958. The insolubility of the problem of homeomorphy. *Doklady Akademii Nauk SSSR* 121: 218–220.

Nagel, E., and J.R. Newman. 1958. *Gödel's Proof.* New York: New York University Press.

Newman, J.R. 1956. *The World of Mathematics. Vols. I–IV*. New York: Simon & Schuster. A small library of the literature of mathematics from A'h-mosé the Scribe to Albert Einstein, presented with commentaries and notes.

Nielsen, J. 1986a. *Jakob Nielsen: Collected Mathematical Papers. Vol. 1. Contemporary Mathematicians*. Boston: Birkhäuser Boston, Inc. Edited and with a preface by Vagn Lundsgaard Hansen.

Nielsen, J. 1986b. *Jakob Nielsen: collected mathematical papers. Vol. 2. Contemporary Mathematicians*. Boston: Birkhäuser Boston, Inc. Edited and with a preface by Vagn Lundsgaard Hansen.

Poincaré, H. 1985. *Papers on Fuchsian Functions*. New York: Springer-Verlag. Translated from the French and with an introduction by John Stillwell.

Poincaré, H. 2010. *Papers on Topology*, Volume 37 of *History of Mathematics*. Providence/London: American Mathematical Society/London Mathematical Society. *Analysis situs* and its five supplements, translated and with an introduction by John Stillwell.

Post, E. L. 1944. Recursively enumerable sets of positive integers and their decision problems. *Bulletin of the American Mathematical Society* 50: 284–316.

Quine, W.V.O. 1951. *Mathematical Logic*. Cambridge: Harvard University Press. Revised ed.

Riemann, G.F.B. 1851. *Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse*. *Werke*, 2nd ed., 3–48.

Sawyer, W. W. 1943. *Mathematician's Delight*. Penguin Books.

Sawyer, W. W. 1955. *Prelude to Mathematics*. Penguin Books.

Seifert, H., and W. Threlfall. 1934. *Lehrbuch der Topologie.* IV + 353 S. 132 Abb. Leipzig: B. G. Teubner.

Seifert, H., and W. Threlfall. 1980. *Seifert and Threlfall: A Textbook of Topology*, Volume 89 of *Pure and Applied Mathematics*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London. Translated from the German edition of 1934 by Michael A. Goldman, With a preface by Joan S. Birman, With "Topology of 3-dimensional fibered spaces" by Seifert, Translated from the German by Wolfgang Heil.

Stillwell, J. 1980. *Classical Topology and Combinatorial Group Theory*, Volume 72 of *Graduate Texts in Mathematics*. New York: Springer-Verlag.

Stillwell, J. 1989. *Mathematics and Its History*. *Undergraduate Texts in Mathematics*. New York: Springer-Verlag.

Stillwell, J. 1996. *Sources of Hyperbolic Geometry*, Volume 10 of *History of Mathematics*. Providence/London: American Mathematical Society/London Mathematical Society.

Stillwell, J. 2004. Emil Post and his anticipation of Gödel and Turing. *Mathematics Magazine* 77 (1): 3–14.

Stillwell, J. 2012. Poincaré and the early history of 3-manifolds. *Bulletin of the American Mathematical Society (N.S.)* 49 (4): 555–576.

Topsøe, F. 1990. *Spontaneous Phenomena*. Boston: Academic Press, Inc. A mathematical analysis, Translated from the Danish by John Stillwell.

Turing, A. M. 1936. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society (2)* 42 (3): 230–265.

Part III
Mathematics and Natural Sciences

Chapter 12
Mathematical Practice: How an Astronomical Table Was Made in the *Yuanjia li* (443 AD)

Anjing Qu

Abstract Simply according to the analysis and mathematical modeling of the structure of constants, without the support of any written historical materials, the procedure by which an astronomical table of solar shadow in the *Yuanjia li* (443 AD) was made is reconstructed. The results of this paper show that the sources of some important basic constants in traditional Chinese calendars are quite unexpected, since they may not come from practical measurement, or simple adjustment of the measured data. The constant of length of solar shadow in the *Yuanjia li* was deduced from an indeterminate equation, which was artificially constructed from a mathematical model. This research is an example of the paradigm of *recovery*, and the paradigm is analogous to the focus on *mathematical practice* in history of exact science.

12.1 Paradigm of *Recovery* and *Mathematical Practice*

For the past few years, *mathematical practice* as a new focus of the history and philosophy of mathematics has become more and more popular in Europe and other places. One of the origins for *practice* comes from a renewal of the word *pratique* in French history and philosophy of science, in particular through the influence of Michel Foucault, in the 1970s and 1980s (Ritter 1989). As a phrase, *mathematical practice* appeared in the title of an article of the *Journal of Symbolic Logic* in the 1980s (Shapiro 1985). There were works on the history and philosophy of mathematics written with this paradigm in the 1990s (Fraser 1999; Bowers et al. 1999). Since the beginning of the new century, especially in the last decade, more and more historians and philosophers of mathematics have reported their results based on the topic of *mathematical practice* in various journals and academic

A. Qu (✉)
Institute for Advanced Studies in History of Science, Northwest University, Xi'an, China
e-mail: qaj@nwu.edu.cn

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_12

conferences on the history of mathematics (Mancosu 2008; Soler et al. 2014; Chemla 2013) .

What is *mathematical practice*? Based on the *Association for the philosophy of mathematical practice*, we can see that "It includes the study of a wide variety of issues concerned with the way mathematics is done, evaluated, and applied, and in addition, or in connection therewith, with historical episodes or traditions, applications, educational problems, cognitive questions, etc."

Why is *mathematical practice* a new topic? Because the two international mainstreams, history of *mathematical practice* and philosophy of *mathematical practice*, are both trying to open up the field of historical and philosophical studies, avoiding what they perceive to be an overly narrow focus on written texts (Ferreirós 2016). The latter is a traditional paradigm to study the history of mathematics and astronomy, characterized by *discovery.*

What does *discovery* mean? Or equivalently, what's the problem field of the paradigm of *discovery*? When we have full and concrete historical materials, the interpretation of these materials, and the explanation of the mathematical thoughts and methods understood by ancient mathematicians would go with the paradigm of *discovery*.

However, this paradigm has its limitations. First of all, only explicit text information can be used in the research. Secondly, it's not easy to find fresh or ignored historical materials. Against this backdrop, Wu Wen-Tsun put forward a new paradigm, gave a few principles to characterize it as *recovery* (Wu 1986). Historians of mathematics in China have been familiar with this kind of paradigm since the 1980s (Qu 2002). It must be mentioned that broken chains of evidence or circumstantial evidence are allowed in the paradigm of *recovery*. In this way, the original problem-field for the paradigm of *discovery* has been broken through.

More explicitly, the traditional paradigm of the history of mathematics focuses on "what", while the paradigm of *recovery* focuses on "how". Moreover, the paradigm of *recovery* allows historians to analyze the data structures, investigate the illustration features, study the algorithm procedure, based on the limited documents, to dig out the underlying information, such as formulae, algorithms, figures and so on, in order to logically deduce the way in which the ancient mathematicians created knowledge. This research idea is similar to the history of *mathematical practice*.

Therefore, due to the lack of direct support from the mathematical text one wants to study, the way in which the historian understands how knowledge was created through a particular *mathematical practice* is not *discovery* but *recovery*. It shows that the problem-field of *recovery* is different from that of the traditional paradigm of *discovery*.

The shift of paradigm for the history of exact science expanded the problem-field of traditional historiography. There will be a great quantity of new problems arising from this new paradigm. By means of expanding the problem-field, we aim to come up with more new problems to reinvigorate the study of the history of exact science, such as the history of mathematical astronomy, and the history of modern mathematics (Qu 2018).

Of course, the problem raised and solved in this paper, focusing on *mathematical practice* by means of the paradigm of *recovery*, is very interesting.

12.2 The Problem

In Chinese, the *li* is a calendar-making system which contains astronomical constants and algorithms dealing with the movements of the sun, the moon, and five planets. The *Yuanjia li* (443 AD) is an important calendar-making system made by He Chengtian (370–447 AD) in the Liu Song Dynasty. In this paper, we take an astronomical table as an example, to discuss how He Chengtian obtained the constants of length of solar shadow in the *Yuanjia li* with the paradigm of *recovery* (The Compilation of Chapters of Astronomy, Music and Calendar in the Official History of China 6 歷代天文律曆等志匯編(六) 1976b).

Normally, the length of solar shadow is measured by a gnomon with 8 *chi* height at noon.[1] The solar shadow table at 24 solar terms in a year is the basic data of many astronomical calculations. So, it is very important.

From the Eastern Han Dynasty to the Liu Song Dynasty (first to fifth century AD), we can find the tables of the solar shadow in 3 calendars, *Sifen li*, *Jingchu li* and *Yuanjia li*. To recover the construction process of the table in the *Yuanjia li*, we put the constants in these 3 calendars into Table 12.1 as a reference. The tables are the same in the *Sifen li (Quaternary li)* (85 AD) (The Compilation of Chapters of Astronomy, Music and Calendar in the Official History of China 5 歷代天文律曆等志匯編(五) 1976a) and the *Jingchu li* (237 AD) (The Compilation of Chapters of Astronomy, Music and Calendar in the Official History of China 5 歷代天文律曆等志匯編(五) 1976a, 1632–1634). By symmetry, we just need to list the 13 constants from the Winter Solstice to the Summer Solstice. The unit of the constant in Table 12.1 is *fen*, 1 foot (*chi*) = 100 *fen*. The *Difference* is the first order difference of two adjacent solar terms (Fig. 12.1).

With the paradigm of *discovery*, we only need to check that the constants in Table 12.1 are the lengths of the solar shadow at 24 solar terms in a year. In general, people don't discuss how the astronomers obtained these constants without the support of original written materials. However, with the paradigm of *recovery*, the question "how to obtain this group of data?" is legitimate and worth discussing.

In this paper, without any original written description, we analyze the structures and characteristics of the data to answer the question: how did He Chengtian obtain the constants of length of the solar shadow at 24 solar terms?

This is the problem in this paper.

[1] *fen,* as well as *chi*, is a unit of length. One *chi* approximately equals to one foot.

Table 12.1 The solar shadow tables in three calendars[a]

	Sifen li/Jingchu li		*Yuanjia li*	
Solar Term	Solar Shadow	Difference	Solar Shadow	Difference
Winter solstice	1300	70	1300	52
Minor cold	1230	130	1248	114
Major cold	1100	140	1134	143
Start of spring	960	165	991	169
Rain water	795	145	822	150
Awakening of insects	650	125	672	133
Spring equinox	525	110	539	114
Clear and bright	415	95	425	100
Grain rain	320	68	325	75
Start of summer	252	54	250	53
Grain buds	198	30	197	28
Grain in ear	168	18	169	19
Summer solstice	150		150	

[a]The Solar Shadow here means the constant of the length of solar shadow at the moment of each solar term. These are the original data quoted from the calendars

Fig. 12.1 The Solar shadow table of the *Yuanjia li.* (Shen 1739)

12.3 Characteristics of Solar Shadow Table in the *Yuanjia li*

It's easy to understand the creation process of algorithms, theorems, concepts and theories if there are sufficient written materials. The first one to explain it clearly would go with the paradigm of *discovery.* It's one of the problems in the field of the traditional paradigm of historical research.

However, when facing indirect or broken chains of evidence, we can only restore the construction process according to circumstantial evidence. This kind of research would deal with a problem in the paradigm of *recovery*.

Recovering the construction process of the solar shadow table in the *Yuanjia li* is an example of the problems above, because there is no literature left to tell us He Chengtian's construction process of the table in the *Yuanjia li*.

Then, how did He Chengtian obtain these constants in his table? The simplest explanation, of course, is that all the data comes from practical measurement. In order to explore whether the data come from practical measurement or an algorithm, we need to find out the rules behind the data.

Dividing the *Difference* of the solar shadow table in the *Yuanjia li* into two factors, we find that except for the last three numbers (53, 28, 19), one takes into account only the first nine digits, they can be divided into three groups with the same factor, namely 19, 13, 25 in each group:

$$114 = 6 \times 19, \quad 133 = 7 \times 19, \quad 114 = 6 \times 19, \tag{12.1}$$

$$52 = 4 \times 13,\ 143 = 11 \times 13,\ 169 = 13 \times 13, \tag{12.2}$$

$$150 = 6 \times 25,\ 100 = 4 \times 25,\ 75 = 3 \times 25. \tag{12.3}$$

As shown in Table 12.2, if we add the three factors that are not in common in group (1), the sum will be

$$6 + 7 + 6 = 19,$$

This number is the same as the *Difference* of Grain in Ear. If we add the three factors that are not in common in group (2), the sum will be

$$4 + 11 + 13 = 28,$$

This number is the same as the *Difference* of Grain Buds. Therefore, the sum of the three numbers of group (2) is 28×13, while the sum of the numbers in group (3) is 13×25. From this, it can be seen from this that the sum of the six numbers of groups (2) and (3) is

$$28 \times 13 + 13 \times 25 = 53 \times 13,$$

in which the sum of the factors is $28 + 25 = 53$, which is the *Difference* of Start of Summer.

These perfectly matched data suggest that the solar shadow table in the *Yuanjia li* was calculated meticulously by He Chengtian. In this way, we find that these data were the results of an algorithm not practical measurements.

How did He Chengtian calculate these constants? To solve this problem, we can build a mathematical model according to the analysis of structural characteristics above.

12.4 Mathematical Model of the Solar Shadow Table in the *Yuanjia li*

According to the structural analysis of the data, the twelve *Difference* constants in the *Yuanjia li* can be divided into two groups. We take x, y, z to express the *Differences* of Grain in Ear, Grain Buds, Start of Summer, respectively. See Table 12.2. They are basic constants, playing an important role in constructing the model.

The difference of the solar shadow lengths of Winter Solstice and Summer Solstice is 1300 – 150, which equals to the sum (1150) of all *Differences* from Winter Solstice to Grain in Ear. Thus, the sum of the nine *Differences* from Winter Solstice to Grain Rain should be

$$1150 - (x + y + z)\,.$$

Therefore, if we can determine $x + y + z$, the next problem is to distribute the $1150 - (x + y + z)$ into the nine *Differences* from Winter Solstice to Grain Rain.

For the construction process of the model, please refer to Tables 12.2 and 12.3.

According to the analysis in the previous section, we find that He Chengtian divided the nine *Differences* into three groups, in which the group (1) contains the factor x, and the sum of the three numbers can be expressed as

$$x \cdot m_1 + x \cdot m_2 + x \cdot m_3 = x \cdot m.$$

Table 12.2 Structures of the solar shadow table in the *Yuanjia li*

Solar Terms	Solar Shadow	Difference	(1)	(2)	(3)
Winter solstice	1300	52		4·13	
Minor cold	1248	114	19·6		
Major cold	1134	143		11·13	
Start of spring	991	169		13·13	
Rain water	822	150			25·6
Awakening of insects	672	133	19·7		
Spring equinox	539	114	19·6		
Clear and bright	425	100			25·4
Grain rain	325	75			25·3
Start of summer	250	53			
Grain buds	197	28			
Grain in ear	169	19			
Summer solstice	150				
				28·13	25·13
		1150-100	19·19	53·13	

Table 12.3 The mathematical model of the solar shadow table in the *Yuanjia li*

Difference	(1)	(2)	(3)
+		$y_1 \cdot n$	
+	$x \cdot m_1$		
+		$y_2 \cdot n$	
+		$y_3 \cdot n$	
+			$(z-y) \cdot n_1$
+	$x \cdot m_2$		
+	$x \cdot m_3$		
+			$(z-y) \cdot n_2$
+			$(z-y) \cdot n_3$
z			
y			
x			
$1150-(x+y+z)$	$x \cdot m$	$y \cdot n$	$(z-y) \cdot n$
		$z \cdot n$	

The three numbers in group (2) all contain the same factor y. Dividing the *Difference* of Grain Buds properly, the sum of these three data can be expressed as

$$y_1 \cdot n + y_2 \cdot n + y_3 \cdot n = y \cdot n.$$

The three numbers in group (3) all contain the factor $z - y$, so the sum is

$$(z-y) \cdot n_1 + (z-y) \cdot n_2 + (z-y) \cdot n_3 = (z-y) \cdot n.$$

Thus, the sum of the six numbers in groups (2) and (3) is

$$y \cdot n + (z-y) \cdot n = z \cdot n.$$

Therefore, the main matter is reduced to an indeterminate equation:

$$1150 - (x+y+z) = x \cdot m + z \cdot n. \tag{12.4}$$

Once the *Differences* of Grain in Ear, Grain Buds and Start of Summer, namely x, y, z, are selected, the indeterminate Eq. 12.4 is established. Then, solving the Eq. 12.4, we obtain m, n, and distribute them properly into

$$m_1 + m_2 + m_3 = m;$$
$$n_1 + n_2 + n_3 = n.$$

In this way, the *Differences* of solar shadow from Winter Solstice to Grain in Ear are obtained, and the solar shadow table in the *Yuanjia li* is constructed.

12.5 Recovery of the Procedure of the Solar Shadow Table in the *Yuanjia li*

Now we can recover He Chengtian's construction process of the solar shadow table in the *Yuanjia li*. According to the previous discussion, the whole matter is reduced to the indeterminate Eq. 12.4. Before determining the coefficients and final results of the Eq. 12.4, a hypothesis needs to be made clear:

He Chengtian was a mathematical astrologer (Chen 2003). In terms of mathematics, he invented an algorithm named *tiao rifa* which means a method of denominator selection. The essence of this algorithm is to select an appropriate rational number to approximate the constant of the synodic month. He Chengtian reduced this problem to solving an equation similar to Eq. 12.4 (Li 2007). The construction of this equation contains the pursuit of some beautiful numbers, which include some special prime factors. This is the inheritance of the digital mysticism from the Qin and Han Dynasties (from second century BC on) (Chen 1984). What's more, in the era of He Chengtian, the *Chinese Remainder Theorem* had already been used to solve linear congruences or indeterminate equations. This background laid the necessary mathematical foundation for solving the Eq. 12.4.

As we have seen, He Chengtian's first step was to transform the Eq. 12.4 into an indeterminate or congruence equation with m and n as unknown quantities, by determining the *Differences* x, y, z of Grain in Ear, Grain Buds and Start of Summer, respectively.

Referring to the data in the *Sifen li (Quaternary li)* and *Jingchu li*, the *Differences* of Grain in Ear and Start of Summer, namely x, z, are 18 and 54, respectively. In order to make a perfect Eq. 12.4, it's better to choose the nearest prime numbers as x, z, which were taken for granted by digital mysticism at that time. The nearest prime numbers to 54 and 18 are $z = 53$, and $x = 17$ or 19.

If we choose the results found in the *Yuanjia li*, $x = 19$, $z = 53$, and $x + y + z = 100$, then according to the Eq. 12.4, there is an indeterminate equation:

$$1050 = 19m + 53n. \tag{12.5}$$

We can easily obtain the general solution of this equation by means of the *Chinese Remainder Theorem:*

$$m = 19 + 53k;$$
$$n = 13 - 19k.$$

in which k is an arbitrary integer. It can be seen that there is only one set of positive integer solutions of the Eq. 12.5:

$$m = 19, n = 13.$$

This is exactly the result that He Chengtian used in the *Yuanjia li*.

Table 12.4 The possible options of the solar shadow table in the *Yuanjia li*

x	z	$x+y+z$	m	n	
19	53	100	19	13	OK
19	53	101	5	18	Non prime number, $m$is too small
19	53	102	44	4	Non prime number, n is too small
19	53	103	30	9	Non prime number, n is too small
19	53	104	16	14	Non prime number
17	53	100	15	15	Non prime number
17	53	101	43	6	Non prime number, n is too small
17	53	102	18	14	Non prime number
17	53	103	46	5	Non prime number, n is too small
17	53	104	21	13	Non prime number

In fact, we have found out that $x + y + z = 102$ in the *Sifen li (Quaternary li)* and the *Jingchu li*, as shown in Table 12.1. Then, we limit the value range of $x + y + z$ to 102 ± 2, so, no matter whether $x = 19$ or 17, the indeterminate Eq. 12.4 has only one set of positive integer solutions. Limiting m, n to be prime numbers, the solution used in the *Yuanjia li* is the unique suitable result. The possible options are shown in Table 12.4.

In other words, according to the mathematical model made by He Chengtian, the solar shadow table in the *Yuanjia li* is obtained by solving the indeterminate Eq. 12.4. In this equation, if we limit the coefficients x, y, the solutions m, n are all prime numbers, then the data in the *Yuanjia li* give the only proper option in a very wide range of possible options.

This is the procedure of construction of the solar shadow table in He Chengtian's *Yuanjia li*.

12.6 Why Does He's Method Matter?

We have recovered He Chengtian's method for selecting the constants of the length of the solar shadow by the paradigm of *recovery*. He Chengtian transformed the simple astronomical problem into a linear indeterminate Eq. 12.4 by a mathematical model.

This is unexpected!

Did He Chengtian use this method on purpose? Or just as a special case? Why does He's method matter?

In fact, the method He Chengtian used here is not a unique case. As mentioned before, as an astrologer, He Chengtian's most famous invention in the history of mathematics is *tiao rifa*, i.e., a method of denominator selection. In traditional Chinese calendars, there are many fractions selected properly as the astronomical constants used in calendars, *rifa* is the denominator of the constant of the synodic month.[2] *Tiao rifa* is actually a method for determining the constant of the synodic month used in calendars.

From the literature of the Song Dynasty (eleventh to thirteenth century), there are three historical records about He Chengtian's method of denominator selection (*tiao rifa*). The earliest record was given by Zhou Cong, the author of the *Mingtian li* (1064 AD). In the first article of *The Commentary of Mingtian li*, Zhou Cong wrote:

> 調日法
> 造曆之法, 必先立元, 元正然後定日法, 法定然後度周天分以定分、至, 三者有程, 則曆可成矣。(The Compilation of Chapters of Astronomy, Music and Calendar in the Official History of China 8 歷代天文律曆等志匯編(八) 1976c)
> The method of denominator selection
> For making a calendar, must determine the epoch first. Then select a proper denominator of the constant of the synodic month. Next, determine the constant of the tropical year and then to determine the 24 solar terms. Once these three steps above are finished, the calendar is ready to compile.

This statement tells us the three steps of compiling a calendar: determine the epoch, select the denominator of the constant of the synodic month, and determine the constant of the tropical year. But how to select a proper denominator? Zhou Cong continued that:

> 宋世何承天更以四十九分之二十六為强率, 十七分之九為弱率, 于强弱之際以求日法。承天日法七百五十二, 得一十五强, 一弱。自後治曆者, 莫不因承天法...... *He Chengtian, in the Liu Song Dynasty, took 26/49 as the bigger ratio, 9/17 as the smaller ratio, and selected the rifa (denominator) between the bigger and smaller ratios. Chengtian's rifa was 752, and he obtained the number of bigger ratios (hereinafter called "bigger number") is* $n = 15$*, the number of smaller ratios (hereinafter called "smaller number") is* $m = 1$*. Later calendar makers all followed He Chengtian's approach...*

This is the earliest expression we know for the method of denominator selection in history. This paragraph states in effect that if we suppose the fractional part of the constant of the synodic month is B/A, A is the denominator (*rifa*), B is the numerator (*shuoyu*). He Chengtian knew that:

$$\frac{9}{17} < \frac{B}{A} < \frac{26}{49},$$

in which 26/49 is the bigger ratio, 9/17 is the smaller ratio. Therefore, find the positive integers m, n that satisfy

$$\frac{B}{A} = \frac{9m + 26n}{17m + 49n}, \tag{12.6}$$

[2] The calendars after the Sui Dynasty (from the seventh century on) cancelled the intercalary circle, so the *rifa* was also selected as the denominator of the constant of the tropical year. About the method of *tiao rifa*, one may find a discussion in the book of Jean Claude Martzloff: *Astronomy and Calendars – The Other Chinese Mathematics, 104 BC-AD1644*. Springer.

in which n is the bigger number, m is the smaller number. The method of denominator selection is: select a proper denominator A to solve the indeterminate equation

$$A = 17m + 49n. \tag{12.7}$$

Obtain the bigger number n and smaller number m to determine B/A.

For example, as Zhou Cong said, given the denominator is $A = 752$ in the *Yuanjia li*, we can obtain that the bigger number is $n = 15$ and smaller number is $m = 1$. Therefore, the fractional part of the constant of the synodic month in the *Yuanjia li* is

$$\frac{B}{A} = \frac{9 + 26 \times 15}{17 + 49 \times 15} = \frac{399}{752}.$$

This process was demonstrated in detail in the algorithm for deducing the epoch in Qin Jiushao's *Mathematical Treatise in Nine Chapters* (1247 AD). At the beginning of the algorithm for deducing the epoch, Qin Jiushao (1208–1268) wrote:

> 以暦法求之, 大衍入之。調日法, 如何承天術。用强弱母子互乘, 得數, 並之, 為朔余。(Chen 1984, 263)
>
> *To determine it (the epoch), apply the algorithm related to the Chinese Remainder Theorem. The method of denominator selection applied here is just like He Chengtian's method. Multiply the bigger number n of rifa and the numerator of the bigger ratio, multiply the smaller number m of rifa and the numerator of the smaller ratio, add these two products together, the sum is the numerator* (*shuoyu*) (*of the fractional part of the synodic month*).

Qin's algorithm is a method to determine the superior epoch[3] by solving a set of linear congruence equations. However, how to operate the method of denominator selection is an unsolved problem. According to the expression of the algorithm by Qin Jiushao (Chen 1984, 470–474), the above algorithm operates roughly as follows:

According to He Chengtian's algorithm, select a proper denominator (*rifa*) A and obtain the indeterminate Eq. 12.7. Solve it, obtain the smaller number m and bigger number n, multiply the smaller ratio 9 and bigger ratio 26 in the Eq. 12.6, respectively, and add them together to obtain the numerator of the fractional part of the constant of the synodic month:

$$B = 9m + 26n.$$

Thus, it can be seen that He Chengtian's method of denominator selection is to transform the selection of the constant of the synodic month B/A into the solution of the indeterminate Eq. 12.7. This method is completely consistent with the selection

[3] The Superior Epoch is an ideal epoch which is a special moment when the sun passes through the point of winter solstice, and five planets gather in the same longitude line with the sun. It is at a mid-night.

of the length of the solar shadow at 24 solar terms in the *Yuanjia li*, which is reduced to solving the indeterminate Eq. 12.5.

Transforming the selection of astronomical constants into the solution of an indeterminate equation is also the transformation of an astronomical problem into a simple mathematical problem by mathematical modeling.

That's what He Chengtian did!

It's generally accepted that Chinese astrologers in old time were all pragmatists who put precision first, lacking the construction of theoretical models and the application of mathematical means. In reality, He Chengtian's case shows how to determine the astronomical constants by establishing the mathematical model, by constructing and solving a linear indeterminate equation. This kind of process reveals the sanctity and mystery of calendar-making.

If we did not study the method behind the astronomical constants by focusing on *mathematical practice*, by the means of the paradigm of *recovery*, we would not know Chinese mathematicians' special requirements for the analysis of linear indeterminate problems. Furthermore, it's impossible to understand why some achievements like the *Chinese Remainder Theorem* first appeared in China.

This is the significance of constructing of the solar shadow table in the *Yuanjia li* by the paradigm of *recovery*. In this way, we not only come up with new questions, but also come to understand the ancients' mode and process of knowledge creation more deeply.

12.7 Conclusion

In this paper, we reconstruct He Chengtian's algorithm to calculate the constants in the table of solar shadow in the *Yuanjia li*, and obtain his creation process of this table, by the paradigm of *recovery*, in the case that only the table is recorded without any description of the construction process. The result in this paper shows that, some basic data in a traditional Chinese calendar may surprisingly not come from practical measurement or simple adjustment of measured data, but derive from some interesting mathematical method and artificial model.

He Chengtian's era inherited the digital mysticism from the Qin and Han Dynasties (from the second century BC onwards). In some sense, it shows the pursuit of the integers contains prime factors. The recovery of the construction process of the solar shadow table in the *Yuanjia li* further confirms this opinion. This is very significant.

With the paradigm of *recovery*, the model construction and mathematical derivation are very interesting. In fact, the solution of this kind of problem is not only based on limited written records, such as the table of solar shadow in the *Yuanjia li*, but also on He Chengtian's other mathematical works, his social status and so on. So, it's different from the traditional research paradigm characterized by *discovery*. In other words, with the paradigm of *discovery*, we focus on "what" are the numbers in the table of solar shadow in the *Yuanjia li*, while with the paradigm of *recovery*, we focus on "how" to obtain the constants in the table. This deepens our understanding of the ancients' knowledge-creation process.

If we don't accept the paradigm of *recovery*, it's hard to come up with this sort of problems. In this way, accepting the paradigm of *recovery* will help us to update and enlarge the problem field of history of mathematics. At the same time, the paradigm of *recovery* is consistent with the international trend of study of *mathematical practice* on the history of exact science. Therefore, introducing *recovery* and the method of studying *mathematical practice* into research of the history of exact science is in line with the international mainstream.

Acknowledgement The author is grateful to the valuable comments and suggestions of Catherine Goldstein, Christopher Hollings, and Xi Liu, who polished the article for several times.

References

Bowers, J., P. Cobb, and K. McClain. 1999. The Evolution of Mathematical Practices: A Case Study. *Cognition and Instruction* 17 (1): 25–64.

Chemla, K. 2013. Shedding Some Light on a Possible Origin of Concept Of Fractions in China. *Sudhoffs Archiv Zeitschrift für Wissenschaftsgeschichte* 97 (2): 174–198.

Chen, J.J. 陳久金. 1984. A Research of the Method of Denominator Selection 調日法研究. *Studies in the History of Natural Sciences* 3 (3): 245–250.

Chen, M.D. 陳美東. 2003 *A History of Chinese Science and Technology (Astronomy)*中國科學技術史(天文學卷). Beijing: Science Press.

Ferreirós, J. 2016. *Mathematical Knowledge and the Interplay of Practices*. Princeton: Princeton University Press.

Fraser, C. 1999. Philosophy of Mathematics and Mathematical Practice in the Seventeenth Century (Book Review). *Notre Dame Journal of Formal Logic* 40 (3): 447–454.

Li, J.M. 李繼閔. 2007. *The Origin and Development of Algorithm*算法的源流. Beijing: Science Press.

Mancosu, P., ed. 2008. *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press.

Qu, A.J. 曲安京. 2002. Third Approach to the History of Mathematics in China. In *Proceedings of the International Congress of Mathematicians (III)*, ed. LI Tatsien (LI Daqian), 947–958. Beijing: Higher Education Press.

———. 2018. A New Approach to the History of Modern Mathematics – A Case Study on the Theories of Algebraic Equation of Lagrange and Gauss. *Studies in Philosophy of Science and Technology* 35 (6): 67–85.

Ritter, J. 1989. Babylone–1800. In *Eléments d'histoire des sciences*, ed. Michel Serres, 33–61. Paris: Bordas.

Shapiro, S. 1985. Second–Order Languages and Mathematical Practice. *Journal of Symbolic Logic* 50 (3): 714–742.

Shen, Y. 沈約. 1739. *The Book of Song*宋書100卷(卷十三曆志). Block–printed edition in the 4th year of Qianlong清乾隆四年刻本.

Soler, L., S. Zwart, M. Lynch, and V. Israel-Jost, eds. 2014. *Science After the Practice Turn in the Philosophy, History, and Social Studies of Science*. New York: Routledge.

Wu, W.T. 吳文俊. 1986. Recent Studies of the History of Chinese Mathematics. In *Proceedings of the International Congress of Mathematicians*, ed. Andrew M. Gleason, 1657–1667. Providence: American Mathematical Society.

Zhonghua Book Company 中華書局. ed. 1976a. The Compilation of Chapters of Astronomy, Music and Calendar in the Official History of China 5 歷代天文律曆等志匯編(五). Beijing: Zhonghua Book Company.

———. ed. 1976b. The Compilation of Chapters of Astronomy, Music and Calendar in the Official History of China 6 歷代天文律曆等志匯編(六). Beijing: Zhonghua Book Company.

———. ed. 1976c. The Compilation of Chapters of Astronomy, Music and Calendar in the Official History of China 8 歷代天文律曆等志匯編(八). Beijing: Zhonghua Book Company.

Chapter 13
On "Space" and "Geometry" in the Nineteenth Century

Jesper Lützen

Abstract What did mathematicians mean by the words "space" and "geometry" in the nineteenth century? This chapter will try to answer this question, starting with an analysis of Hertz's, Lipschitz' and Darboux's geometrization of mechanics and continuing with a discussion of the use of the words by mathematicians who are usually credited as the principal inventors of non-Euclidean and higher dimensional geometries. The conclusion is that most mathematicians prior to 1880 used the words to denote (intuited) physical space and the geometry describing that space. This is the background against which one should evaluate the sometimes confusing nineteenth century discussions about the existence of geometries other than Euclid's geometry. The question was radically changed with the advent of modernist structuralist mathematics, as described in (Gray, Plato's Ghost: the modernist transformation of mathematics. Princeton University Press, Princeton, 1994).

13.1 Introduction

During the nineteenth century, geometry underwent revolutionary changes. The developments leading to these changes have been described and analyzed by Jeremy Gray in many well-written, informative and thoughtful books and papers. While *Ideas of Space* (Gray 1989) deal with the emergence of non-Euclidean or hyperbolic geometry, *Worlds Out of Nothing* also deals with the development of projective geometry. The two novel geometries were considered rather differently from a philosophical point of view. Projective geometry was considered as an extension of classical Euclidean geometry, while non-Euclidean geometry as well as geometry of 4- and higher dimensional spaces were considered as alternatives to 2- and 3-dimensional Euclidean geometry. In some of my earlier publications

J. Lützen (✉)
Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark
e-mail: lutzen@math.ku.dk

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_13

I have discussed the "application" of high-dimensional Riemannian differential geometry to mechanics by nineteenth century authors such as Rudolf Lipschitz, Gaston Darboux and Heinrich Hertz. For example, I wrote about (Darboux 1888):

> With this approach Darboux had made dynamics an integral part of n-dimensional Riemannian geometry, at least insofar as only trajectories with the same constant energy are concerned. (Lützen 1995, 47)

I still think this is an accurate mathematical description of Darboux's approach to mechanics, as well as that of Lipschitz and Hertz. However, in this paper I want to point out that the actors themselves considered the relation between geometry and mechanics in a different way. For them the formalism they developed for dealing with mechanics was not an *application* of high dimensional geometry. Rather they considered their treatment of mechanics to be *analogous* to the new geometric methods of high-dimensional non-Euclidean geometry. The use of the concept of analogy rather than the concept of application circumvented the problem that these authors considered geometry as a science of space, and considered space to be physical space. Configurations of a mechanical system, on the other hand dealt with something different. The ontology of the two domains was different and thus the one could not be an application of the other.

This illustrates that in order to understand the nineteenth century ideas about the connections between geometry and mechanics it is important to keep in mind that the nineteenth century understanding of the words "space" and "geometry" was different from our modern very broad meaning attributed to these terms. Today we speak of all sorts of spaces: vector spaces, Hilbert spaces, topological spaces etc. and many different kinds of geometry such as hyperbolic and elliptic geometry, *n*-dimensional geometry, algebraic geometry, non-commutative geometry etc. Nineteenth century mathematicians had a much more limited idea of what "space" and "geometry" meant.

In this paper, I shall begin by discussing the above-mentioned connections between geometry and mechanics in order to investigate what they can teach us about the meaning of the words "space" and "geometry" in the nineteenth century. After that I shall turn to some of the mathematicians who are famous for their introduction of alternative geometries and investigate what they meant by these words. I shall conclude that most of them also reserved these words to refer to physical space although they had differing views on the necessary, a priori, empiric or conventional nature of our knowledge of this space. At the end I shall discuss how David Hilbert's axiomatic formalistic approach to mathematics, distinguished between the problem of the nature of physical space and the admissibility of a geometry from a mathematical point of view.

13.2 Heinrich Hertz's Geometry of Systems of Points

When I wrote my book (Lützen 2005) on Heinrich *Hertz's Principles of Mechanics* (Hertz 1894), I was struck by Hertz's ambivalent view of the recent developments of geometry. On the one hand, he described the motion of a mechanical system of a finite number of mass points in terms of what he called a "geometry of systems of points". Today we use the term configuration space. We consider it as a high dimensional space described by a non-Euclidean Riemannian metric. In this geometry, the trajectory of a mechanical system is a geodesic on a particular submanifold, at least if all connections in the system are holonomic. Indeed, physicists have learned to appreciate this geometric approach to mechanics from Hertz's book, or rather from the simplified version of it published by Hendrik Antoon Lorentz 8 years later (Lorentz 1902). On the other hand, in the same book, Hertz described such high dimensional spaces as perplexing, unnatural and supra-sensible.

In 1877 when he was still a student Hertz wrote to his parents:

> The entire new mathematics (from about 1830 on) is, I think, of no great value to the physicist, however beautiful it may be intrinsically, for I find it so abstract, at least in parts, that it no longer has anything in common with reality; for instance, the non-Euclidean geometry, which is based on the assumption that the sum of the angles in a triangle need not be always equal to 2 right angles, or the geometry dealing with space of four, five, or more dimensions etc. Even the elliptical functions are, I think, of no practical value. But perhaps I am mistaken. (Hertz 1977, 71–72)

It is not surprising that the young Hertz found no application for the new mathematical ideas. What is surprising is that he continued his negative evaluation of the new geometries even after he had used them in his *Principles of Mechanics*. In order to understand that, we need to consider how Hertz in the introduction to his book, defended that he had cast his mechanics in a geometric form:

> . . . the geometry of systems of points. The development of this geometry has a peculiar mathematical attraction; but we only pursue it as far as is required for the immediate purpose of applying it to physics. A system of n points presents a $3n$-manifold of motion . . . . Hence, there arise many analogies with the geometry of space of many dimensions; and these in part extend so far that the same propositions and notations can apply to both. But we must note that these analogies are only formal, and that, although they occasionally have an unusual appearance, our considerations refer without exception to concrete images of space as perceived by our senses. Hence all our statements represent possible experiences; if necessary they could be confirmed by direct experiments . . . . Thus we need not fear the objection that in building up a science dependent upon experience, we have gone outside the world of experience. (Hertz 1894, 36/30)

Concerning the Hamilton-Jacobi formalism Hertz wrote:

> It has long since been remarked by mathematicians that Hamilton's method contains purely geometrical truths, and that a peculiar mode of expression, suitable to it, is required in order to express these clearly. But this fact has only come to light in a somewhat perplexing (verwirrender) form, namely in the analogies between ordinary mechanics and the geometry of space of many dimensions, which have been discovered by following out Hamilton's thoughts. Our mode of expression gives a simple and intelligible explanation of these analogies. It allows us to take advantage of them, and at the same time it avoids the unnatural admixture of supra-sensible abstractions with a branch of physics. (Hertz 1894, 39/32–33)

So apparently, the reason why Hertz objected to the new geometries of the mathematicians was that they were supra-sensible in the sense that they lay outside the world of experience. Indeed, according to Hertz's introduction of "geometry" in his *Principles of Mechanics*, "space" is Euclidean and 3-dimensional. He introduced the notion of space and geometry in two steps. The first one can be found in the beginning of the first book (the *Principles of Mechanics* is divided into two "books"):

> The subject-matter of the first book is completely independent of experience. All the assertions made are *a priori* in Kant's sense. They are based upon the laws of the internal intuition of, and upon the logical forms followed by, the person who makes the assertions; with his external experience they have no other connection than these intuitions and forms may have. (Hertz 1894, §1)

In particular Hertz wrote about space:

> The space of the first book is space as we conceive it (der Raum unserer Vorstellung). It is therefore the space of Euclid's geometry, with all the properties which this geometry ascribes to it. (Hertz 1894, §2)

Here Hertz followed Immanuel Kant who in the *Critique of Pure Reason* had argued, that one could not sense space. On the contrary, we order all our sensations in space. Thus, space must necessarily come before any sensation (it is a priori). According to Kant, our a-priori knowledge of space is constructed in our intuition.[1] To exemplify this, Kant argued that we do not have to make any measurement in order to know that the angle sum of a triangle is equal to two right angles. Indeed, by following Euclid's arguments and constructions, not on paper, but in our intuition, we can establish this theorem a-priori (see Friedman 1985).

Such a Kantian view of space was not uncommon in 1894 but it is surprising to find it expressed by Hertz after he had worked for 3 years in Hermann von Helmholtz's laboratory in Berlin. Indeed, Helmholtz was famous for his successful popularization of non-Euclidean and Riemannian geometry. In most of his writings on the question, Helmholtz argued that the nature of space was an empirical question. In a few places, he expressed ideas that were closer to Henri Poincaré's later conventionalism, but in any case, he clearly rejected that the nature of space could be determined a priori, as Kant had claimed.

Hertz's discussion of the nature of space in the second book of his *Principles of Mechanics* is more in tune with the empiricist views of his mentor:

> In this second book we shall understand times, spaces and masses to be symbols of objects of external experience; ... Our statements concerning the relations between times, spaces and masses must therefore also be in accordance with possible, and in particular, future

[1] Kant described our intuition of space as an external intuition whereas time according to him is an internal intuition. As we can see in the above quote, Hertz did not make such fine distinctions. In general when I refer to Kantian views, I refer to the more or less simplified Kant-like ideas that were around in the nineteenth century according to which geometry is a priori, synthetic and constructed in our intuition.

experiences. These statements are based, therefore, not only on the laws of our intuition and thought, but in addition on experience. (Hertz 1894, §296)

Hertz claimed that if we measure lengths according to the methods of practical geometry by way of a scale, we know by experience that Euclidean geometry always gives the correct results (Hertz 1894, §299). I shall not go into the difficult discussion of the mixture of Kantian and empiricist ideas in these quotes (See Lützen 2005, 127–133). What is important here is that for Hertz, “space” meant physical space and this space is according to Hertz described accurately by Euclidean geometry. Therefore, for Hertz any talk about a high dimensional non-Euclidean *space* is a perplexing, unnatural and supra-sensible phantasy that is not in accordance with the true nature of space.

However, according to Hertz that is not a problem, because his own geometry of system of points is in accordance with experience. It does not deal with points in a high dimensional space but with configurations of a system of points in ordinary Euclidean space. In other words, the ontology of the objects in his geometry is entirely different from those of the objectionable objects in a high dimensional geometry. To be sure, Hertz developed a geometric vocabulary for mechanical systems and their trajectories such that the same propositions apply both to his geometry of systems of points and to the high dimensional phantasies of the mathematicians. But that does not make his geometry of systems of points an *application* or a *special case* of the high dimensional spaces of the mathematicians.

Instead, as can be seen in the quotes above, Hertz considered the relation between the two geometries as an *analogy*. Here Hertz seem to use the word analogy in the same way it was used by William Thomson (Lord Kelvin) and other mathematical physicists at the time: Two theories about entirely different things are analogous if they can be described by the same mathematical analytical formalism. For example, Thomson (and Michel Chasles before him) noticed that potentials and stationary heat conduction were both described by the same differential equation, the Laplace equation (See Knudsen 1985). This allowed him to transfer theorems, such as the maximum principle from one domain (heat conduction) where they were physically obvious to the other domain (potential theory) where they were less obvious.

Hertz was in a similar situation. His analytical formalism of differential forms could describe both his physically real systems of points as well as the mathematical phantasy of non-Euclidean spaces of many dimensions. And if we generalize our intuitions about 3-dimensional space to higher dimensions, we can obtain a geometric intuition about the motion of mechanical systems. In particular, it allows us to see that the Hamilton-Jacobi formalism is analogous to the geometric methods used by Carl Friedrich Gauss when he studied geodesics on surfaces.

13.3 Darboux's Geometrization of Mechanics

One might think that this ambivalent view of the status of high dimensional and curved spaces was peculiar to Hertz. After all, he was not a mathematician but a physicist.[2] So let us now turn to the views expressed explicitly or implicitly by nineteenth century mathematicians and let us begin with mathematicians who also geometrized mechanics. In particular, let us begin with Lipschitz and Darboux, to whom Hertz referred explicitly in the preface to his book. According to his own testimony (which is corroborated by his early drafts of his book), Hertz did not know of the works of these two mathematicians when he began working on his geometry of systems of points, But he later found them "very suggestive" (Hertz 1894, Preface).

From a physical point of view, there is a great difference between Hertz's approach to mechanics and the approach of his two predecessors. In Hertz's image of mechanics, force and potential energy were just epiphenomena resulting from rigid connections (constraints) between the visible system and a hidden system of masses performing adiabatic cyclic motions. In this way, Hertz got rid of the notion of force as a fundamental concept, a concept that he considered physically and logically problematic. This elimination of forces is the main physical message of Hertz's book. His geometric theory of systems of points changed the mathematical form of the presentation, but was logically independent of his new explanation of interactions in the system.

For Lipschitz and Darboux, on the other hand, the main objective of their geometrizations of mechanics was to show how one can include the forces (or the potential) into the geometry, in such a way that the motion of the mechanical system can be described as a geodesic motion along a shortest path in the new geometry. Thus, the geometrization carried a more fundamental burden than it did in Hertz's image. However, where Hertz emphasized the physical importance of his new explanation of forces, Darboux and Lipschitz only considered their geometrical presentation of mechanics as a mathematical trick that could enhance our intuition about the formalism. In the preface to the second volume of his *Lectures on the General Theory of Surfaces* Darboux wrote:

> In particular, I have stressed the connection one can find here between the methods employed by Gauss in the study of geodesics and those that Jacobi later applied to the problems of analytical mechanics. In this way, I have been able to show the great interest of Jacobi's beautiful discoveries when those are considered from a geometrical point of view. (Darboux 1888, Preface)

Darboux's and Lipschitz's geometrization of mechanics is based on the principle of least action which in Jacobi's formulation says: Given a conservative mechanical system with total energy h, whose configuration can be described by n independent

[2] In a letter to his parents Hertz called his study of the principles of mechanics "this mathematical work".

generalized coordinates q_i in terms of which the kinetic energy T of the system can be expressed by the quadratic form

$$T = \frac{1}{2} \sum_{i,j} a_{i,j} \dot{q}_i \dot{q}_j.$$

Moreover let U denote the potential energy as a function of the generalized coordinates and consider the quadratic differential form

$$ds^2 = (h - U) \sum_{i,j} a_{ij} dq_i dq_j. \tag{13.1}$$

Then the trajectory followed by the mechanical system between two configurations A and B will at least locally minimize the action integral

$$\int_A^B ds \tag{13.2}$$

among all "curves" between A and B.

Formulated in a modern way the system will follow a geodesic in the n-dimensional Riemannian manifold defined by the Riemannian metric (13.1). Darboux formulated this insight differently:

> The general problem of mechanics is nothing but the generalization to an arbitrary number of variables of the problem of the study of geodesics. (Darboux 1888, 500)

We notice, that he did not speak of an n-dimensional geometry or space, but of "an arbitrary number of variables". Indeed, he never introduced general n-dimensional Riemannian geometry in his great four volume lectures. This is surprising. Indeed, he knew of Bernhard Riemann's Habilitationsvortrag as well as Beltrami's development of the analytical formalism for Riemannian manifolds. One would think that it would be natural for Darboux to generalize his investigations of 2- and 3-dimensional differential geometry to higher dimensional Riemannian spaces, and then *apply* this formalism to mechanics. However, as it were, Darboux only introduced a concrete theory of the n free coordinates of a mechanical system. In this way he circumvented the problem of the existence of an n-dimensional space and the nature of objects in it. In this sense, his approach was similar to Hertz's later approach. And like Hertz, Darboux saw the connection between geometry and mechanics as an analogy.

This analogy was presented in a very pedagogical way in three consecutive chapters of his book. The first one (Chapter VI) was entitled: "Analogy between dynamics of motions in a plane and the theory of geodesic lines". In this chapter, the mechanical system had only two degrees of freedom, and thus he could without problem interpret formula (13.1) in two variables as a line element on a surface. The subject of the next chapter was entitled: "Application of the previous methods to the study of motions in space". Here the configurations were three dimensional and thus they could be described in ordinary geometric terms. However, Darboux consciously did not interpret the three dimensional quadratic form (13.1) as the line element in a (non-Euclidean) space of three dimensions. But the analytic formalism carried over from the two-dimensional case.

Finally, in chapter VIII, Darboux turned to "The general problem of dynamics". Here he dealt with a general conservative system of an arbitrary number of points. First, he generalized the analytic formalism developed in the previous chapter. However, having derived the general principle of least action in the form presented above, and having emphasized that the problem of general dynamics is therefore a generalization of the problem of geodesics, to an arbitrary number of variables (see quote above) he began introducing some geometric notions in this general setting. He defined the angle between two directions given by two infinitesimal displacements of the system, a notion of a line (curve) and a (hyper) surface, and he attributed the name "geodesic" to a curve that minimizes the integral (13.2) where ds is given by formula (13.1). In this way, he could formulate the Jacobi theory in a beautiful geometric form that showed the analogy with Gauss's theory of geodesics on a surface.

We notice, that in this way, Darboux avoided any consideration of the ontology of general curved spaces and spaces with more than three dimensions. As in Hertz's later mechanics, the ontology in his presentation was unproblematic. The theory dealt with mechanical systems of any number of points not with points in a high dimensional curved space. It should also be emphasized, that Darboux did not use the word "point" for a configuration of a mechanical system with more than one mass point. He used words like "system of values" or "elements" or "positions".

13.4 Lipschitz's Geometrization of Mechanics

Darboux's treatment of the analogy between Gauss's theory of surfaces and the Hamilton-Jacobi formalism of mechanics was a simplified and pedagogically improved version of a treatment of the same subject by Lipschitz (1872) to whom Darboux referred. Like Darboux, Lipschitz in general used the word "space" (Raum) to refer to physical space. However, unlike Darboux, he explicitly expressed the possibility that space could a priori have any number of dimensions and could be equipped with a non-Euclidean metric:

> The new speculations about the nature of space have shown that it is not necessary to assume that the element of a line from a given point is representable by the square root of the sum of the squares of the differentials of suitable coordinates of the point. If one disregards certain conditions that are in fact (thatsächlich) satisfied in real space, it is permissible to assume the line element to be the square root of an arbitrary essentially positive quadratic form, or more generally the p'th root of an essentially positive form of degree p in the differentials of arbitrary coordinates of the point. This more general hypothesis about the nature of space can be adjusted to the concepts of mechanics. (Lipschitz 1872, 116–117)

Lipschitz’s remark about the actual nature of real space shows that he was still convinced that physical space was empirically known to be 3-dimensional and Euclidean, but he was willing to explore a more general assumption. His more general assumption about physical space made his treatment of mechanics more general but also more complicated than Darboux’s later treatment. However, like Darboux, Lipschitz did not use the word “space” for the configuration space of a mechanical system. In most of the paper, Lipschitz only used analytic vocabulary when writing about mechanical systems. Only at the very end, he introduced a few geometric notions, in particular a concept of orthogonality in configuration space that allowed him to formulate the main theorems of the Hamilton-Jacobi formalism in a way that is entirely analogous to Gauss’s theorems concerning geodesics on a surface.

Only once in his paper did Lipschitz refer to an n-dimensional “space”, namely when he referred to a paper by Eugenio Beltrami “The general theory of differential parameters” (1869), from which he borrowed some of the analytic formalism. Having summarized the main ideas of Gauss’s study of geodesics, Lipschitz wrote: “This intuition has been extended by Mr. Beltrami to a space (Raum) of n dimensions” (Lipschitz 1872, 119). Beltrami’s paper did not deal with mechanics but with the transformation of quadratic forms. But like Lipschitz and Darboux after him, Beltrami used some geometric vocabulary. But he felt obliged to defend this use of geometric language also in cases where the number of variables exceeds 3. He did so by referring to a casual remark by Gauss to which we shall return below.

As I have pointed out in previous publications (Lützen 1990, 679–686; 1995, 28–34), Liouville anticipated several aspects of Lipschitz’s and Darboux’s geometric approach to mechanics by several decades. However, Liouville only used geometric language when dealing with situations, that could be described in terms of 3-dimensional Euclidean space and 2-dimensional surfaces embedded in such a space. In the case of mechanical systems with a higher number of degrees of freedom, he only generalized the analytic formalism but did not use any geometric vocabulary. That is not surprising, considering that he conducted his research before Riemann.

Having considered analogies between mechanics and geometry, let us now turn to those authors who are usually regarded as the inventors or propagators of geometries other than two and three dimensional Euclidean geometry.

13.5 Non-Euclidean Geometry

Sometimes the history of non-Euclidean geometry is told as a purely mathematical story of how it was discovered that a particular axiom, the parallel postulate, is independent of the other axioms in the axiom system of geometry. However, as it clearly transpires from Jeremy Gray’s *Ideas of Space* the early pioneers, Gauss, Nikolai Ivanovich Lobachevsky and János Bolyai, considered the question of the parallel postulate as a problem about the nature of the physical space we inhabit. They did not buy Kant’s argument that we construct geometry a priori in our intuition but believed that they could intuit a geometry in which the negation of the parallel postulate holds. They came to the conviction that the truth or falsity of the parallel postulate was an empirical problem. As Gauss formulated it in a letter to Bessel from 1830:

> According to my deepest conviction, the theory of space has a completely different position in our a-priori knowledge from that of the theory of pure quantity; to our knowledge of the latter belongs that complete conviction of its necessity (therefore also of its absolute truth) which is peculiar to the latter; we must humbly admit that if the number is merely a product of our minds, space also has a reality outside of our minds, to which we cannot completely prescribe its laws a priori. (Gauss 1900, 201)

Gauss occasionally in jest expressed the wish that Euclid's geometry "were not true" (Gauss 1900, 169, 187). Thus for Gauss, there is only one space, namely physical space and only one true geometry, namely the geometry of space. However, since the properties of this space and its geometry is not a-priori according to Gauss and since no empirical observation has revealed whether the parallel postulate is true or false, Gauss was willing to consider several a-priori possible geometries.

Lobachevsky went even further in an empirical direction giving new "material" definitions of the elementary objects of geometry like plane, line and point. Moreover, he tested the empirical exactitude of the parallel postulate by astronomical observations establishing that "in triangles whose sides are attainable for our measurement, the sum of the three angles is not indeed different from two right angles by the hundredth part of a second". (Lobachevsky 1840, 45).

A generalization of such ideas can be found in Riemann's famous Habilitationsvortrag. However, before we turn to that very deep analysis of space, we shall have a look at the question of high dimensional space.

13.6 Spaces of More Than 3 Dimensions

The idea of a manifold or a space of more than three dimensions has been the subject of philosophical and mathematical discussion since Aristotle.[3] However, it only became the subject of real mathematical investigation during the nineteenth century. During the second half of that century it even enjoyed some popularity among amateur mathematicians and lay people. The fascinating story of geometry of 4- and higher dimension has been the subject of several recent books (e.g. Volkert 2018; Throesch 2017)[4]. One of the first introductions and investigations of linear n-dimensional manifolds was Die *Lineale Ausdehnungslehre* (the linear theory of extension) published by Hermann Grassmann in 1844. In the introduction to that book, Grassmann explained the historical development of his ideas: Having

[3] For a brief introduction to the subject see (Cajori 1926).

[4] Also Richards (1988) deals with this subject in the English setting where it enjoyed special popularity.

initially introduced certain algebraic ideas about geometry, he became acquainted with August Ferdinand Möbius’ Barycentric calculus. Möbius had contemplated the reason why in space there are symmetrical figures that cannot be brought into coincidence. He had concluded that “The reason may be looked for in this, that beyond the solid space of three dimensions there is no other, none of four dimensions” (Möbius 1827, quoted from Smith 1959, 526).

Grassmann then began to systematize his ideas and then

> it turned out that the analysis I had found did not, as I had believed at first, deal exclusively with the field of geometry; rather, I soon realized that I had entered a new science of which geometry itself was only a special application. In fact I had for some time been aware that contrary to arithmetic and combinatorics, geometry could by no means be considered as a branch of mathematics since it already refers to something given in nature (namely space), and that there must therefore be a branch of mathematics, that in a purely abstract way produce laws similar to those that in geometry appear to be bound to space. (Grassmann 1844, Vorrede)

Thus, Grassmann also considered the science of space and geometry to be a natural science. His new theory of extension, on the other hand, was a purely analytical mathematical theory that was not hampered by the limitations of space. In particular, the limitations to three dimensions disappeared in his linear theory of extension.

The second extensive treatment of n-dimensional geometry was written by Ludwig Schläfli in 1850–1852. However, his so-called theory of multiple continuity remained unpublished until 1901. When he presented his long manuscript to the Imperial Academy in Vienna in 1852 he characterized it as a

> new branch of analysis ... which is also an analytic geometry of n dimensions that includes the plane and space as special cases for $n = 2, 3$. (Schläfli 1852, 171)

In the rest of the book, the word “space” is reserved for 3-dimensional Euclidean space. In the general case for arbitrary n, he used phrases like “continuum with dimension number n”. In the introduction, Schläfli reminded the reader of the benefits that can be gained when one considers a problem with two independent variables both from a geometric and from a purely analytical point of view. However, he pointed out, when the number of variables exceed three, the useful geometric intuition and expressions are no longer available. His new theory of multiple continuity was intended to supply an analytic theory that would still provide an intuition and a geometrical vocabulary in the event of more than three variables.

In the main part of the book Schläfli worked in an entirely analytic way, but by introducing geometric names that in dimension 2 and 3 corresponded with the usual names in 3-space, he was able to express many analytic theorems in a geometric way. In particular, he described all regular polytopes in 4-space. He published a few of the results from his book in separate papers, but since the results were expressed in purely analytic ways, the reader might easily have missed the novel geometric aspect.

13.7 Bernhard Riemann

Riemannian geometry. This is the modern designation for a general type of geometry on a space of n dimensions where distances are determined by a metric expressed as a quadratic differential form. Such spaces were introduced by Bernhard Riemann in his visionary Habilitation lecture in 1853. However, Riemann himself reserved the words geometry and space for physical space. As the title of his manuscript indicates, its aim was to investigate the "Hypotheses which lie at the Foundation of Geometry". According to Riemann, this question had been left in the dark despite many attempts by philosopher and mathematicians to settle it.

> The reason of this lies perhaps in the fact that the general concept of multiple extended magnitudes, in which spacial magnitudes are comprehended, has not been elaborated at all. Accordingly I have proposed to myself at first the problem of constructing the concept of a multiply extended magnitude out of general notions of quantity. From this it will result that a multiply extended magnitude is susceptible of various metric relations and that space accordingly constitutes only a particular case pf a triply extended magnitude. A necessary sequel of this is that the propositions of geometry are not derivable from general concepts of quantity, but that those properties by which space is distinguished from other conceivable triply connected magnitudes can be gathered only from experience. There arises from this the problem of searching out the simplest facts by which the metric relations of space can be determined, a problem which in nature of things is not quite definite; for several systems of simple facts can be stated which would suffice for determining the metric relations of space; the most important for present purposes is that laid down for foundations by Euclid. These facts are like all facts not necessary but of a merely empirical certainty; they are hypotheses; one may therefore inquire into their probability, which is truly very great within the bounds of observation, and thereafter decide concerning the admissibility of protracting them outside the limits of observation, not only toward the immeasurably large, but also toward the immeasurably small. (Riemann 1854, quoted from Smith 1959, 411–412)

Thus, as for Gauss and Lobachevsky, "space" for Riemann means physical space and geometry describes the metric relations of this space. And like Gauss and Lobachevsky, Riemann was willing to entertain the possibility that Euclidean geometry does not give the best description of physical space. Indeed, compared to Gauss and Lobachevsky, Riemann extended the a priori possible manifolds to manifolds with other dimensions than three and to manifolds that are described by a Riemannian metric that has varying curvature.

He argued that the hypotheses concerning relations of extension (topological properties) such as the number of dimensions and the unlimitedness of space could be decided with a greater certainty than the metric properties of space such as its curvature and its infinite extent. Since Gauss and Lobachevsky assumed the validity of the Euclidean theorems (implicit axioms) of congruence, (in Riemannian terms the constant curvature of space) their empirical tests of the nature of space involved astronomically large figures. Riemann, on the other hand was more interested in the immeasurably small. His general Riemannian manifolds could have a curvature that varied over immeasurably small distances in such a way that the curvature of a measurable portion of space would average out to zero within the accuracy of measurement.

> Now however the empirical notions on which spatial measurements are based appear to lose their validity, when applied to the indefinitely small, namely the concept of a fixed body and that of a light-ray; accordingly it is entirely conceivable that in the indefinitely small the spatial relations of size are not in accord with the postulates of geometry, and one would indeed be forced to this assumption as soon as it would permit a simpler explanation of the phenomena. (Riemann 1854, 424)

In the above quote, "postulates of geometry" seem to refer to Euclid's postulates. Otherwise Riemann used the words "geometry" and "space" to refer to physical space and the geometry (Euclidean or otherwise) that best describe it. In the Habilitationsvortrag, Riemann never used the words space and geometry to denote general n-dimensional Riemannian manifolds. They are called multiply extended magnitudes or n-fold extended magnitudes or Manifolds (Manigfaltigkeiten) of n dimensions.

Only in his prize essay on heat conduction (Riemann 1861) written 7 years later, Riemann in one place wrote about a "general *space* of n dimensions" (Riemann 1861, 403). He introduced this concept in order to give a "geometric interpretation" of his analytical formulas. Although this geometric interpretation according to Riemann transcends our intuition and is based on unusual concepts, he found it useful to point it out in passing. As far as I know, this is the only occurrence in Riemann's works where the words space and geometry represent something other than physical space and its geometry. He may have taken the liberty of using the words in a more generalized sense in his prize essay because the paper does not deal with the nature of physical space and therefore does not create confusion about the meaning of the word.

13.8 Beltrami's Real Substrate

Beltrami is often given the honor of having proved the consistency of non-Euclidean geometry and thus the impossibility of proving the parallel postulate. This is true in the following sense: On the basis of Gauss' intrinsic theory of surfaces, Beltrami showed in 1868 that the geometry on a surface with constant negative Gauss curvature is the same as Lobachevskian non-Euclidean geometry. More specifically, he showed that if we measure lengths on the surface with the metric that the surface inherits from the ambient 3-dimensional Euclidean space and if we let geodesics play the role of straight lines, then all of Lobachevsky's theorems of non-Euclidean geometry hold true. In modern terms: A surface with constant negative curvature is a model of non-Euclidean geometry.

From Beltrami's observation one can argue as follows: If there were a contradiction in non-Euclidean geometry, this would turn up as a contradiction about the geometry of the surface, i.e. as a contradiction in Euclidean geometry. So if Euclidean geometry is consistent, non-Euclidean geometry must be consistent as well. In other words, non-Euclidean geometry is consistent relative to Euclidean geometry. Since the consistency of Euclidean geometry was evident to Beltrami and his contemporaries, this proved that non-Euclidean geometry is consistent. In particular, it is impossible to prove the parallel postulate from the remaining axioms common to both Euclidean and non-Euclidean geometry, because it contradicts the negation of that postulate which is assumed as an axiom in non-Euclidean geometry.

With this argument in mind, it is surprising to read what Beltrami wrote to Genocchi 1 year after having published his two famous papers on non-Euclidean geometry:

> Nevertheless, I want to declare that I am not even yet persuaded about the impossibility of proving the Euclidean geometry and I hope that no passage in my writings have been formulated in a way that leads one to suppose the opposite. (Voelke, 2005)

Indeed in his two papers he did not discuss the question of consistency of non-Euclidean geometry or the improvability of the parallel postulate. What then was the aim of Beltrami's paper? He explained it as follows:

> We have thought, to the extent of our ability, to convince ourselves of the results of Lobachevsky's doctrine; then following the tradition of scientific research, we have tried to find a real substrate for this doctrine, rather than admit the necessity for a new order of entities and concepts. We believe we have attained this goal for the planar part of the doctrine, but we believe that it is impossible to proceed further. (Beltrami 1868a, 7)

Thus, the problem addressed by Beltrami was not a problem of consistency or provability, but an ontological question about finding a real substrate, i.e. real things that are correctly described by the theorems of non-Euclidean geometry. But what did he mean by "real"? This becomes clear in the second of his papers on non-Euclidean geometry, where he described a model of 3- or even n-dimensional non-Euclidean geometry, in the spirit of Riemann, whose Habilitationsvortrag he had now been acquainted with. In this case he did not claim that his model was a real substrate:

> Thus, all the non-Euclidean concepts find a perfect correspondent in the geometry of space of constant negative curvature. One only has to observe, that where the concepts of planimetry receives a true and proper interpretation because they are constructible on a real surface, those that embrace three dimensions are only susceptible of an analytic representation, because the space in which such a representation can be realized is different from the one we usually call *space*. (Beltrami 1868b)

Thus, for Beltrami a model is only real, if it is realized in what we usually call space. By that he clearly meant physical space that he seemingly assumed to be Euclidean. This view of the ontology of non-Euclidean geometry is very different from that of Gauss, Lobachevsky and Riemann. For the latter, the problem of a real substrate did not arise, or rather it was trivial. They thought that the usual points, lines planes etc. of physical space might satisfy the axioms or hypotheses of non-Euclidean geometry. These objects were considered real in themselves and so it was not necessary to "admit the necessity for a new order of entities and concepts" as Beltrami expressed it.

Beltrami, on the other hand, seems to have believed that physical space was Euclidean. On the other hand he believed that non-Euclidean geometry was a valuable new "doctrine" and apparently he also believed that such a doctrine should deal with something real. This real substrate he found in the geometry of a real surface of constant negative curvature. Non-Euclidean 3-space on the contrary, could be described by a Riemannian manifold, but that was not a real substrate, because it could not be imbedded in physical (i.e. Euclidean) 3-space.

As shown by Voelke (2005), it was Jules Houël who convinced Beltrami that his model in fact demonstrated the impossibility of proving the parallel postulate and thus that non-Euclidean geometry was consistent. The argument was subsequently popularized by Poincaré who discovered another model of non-Euclidean geometry about a decade after Beltrami.

13.9 Discussions About the Existence of Hyperspace

During the last three decades of the nineteenth century, high-dimensional and non-Euclidean geometries became rather popular among mathematicians, both first rate mathematicians like Poincaré, Camille Jordan, Arthur Cayley, James Joseph Sylvester and Felix Klein as well as a great number of minor figures. Four- and higher-dimensional spaces and geometric notions in such spaces were introduced analytically. The use of geometric vocabulary was mostly defended as a suggestive intuitive tool. For example Jordan (1872, 1875) illustrated the procedures of linear algebra in terms of *n*-dimensional geometry considering “a point as defined in the space of *n* dimensions by the values of *n* coordinates”. (Jordan 1872, 50). Other less famous mathematicians studied 4- dimensional geometry and in particular polytopes in such spaces. But alongside such no-nonsense use of higher dimensional geometry there arose a more popular and more speculative literature about the existence of *n*-dimensional hyperspaces including curved (non-Euclidean) spaces.

These discussions were marred by a lack of agreement about what it means for such a space to exist. Many actors ascribed to the empirical point of view, others revived Kant’s ideas: a space and a geometry exists if it can be intuited. Yet others began to defend the position that one could legitimately study such spaces, if they were only logically consistent. In the twentieth century the latter point of view won the day. However, in order to understand the rather chaotic discussions in the mathematical and in the popular literature at the end of the nineteenth century, it is important to keep these different criteria of existence in mind. It also added to the confusion that all three ideas of existence were unclear: It was clear to most empiricists that empirical space was 3 dimensional and close to Euclidean, but one could speculate about the possibility of new physical phenomena describable by some sort of hyperspace. The Kantian stance begged the question: what can be intuited? Many claimed that only 3-dimensional geometry could be intuited, but other authors claimed (following Helmholtz and Sylvester) that one could train one’s mind to intuit higher- dimensional curved spaces. Moreover, the question of consistency was difficult because until the popularization of Beltrami’s method of models, one had no certain way to argue that a mathematical theory did not contain contradictions.

Finally, several actors ascribed to various mixtures of the notion of existence of a space or a geometry. For example we have seen that Gauss expressed very outspoken empiricist ideas about the validity of the parallel postulate. On the other hand he mentioned an "extension to a geometry of more than 3 dimensions for which we human beings have no intuition (Anschauung) but which abstractly considered is not contradictory and consequently could belong to higher beings" (Gauss 1900, 241–242). This is an interesting combination of the consistency criterion and a Kantian view that 3-dimensionality may be a precondition of our way of perceiving phenomena, but that real world might be higher dimensional. Similarly Helmholtz and Riemann expressed both empiricist and conventionalist points of view.

I shall not go into details with the fascinating story of the gradual acceptance of non-Euclidean and higher dimensional geometries. Interested readers can be referred to the very interesting treatment in Volkert (2013, 2018), Throesch (2017), Voelke (2005) and Richards (1988). I only want to mention two scandals that made hyperspace a dangerous subject for "sober thinkers" as Simon Newcomb put it (see below). The first scandal was caused by the German astrophysicist Karl Friedrich Zöllner in 1878. He had attended a séance with the spiritualist medium Henry Slade who could perform various tricks, that could be explained by assuming that he had access to a fourth dimension. For example he could untie an irreducible trefoil knot. Based on these experiments, Zöllner believed that he had demonstrated that space is four dimensional, a claim he had earlier made on astronomical grounds. This gave rise to a long and bitter debate that disgraced the problem of higher spaces.

The other scandal was created by the British mathematician Charles Howard Hinton who popularized the fourth dimension in his Scientific Romances. The scandal arose when he was convicted of bigamy and his ideas about the fourth dimension were connected to his loose morals and his father's notorious lawbreaking free-love philosophy (Throesch 2017, 32).

It is on this background we should see Hertz's and Darboux's careful way of dealing with hyperspace.

13.10 Newcomb's Address 1897

As an example of the view of higher dimensional and possibly curved spaces at Hertz's time I shall briefly analyze the American mathematician Simon Newcomb's Address *The Philosophy of Hyperspace* given in 1897. It was addressed primarily to philosophers and laymen. Newcomb emphasized that

> The question whether a fourth dimension may possibly exist, and whether it can be legitimately employed for any mathematical purpose, is one on which ideas are not universal. (Newcomb 1897, 187)

Calling attention to the way a student can generalize plane geometry to geometry of 3 dimensions he asked:

> Why should he stop there? You reply, perhaps, because there are only three dimensions in actual space. But in making hypotheses we need not limit ourselves to actualities; we can improve our methods of research, and gain clearer conceptions of the actual by passing outside and considering the possible. (Newcomb 1897, 188)

Recall that in the defense of his geometry of systems of points Hertz had emphasized that "we need not fear the objection that in building up a science dependent upon experience, we have gone outside the world of experience" (Hertz 1894, 36/30). This assurance sounds as a direct response to Newcomb's argument had it not been for the fact that Hertz's book was published 4 years before Newcomb's address. Hertz clearly did not like the idea of using the merely possible to enlighten the actual. And he assured his readers that this was not what happened in his geometry of system of points.

Newcomb then mentioned a favorite argument[5] for introducing a fourth dimension.

> Euclid proves by superposition that the two triangles in a plane having two angles and the included side equal are equal to each other. In the demonstration it is assumed that the triangles can be made congruent by simply placing one upon the other without taking it out of the plane. From this the conclusion is drawn that the same conclusion holds true if one of the triangles be obversed. But in this case they cannot be brought into congruence without taking one of them out of the plane and turning it over. The third dimension is thus assumed in geometry involving only two dimensions.
>
> Now consider the analogous case in space. Two pyramids upon congruent bases may be proved equal by bringing them into congruence with each other. But suppose they differ only in that one is the obverse of the other, so that they could be brought into congruence only by looking at one of them in a mirror and then placing the other into congruence with the image of the first as seen in the mirror. Would we detract from the rigor of the demonstration by assuming the possibility of such an obversion without changing the volume of the pyramid? With a fourth dimension we should have no detraction from rigor. We would simply obvert the pyramid as we would turn over the triangle. (Newcomb 1897, 189–190)

Newcomb then continued to distinguish between the empirical and the Kantian view of existence of four dimensional space:

> The question of the fourth dimension as a reality may be considered from two points of view, its conceivability and its possible objective reality. If by conceivability we mean the power of being imagined in the mind it must be admitted that it is absolutely inconceivable. ... This (our inability to conceive four lines orthogonal to each other) clearly transcends all possibility even of imagination. The fourth dimension in this sense is certainly inconceivable. (Newcomb 1897, 190)

The rejection of the possibility of conceiving the fourth dimension was widely accepted, but as mentioned above it was contested by many for example by G.F. Rodwell:

> Space of four dimensions is transcendental space: It is beyond the limit of our experience, but not beyond the limit of our imagination. (G.F. Rodwell 1873 quoted from Volkert 2018, 27, 37)

[5] As we mentioned above, Möbius and Grassmann also referred to this argument. For a more comprehensive history of it see (Volkert 2018, 1–37).

As far as the objective reality was concerned, Newcomb was less assertive.

> Those who speculate on the possible have taken great pleasure in imagining another universe alongside of our own and yet distinct from it. The mathematician has shown that there is nothing absurd or contradictory in such a supposition. But when we come to the question of physical fact we must admit that there appears to be no evidence of such a universe. If it exists, none of its agencies intrude into our own universe, at least in the opinion of sober thinkers. The intrusion of spirits from without into our world is a favorite idea among primitive men, but tends to die out with enlightenment and civilization. Yet there is nothing self-contradictory or illogical in the supposition. (Newcomb 1897, 190)

Newcomb's allusion to primitive ideas concerning other worlds is phrased in a general way, but he probably also had the Zöllner-Slade and the Hinton scandals in mind. Toward the end of his address, Newcomb mentioned that recent physical phenomena in the molecular range might turn out to be explainable by the assumption of a fourth dimension. Other mathematicians like William Kingdon Clifford had also speculated about the possibility of explaining new physical phenomena by way of curved spaces. Newcomb thus concluded:

> We have no experience of the motion of molecules; therefore we have no right to say that those motions are necessarily confined to three dimensions. Perhaps the phenomena of radiation and electricity might yet be explained by vibrations in a fourth dimension . . . .. We must leave it to our posterity to determine whether, in either way, the hypothesis of hyperspace can be used as an explanation of observed phenomena. (Newcomb 1897, 195)

Although Newcomb ended on an empirical note, we can notice that in the quote above from page 190, he twice mentioned that there is nothing contradictory in the supposition of a four dimensional space and even claimed that mathematicians had proved this consistency. Here he may refer to the usual analytic "models" of n-dimensional linear spaces, as for example described by Jordan, or to Beltrami's model of n-dimensional space with constant negative curvature. However, it is interesting to notice that the question of consistency only comes up here as a necessary property of the possible objective reality of space. The idea seems to be that inconsistent theories of space are not possible. He also vaguely alluded to the idea that consistency implied mathematical existence:

> Our conclusion is that space of four dimensions, with its resulting possibility of an infinite number of universes alongside our own, is a perfectly legitimate mathematical hypothesis. (Newcomb 1897, 191)

13.11 The Axiomatization of Geometry

Newcomb's address illustrates that even during the 1890s there was no clear answer to the question: what does it mean for a geometry to be true. Empirical arguments about the nature of physical space were mixed with philosophical, psychological and even physiological considerations concerning our ability to conceive or intuit space. Logical arguments concerning consistency were also put forward and gradually they became more pronounced. Indeed, as Volkert has pointed out (Volkert 2018, 166), the late nineteenth century debates concerning the nature of space and geometry prepared the ground for a purely axiomatic view of mathematics that we usually associate with Hilbert.

With Hilbert the mathematical idea of geometry became in principle entirely separated from the empirical questions of the nature of physical space and our intuition of it. This clarification and emancipation went hand in hand with Hilbert’s view of the objects of geometry and mathematics in general. Before Hilbert, a system of axioms had been considered as a system of truths about some clearly defined objects (points, lines, planes etc.). However, in his *Grundlagen der Geometrie* Hilbert made a point out of not giving explicit definitions of the basic objects and relations of geometry avoiding thereby any reference to a physical space or anything else outside of the mathematical axiom system itself. The basic notions of geometry were only defined implicitly by the requirement that they satisfy the axioms. Since there is no predetermined subject matter about which the axioms are true, the question of the consistency of the axiom system became pertinent. And with Hilbert, consistency of an axiom system was not only considered a necessary condition for the truth of the mathematical theory it described, it also became a sufficient condition for the mathematical existence of the objects the theory deal with. Indeed, for Hilbert and the formalists following his lead in the twentieth century, an axiomatically defined system exists from a mathematical point of view if and only if it is consistent.[6]

Beltrami’s method of models provided Hilbert with the method to prove the consistency of his axiom system of geometry relative to arithmetic of the real numbers, as well as a method of proving the independence of crucial axioms like the parallel postulate and the continuity axioms. It is ironic that Beltrami’s idea of a model that he developed as a means to supply a real substrate to non-Euclidean geometry, ended up as a crucial method to prove consistency in Hilbert’s axiomatic approach that explicitly avoided any real substrate at all.

I began this paper by wondering about some apparently odd utterances by nineteenth century mathematicians concerning geometry and space. One of the reasons why they appear weird to a modern reader is that we are so steeped in the modern structural axiomatic thinking. However, as pointed out by Volkert, we

[6] Blanchette (2018) gives a fine analysis of the discussion between Hilbert and Gottlob Frege. This discussion shows how fundamentally Hilbert’s new foundation of mathematics broke with the older nineteenth century (and older) approach to the foundation of mathematics defended by Frege. However, as pointed out by Rowe, Gray (2007, 254) and others, Hilbert’s axiomatization build on earlier works by Hermann Wiener, Friedrich Schur, Moritz Pasch, Giuseppe Peano and Giuseppe Veronese. In particular Gray points out that “The point (that the objects of geometry should be purged of their intuitive geometrical meanings) was not original with Hilbert, but he grasped its significance more profoundly” (Gray 2007, 254). Moreover, Corry (2004) has warned us not to consider Hilbert a full-fledged formalist. Hilbert emphasized that the fruitful axiomatic systems studied by mathematicians have come about by *axiomatizing* informal intuitive mathematical theories that we use to describe the real world. For example in geometry, Hilbert had historic and practical reasons for choosing the axioms the way he did. In that sense the axioms are not arbitrary. They and the objects of the theory correspond to our intuitive idea of space. But the mathematical arguments within the theory should be independent of the historic or intuitive origin of the axioms.

should be careful not to project our modern concepts back into the nineteenth century:

> The distinction between an abstract-axiomatic geometry and an interpreting geometry – as for example the geometry of light rays – does not seem to have found general attention .... until the 20th century. This distinction depended on an appropriate development concerning axiomatization of geometry as well as a clear idea of what a model is. It is dangerous to interpret these ideas back into the discussions of the 19th century. (Volkert 2018, 35)

13.12 Conclusion

In this paper I have tried to argue that when 19th century mathematicians wrote about "space" and "geometry" they meant physical space and the geometry that correctly describe this space. The objects of geometry were points, lines, planes etc. that were mostly considered as well-understood concepts that need no special introduction. Most mathematicians believed that this geometry was 3-dimensional and Euclidean, but after 1870 non-Euclidean geometries and higher dimensions were gradually considered as possible descriptions of physical space.

Non-Euclidean geometry was from the start developed explicitly as a possible description of physical space. *N*-dimensional geometries (as we call them), on the other hand, were first developed as a way to deal with analysis of more than three variables in a way that would allow the reader to use his geometric intuition to clarify the situation (see e.g. Schläfli 1852, 175). Still the early inventors of such "geometries" were careful to use the word "space" and "geometry" only about three dimensional physical space. For their more general concepts they used words as "multiple continuity" (Schläfli), "linear theory of extension" (Grassmann) and "multiply connected manifold" (Riemann). However after 1870 the phrase higher space or n-dimensional geometry slowly prevailed and some mathematicians like Clifford, Ball, Newcomb and others began contemplating if some recently discovered physical phenomena could perhaps be explained by the hypothesis of a fourth dimension. Still, when some mathematicians and in particular philosophers, amateurs and charlatans began to entertain the idea that spiritual phenomena took place in the fourth dimension, more sober mathematicians became careful to formulate themselves in ways that would not associate them with disreputable "primitive men".

There were two methods to argue about the nature of space: In the beginning of the century an empirical view of geometry gradually prevailed but toward the end of the century there was a resurgence of neo Kantian views. According to the latter, a geometry exist if it can be intuited. However, since there was no general agreement as to what can be intuited, the Kantian view of what is a true and useful geometry was particularly slippery. On the one hand many used the argument to exclude higher dimensions as well as non-Euclidean geometries. On the other hand, even those who stuck to the traditional idea of space being a priori 3-dimensional and

Euclidean could argue for the introduction of analytical spaces of higher dimension, on the ground that they enhanced the *intuitive* understanding of the situation.

Consistency was considered as a necessary condition for the possibility of the new geometries. However, in general, consistency was not accepted as a sufficient condition for their truth or existence. Moreover, before 1870, no one seems to have contemplated the possibility of *proving* consistency of a mathematical theory. And when Beltrami discovered the model-method, he considered it as a method of giving non-Euclidean geometry a real substrate. Only with Houël, Poincaré and finally Hilbert the method was shown to provide the key to proofs of relative consistency.

Thus when mathematicians such as Lipschitz and Darboux and physicists like Hertz described mechanics in a geometric form, they emphasized the intuitive understanding gained by the geometrical form. However they did not present their new approach as an *application* of an *n*-dimensional *geometry* of curved *space*. Rather they saw their new formalism for mechanics as *analogous* to *n*-dimensional geometry. For them the two theories were different because they deal with different objects. Mechanics deals with mechanical systems of a collection of points (in space), geometry, on the other hand, deals with single points in space. Both Hertz and Darboux were probably aware that the idea of an *n*-dimensional geometry of a curved space was by many considered a "wild flight of an unbridled imagination" (Newcomb 1897, 187). In particular after the Zöllner affair, it was therefore important to signal, that their geometric formalism for mechanical systems was not subject to the objections raised against the imaginary spaces of the mathematicians. Hertz explicitly declared that "we need not fear the objection that in building up a science dependent upon experience, we have gone outside the world of experience". Darboux was less explicit, but the structure of his presentation strongly suggests that he had a similar agenda.

The development of higher-dimensional and curved "spaces" were instrumental in the development of the modern axiomatic, structural and formalistic view of mathematics. Both its lack of external reference for its objects, the free choice of axioms, the identification of consistency and existence, and the model-method for establishing relative consistency have their roots in the disputes about the new geometries. Conversely, this paper has highlighted the fact that one cannot understand the development of geometry in the nineteenth century without leaving the modern philosophy of mathematics aside. In particular it is important to keep in mind that in the nineteenth century the objects of geometry were objects in space either in physical space out there or in our imagination of physical space.

References

Beltrami, Eugenio. 1868a. Saggio di interpretazione della geometria non-euclidea. *Giornale di matematiche* 6, 284–312. English translation in John Stillwell: Sources of Hyperbolic Geometry. American Mathematical Society, London Mathematical Society, 1996, 1–28.

———. 1868b. Teorie fundamentale degli spazii di curvatura constante. *Annali di matematica pura ed applicata* 2 (2), 232–255. Here translated from Voelke 205, 168.

———. 1869. Sulla teorica generale dei parametric differenziali. *Memorie dell'Accademia delle scienzedell'Istituto di Bologna*, series 2, Vol. 8.

Blanchette, Patricia. 2018. The Frege-Hilbert Controversy. In *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), ed. Edward N. Zalta. https://plato.stanford.edu/archives/fall2018/entries/frege-hilbert/.

Cajori, Florian. 1926. Origins of Fourth Dimension Concepts. *The American Mathematical Monthly* 33: 397–406.

Corry, Leo. 2004. *David Hilbert and the Axiomatization of Physics (1898–1918)*. Dordrecht: Kluwer.

Darboux, Gaston. 1888. *Leçons sur la théorie générale des surfaces*, 2. Partie. Paris: Gauthier-Villars.

Friedman, Michael. 1985. Kant's Theory of Geometry. *The Philosophical Review* 94: 455–506.

Gauss, Carl Friedrich. 1900. *Werke* Vol. VIII. Gesellschaft der Wissenschaften zu Göttingen.

Grassmann, Hermann. 1844. *Die Lineale Ausdehnungslehre*. Leipzig: Otto Wigand.

Gray, Jeremy. 1989. *Ideas of Space: Euclidean, Non-Euclidean, and Relativistic*. 2nd ed. Oxford: Clarendon Press.

———. 1994. *Plato's Ghost: The Modernist Transformation of Mathematics*. Princeton: Princeton University Press.

———. 2007. *Worlds Out of Nothing. A Course in the History of Geometry in the 19th Century*. London: Springer.

Hertz, Heinrich. 1894. *Die Prinzipien der Mechanik in neuem Zusammenhange dargestellt. Barth, Leipzig*. Gesammelte Werke 3 (1910). Reprinted Vaduz Sändig. 1984. English translation: *The Principles of Mechanics Presented in a New Form*. Macmillan, 1900. Reprinted Dover, New York (1950).

Hertz, Heinrich. 1977. *Memoirs, Letters, Diaries*. San Francisco: San Francisco Press.

Jordan, Camille. 1872. Essai sur la géométrie à *n* dimensions. *Comptes rendus de l'Académie des Sciences* 75: 1614–1616.

———. 1875. Essai sur la géométrie à *n* dimensions. *Bulletin de la Société Mathématique de France* 3: 103–174.

Knudsen, Ole. 1985. Mathematics and Physical Reality in William Thomson's Electromagnetic Theory. In *Wranglers and Physics. Studies on Cambridge Physics in the 19th Century*, ed. P. Harman. Manchester: Manchester University Press.

Lipschitz, Rudolf. 1872. Untersuchung eines problems der Variationsrechnung, in welchem das problem der Mechanik enthalten ist. *Journal für die reine und angewandte Mathematik* 74: 116–149.

Lobachevsky, Nicholai I. 1840. Geometrical Researches on the Theory of Parallels. Translated into English by George Bruce Halsted. In Roberto Bonola*: Non-Euclidean Geometry*. New York: Dover, 1955.

Lorentz, Hendrik Antoon (1902) Some considerations on the principles of mechanics, in connection with Hertz's Prinzipien der Mechanik. *Verslagen der Zittingenvan de Wis- en NaturkundigeAfdeeling der Koniklizke Akademie van Wetenschappen*, 10, 876. *Abhandlungen über Theoretische Physik* I (1907), 1–22.

Lützen, Jesper. 1990. *Joseph Liouville 1809–1882: Master of Pure and Applied Mathematics*. New York: Springer.

———. 1995. Interactions Between Mechanics and Differential Geometry in the 19th Century. *Archive for History of Exact Sciences* 49: 1–72.

———. 2005. *Mechanistic Images in Geometric Form: Heinrich Hertz’s Principles of Mechanics*. Oxford: Oxford University Press.

Möbius, August Ferdinand. 1827. *Der barycentrische Calcul*. Leipzig: Barth.

Newcomb, Simon. 1897. The Philosophy of Hyperspace. Presidential Address Delivered Before the American Mathematical Society at Its Fourth Annual Meeting, December 29, 1897. *Bulletin of the American Mathematical Society* 4 (5): 187–195 (1898).

Richards, Joan L. 1988. *Mathematical Visions. The Pursuit of Geometry in Victorian England*. Boston: Academic.

Riemann, Bernhard. 1854. Ueber die Hypothesen, welche der Geometrie zu grunde liegen. In *Bernhard Riemann’s Gesammelte Mathematische Werke*. 2. Ed., 272–287. Teubner, Leipzig 1892.

———. 1861. Commentatio mathematica, qua respondere tentatur quaestioni ab Ill^{ma} Academia Parisiensi propositae. In *Bernhard Riemann’s Gesammelte Mathematische Werke*. 2. Ed., 391–404. Teubner, Leipzig 1892.

Schläfli, Ludwig. 1852. Theorie der vielfachen Kontinuität. *In Schläfli Gesammelte Mathematische Abhandlungen* Band 1. Birkhäuser, Basel 1950.

Smith, David Eugene. 1959. *A Source Book in Mathematics*. New York: Dover.

Throesch, Elizabeth L. 2017. *Before Einstein. The Fourth Dimension in Fin-de-Siècle Literature and Culture*. London: Anthem Press.

Voelke, Jean-Daniel. 2005. *Renaissance de la géométrie non euclidienne entre 1860 et 1900*. Bern: Peter Lang.

Volkert, Klaus. 2013. *Das Undenkbare denken. Die Rezeption der nichteuklidischen Geometrie im deutschsprachigen Raum (1860–1900)*. Berlin: Springer.

———. 2018. *In Höheren Räumen. Der Weg der Geometrie in die vierte Dimension*. Berlin: Springer.

Chapter 14
Gauging Potentials: Maxwell, Lorenz, Lorentz and Others on Linking the Electric Scalar and Vector Potentials

Jed Z. Buchwald

Abstract The development of tractable relationships for calculating the radiation produced by electric oscillators relied on a particular connection between the scalar potential, whose gradient is determined by charge density, and the vector potential, whose curl determines magnetic induction. Other such connections are permissible in electrodynamic field theory, including one used by Maxwell, but only one among them leads easily to radiation. Forms of electrodynamics other than Maxwell's, including those by Kirchhoff, Helmholtz, Riemann and others, were not free to choose among such relationships, though none among these theories, with the limiting exception of Helmholtz's, led to wavelike equations of propagation without assuming propagation *a priori*. One theory besides Maxwell's that did was developed by the Danish physicist Lorenz. Although Lorenz derived the link between potentials later deployed for radiation in classical electrodynamics, his theory allowed no other choice. Three decades after Lorenz, the Dutch physicist Lorentz produced the first general understanding that the connection between the potentials can take various forms. Such a link was subsequently denominated a "gauge", governed by the understanding that a particular choice, called a "gauge transformation", must leave the measurable quantities of the theory unaltered. What follows traces the ways in which connections between the potentials were deployed in the nineteenth and early twentieth centuries.

By the end of the twentieth century's first decade texts on electrodynamics generally presented substantially equivalent forms of field theory supplemented by what became known as the Lorentz force on moving, charged particles. The measurable quantities of the theory comprised the electric and magnetic fields together with charge and current. Two auxiliary variables were generally used, both of which had roots dating to the early to mid 1800s: the scalar and vector potentials, associated

J. Z. Buchwald (✉)
Doris and Henry Dreyfuss Professor of History, California Institute of Technology, Pasadena, CA, USA
e-mail: buchwald@caltech.edu

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_14

respectively with charge and current. The fundamental equations of the theory placed constraints on the relationship between these two variables, but within these limitations different expressions for them could be adapted while maintaining unaltered the measurable variables of the theory. Later in the twentieth century a particular choice of such auxiliaries came to be called a *gauge.* Two such gauges were, and to this day remain, central: one in which the sum of the vector potential's divergence with the time derivative of the scalar potential vanishes, while, in the other, the divergence vanishes. Both gauges require an associated scalar function to satisfy specific conditions.

Prior to the early twentieth century, theories for electrodynamics based on the interactions between charged particles did deploy a connection between the potentials, but we will see that none of them was free to choose any other than a particular one because of the model used for the electric current. Maxwell's theory, in which charges are an epiphenomenon of field processes, was free to choose, but we will also see that he did not recognize the full scope of that freedom.[1] Compactly to represent the difference between theories that must assume a certain relation and other theories that may, but need not, do so, what follows introduces the distinction between a *condition* and a *relation.* If a given theory cannot avoid a specific connection between the potentials, then we will call such a connection a *condition*. If it is not required by theory, we will call it a *relation*. Only *relations* are gauges because *conditions* lock the connection down.

When, one might ask, did *conditions* become *relations*? That recognition was for many years attributed to H. A. Lorentz's (1853–1928) exposition of electrodynamics in 1903.[2] However, in 2001 J. D. Jackson and L. B. Okun argued that the first proper use should instead be attributed to the Danish physicist Ludvig Lorenz (1829–1891) since he had employed the equation previously associated with Lorentz in an 1867 theory based on retarded interactions.[3] In the years since, physics articles and texts have generally adopted their suggestion. In what follows we'll explore the ways in which several theories deployed such connections in order to tease out which among them truly recognized the *relational* character of the connection.

14.1 Invariance in "Classical" Electrodynamics

In order properly to grasp the difference between historical conceptions of the links between the scalar and vector potentials, we need to begin with the manner in which a *gauge transformation* – one that leaves unaltered the theory's measurable

[1] On Maxwell's field understanding of charge see (Buchwald 1985, 23–34).

[2] (Lorentz 1903b, 157). In earlier work Lorentz had instead imposed zero-divergence on the vector potential without discussion (Lorentz 1892, 14).

[3] (Jackson and Okun 2001). See (Zangwill 2013) and (Thorne and Blandford 2017) for contemporary presentations of electrodynamics.

variables – is construed in what is conventionally referred to as “classical” electrodynamics. We begin with the four basic “Maxwell” equations in what is known as their “Hertz-Heaviside” form:

$$\nabla \cdot \varepsilon \boldsymbol{E} = 4\pi\rho \tag{14.1}$$

$$\nabla \cdot \boldsymbol{B} = 0 \tag{14.2}$$

$$\nabla \times \boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} = 0 \text{``}\mathit{Faraday}\text{"} \tag{14.3}$$

$$\nabla \times (\boldsymbol{B}/\mu) = \frac{4\pi}{c}\boldsymbol{C} + \frac{1}{c}\frac{\partial \varepsilon \boldsymbol{E}}{\partial t}\text{``}\mathit{Ampere}\text{"} \tag{14.4}$$

Here **E, B** are respectively electric and magnetic fields, while ρ is the volume density of electric charge and $\boldsymbol{C}$ the current area density. Taking the divergence of 14.4 and then applying 14.1 produces the “equation of continuity” between charge and electric current:

$$\nabla \cdot \boldsymbol{C} + \frac{\partial \rho}{\partial t} = 0 \qquad \mathit{Equation\ of\ Continuity} \tag{14.5}$$

These equations accordingly involve four fundamental quantities: the fields **E**, **B** and the material densities ρ, **C**, with the continuity equation being an implication of the field equations and not a distinct requirement. These four are the only observable quantities; any additional variables that may be introduced must leave them unaltered.

The requirement on the divergence of the **B** field 14.2 permits the latter’s expression as the curl of such an auxiliary quantity, namely the vector potential **A**. Introducing that vector yields a link between its time derivative and the electric field, leading to a second auxiliary function, the scalar potential φ:

Introduction of the vector potential: $\boldsymbol{B} = \nabla \times \boldsymbol{A}$.
The Faraday Law 14.3 *then becomes:* $\nabla \times \left(\boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}\right) = 0$
which allows

$$\boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} = -\nabla\varphi \ \mathit{Faraday\ Law\ in\ terms\ of}\ \boldsymbol{A}, \varphi \tag{14.6}$$

Inserting these results into 14.1 and 14.4 produces two coupled, time-dependent equations involving the auxiliary potentials and the measurable quantities representing charge and current:

$$\nabla^2\varphi = -4\pi\rho - \frac{1}{c}\frac{\partial\left(\nabla\cdot\boldsymbol{A}\right)}{\partial t} \tag{14.7}$$

$$\nabla^2\boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} = \nabla\left(\nabla\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial\varphi}{\partial t}\right) - \frac{4\pi}{c}\boldsymbol{C} \tag{14.8}$$

The fact that **A** is parasitic upon the **B** field, while φ depends on **E** and **A**, opens a path to uncoupling these equations. In a first step, the freedom entailed by the introduction of **A** allows the addition to the latter of the gradient of an otherwise arbitrary scalar function Λ without affecting the value of the magnetic field. To maintain the Faraday Law in form 14.6, the scalar φ must be correspondingly altered:

Adding the gradient of the scalar function Λ to the vector potential

$$\boldsymbol{A}_{HH} \equiv \boldsymbol{A} + \nabla\Lambda \tag{14.9}$$

requires changing the scalar potential to

$$\varphi_{HH} \equiv \varphi - \frac{1}{c}\frac{\partial\Lambda}{\partial t} \tag{14.10}$$

to maintain the Faraday Law

Such a change depends directly upon the function Λ. Alternatives to field theory in which the potentials are specified directly do not have the freedom to effect a *relation* of this kind. Different equations linking the potentials can be obtained, with the function Λ then satisfying a specific equation. A given choice came to be called a "gauge transformation" with the understanding that the measurable variables of the theory (**E, B,** ρ**, C**) always remain unaltered

We will label the relation that came to be most widely chosen for field theory in its Hertz-Heaviside form R_{HH} for the moment to avoid a particular attribution. By means of it 14.7 can be uncoupled from 14.8 provided that the scalar Λ satisfies a restriction, thereby yielding wave equations for the modified potentials with the charge and current densities as sources:

Relation R_{HH}: $\nabla\cdot\boldsymbol{A}_{HH} + \frac{1}{c}\frac{\partial\varphi_{HH}}{\partial t} = 0$
the definitions of $\boldsymbol{A}_{HH}$, φ_{HH} *then require*

$$\nabla^2\Lambda - \frac{1}{c^2}\frac{\partial^2\Lambda}{\partial t^2} = -\left(\nabla\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial\varphi}{\partial t}\right) \tag{14.11}$$

thereby uncoupling 14.7 *and* 14.8 *to become*

$$\nabla^2\varphi_{HH} - \frac{1}{c^2}\frac{\partial^2\varphi_{HH}}{\partial t^2} = -4\pi\rho \tag{14.12}$$

$$\nabla^2\boldsymbol{A}_{HH} - \frac{1}{c^2}\frac{\partial^2\boldsymbol{A}_{HH}}{\partial t^2} = -\frac{4\pi}{c}\boldsymbol{C} \tag{14.13}$$

The function Λ correspondingly satisfies an inhomogeneous wave equation whose source is $\nabla \cdot \mathbf{A} + \frac{1}{c}\frac{\partial\varphi}{\partial t}$. If, from the set of untransformed potentials, only those are chosen that themselves satisfy an R_{HH} relation, then Λ will satisfy a homogeneous wave equation. Given R_{HH}, the propagation Eqs. 14.12, 14.13 taken together satisfy the equation of continuity 14.5, as of course must be the case since they devolve from the fundamental field equations. The R_{HH} gauge is thereby tied to the satisfaction of wave equations by the scalar and vector potentials. Provided that Λ is chosen to satisfy either the homogeneous or inhomogeneous wave equation, then 14.14 for the electric field **E** follows from the Faraday Law in the form 14.6:

$$\nabla \cdot \boldsymbol{A}_{HH} + \frac{1}{c}\frac{\partial\varphi_{HH}}{\partial t} = 0$$

with

$$\boldsymbol{E} = -\frac{1}{c}\frac{\partial\boldsymbol{A}_{HH}}{\partial t} - \nabla\varphi_{HH}$$

eliminates the transformed scalar potential to yield

$$\boldsymbol{E} = -\frac{1}{c}\frac{\partial\boldsymbol{A}_{HH}}{\partial t} + c\nabla\left(\nabla \cdot \int \boldsymbol{A}_{HH}dt\right) \tag{14.14}$$

Note that the **E** field can now be expressed entirely in terms of the vector potential because the R_{HH} relation replaces the term in the scalar potential with a time integral over the vector potential's divergence. Equation 14.14 can be applied directly to problems involving radiation since the vector potential satisfies the d'Alembertian 14.13 with current **C** as source, i.e.given $R_{HH,}$ the vector potential is $\frac{1}{4\pi}\int \frac{\boldsymbol{C}\left(\boldsymbol{r}',t-\frac{|r-r'|}{a}\right)}{|\boldsymbol{r}-\boldsymbol{r}'|}d^3r'$. This provides a direct route for solving the radiation fields emitted by an electric oscillator by working directly with a vector potential.

Relation R_{HH} is not uniquely required given the freedom provided by the ability to add any gradient to the vector potential. One alternative sets the latter's divergence to zero. After the early 1900s this alternative was termed the "Coulomb gauge" because under it the scalar potential satisfies the time-independent Poisson equation, while its time derivative couples to the current to act as sources in a d'Alembertian (14.16) for the vector potential:

The "Coulomb" relation in Hertz-Heaviside

$$R_{coul} : \nabla \cdot \boldsymbol{A}_{coul} = 0$$

the definitions of $\boldsymbol{A}$, φ *then require*

$$\nabla^2 \Lambda - \frac{1}{c^2}\frac{\partial^2 \Lambda}{\partial t^2} = -\frac{1}{c}\frac{\partial \varphi_{coul}}{\partial t}$$

thereby uncoupling 14.7 from 14.8 to become

$$\nabla^2 \varphi_{coul} = -4\pi\rho \tag{14.15}$$

$$\nabla^2 \boldsymbol{A}_{coul} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}_{coul}}{\partial t^2} = -\frac{4\pi}{c}\boldsymbol{C} + \frac{1}{c}\nabla\left(\frac{\partial \varphi_{coul}}{\partial t}\right) \tag{14.16}$$

In this gauge the d'Alembertian for Λ has only the time derivative of the scalar potential as source since the divergence of the vector potential is zero (*cf* 14.11). To understand the physical consequences of 14.16 requires introducing the Helmholtz decomposition theorem to uncouple the potentials. According to the latter, which Hermann Helmholtz (1821–1894) developed in his path-breaking 1858 paper on vortices (not, of course, in vector form),[4] a directed function can be fully specified by the difference between the gradient of the divergence of its potential throughout space and the curl of the curl of the same integral provided that the field vanishes at infinity:

Helmholtz decomposition

$$\boldsymbol{F}(\boldsymbol{r}) \equiv -\frac{1}{4\pi}\int \frac{\boldsymbol{C}(\boldsymbol{r}')}{\left|\boldsymbol{r}-\boldsymbol{r}'\right|} d^3r' \ so\ \boldsymbol{C} = \nabla(\nabla\cdot\boldsymbol{F}) - \nabla\times(\nabla\times\boldsymbol{F})$$

After simple manipulation it follows that the first term in this expression for the current **C** is canceled by the scalar potential term in 14.16 by means of 14.15 and the continuity Eq. 14.5, leaving only the second term, $-\nabla\times(\nabla\times\mathbf{F})$, as source. **C** is then known for obvious reasons as the "transverse current," while the expression for the vector potential now contains $-\nabla\times(\nabla\times\boldsymbol{F})$ instead of simply **C**. In this

[4] (Helmholtz 1858).

gauge the **E** field retains the term in the gradient of the scalar potential, whereas that potential disappears from **E** in the R_{HH} alternative. This poses the apparent problem for radiation calculations that the scalar potential does not propagate. However, neither potential is directly measurable in classical electrodynamics, with the electric and magnetic fields depending in the end on retarded expressions for charge and current.[5] Maxwell and, for a time, such deployers of his theory as George Francis FitzGerald (1851–1901) and Joseph John Thomson (1856–1940) presumed the form R_{coul}.

14.2 Potentials Before Maxwell, in His *Treatise*, and in Helmholtz's Alternative

The first appearance of a link between current and charge potentials arose years before Maxwell's electrodynamics in an 1857 article by Gustav Robert Kirchhoff (1824–1887) on conduction, where it occurs as a *condition*. To capture the ponderomotive and electromotive forces due to currents that Wilhelm Weber (1804–1891) had developed, Kirchhoff modified the function that Franz Ernst Neumann (1798–1895) had introduced in 1845 and 1847 to express both sorts of forces, ponderomotive through space and electromotive through time derivatives. To do so he, like Weber, presumed a specific hypothesis concerning the electric current that had been developed by Gustav Theodor Fechner (1801–1887), namely that it consists of equal and oppositely-charged particles flowing with equal but opposite velocities through conductors. This entails a continuity equation containing a factor of 1/2. A fixed link R_k between the vector and scalar potentials resulted.[6]

None of our subjects, excepting Lorentz and (in part) Maxwell, worked with vectors. Nevertheless, all were well-versed in the transformations that directed quantities may undergo and easily represented their results in component form. In what follows we will employ vector notation in order to bring out the central points at issue. In one case, however, we will see that a particular expression may have been missed in part at least because the complexity of transformations done through components can be misleading.

[5] For a demonstration that causality holds in the Coulomb gauge see (Heras 2011) and (Jackson 2002). Both demonstrate that an unretarded term which appears in the Coulomb gauge expression for the vector potential exactly cancels the one in the scalar potential.

[6] (Kirchhoff 1882a, 139) and (Kirchhoff 1882b, 159). On Weber see (Archibald 1989b). A modern account is (Assis 1994). See (Jackson and Okun 2001, 9–10) on Kirchhoff et al. Note that the difference in form between the Neumann and Kirchhoff potential functions vanishes for the interaction between closed circuits. The difference also vanishes if the divergence of the potential is zero.

Kirchhoff's Potentials and Continuity Equation

$$\begin{array}{c} \boldsymbol{A}_k\left(\boldsymbol{r}\right)=\int\frac{(\boldsymbol{r}-\boldsymbol{r}')(\boldsymbol{J}(\boldsymbol{r}')\cdot(\boldsymbol{r}-\boldsymbol{r}'))}{|\boldsymbol{r}-\boldsymbol{r}'|^3}d^3r' \\ \left[the\ Neumann\ form\ is\ \boldsymbol{A}_N\left(\boldsymbol{r}\right)=\int\frac{\boldsymbol{J}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}d^3r'\right] \end{array} \tag{14.17}$$

$$\nabla^2\varphi=-4\pi\rho \tag{14.18}$$

$$\nabla\cdot\boldsymbol{J}+\frac{1}{2}\frac{\partial\rho}{\partial t}=0\ [continuity\ given\ the\ Fechner\ current\ model] \tag{14.19}$$

Kirchhoff's 14.17, 14.18 and 14.19 Link the Vector and Scalar Potentials

$$R_k:\nabla\cdot\boldsymbol{A}_k=\frac{1}{2}\frac{\partial\varphi}{\partial t}$$

R_k is clearly a *condition* since it is entailed jointly by a Poisson Eq. (14.18) and continuity (14.19), with the one-half constant reflecting the Fechner model.[7]

Kirchhoff then used the potentials as follows to formulate Ohm's law, in which σ stands for conductivity:

Kirchhoff's version of Ohm's Law

$$\boldsymbol{C}_k=-2\sigma\left(\nabla\varphi_k+\frac{4}{c^2}\frac{\partial\boldsymbol{A}_k}{\partial t}\right) \tag{14.20}$$

Finally, through a series of approximations, he generated a differential equation for current i in a thin, straight, cylindrical wire of length l and radius a that may be written in the following way:[8]

$$\boldsymbol{C}_k=\frac{1}{2}\nabla^2 i-\frac{1}{c^2}\frac{\partial^2 i}{\partial t^2}-\frac{1}{4\sigma\pi a^2}\frac{\partial i}{\partial t}=0$$

In wires of very high conductivity (and small radius, since the latter was assumed in the deduction), the current propagates as a wave with speed $\frac{c}{\sqrt{2}}$.[9]

---

[7] R_k results either from Kirchhoff's form $\boldsymbol{A}_k$ of the vector potential or from the Neumann form $\mathbf{A}_{\mathrm{N}}$ given the same continuity equation. These two potentials have the same curl.

[8] (Kirchhoff 1882a, 140–41).

[9] In 1856 Weber and Rudolf Kohlrausch (1809-1858) found the constant c to have the value $4.39 * 10^8 m/s$, which implied that the current propagates in such wires at speed $3.1042 * 10^8 m/s$, which was close to the extant value for the speed of light.

Turn next to Maxwell. We begin with the fundamental equations for the theory as presented in his 1873 *Treatise*:[10]

Maxwell's equations in the order set out in the *Treatise*

A "electromagnetic momentum at a point"
B "magnetic induction"
J "(total) electric current"
D "electric displacement"
E "electromotive force"
F "mechanical force"
v "velocity of a point"
H "magnetic force"
M "intensity of magnetization"
C "current of conduction"
ψ "electric potential"
e "electric density"
C "conductivity for electric currents"
K "dielectric inductive capacity"
μ "magnetic inductive capacity"
T "electrokinetic energy"
W "electrostatic energy"

$\mathbf{B} = \nabla \times \mathbf{A}$
$\nabla \bullet \mathbf{A} = 0$ [assumed]
$\mathbf{E} = \mathbf{v} \times \mathbf{B} - \partial \mathbf{A}/\partial t - \nabla \psi$
$\mathbf{F} = \mathbf{J} \times \mathbf{B} - e \nabla \psi$ [note Maxwell's inclusion here of the displacement current]
$\mathbf{B} = \mathbf{H} + 4\pi \mathbf{M}$
$4\pi \mathbf{J} = \nabla \times \mathbf{H}$
$\mathbf{C} = C\mathbf{E}$
$\mathbf{D} = (1/4\pi) K \mathbf{E}$
$\mathbf{J} = \mathbf{C} + \partial \mathbf{D}/\partial t$
$\mathbf{B} = \mu \mathbf{H}$
$e = \nabla \bullet \mathbf{D}$
$T = (1/2)\iiint (\mathbf{A} \bullet \mathbf{J}) d^3 r \rightarrow T = (1/8\pi)\iiint (\mathbf{H} \bullet \mathbf{B}) d^3 r$
$W = (1/2)\iiint (\mathbf{E} \bullet \mathbf{D}) d^3 r$

Here we see that the vector (**A**) and scalar (ψ) potentials appear in the *Treatise*'s final set of equations as fundamental elements. Note also that Maxwell's **J** includes the conduction and displacement currents. Both are to be incorporated in a derived expression for **A**, though Maxwell did not provide one in his basic set. The vector potential had appeared from the outset in his 'dynamical' theory of the electromagnetic field because the theory's equations were constructed on the basis of Lagrange's (or Hamilton's) equations, with **A** functioning dynamically as the field's "electrokinetic momentum."[11] This is why no separate equation directly expresses the absence of the divergence of **B**: that follows from the latter's expression as the curl of the momentum. It remained necessary to deduce a computable expression for **A**.

Maxwell offered two arguments for abolishing the potential's divergence in order to do so. His first argument introduced auxiliary variables $\mathbf{A}'$ and χ in the equation for **A** that results from the combination of the Ampere law in terms of the total current with his basic expression for **B** as $\nabla \times \mathbf{A}$:

[10] (Maxwell 1873), vol. 2, chap. 9.

[11] (Maxwell 1873, sectns. 585-591). Maxwell's electrokinetic momentum per unit area is $\nabla \times \mathbf{A}$, derived from his expression for the circuit's contribution to the field momentum as $\oint \mathbf{A} \cdot d\mathbf{l}$, where **A** is introduced directly as such. He then asserted that $\nabla \times \mathbf{A}$ represents "what we are already acquainted with as the magnetic induction" because the variation of this momentum density through the area surrounded by a circuit results in a force, and on "Faraday's theory, the phenomena of electromagnetic force and induction in a circuit depends on the variation of the number of lines of magnetic induction which pass through the circuit" (sectn. 592).

Maxwell's $\mathbf{J}$ *is the total current* $\mathbf{C} + \frac{\partial \mathbf{D}}{\partial t}$

$$\left\{ \begin{array}{c} \mathbf{B} = \nabla \times \mathbf{A} \\ 4\pi \mathbf{J} = \nabla \times \mathbf{H} \end{array} \right\} \Longrightarrow 4\pi \mu \mathbf{J} = \nabla \left(\nabla \cdot \mathbf{A} \right) - \nabla^2 \mathbf{A}$$

which Maxwell solves as

$$\begin{array}{l} \mathbf{A}\left(\mathbf{r}\right) = \mu \int \frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3 r' - \frac{1}{4\pi} \nabla_r \int \frac{\nabla_{r'} \cdot \mathbf{A}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3 r' \\ \equiv \mathbf{A}' - \nabla \chi \\ \Rightarrow \nabla^2 \chi = -\nabla \cdot \mathbf{A} \;\; and \;\; so \\ R_{max} : \nabla \cdot \mathbf{A}' = 0 \end{array}$$

Since $- \nabla \cdot \mathbf{A}$ is the source for a Poisson equation inχ, it follows that $\nabla \cdot \mathbf{A}'$ vanishes by taking the divergence of both sides of Maxwell's vector potential. Note that $\nabla \chi$ is not an arbitrary addition to the vector potential justified solely by the link of its curl to the **B** field. Not at all: Maxwell's $\nabla \chi$ is just one of the two terms in his solution to the current equation using $\nabla \times \mathbf{A}$ for the magnetic force **H**. The other term contains the total current **J**. This is entirely different from the later procedure, which allows for any gradient to be added to the vector potential on the grounds that the latter is merely an auxiliary quantity that derives ultimately from the **B** field's absence of divergence.

The "quantity χ disappears" from the expression for **B**, Maxwell however noted, and "is not related to any physical phenomenon." Were it a question solely of the links between **J**, **A**, and **B** then χ would simply be irrelevant, and Maxwell's statement concerning its absence from "any physical phenomenon" unnecessary. He likely had more in mind because his basic expression for the electric field then includes the time derivatives of both $\mathbf{A}'$ and the gradient of χ . The latter might have a detectable result were it allowed to stand. Since no evidence for such an addition had ever been found, the simplest procedure would be to discard χ altogether.

Two useful results then follow: first, the divergence of **A,** like that of $\mathbf{A}'$, will "also be zero everywhere," in which case **A** can be expressed by just the first term – and that expression (ignoring magnetic permeability and any constant factor) was well-known because it is formally similar to those in the works of Weber and Kirchhoff – with one critical exception: Maxwell's inclusion of the displacement current in **J**. Nevertheless, a difficulty plagued this route to zero divergence. As Maxwell pointed out in the next section, his solution requires **A** to vanish at infinity. This could not hold for unlimited waves. To treat these, he attempted to deduce the condition in a different manner.[12]

[12] As noted by (Yaghjian 2014, 240).

Maxwell's second route to zero-divergence relied upon the propagation that results from his basic field equations:[13]

Maxwell's "general equations for electromagnetic disturbances"

$$\mu\left(4\pi C + K\frac{\partial}{\partial t}\right)\left(\frac{\partial \boldsymbol{A}}{\partial t} + \nabla\psi\right) - \nabla^2\boldsymbol{A} + \nabla\left(\nabla\cdot\boldsymbol{A}\right) = 0$$

of which Maxwell took the divergence to obtain

$$\mu\left(4\pi C + K\frac{\partial}{\partial t}\right)\left(\frac{\partial\nabla\cdot\boldsymbol{A}}{\partial t} + \nabla^2\psi\right) = 0$$

Limiting his analysis to a "non-conductor", Maxwell set C (the conductivity) to zero, in which case, he asserted, the Laplacian of ψ must be "independent of t." Why so? Because according to him the Laplacian is proportional to what Maxwell termed "the volume-density of free electricity" (his e, which is the divergence of the displacement vector **D**, or what would later be termed conduction charge). Consequently, this must vanish in a non-conducting medium such as free ether. In that case the divergence of the vector potential "must be a linear function of *t,* or a constant, or zero, and we may therefore leave [its divergence] and ψ out of account in considering periodic disturbances."[14]

This second argument is problematic as it stands because Maxwell seems to have regarded the zero-divergence of the vector potential as a *condition* necessitated by the absence of any evidence for $\nabla\chi$. The difficulty is readily seen. Since his fundamental equations have **E** as $-\frac{\partial\mathbf{A}}{\partial t} - \nabla\psi$ with $\nabla\cdot\mathbf{E}$ proportional to the density of "free electricity," the *sum* $\frac{\partial\nabla\cdot\mathbf{A}}{\partial t} + \nabla^2\psi$ is also proportional to it. Consequently, the presumptive constancy of ρ entails that the time derivative of this sum, and not just that of the Laplacian of ψ, must be zero. Although the defect can be remedied, doing so requires Maxwell to have thoroughly appreciated the *relational* character of any such modification, which he apparently never did.[15]

[13] (Maxwell 1873, secs. 783–784). Note that in the *Treatise* Maxwell's expression for ∇^2 has the opposite sign to the modern one "in order to make our expressions consistent with those in which Quaternions are employed" (sec. 616). I follow modern conventions.

[14] (Maxwell 1873 sec. 783).

[15] (Yaghjian 2014, 244) points out that since the addition of the gradient of any function of $\mathbf{r}$,t to **A** leaves **B** unaltered, it's possible to add to ψ the term $\frac{\partial - \int\psi(r,t')dt' + t\int(\rho/|\mathbf{r}-\mathbf{r}'|)d^3r'}{\partial t}$. The time derivative of the redefined function's Laplacian then vanishes identically, leaving $\frac{\partial^2\nabla\cdot\mathbf{A}}{\partial t^2} = 0$, and so the time derivative of $\nabla\cdot\mathbf{A}$ must be independent of time. Consequently, either **A** itself is independent of time, or else its divergence vanishes. To see the difference between the result that Maxwell drew from zero divergence in comparison to the one later drawn for the Coulomb gauge, it suffices to compare their respective expressions for the vector potential. Maxwell's reduces to the first term in **J**, including the displacement current. The Coulomb gauge expression eschews the

The propagation equation for the vector potential accordingly becomes just $\mu K \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} = 0$. This implies that the ether's equations of motion are "similar to those of the motion of an incompressible elastic solid" capable of sustaining finitely propagating transverse oscillations.[16] His ψ thereby became analogous to pressure in an incompressible fluid and had to propagate infinitely rapidly. For some of his followers the implication may not have been troublesome since it was common among British ether theorists including non-Maxwellians like William Thomson, Lord Kelvin (1824–1907) to think the ether incompressible.

Both J. J. Thomson and FitzGerald for example deployed the zero divergence condition in analyzing the electrodynamics of a moving, charged sphere, where it raised immediate difficulties. In particular, FitzGerald in 1881 calculated the resulting displacement currents in the surrounding field due to charge convection, only to discover on requiring the vector potential's divergence to vanish that the displacement currents disappear from the result, paradoxically leaving only the product of charge by velocity to constitute the effective current.[17] In trying to grasp Maxwell's scheme through a model of the field 4 years later, FitzGerald however could find "no instantaneous propagation of anything."[18]

Turn next to the most widely recognized alternative to Maxwell's system that also yielded propagation. In 1870 Helmholtz generalized Kirchhoff's form of the vector potential to include an undetermined constant k and built a novel system that implied propagation without adopting Maxwellian fields. To do so Helmholtz postulated an ether consisting of electrically and magnetically polarizable elements that interacted instantaneously with one another. These elements were not to be assimilated to moving electric particles in the Weber-Fechner fashion (which Helmholtz regarded as both physically and even morally objectionable):[19] they stood as unreduced, fundamental physical structures. Propagation arose out of the pattern of instantaneous interactions among the distributed polarizable elements that constitute the medium. Both longitudinal and transverse waves result by treating the electric polarization's time-derivative as a current to be included in $\mathbf{J}$ and by incorporating the corresponding polarization charge in the continuity equation.

Because Helmholtz's system was based on potentials that do not in themselves propagate, his link between them differed from Kirchhoff's in only one respect, namely in allowing a more general form to the vector potential, restricted only by

displacement current but involves the curl of an integral that depends upon the retarded current, itself integrated through the time it takes emission to reach the observation point: see (Jackson 2002), 3.17.

[16] Maxwell demonstrates transversality for the vector potential independently of the value of its divergence by combining his equations for $\mathbf{E}$ and $\mathbf{B}$ in terms of $\mathbf{A}$ (Maxwell 1873, vols. 2, sectn. 790).

[17] See (Buchwald 1985, 269–76). Oliver Heaviside avoided the use of either the scalar or vector potentials, but he never analyzed the origin of radiation.

[18] (Hunt 1991, 117–18).

[19] (Buchwald 1993).

the requirement that it produce the proper result for closed currents. In the following expression, the additive term in *1−k* integrates out for closed currents:[20]

Helmholtz's Generalized Vector Potential

$$\mathbf{A}_h(\mathbf{r}) = A^2 \left(\int \frac{\mathbf{J}(\mathbf{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} \mathrm{d}^3\mathrm{r}' + \frac{1}{2}(1-k)\nabla_r \int \mathbf{J}(\mathbf{r}') \cdot \nabla_{r'} |\boldsymbol{r}-\boldsymbol{r}'| \, \mathrm{d}^3\mathrm{r}' \right)$$

$$\nabla^2 \varphi = -4\pi\rho$$

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0$$

which lead directly to

$$R_\mathrm{h} : \nabla \cdot \mathbf{A}_h = -k \frac{\partial \varphi}{\partial t}$$

Helmholtz's R_h is accordingly a *condition* and not a relation since it can have no other form in the original formulation of his theory. Insofar as interactions between closed currents are concerned, the constant k can have any non-infinite value. If it were to be zero then Helmholtz's potential would lack divergence, as Maxwell had required in his scheme, while setting it to one produces the Neumann form. Helmholtz and others thought that if Maxwell's system held true then k had to vanish and the polarizability of the ether had to be effectively infinite. Poincaré later argued that only the latter condition was necessary, but this was not understood for some time.[21]

To summarize, we have seen that both Kirchhoff and Helmholtz have R conditions that result ineluctably from the combination of their definitions of the vector potential with the Poisson equation for the scalar potential and current-charge continuity. Maxwell, working with fields, certainly did recognize that the vector potential could have a gradient added to it without affecting his basic equations, but he did not follow through to examine the allowable scalar functions and instead required the zero divergence of his potential.

[20] For a derivation of the Helmholtz condition R_h see (Buchwald 1985, 314).

[21] For a succinct demonstration that the Helmholtz system produces the same equations as Maxwell's if the polarization is infinite see (Darrigol 2000, 417–19). In the limit of infinite polarizability the constant k drops out of consideration.

14.3 Ludvig Lorenz's Link Between Potentials

In 1867, three years before Helmholtz published his system, the Danish physicist Lorenz developed a novel theory that argued for the existence of propagating "currents" in free space as well as in wires under the assumption that the effect of charge and current at a given point is delayed by a time equal to the distance to the point divided by a constant *a* that represents, in effect, the speed at which the disturbance travels.[22] Lorenz wrote before the publication of Maxwell's 1873 *Treatise*, and though he was likely to have been at least familiar with the mechanical model for the field that Maxwell had generated in 1861–2, he disliked such hypotheses, aiming instead to work without any physical hypothesis concerning the character of the current. Maxwell's further development of his theory, absent a model for the ether, had been published in 1864 but Lorenz did not mention it.

The expressions for the vector and scalar potentials with which Lorenz began were the same as Kirchhoff's 14.17, 14.18, 14.19, and he also took over Kirchhoff's version of Ohm's Law (14.20). The factor of ½ in his continuity Eq. (14.19) implies that Lorenz tacitly presumed Fechner's double-flow model for current. In a critical step, he assumed that the scalar potential generated at some point by a given charge takes time to reach the point. To that end he introduced a "new function" for the potential in which retarded replaces unretarded charge:[23]

$$\varphi_{lnz}(t,r) \equiv \int \frac{\rho\left(\mathbf{r}', t-\left|\boldsymbol{r}-\boldsymbol{r}'\right|/a\right)}{\left|\boldsymbol{r}-\boldsymbol{r}'\right|} d^3r'$$

Here $\boldsymbol{r}$ specifies the point at which the potential is required while $\boldsymbol{r}'$ locates the originating charge density; a is "a constant".[24]

Lorenz then expanded the retarded charge in a series about the emission time $|\boldsymbol{r}-\boldsymbol{r}'|/a$, inserted the series into φ_{lnz} and took the gradient of the result with respect to the observation point r, obtaining:

$$\rho\left(\mathbf{r}', t-\left|\boldsymbol{r}-\boldsymbol{r}'\right|/a\right) = \rho\left(\mathbf{r}', t\right) - \frac{\left|\boldsymbol{r}-\boldsymbol{r}'\right|}{a}\frac{\partial\rho\left(\mathbf{r}', t\right)}{\partial t} + \frac{\left|\boldsymbol{r}-\boldsymbol{r}'\right|^2}{2a^2}\frac{\partial^2\rho\left(\mathbf{r}', t\right)}{\partial t^2} - \ldots$$

$$\nabla\varphi_{lnz} = \nabla\varphi_k + \frac{1}{2a^2}\frac{\partial^2}{\partial t^2}\int \frac{\left(\boldsymbol{r}-\boldsymbol{r}'\right)\rho\left(\mathbf{r}', t\right)}{\left|\boldsymbol{r}-\boldsymbol{r}'\right|} d^3r' - \ldots$$

[22] On Lorenz see (Kragh 2018).

[23] (Lorenz 1867b), translated as (Lorenz 1867a). On Lorenz see (Kragh 2018).

[24] Lorenz used just r to denote $|\boldsymbol{r}-\boldsymbol{r}'|$ while integrating over $dx'dy'dz'$.

Note that the charge density on the right-hand in the expansion for $\nabla\varphi_{lnz}$ is to be evaluated at time t and not at the retarded time t-$|\boldsymbol{r}-\boldsymbol{r}'|/a$.

Lorenz's next step was crucial. He used Kirchhoff's (unretarded) continuity Eq. (14.19) to replace $\partial\rho/\partial t$ in his expansion for $\nabla\varphi_{lnz}$ with $-2\,\nabla\cdot\mathbf{C}$ and performed a partial integration to obtain:

$$\nabla\varphi_{lnz} = \nabla\varphi_k - \frac{1}{a^2}\frac{\partial}{\partial t}\int \frac{\mathbf{C}\left(\mathbf{r}', t\right)}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3r' + \frac{1}{a^2}\frac{\partial \mathbf{A}_k\left(\mathbf{r}, t\right)}{\partial t}\dots$$

Here *all terms* on the right-hand side are functions of present, not retarded, time because the charge and current densities obtain at time t. Nevertheless, Lorenz simply moved the term in $\mathbf{C}$ to the left-hand side and without comment replaced what should be $\mathbf{C}(\mathbf{r}', t)$ with $\mathbf{C}(\mathbf{r}', t-|\boldsymbol{r}-\boldsymbol{r}'|/a)$, the current at the *retarded* time to write:

$$\nabla\varphi_{lnz} + \frac{1}{a^2}\frac{\partial}{\partial t}\int \frac{\mathbf{C}\left(\mathbf{r}', t-\left|\boldsymbol{r}-\mathbf{r}'\right|/a\right)}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3r' = \nabla\varphi_k + \frac{1}{a^2}\frac{\partial \mathbf{A}_k\left(\mathbf{r}, t\right)}{\partial t}\dots$$

This had the noteworthy advantage of setting the sum of two retarded functions on the left equal to an infinite series in unretarded ones on the right. That allowed the introduction of a new, *retarded* function, $\mathbf{A}_{lnz}$, for the vector potential such that:

$$\boldsymbol{A}_{lnz}\left(t, \boldsymbol{r}\right) \equiv \int \frac{\boldsymbol{C}\left(\boldsymbol{r}', t-\frac{|\boldsymbol{r}-\boldsymbol{r}'|}{a}\right)}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3r' \tag{14.21}$$

$$\nabla\varphi_{lnz} + \frac{1}{a^2}\frac{\partial \boldsymbol{A}_{lnz}}{\partial t} = \nabla\varphi_k + \frac{1}{a^2}\frac{\partial \boldsymbol{A}_k}{\partial t} - \dots \tag{14.22}$$

If the constant a is very large then the terms in the time derivatives of $\mathbf{A}_k$ in the unretarded right-hand series would be negligible, in which case contemporary experiments allowed replacing Kirchhoff's potentials in Ohm's law with Lorenz's retarded functions:

Lorenz's Retarded Version of Ohm's Law

$$\boldsymbol{C}\left(\boldsymbol{r}, t\right) = -2\sigma\left(\nabla\varphi_{lnz} + \frac{4}{c^2}\frac{\partial \boldsymbol{A}_{lnz}}{\partial t}\right) \tag{14.23}$$

Lorenz's route to 14.23 is hardly unproblematic because it depends on the reconfiguration that produced 14.22 in which an unretarded integral in the current $\mathbf{C}$ is unjustifiably reinterpreted as a retarded one.

Through his earlier work on the theory of light Lorenz knew that the integral over distance of a retarded function satisfies 14.24:

Lorenz's Equation for an Arbitrary Retarded Function

$$\left(\nabla^2 - \frac{1}{a^2}\frac{\partial^2}{\partial t^2}\right)\int \frac{F_{ret}}{|\boldsymbol{r}-\boldsymbol{r}'|}d^3r' = -4\pi F_{unret}(r,t) \tag{14.24}$$

Expression 14.24 can be applied directly to both φ_{lnz} and $\mathbf{A}_{lnz}$. Lorenz thereby transformed 14.23 into 14.25 on dropping the integrals over d^3r':

$$\left(\nabla^2 - \frac{1}{a^2}\frac{\partial^2}{\partial t^2}\right)\int \frac{\mathbf{C}\left(\boldsymbol{r}', t - \left|\boldsymbol{r}-\boldsymbol{r}'\right|/a\right)}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}^3r' = -4\pi\mathbf{C}(\mathbf{r},t)$$
$$= 8\pi\sigma\left(\nabla\varphi_{lnz} + \frac{4}{c^2}\frac{\partial \boldsymbol{A}_{lnz}}{\partial t}\right)$$

$$\left(\nabla^2 - \frac{1}{a^2}\frac{\partial^2}{\partial t^2}\right)\boldsymbol{C} = 8\pi\sigma\left(\nabla\rho + \frac{4}{c^2}\frac{\partial \boldsymbol{C}}{\partial t}\right) \tag{14.25}$$

With 14.25 Lorenz had an apparent transverse wave equation for current, albeit one with a source in the gradient of the charge density and a term that produces absorption. In his words, "*periodical* electrical currents are possible, … such ones travel like a *wave-motion* … and like light, make vibrations which are at right angles to the direction of propagation."[25] Longitudinal waves have been avoided, which presumably means that the divergence of the current must be zero, in which case the charge density and so its gradient are independent of time given continuity. That perhaps suggested a way to eliminate the source term in $\nabla\rho$ in order to reach equations with the same form as the ones he had previously generated for light.

To reach his goal, Lorenz first created a linking function R_{lnz} by interpreting charge and current in the continuity Eq. 14.19 as retarded quantities. For simplicity we use square brackets to denote retardation:

$$\nabla_{r'}\cdot[\boldsymbol{J}] + \frac{1}{2}\frac{\partial[\rho]}{\partial t} = 0 \ [\mathit{retarded\ continuity}]$$

from which he obtained "by partial integration"

$$R_{lnz} : \frac{\partial\varphi_{lnz}}{\partial t} + 2\left(\nabla_r \cdot \boldsymbol{A}_{lnz}\right) = 0$$

[25] (Lorenz 1867a, 293). Lorenz did not complicate matters by allowing the charge or current sources to move, so that the distances $\boldsymbol{r}-\boldsymbol{r}'$ are not themselves dependent on time.

Lorenz properly noted that the divergence in the continuity equation should be taken at the current loci, r'. However, we shall see that in performing the "partial integration" he in the end used the observation point r in calculating the divergence of $\boldsymbol{A}_{lnz}$.[26] Next apply the general formula (14.24) for retarded functions to the definition (14.21) of $\boldsymbol{A}_{lnz}$ to generate 14.26. Combine that result with R_{lnz} and Lorenz's retarded Ohm law (14.23) to produce the requisite propagation Eq. 14.27 for waves of electric "current" without charge appearing as a source:

$$\frac{1}{a^2}\frac{\partial^2 \boldsymbol{A}_{lnz}}{\partial t^2} = \nabla^2 \boldsymbol{A}_{lnz} + 4\pi \boldsymbol{C} \tag{14.26}$$

Lorenz's Propagation Equation for Currents

$$\nabla \times (\nabla \times \boldsymbol{C}) = \frac{1}{a^2}\frac{\partial^2 \boldsymbol{C}}{\partial t^2} + \frac{16\pi\sigma}{a^2}\frac{\partial \boldsymbol{C}}{\partial t} \tag{14.27}$$

If the conductivity is at least very small, as it would have to be in otherwise empty space, then setting a to $c/\sqrt{2}$ transforms the equations formally into the very ones that Lorenz had "already found for the components of light."[27] For transversality to hold, the current must again lack divergence.

Lorenz's R_{lnz} is clearly not a *relation* because it follows directly from the application of continuity to retarded charge and current. It is a *condition.* In principle this is hardly surprising. We saw above that the typical modern route proceeds by using the permissible constraint R_{HH} to uncouple a pair of equations generated by the basic ones for the classical fields, yielding d'Alembertians for the potentials with charge and current as sources. If we *start* with these potentials and continuity in the *unretarded form* 14.19 then R_{HH} certainly does follow. Lorenz however reached R_{lnz} without directly employing the associated d'Alembertians.

Lorenz did not provide full details, remarking only that he obtained R_{lnz} by applying "partial integration" to the *retarded* continuity equation. Working in that manner poses a problem due to what seems to have been an incorrect transformation of divergence under retardation. A correct application is as follows:

[26] Lorenz's coordinates for the observation point are *x,y,z* and for the current locus *x',y',z'*, while his expression for the divergence in R_{lnz} is taken with respect to *x,y,z*.

[27] The speed of propagation would then be $3.1 * 10^8 m/s$ given Weber's and Kohlrausch's value for c. Lorenz remained uncommitted about just what occupied space in order to have "currents", remarking that "there is scarcely any reason for adhering to the hypothesis of an aether; for it may well be assumed that in the so-called vacuum there is sufficient matter to form an adequate substratum for motion" (Lorenz 1867a, 301).

Retarded continuity equation for sources at r'

$$[\nabla_{r'} \cdot \mathbf{C}] + \frac{1}{2}\left[\frac{\partial \rho}{\partial t}\right] = [\nabla_{r'} \cdot \mathbf{C}] + \frac{1}{2}\frac{\partial\,[\rho]}{\partial t} = 0$$

Then deploy the correctly-retarded divergence expression[28]

$$[\nabla_{r'} \cdot \mathbf{C}] = \nabla_{r'} \cdot [\mathbf{C}] + \nabla_r \cdot [\mathbf{C}] \Rightarrow \nabla_{r'} \cdot [\mathbf{C}] + \nabla_r \cdot [\mathbf{C}] = -\frac{1}{2}\frac{\partial\,[\rho]}{\partial t}$$

Next take the retarded time derivative of the scalar potential

$$\left[\frac{\partial \varphi}{\partial t}\right] = \frac{\partial\,[\varphi]}{\partial t} = \int \frac{\left[\frac{\partial \rho}{\partial t}\right]}{|\boldsymbol{r} - \boldsymbol{r}'|} d^3 r' = \int \frac{\frac{\partial [\rho]}{\partial t}}{|\boldsymbol{r} - \boldsymbol{r}'|} d^3 r'$$

and so

$$R_{lnz}^{correct} = \frac{\partial\,[\varphi]}{\partial t} + 2\int \frac{\nabla_{r'} \cdot [\mathbf{C}] + \nabla_r \cdot [\mathbf{C}]}{|\boldsymbol{r} - \boldsymbol{r}'|} d^3 r' = 0$$

To obtain R_{lnz} requires discarding $\nabla_{r'} \cdot [\mathbf{C}]$. Doing so breaks the character of his system and, as a result, the possibility of establishing a formal, if retrospective, connection between his scheme and classical electrodynamics.[29] That Lorenz missed the form $R_{lnz}^{correct}$ is hardly surprising. The question of how to take retardation into account, for example, prompted an exchange between Carl Neumann (1832–1925) and Rudolf Clausius (1822–1888) in 1868, for Neumann had essayed retardation that year, albeit without going further in Lorenz's fashion.[30] In the end, Lorenz's propagation Eq. (14.27) for current depends on two problematic steps: first, the

[28] (Heras 2007, 656) provides the correct expression for the divergence under retardation. He notes that operating on functions of $\boldsymbol{r}'$, t' can be misleading because they depend explicitly on source coordinates through $\boldsymbol{r}'$ and implicitly through $t' = t - |\boldsymbol{r} - \boldsymbol{r}'|/a$.

[29] (Heras 2007) demonstrates that starting from a continuity equation (absent the factor ½) between otherwise-unspecified functions ρ_{fml},$\mathbf{C}_{fml}$, it is possible to define retarded scalar and vector fields that satisfy equations formally similar to the Maxwell four of classical electrodynamics provided that the requisite care is taken in developing identities with retarded functions.

[30] Neumann followed Bernhard Riemann (1826-1866), who had "[in a posthumous publication] presented a (faulty) derivation of the Weber law by assuming a conservative electrodynamics force, and by positing that the potential associated with this force was propagated with constant velocity" (Archibald 1989a, 787). Neumann's calculation took the distance to be used in computing the potential when it is received at a given point as the distance *at that time* between the point and

feasibility of replacing $\mathbf{A}_k$ with $\mathbf{A}_{lnz}$ in Ohm's law, which follows from his refiguring of a term as a retarded variable in order to obtain Eq. 14.22, and second, an apparently incorrect deployment of the continuity equation under retardation.

Lorenz's paper was known shortly after its publication, not least by Maxwell, whose developed account of the electromagnetic field had appeared 4 years earlier.[31] In an addendum to a paper on electric units, he briefly argued that Lorenz's (or any other) theory based on the retarded propagation of potentials would necessarily violate action and reaction.[32] Lack of familiarity with the conditions that obtain under retardation vitiates Maxwell's argument – as, for different reasons, it had Lorenz's own derivations. Years later Heinrich Hertz (1857–1894) also referred to theories that presuppose the propagation of potentials.[33]

14.4 The Emergence of Gauge

In 1900 Emil Wiechert (1861–1928) published an article on the elementary laws of electrodynamics that broached the gauge freedom permitted in field theory.[34] He began with a set of initially uninterpreted equations designed to capture the propagation of transverse waves in free space. Only later in the article did Wiechert draw a connection to electrodynamics. Nevertheless, for clarity the following express his equations using electric and magnetic field vectors.

$$\frac{\partial^2 \boldsymbol{E}}{\partial t^2} = c^2 \nabla^2 \boldsymbol{E} \tag{14.28}$$

$$\nabla \cdot \boldsymbol{E} = 0 \tag{14.29}$$

$$\frac{\partial \boldsymbol{H}}{\partial t} = -c\, \nabla \times \boldsymbol{E} \tag{14.30}$$

$$\nabla \cdot \boldsymbol{H} = 0 \tag{14.31}$$

---

a moving emitter, rather than the distance traveled between the two from emission to reception, which was the essence of Clausius' critique (*ibid.,* 789).

[31] (Maxwell 1890a).

[32] (Maxwell 1890b, 137–38).

[33] (Hertz 1884).

[34] (Wiechert 1900).

Wiechert remarked that his 14.28 and 14.29 allow the introduction of 14.32 as an "analog" to 14.30, and thence 14.33 (given 14.30 and 14.31):

$$\frac{\partial \boldsymbol{E}}{\partial t} = c\,\nabla \times \boldsymbol{H} \tag{14.32}$$

$$\frac{\partial^2 \boldsymbol{H}}{\partial t^2} = c^2 \nabla^2 \boldsymbol{H} \tag{14.33}$$

He then introduced the vector potential **A** (as usual given 14.31):

$$\boldsymbol{H} = \nabla \times \boldsymbol{A} \tag{14.34}$$

Equation 14.30 then links **A** to **E** *via* 14.35:

$$c\nabla \times \boldsymbol{E} = -\partial \nabla \times \boldsymbol{A}/\partial t \tag{14.35}$$

At this point Wiechert asserted (presumably on the basis of 14.35) that "the entire system shows that c**E** can only differ from the vector $-\partial\mathbf{A}/\partial t$ by a vector part that has a scalar potential," in which case **E** satisfies 14.36, the Faraday Law in the form 14.6 used today:

$$\boldsymbol{E} = -\nabla\varphi - \partial \boldsymbol{A}/\partial t \tag{14.36}$$

Putting all of this together, he obtained the following propagation equation for the vector potential:

$$\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = c^2 \nabla^2 \boldsymbol{A} - c\nabla\left(\frac{\partial \varphi}{\partial t} + c\,\nabla \cdot \boldsymbol{A}\right) \tag{14.37}$$

Finally, Wiechert claimed that the "indeterminacy" in **A** permits dropping the second term on the right in 14.37, thereby yielding homogeneous propagation equations for both **A** and φ. Later in the paper he extended the result to include charge and current, whereby the d'Alembertians for the potentials become inhomogeneous (*viz* 14.12 and 14.13).[35]

Wiechert had certainly shown that propagating potentials follow on the basis of equations that encompass electrodynamics provided that the expression $\frac{\partial \varphi}{\partial t} + c\nabla \cdot \mathbf{A}$ vanishes. However, to do so he merely asserted that "indeterminacy" permits that sum to vanish. Which would be permissible provided that Eq. 14.36 for **E** had itself

[35] (Wiechert 1900, 557–59).

been justified by appropriate modifications of the two potentials. But Wiechert had not done so and had therefore not developed a scheme that fully deploys the system's inherent gauge freedom. What's missing is the introduction of a scalar function equivalent to Λ in the modern expressions 14.9 and 14.10 which generalizes the potentials in a manner that retains the form 36 of the Faraday law. Absent such an addition, the system's gauge freedom remains obscure since it has not been shown that 14.36 remains tenable under the requirement that $\frac{\partial\varphi}{\partial t}+\mathrm{c}\nabla\cdot\mathbf{A}$ vanishes. Wiechert had instead *assumed* 14.36 as reasonably justified by 14.35 without investigating what that generally requires of the potentials and had then moved directly from it to 14.37. A crucial step is missing.

Three years later, Lorentz published a detailed pair of articles developing electrodynamic field theory. The first of the two presented the essentials of the theory, while the second extended it to the case of moving, charged particles governed by what came to be known as the Lorentz force.[36] Lorentz had read Wiechert's 1900 piece, which he referred to, and in the second article he developed in full the general condition that gauge freedom permits through the introduction of a function (his χ) equivalent to Λ. Here we find for the first time a complete recognition that the scalar and vector potentials are both purely auxiliary quantities that can be introduced on the bases of the zero divergence of the magnetic field and the proportionality of its time derivative to the curl of the electric field. Imposing the R_{HH} relation then produces a d'Alembertian for χ, maintaining thereby the Faraday Law in the form 14.36 and generating propagation equations for the modified potentials (Fig. 14.1). Both articles appeared in the widely-read *Encyclopädie der Mathematischen Wissenschaften*.

We remarked above that for many years the R_{HH} relation was (correctly) attributed to Lorentz. Jackson and Okun surveyed prior appearances of similar equations and concluded that what is "universally called the "Lorentz condition," is seen to originate with Lorenz more than 25 years before Lorentz."[37] Yet two aspects of Lorenz's work militate against such an attribution. First, Lorenz's route to a central element of his theory, namely Ohm's law in terms of retarded functions, was itself problematic. But second, Lorenz derived his version R_{lnz} as a *condition* that arises directly out of the continuity equation under the assumption that charge and current are retarded functions. In doing so complexities that arise in taking divergences under retardation raised issues. That result was critical for his theory, because with it he was able to obtain equations of the same form that he had previously generated for light.

Although a non-retarded equation of continuity combined with d'Alembertians for the scalar and vector potentials does entail R_{HH}, Lorenz's route to an equation (R_{lnz}) formally similar to the latter is problematic. Even were it not, the resulting expression is not a gauge transformation because it cannot be otherwise since the theory is grounded on retarded potentials. Lorenz had therefore neither introduced

[36] (Lorentz 1903a, b).

[37] (Jackson and Okun 2001).

(VII) $$\Delta\varphi - \frac{1}{c^2}\ddot{\varphi} = -\varrho,$$

(VIII) $$\Delta\mathfrak{a} - \frac{1}{c^2}\ddot{\mathfrak{a}} = -\frac{1}{c}\varrho\mathfrak{v},$$

und es ist

(IX) $$\mathfrak{d} = -\frac{1}{c}\dot{\mathfrak{a}} - \operatorname{grad}\varphi,$$

(X) $$\mathfrak{h} = \operatorname{rot}\mathfrak{a}.$$

Zwischen den beiden Potentialen besteht die Relation

(2) $$\operatorname{div}\mathfrak{a} = -\frac{1}{c}\dot{\varphi}.$$

Man gelangt zu diesen Formeln in folgender Weise. Aus (V) geht sofort hervor, daß die magnetische Kraft sich in der Form (X) darstellen läßt, und es folgt dann weiter aus (IV)

$$\operatorname{rot}\left(\mathfrak{d} + \frac{1}{c}\dot{\mathfrak{a}}\right) = 0,$$

was auf (IX) führt. Hierbei bleiben, obgleich bei jeder elektromagnetischen Erscheinung $\mathfrak{d}$ und $\mathfrak{h}$ bestimmte Funktionen von x, y, z, t sind, die Potentiale $\mathfrak{a}$ und φ teilweise unbestimmt. Diesen Übelstand heben wir, indem wir dieselben der weiteren Bedingung (2) unterwerfen. Sind nämlich $\mathfrak{a}_0$ und φ_0 irgend welche mit den Gleichungen (IX) und (X) verträgliche Potentiale, so ist jedes andere zulässige Funktionenpaar $\mathfrak{a}$, φ von der Form

$$\mathfrak{a} = \mathfrak{a}_0 - \operatorname{grad}\chi, \quad \varphi = \varphi_0 + \frac{1}{c}\dot{\chi}.$$

Hier läßt sich nun die skalare Funktion χ so bestimmen, daß der Gleichung (2), oder

$$\Delta\chi - \frac{1}{c^2}\ddot{\chi} = \operatorname{div}\mathfrak{a}_0 + \frac{1}{c}\dot{\varphi}_0$$

genügt wird. Man gelangt schließlich zu (VII) und (VIII), wenn man die Werte (IX) und (X) in (I a) und (III) einführt und dabei (2) berücksichtigt.

Fig. 14.1 Lorentz's introduction of gauge freedom. (Lorentz 1903b, 157)

gauge as a *relation* nor was his analysis free from internal difficulties. The virtue of his condition, setting aside the problematic route to it, consists in Lorenz having connected it to a system that could yield the same form of wave equation that he had earlier produced for light. Wiechert reached further relying on electrodynamics, but his introduction of a gauge relation remained incomplete since it lacked the necessary scalar and its resulting d'Alembertian. Only H. A. Lorentz achieved full generality, with propagation for the scalar and vector potentials emerging naturally and easily. As questions concerning radiation became technologically significant following Guglielmo Marconi's (1874–1937) and John Ambrose Fleming's (1849–1945) efforts, engineer-physicists adopted what was widely known as the relation developed by Lorentz in order to generate malleable equations for the field generated by arrays of antennas.[38] Lorentz's is the first theory to have recognized the full generality allowed by the link between scalar and vector potentials, and so it seems appropriate that R_{HH} ought to return to its previous characterization as R_{LNTZ}, the "Lorentz gauge." The development of gauge as a general property of physical theories evolved into a profound mathematical property connected to conservation principles in the hands of Emmy Noether (1882–1935).[39]

Nevertheless, if we separate Lorenz's system altogether from electrodynamic field theory, there is a reasonable sense in which he was the first to attempt development of a scheme based directly on presumptive retarded interactions between charged particles, though his was flawed in several significant ways. Field theory can be developed on such a basis – with a major caveat: namely, that the introduction of a measure for what is otherwise field energy enters as a separate structure for purposes of practical application. A pure theory dispenses altogether with field ontology and presumes only point charges and both retarded and advanced interactions along the lines argued for by John Wheeler (1911–2008) and Richard Feynman (1918–1988).[40]

Acknowledgments I thank, Karine Chemla, Anne Kox, Andrew Zangwill, and two anonymous reviewers for their suggestions on drafts of this piece.

Bibliography

Archibald, Thomas. 1989a. Carl Neumann versus Rudolph Clausius on the Propagation of Electrodynamic Potentials. *American Journal of Physics* 54: 786–790.

———. 1989b. Physics as a Constraint on Mathematical Research: The Case of Potential Theory and Electrodynamics. In *The History of Mathematics: Institutions and Applications*, ed. D.E. Rowe and J.M. Cleary, 28–75. Academic.

Assis, A.K.T. 1994. *Weber's Electrodynamics*. Kluwer Academic Publishers.

Buchwald, J. 1985. *From Maxwell to Microphysics. Aspects of Electromagnetic Theory in the Last Quarter of the Nineteenth Century*. Chicago: The University of Chicago Press.

———. 1993. Helmholtz's Electrodynamics in Context: Object States, Laboratory Practice and Anti-Idealism. In *Hermann von Helmholtz and the Foundations of Nineteenth-Century Science*, ed. D. Cahan, 334–373. University of California Press.

Darrigol, O. 2000. *Electrodynamics from Ampère to Einstein*. Oxford: Oxford University Press.

Helmholtz, Hermann. 1858. Ueber Integrale Der Hydrodynamischen Gleichungen. *Journal für die reine und angewandte Mathematik* 55: 25–55.

Heras, J.A. 2007. Can Maxwell's Equations Be Obtained from the Continuity Equation? *American Journal of Physics* 75: 652–657.

[38] On Marconi and Fleming see (Hong 2001). The early development of radio wave transmission is analyzed in (Yeang 2013). For an example of the manner in which the radiation gauge has been used by engineering physicists see, e.g., (Skilling 1942).

[39] On Noether see (Rowe 2021).

[40] (Wheeler and Feynman 1949). (Lazarovici 2018) argues for the irremediable inconsistency of an electrodynamics grounded on fields.

———. 2011. A Short Proof That the Coulomb-Gauge Potentials Yield the Retarded Fields. *European Journal of Physics* 32: 213–216.
Hertz, Heinrich. 1884. Über Die Beziehungen Zwischen Den Maxwell'schen Elektrodynamischen Grundgleichungen Der Gegnerischen Elektrodynamik. *Annalen der Physik und Chemie* 23: 84–103.
Hong, S. 2001. *Wireless: From Marconi's Black-Box to the Audion*. Cambridge, MA: MIT Press.
Hunt, B. 1991. *The Maxwellians*. Ithaca/London: Cornell University Press.
Jackson, J.D. 2002. From Lorenz to Coulomb and Other Explicit Gauge Transformations. *American Journal of Physics* 70: 917–928.
Jackson, J.D., and L.B. Okun. 2001. Historical Roots of Gauge Invariance. *Reviews of Modern Physics* 73: 663–680.
Kirchhoff, G. 1882a. Ueber Die Bewegung Der Elektricität in Drähten (1857). In *Gesammelte Abhandlungen von G. Kirchhoff*, 131–153. Leipzig: Barth.
———. 1882b. Ueber Die Bewegung Der Elektricität in Leitern (1857). In *Gesammelte Abhandlungen von G. Kirchhoff*, 154–167. Leipzig: Barth.
Kragh, H. 2018. *Ludvig Lorenz. A Nineteenth-Century Theoretical Physicist*. The Royal Danish Academy of Sciences and Letters.
Lazarovici, Dustin. 2018. Against Fields. *European Journal for Philosophy of Science* 8: 145–170.
Lorentz, H.A. 1892. *La Théorie Électromagnétique de Maxwell et Son Application Aux Corps Mouvants*. E. J. Brill.
———. 1903a. Maxwells Elektromagnetische Theorie. In *Encyklopädie Der Mathematischen Wissenschaften*, 63–144. Leipzig und Berlin: B. G. Teubner.
———. 1903b. Weiterbildung Der Maxwellschen Theorie. Elektronentheorie. In *Encyklopädie Der Mathematischen Wissenschaften*, 145–280. Leipzig und Berlin: B. G. Teubner.
Lorenz, L. 1867a. On the Identity of the Vibrations of Light with Electrical Currents. *The London, Edinburgh, and Dublin Philosophical Magazine* 34: 287–301.
———. 1867b. Ueber Die Identität Der Schwingungen Des Lichts Mit Den Elektrischen Strömen. *Annalen der Physik und Chemie* 131: 243–262.
Maxwell, James Clerk. 1873. *A Treatise on Electricity and Magnetism*. 3rd ed. Oxford: Clarendon Press.
———. 1890a. A Dynamical Theory of the Electromagnetic Field. In *The Scientific Papers of James Clerk Maxwell*, 526–597. Cambridge: Cambridge University Press.
———. 1890b. On a Method of Making a Direct Comparison of Electrostatic with Electromagnetic Force; with a Note on the Electromagnetic Theory of Light. In *The Scientific Papers of James Clerk Maxwell*, 125–143. Cambridge: Cambridge University Press.
Rowe, David. 2021. *Emmy Noether - Mathematician Extraordinaire*. Cham: Springer.
Skilling, Hugh. 1942. *Fundamentals of Electric Waves*. New York: Wiley.
Thorne, Kip S., and Roger D. Blandford. 2017. *Modern Classical Physics*. Princeton: Princeton Univeristy Press.
Wheeler, John Archibald, and Richard Phillips Feynman. 1949. Classical Electrodynamics in Terms of Direct Interparticle Action. *Reviews of Modern Physics* 21: 425–433.
Wiechert, E. 1900. Elektrodynamische Elementargesetze. *Archives Néerlandaises* 5: 549–573.
Yaghjian, A.D. 2014. Reflections on Maxwell's Treatise. *Progress in Electromagnetics Research* 149: 217–249.
Yeang, Chen-Pang. 2013. *Probing the Sky with Radio Waves*. Chicago: The University of Chicago Press.
Zangwill, Andrew. 2013. *Modern Electrodynamics*. Cambridge: Cambridge University Press.

Chapter 15
Ronald Ross and Hilda Hudson: A Collaboration on the Mathematical Theory of Epidemics

June Barrow-Green

15.1 Introduction

In 1916 Ronald Ross published the first of three Royal Society papers on the mathematical study of epidemiology or, as he called it, 'pathometry'. The second and third of these papers appeared the following year co-authored with the mathematician Hilda Hudson. At the time Hudson, who had ranked equivalent to the seventh wrangler in the 1903 Cambridge Mathematical Tripos,[1] was known for her work on Cremona Transformations. So how and why did Hudson, a geometer, end up collaborating with Ross, a medical man, on the theory of epidemics? And what role did she play in the collaboration? In this chapter, I discuss the nature and extent of their collaboration, setting it into the broader context of Ross's mathematical aspirations.

[1] Wranglers were students who were in the first class of the Mathematical Tripos with their results being listed in order of merit. From 1881 women had had the right to sit the Mathematical Tripos but it was not until 1948, when they were admitted as full members of the University, that they could be awarded degrees and ranked with the men.

J. Barrow-Green (✉)
School of Mathematics & Statistics, The Open University, Milton Keynes, UK
e-mail: june.barrow-green@open.ac.uk

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_15

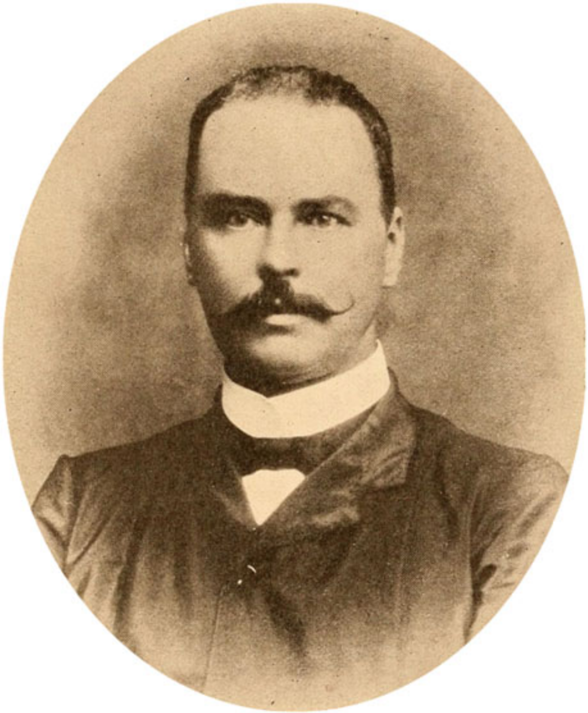


Fig. 15.1 Ronald Ross (1857–1932)

15.2 Ronald Ross Discovers Mathematics

Ronald Ross (Fig. 15.1) is well-known in medical circles for his discovery of the role of mosquitoes in malaria transmission. Trained as a doctor, he had joined the Indian Medical Service in 1881 and it was while he was in India in 1897 that he made the break-through discovery for which he received the Nobel Prize for medicine in 1902.[2] Less well known is the fact that earlier, but also while in India, he discovered an enthusiasm for mathematics. As he recalled:

> One day [in 1882] I commenced to read an old prize book which I had won at Springhill School for mathematics in 1871, called *The Orbs of Heaven*, by O. M. Mitchell, which eloquently described the great mathematical triumphs of the astronomers; and I was so fired by the theme that I determined then and there to study mathematics. I bought nearly

[2] The mosquito-malaria theory, which had been originally developed by the Scottish physician Patrick Manson in 1894, was experimentally proved by Ross in 1897. In his acceptance speech for the Nobel Prize, Ross did not acknowledge Manson's role in the discovery, an omission which led to a notorious falling out between the two men, see Chernin (1988b). Manson, and the French physician Charles Laveran, were also nominated for the 1902 Nobel Prize for their work on malaria.

> all Todhunter's series[3] and, literally, went through them in a few weeks, up to the end of the Calculus of Variations, though I had not advanced beyond Quadratic Equations at school. ... I can scarcely describe my enthusiasm. It was an aesthetic as well as an intellectual enthusiasm. (Ross 1923, 49)

Ross went on to say that it was then that he first began thinking of the "possible application [of mathematics] for explaining why epidemics of disease exist" (Ross 1923, 50). Shortly afterwards, while in Madras, he acquired a copy of Kelland and Tait's *Introduction to Quaternions* (1873) which inspired him to develop his own system of what he called 'vector-geometry'. In 1891, having become friends with Tait's eldest son, who was in the Mysore Education Department, he felt sufficiently emboldened to send his work to Tait for an opinion. Tait, who had by then given up on the study of quaternions, was, according to Ross in his *Memoirs*, "most discouraging", and Ross put the work to one side (Ross 1923, 94). In fact, Tait had told Ross that he had shown the paper to "one of the very greatest authorities on such subjects" who had said that in view of the number of similar geometries, Ross first needed to provide "examples of [the method's] power in attacking well-known problems, or of its fertility in producing new theorems".[4]

Ross returned to England in 1899 to take up a lectureship at the Liverpool School of Tropical Medicine where he continued to work on malaria. He sought help on mathematics from Frank Carey, the professor of mathematics at Liverpool University College (later Liverpool University).[5] Carey directed him to Whitehead's *Universal Algebra* (1898), from which Ross discovered that he had been pre-empted by Grassmann.[6] Undaunted, he revised his work on vector geometry—he described it as a combination of "Grassmann's system with Hamilton's quaternions" (Ross 1923, 415)—before reading it in front of the Liverpool Mathematical Society and publishing it at his own expense (Ross 1901).[7]

15.3 Ross Begins Work on Mathematical Epidemiology

In 1904 Ross was invited to speak at the International Congress of Arts and Sciences in St. Louis. The Congress, which was part of the World's Fair, attracted many

---

[3] Isaac Todhunter was the leading nineteenth century British writer of mathematics textbooks, see Barrow-Green (2001).

[4] Letter from Tait to Ross, 27 January 1892, LSHTM Ross/163/14/04.

[5] Carey, who had graduated as third wrangler in the Cambridge Mathematical Tripos in 1883, had been appointed to the professorship in 1886.

[6] A note penned by Ross on the envelope containing the reply from Tait concludes: "After I returned to England in 1899, I read the *Universal Algebra* of A.N. Whitehead F.R.S. and found the whole of my method set forth in it under the name of Grassmann's work. But of course I cannot prove, & do not even suggest, that Whitehead was the expert to whom Tait sent my paper." LSHTM Ross/163/14/13.

[7] Even though the paper attracted little interest at the Society, Ross persisted with the topic and in 1918 published a new, albeit unfinished, version (Ross 1918), which was reprinted in 1930.

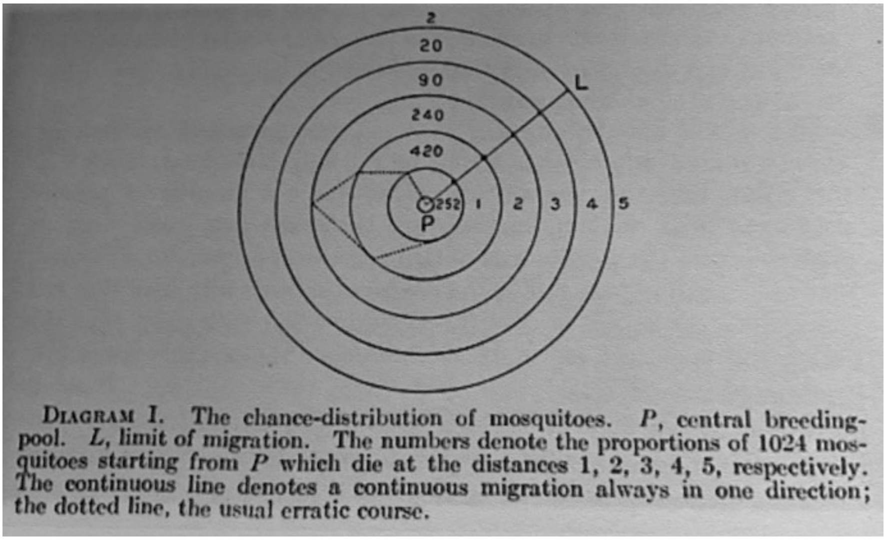


Fig. 15.2 Diagram from Ross (1905a, 1025)

distinguished scientists and mathematicians from Europe,[8] and Ross, who was speaking in the Section of Preventative Medicine on the topic of mosquito reduction, drew a large audience. But those who were expecting a medical talk were in for a surprise. For by now Ross had become convinced that mosquito reduction could not be tackled scientifically without mathematical analysis (Ross 1905a, 1028–1029). His talk was primarily mathematical with the result that there were "hundreds of disappointed doctors who did not understand a word" (Ross 1923, 491).

In his talk, Ross discussed the following problem:

> Suppose a box containing a million gnats were to be opened in the centre of a large plain, and that the insects were allowed to wander freely in all directions, how many of them would be found after death at a given distance from the place where the box was opened? (Ross 1905a, 1026).

Considering the problem to be one in probability, he tentatively proposed that the number of gnats would be greatest near the box and that this number would reduce as a function of the distance away from the box (Fig. 15.2), a principle he named the "law of random migration" (Ross 1905a, 1027).

Although his solution was not in line with current thinking,[9] his results did agree with those of a talented young mathematician in Liverpool, Ronald Hudson

[8] As Gray has described, Henri Poincaré was one of the stars of the Congress. It was there that Poincaré "came as close as he ever would to producing a theory of electrodynamics that would have rivalled Einstein's a year later." (Gray 2008, 173–174). See also Gray (2013, 104–107).

[9] It was generally believed that if mosquitoes were cleared from one spot, then other mosquitoes would rush in from outside to fill the vacuum.

Fig. 15.3 Ronald Hudson (1875–1904). *The Daily Graphic*, 23 September 1904

(Fig. 15.3), who, at Ross's request, had been working on a similar analysis (Ross 1905b, 151).[10] Hudson died in September 1904 and Ross, recognising that he himself did not have the mathematical knowledge required for a complete solution, turned to Karl Pearson for help. Pearson, with the assistance of one of his students, was eventually able to oblige (Pearson and Blakeman 1906).[11]

But it was in 1908, in a *Report on the Prevention of Malaria in Mauritius*, that Ross first began to use mathematics in the study of infectious disease, and, as the epidemiologist Paul Fine wrote, it is "In this document, we find the first clear formulation of Ross's great contribution to epidemiological methodology" (Fine 1975, 2). And it is here that Ross introduced the term 'pathometry' (Ross 1908, 30). By pathometry he meant epidemiology in the case of a priori studies

[10] Letter from Ronald Hudson to Frank Carey, 8 July 1904, LSHTM Ross 163/01/02. Ronald Hudson was Hilda Hudson's brother and will be mentioned again in Sect. 15.5.

[11] Initially Pearson too found the problem beyond him. He told Ross that it would require "a strong mathematical analyst" to solve it, but that to find such a person he [Pearson] would have to "restate it as a chessboard problem or something of that sort in order to get mathematicians to work at it!" He told Ross that he would contact him again after he had "seen what a letter to *Nature* will do on the general problem." Letter from Pearson to Ross, 21 July 1905, LSHTM Ross 163/01/12. Pearson based his estimates of mosquito densities on the value of certain constants which had been assumed by Ross on a basis of general probability (as opposed to being determined). It was in the letter to *Nature* that Pearson, prompted by Ross's principle of 'random migration', introduced the term 'random walk' into mathematics (Pearson 1905a,b). Thus, in a broad sense, Ross can be considered an originator of the term.

although, as he would later write, "like most [of his] other suggestions" the term never caught on.[12] Ross, by attacking the problem using a priori methods—that is assuming a knowledge of the causes, constructing the equations, and testing the calculated results with observed statistics—was striking new ground. Previous studies had begun with the statistics, and worked backwards to the underlying causes. Importantly, by using mathematics Ross had been able to show that malaria transmission could be prevented if the mosquito population could be reduced to below a certain threshold (Ross 1908, 30–37). However, he had an upward struggle to convince medical and public health professionals of this. Fighting shy of mathematics, they believed it was necessary to eradicate the entire mosquito population for malaria to die out.

The mathematical analysis in the *Report*, which was only brief, led Ross to a simple algebraic equation for defining the number of new infections of malaria in a month (Ross 1908, 31–33).[13] The real development came 3 years later in the second edition of his book *The Prevention of Malaria* in a lengthy mathematical addendum entitled 'Theory of Happenings' and which formed a new Section 66 of the book Ross (1911a, 651–686). In Section 28 of the first edition, published the year before, he had used results from the *Report* to deduce an elementary difference equation (Ross 1910, 156–164). Now he elaborated his mathematical theory applying his reasoning to infectious diseases in general (as opposed to malaria in particular). Convinced of the importance of mathematics for epidemiology, he introduced Section 66 as follows:

> The mathematical treatment adopted in Section 28 has been met with some questioning by critics. Some have approved of it, but others think that it is scarcely feasible owing to the large number of variables which must be considered. As a matter of fact all epidemiology, concerned as it is with the variation of disease from time to time or from place to place, must be considered mathematically, however many variables are implicated, if it is to be considered scientifically at all. . . . And the mathematical method of treatment is really nothing but the application of careful reasoning to the problems at issue. . . . I am convinced that many readers will be able to follow the work without difficulty. (Ross 1911a, 651)

Given an epidemic, he constructed a model to determine the numbers of unaffected and affected individuals after a certain period of time, allowing for the natural fluctuations of a population (birth and death, immigration and emigration), and for the variance of those fluctuations as a consequence of the epidemic. Taking the case where a particular event, such as an act of infection, insect-bite, etc., occurs

[12] Unpublished manuscript (p. 3), 31 October 1931, RCPSG 9/M/9/1/48. In the same manuscript Ross expressed his preference for keeping the term 'epidemiology' for a posteriori studies. In 1928 he wrote that "by pathometry I mean the *quantitative* study of disease" (Ross 1928, 154). Thus it seems that for Ross the use of mathematics represented something 'quantitative' connected to causality in epidemics whilst the use of statistics represented something 'qualitative' in the sense that it did not describe the actual mechanism of causing infection.

[13] For a development of the mathematical ideas in Ross (1908), see Waite (1910), although correspondence between Ross and Pearson shows Ross was not satisfied with Waite's paper (LSHTM Ross 163/07/03, 163/07/10).

to a constant proportion of a population in a unit of time, the goal was to ascertain the number of affected individuals on a given date, how many had been affected twice, three times, etc. Ross called this "The Problem of Happenings", where the "happenings may be some kind of accident or disease, birth, death, marriage or anything else we can think of—vaccination, receipt of bequests, conversion to some creed, etc." (Ross 1911a, 655). His model also allowed for the calculation of the frequency of reinfections, something which had not been done before. Extending his earlier method, he arrived at a system of difference equations from which he could calculate the number of affected individuals in successive time intervals, such as a minute or a day. Finally, to remove the dependence on a particular time interval, he made the interval of time infinitesimally small, reformulating the model as a pair of coupled first order differential equations.[14]

Shortly after the publication of the addendum he was asked to provide an account of it for *Nature*. Although initially Ross had been convinced of the clarity of the mathematics in the addendum, in this account he is frank about his mathematical limitations, acknowledging that his work required "verification and completion by better mathematicians" (Ross 1911b, 466).[15] And once again he took the opportunity to have a dig at the medical profession:

> These studies require to be developed much further; but they will already be useful if they help to suggest a more precise and quantitative consideration of the numerous factors concerned in epidemics. At present medical ideas regarding these factors are generally so nebulous that almost any statements about them pass muster, and often retard or misdirect important preventative measures for years. (Ross 1911b, 467)

Many years later, when revisiting the addendum, he admitted in the *British Medical Journal* (*BMJ*) that it "contained many misprints and omissions" and that even he had "considerable difficulty in understanding parts of it" (Ross 1929, 674).[16]

15.4 Study of A Priori Pathometry: Part I

In March 1915, a short article appeared in the *BMJ* in which Ross announced that he had been able to extend his theory to the case where the population varies, and had been able to simplify the equations to the extent that they now gave "an elegant

[14] For a more detailed discussion of Ross (1911a), see Smith et al. (2012).

[15] One result of the *Nature* article was a response from the mathematician/mathematical demographer Alfred Lotka who proposed a closed-form solution to the system of differential equations obtained by Ross (Lotka 1912). Lotka further developed his analysis of Ross's equations in Lotka (1923).

[16] Ross was even more critical of his addendum in an unpublished manuscript, written in the same year as the *BMJ* article: "The Addendum was written in a great hurry ... the whole article was very confusedly written and was almost meaningless to readers and even to myself.", (Ross, Preface to 'Two-Party Aggregates'). Typescript dated 20 May 1929, RCPSG 9/M/17/1/5.

(though tentative) mathematical theory of both epidemic and endemic communicable diseases" (Ross 1915, 546). Having little confidence in the mathematical abilities of the medical profession, he gave only an outline of the mathematics, telling his readers that the full version was more suitable for a mathematical journal. The article caught the eye of the physician and statistician John Brownlee who had been working on epidemiology from an a posteriori perspective, and who, in a positive response, outlined his own most recent conclusions (Brownlee 1915).[17]

The full version of Ross's article was the first of the three Royal Society papers.[18] It began:

> It is somewhat surprising that so little mathematical work should have been done on the subject of epidemics, and, indeed, on the distribution of diseases in general. Not only is the theme of immediate importance to humanity, but it is one which is fundamentally connected with numbers, while vast masses of statistics have long been awaiting proper examination. But, more than this, many and indeed the principal problems of epidemiology on which preventative measures largely depend, such as the rate of infection, the frequency of outbreaks, and the loss of immunity, can scarcely ever be resolved by any other methods than those of analysis. (Ross 1916, 204–205)

He could hardly have made his position clearer. Extending his earlier ideas and clothing them in the language of the calculus, he now presented a more general development of his method, which he called "Theory of Happenings", framing the problem as follows:

> Suppose that we have a population of living things numbering P individuals, of whom a number Z are affected by *something* (such as a disease), and the remainder A are not so affected; suppose that a proportion $h.dt$ of the non-affected become affected in every element of time dt, and that, conversely, a proportion $r.dt$ of the affected become unaffected, that is, revert in every element of time to the non-affected group; and, lastly, suppose that both the groups, the affected and non-affected, are subject also to possibly different birth-rates, death-rates, and immigration and emigration rates in an element of time; then what will be the number of affected individuals, of new cases, and of the total population living at any time t? (Ross 1916, 208)

From this he was led to a system of three first order differential equations, the solution of which, under certain conditions, yield asymmetric bell-shaped curves rising more steeply than they fall, similar in shape to the curves that describe the spread of an epidemic.

There seems to have been little immediate reaction to Ross's paper. Ross did not expect the medical profession to take notice of it but almost no-one else did either. The trouble was that, from a mathematical point of view, it was not the mathematics itself that was interesting but the novel application of it. And the fact that the paper appeared in the middle of the War cannot have helped. The one

[17] In 1914 Brownlee had become the founding director of the statistics department of the UK Medical Research Council, having formerly been the physician-superintendent of the Glasgow Fever Hospital.

[18] Although papers submitted to the Royal Society were refereed and referee reports were usually archived, no report exists for this paper, possibly due to disruption caused by the War.

exception was the medical statistician Major Greenwood who drew attention to it in an article in *Nature* in which he discussed the general application of mathematics to epidemiology (Greenwood 1916). Greenwood, who was already familiar with Ross's earlier work, outlined Ross's method, praising "its simplicity and elegance" and highlighting the advantages of its generality, urging that further attention should be paid to the subject:

> No sensible man doubts the importance of investigations such as these; it is high time that epidemiology was extricated from its present humiliating position as the plaything of bacteriologists and public health officials, or as, at the best, a field for the display of antiquarian research. The work of Sir Ronald Ross ... and of a few others should at least elevate epidemiology to the rank of a distinct science." (Greenwood 1916, 244).

At the end of the preface to his Royal Society paper, Ross gave the section headings for the next part of the paper even though he had not yet finished it: Hypothetical Epidemics; Hypometric Happening; Parameter Analysis; Variable Happening.[19] But, some 15 months later in October 1916, when he submitted the next part to the Royal Society, it did not look exactly as he had predicted. Not only was there an additional section, but it was co-authored with the mathematician Hilda Hudson. What had happened to induce these changes and who was Hilda Hudson?

15.5 Hilda Hudson

Hilda Hudson (Fig. 15.4) was born in Cambridge in 1881 into a family of mathematical talent. Her mother, Mary Hudson (née Turnbull) (1843–1882) had, at the age of 30, attended the 'Lectures for Women in Cambridge' at the fledgling Newnham College where she was taught mathematics by her future husband, William Henry Hoar Hudson (1838–1915), then a mathematics lecturer at St John's College.[20] Her parents married in 1875, and in 1882, the year after Hilda was born and the year in which her mother died, her father was appointed professor of mathematics at King's College, London.

[19] Ross added the section headings in the proofs which indicates that he did the work between submitting the paper in July 1915 and correcting the proofs in October 1915. RCPSG 9/M/8/2/2.

[20] The Lectures for Women in Cambridge, which began in 1871, were organised by the Association for Promoting the Higher Education of Women in Cambridge, one of the two founding organisations of Newnham College. Thus Hilda's mother can be considered one of Newnham's earliest students. Hilda's father, who was on the General Committee of Management for Newnham, lectured on arithmetic and algebra at Newnham during the 1873/1874 session. I am grateful to Frieda Midgley, the archivist at Newnham College, who supplied me with information about Hilda's father.

Fig. 15.4 Hilda Hudson in 1918

Hudson's brother Ronald (mentioned above) was senior wrangler in 1898 and was considered the most gifted geometer of his generation in Cambridge,[21] and her elder sister, Winifred, was ranked equivalent to the eighth wrangler 2 years later. In 1903 Hudson herself went one better than her sister being ranked equivalent to the seventh wrangler.

In 1904, having passed Part II of the Tripos in the First Class—the only woman of her year to sit Part II—Hudson, encouraged by Arthur Berry (1862–1929), went to Berlin where she attended the lectures of Schwarz, Schottky and Landau. Since at the time it was quite unusual for British men, let alone women, to travel abroad for postgraduate mathematical study,[22] it is indicative of her talent that it was considered appropriate for her to do so. It seems likely that it was while she was in Berlin that she developed her deep interest in conformal transformations, a subject first introduced to her when she was at Cambridge and which in due course would become her main area of research, culminating with her treatise *Cremona Transformations in Plane and Space* for which she is now principally remembered

[21] The senior wrangler was the top student of the year in the Cambridge Mathematical Tripos. 1898 was a particularly strong year—James Jeans came second and G.H. Hardy fourth, both Jeans and Hardy having studied for only 2 years rather than the usual three. Ronald Hudson was killed in a climbing accident on Snowdon in 1904. For a discussion of his fine book *Kummer's Quadric Surface* (1905), see Barrow-Green and Gray (2006, 324–325).

[22] In the early 1890s Grace Chisholm had blazed a trail when she went from Cambridge to Göttingen to study for a doctorate with Felix Klein. Although she was successful, being awarded her doctorate in mathematics in 1895 and being the first English woman to accomplish such a feat, it was a trail few had followed.

(Hudson 1927).[23] Her obituarist, J.G. Semple, described her as having "a powerful, almost uncanny, geometrical intuition" (Semple 1969, 358).

After her return from Berlin, Hudson held appointments at Newnham, first as a lecturer and then as an Associate Research Fellow, spending the year 1912–1913 at Bryn Mawr (USA) in the department headed by the British geometer Charlotte Scott.[24] In 1912, just prior to her visit to the USA, Hudson gave a talk at the International Congress of Mathematicians (ICM) at Cambridge, becoming the first woman to give a talk at an ICM.[25] The following year, 1913, she was awarded an ScD degree from Trinity College, Dublin,[26] and 1913 was also the year in which she left Cambridge for London to take up a position as mathematics lecturer at the West Ham Technical Institute.[27] She remained at West Ham until 1917 when she responded to a call from the Admiralty to work in the Technical Section of the Air Department.[28] Meanwhile she was also active in the London Mathematical Society (LMS), collecting another first when, in 1917, she became the first woman to be elected to the LMS Council. After the war she worked for a short time in the aircraft industry before retiring from salaried employment and devoting herself to

[23] The treatise was well-received, being considered "a well-proportioned and carefully elaborated treatise on the whole subject" (Snyder 1927, 488) and "an indispensable authority and a source of inspiration to all workers in the field" (White 1928, 149). In the preface, Hudson mentions that it was Arthur Berry (1862–1929) who first introduced her to the subject.

[24] In 1880 Charlotte Scott was the first woman to be ranked equivalent to a wrangler when she achieved marks equal to that of the eighth wrangler. For details of Scott's work, see Lorenat (2020).

[25] Hudson was not the first woman to be invited to give a talk at an ICM. That honour goes to the Italian Laura Pisati who was invited to give a talk at the Rome ICM of 1908 but who died shortly before the Congress opened so that her paper was read by a male colleague.

[26] The ScD (Doctor of Science) was based on submitted papers, and one of her examiners was the group theorist William Burnside. There was a standard fee of £25 but £15 was waived on account of her assistance with re-editing Volume 2 of the fifth edition of George Salmon's *Analytic Geometry in Three Dimensions* (1915) to which she contributed new sections on Cremona Transformations. Trinity College Dublin Archives, TCD MUN/V/5/20 pp. 294–295, 326, 328. I am very grateful to Aisling Lockhart, the archivist at Trinity College Dublin, for supplying me with this information.

The ScD was not Hudson's first degree from Dublin. In 1906, she was one of the so-called 'Steamboat Ladies' who travelled by steamboat to Dublin to be awarded an ad eundem University of Dublin degree at Trinity College Dublin. These degrees were open to women from Oxford and Cambridge to enable them to prove their degree status. For further information on the Steamboat Ladies, see Parkes (2007).

[27] Located in East London, the West Ham Technical Institute (later West Ham College of Technology) was one of three colleges that merged in 1970 to form the North East London Polytechnic (later Polytechnic of East London) and which in 1992 became the University of East London.

[28] Hudson was honoured for her wartime aeronautical work by being appointed Officer of the Order of the British Empire (OBE). For details of this work, see Royle (2017, 351–357).

geometrical research, and later channelling her energies into social work[29] and the Student Christian Movement (Creese 2004).[30]

Hudson's main collaboration with Ross took place during two periods: the first in 1916–1917 and the second in 1928–1931, although the two kept in touch in between.[31]

15.6 The Ross-Hudson Collaboration 1916–1917

Shortly after sending the first part of the paper to the Royal Society in July 1915, Ross left for the Dardanelles to investigate an outbreak of dysentry among the troops, putting the epidemiological work on hold.[32] He returned to it the following spring but finding himself short of time he applied to the Royal Society for "a grant in aid of the payment of a lady mathematical worker for assistance in finishing the second part of his paper".[33] In early April he learnt that his application had been successful. That he specifically asked for a "lady" mathematician is no surprise given the War and the recent imposition of conscription.[34] The grant was to be paid from the Government Grant Urgency Fund, indicating that the work was considered essential within the context of the War.

Whether Ross had any ideas about whom he would like to assist him or whether he simply circulated the relevant information to likely parties is not known.[35] But

[29] Letter from Hudson to Ross, 21 July 1928, RCPSG 9/M/9/1/6.

[30] Hudson also had political interests. In May 1929 she spent 2 weeks in Birmingham canvassing on behalf of the social economist Louis Anderson Fenn who was standing as a Labour candidate in the General Election, the first General Election in which women aged 21–29 were allowed to vote. Letter from Hudson to Ross, [May 1929], RCPSG 9/M/17/1/30.

[31] See, for example, the letter from Ross to Hudson of 26 July 1924, in which Ross expresses his pleasure at hearing from Hudson, RCPSG 9/M/9/1/4.

[32] Ross made an important contribution to the war effort with his research and treatment of soldiers suffering from malaria and dysentry, tropical diseases being a far greater killer than the enemy in many places beyond Europe. See Cranna (2018).

[33] Minutes of Meetings of the Council of the Royal Society, Volume 25 (20 May 1915–4 July 1918), p. 141, RSA CMO/25. Letter from RWF Harrison (Assistant Secretary of the Royal Society) to Ross, 8 April 1916, RSA NLB/53/90. Ross's application for funding was supported by the Government Grant Mathematics Board; the payment was made from the Government Grant Urgency Fund.

[34] In January 1916 the Military Service Act introduced conscription for single men between the ages of 18 and 41.

[35] At the end of 1915, the Federation of University Women compiled a register of graduates together with their qualifications and training. It is therefore possible Ross found Hudson through this register. The existence of the register was publicised in *The Times* on 28 October 1915 (p. 5), under the headline 'The demand for qualified women'.

at all events word must have got out quickly because on 15 April 1916 Hudson responded, writing to Ross:[36]

> I understand that Miss Thomas[37] has mentioned me to you in connection with mathematical work of importance during the war and I should be very glad to know about it.

Hudson then summarised her mathematical history and personal circumstances, from which it is evident that the two had not met before.

Ross promptly responded, writing on 19 April 1916:[38]

> [I] am very glad that you may be available for the work—especially as your brother, who was unfortunately killed in Wales ten years ago, was a friend of mine and helped me considerably with some old mathematical work which I had been doing.
>
> The work with which I now require assistance consists of the a priori mathematical study of epidemics. This work was commenced by Daniel Bernoulli two hundred years ago.[39] Recently the Royal Society published the First Part of a paper of mine (which I send you herewith). This paper attempts to find out what of curves of epidemics should be on the supposition that we know exactly how such epidemics are caused—as you will see by the paper itself. I have integrated all the equations and have written most of the academical portion of the Second Part. But the Second Part will require the application of my curves to actual epidemics, of which we have many records. For this latter study, I do not have sufficient time available and I therefore asked the Royal Society for a Government Grant to pay for a lady Mathematician to assist me for one year. The Society has immediately been kind enough to give me £150 for this purpose for one year. Of this sum I am to keep about £30 for possible expenses. There will therefore be a remainder of £10 a calendar month for salary to my assistant. I cannot say whether the grant will be continued, but think that this will be likely if the work proves to be of value.
>
> The researches will be carried out officially in this laboratory [Marcus Beck Laboratory, Royal Society of Medicine, London] and in connection with the researches which we are conducting in measles and dysentry; and the large medical library downstairs contains much literature on epidemics. At present the work will lie principally in examining the figures of epidemics statistically and enquiring whether or not they fit my a priori curves; but I should also like my assistant to help as regards the curves of epidemics, especially measles, in general. Under these circumstances part-time work will be quite sufficient, and much of it can be done at home—though I should like to see my assistant most days of the week in order to consider details with her and to show her what to do. ...
>
> I do not know whether you have followed recent literature on statistical methods, but, if not, I dare say you would be willing to spend a little time on the subject. I dare say that Professor Karl Pearson would help us a little in this direction with his advice. ...

Ross also made a point of telling Hudson that any resulting publications would be in joint names. At the time joint publications involving women were unusual which

[36] Letter from Hudson to Ross, 15 April 1916, RCPSG 9/M/8/5/31.

[37] So far Miss Thomas has remained elusive, but she may have been employed at the West Ham Technical Institute (where Hudson was working at the time of her letter). Ross referred to her as 'Dr' rather than 'Miss' Thomas in his letter to Hudson of 19 April 1916, RCPSG 9/M/8/5/32.

[38] Letter from Ross to Hudson, 19 April 1916, RCPSG 9/M/8/5/32.

[39] Ross is here referring to Bernoulli's famous paper on smallpox which was published in 1766. Much later, he stated that he had never actually seen Bernoulli's paper. See, Ross, 'A Priori Epidemiology', unpublished manuscript, 31 October–16 November 1931, RCPSG 9/M/9/1/48. For a discussion of Bernoulli's paper, see Dietz and Heesterbeek (2002).

is presumably why Ross felt the need to spell it out.[40] The fact that he did so could indicate an awareness of Hudson's academic standing. Or perhaps he was simply conscious of the amount of work he needed her to do. Even though he did not know her personally, the family connection may also have played a part in this respect. Certainly, as this letter indicates, Ross's friendship with Hudson's brother helped the collaboration to get off to a good start.

Hudson wrote back to Ross on the 21 April accepting the job and agreeing a start date of 1 May 1916.[41] She thought the subject matter appeared "most interesting, and just as important as if had been the 'war job' that [she] had supposed it to be". She also alerted Ross to the fact that she had never worked on statistics. Her appointment was announced in the *BMJ* on 27 May 1916 under the heading 'Pathometry' (Anon 1916).

Although it is evident from Ross's initial letter to Hudson that he wanted help with applying the statistics of epidemics to his theory, that was not what she ended up doing. The joint paper continued the theoretical development begun in Part I, with the application of statistics promised for the future (Ross and Hudson 1917, 213).

By the time Hudson was on board, Ross, as indicated in the introduction to Part I of his paper, had already prepared an incomplete draft of the second part.[42] This draft relates mostly to Part II of the published version, and provided the starting point for Hudson's contribution. Although the structure in the final version remained as Ross had laid out, it would seem from Hudson's drafts that she reworked much of the material, adding, deleting, and bringing general clarity to the arguments. However, since, as evidenced by the correspondence, the collaboration involved regular meetings in person, it is not possible to know the exact circumstances in which Hudson's drafts were composed.[43]

The content of Part II centres on two groups of individuals—the number of new cases and the number of non-affected cases as proportions of a population—each taken as a function of time. It is shown how the two groups vary as the four "variation elements" in each group—births, deaths, immigration, emigration—are given different values. The effect of different rates of infection is investigated, beginning with the assumption that each affected individual daily infects one other individual and then looking at what happens as the infection rate is reduced. In the

[40] One of the few men who did co-author papers with women was Karl Pearson, although his situation was rather particular since he ran laboratories in which he employed several women to work as assistants.

[41] Letter from Hudson to Ross, 21 April 1916, RCPSG 9/M/8/5/33.

[42] 'Studies on A Priori Pathometry. Part II'. RCPSG 9/M/8/3/1. The draft is dated 12 July 1915. There is a note on the bottom of the first page which states: 'Omitted from Part II. October 1916. The following pages belonged to the original draft, and have been replaced by m.s. in the preceding'.

[43] For an example of Hudson's drafts, see the mathematical manuscript RCPSG 9/M/8/1/3.

case when the infection rate is high and the reversion rate is low,[44] the model leads to a nearly symmetrical bell-shaped epidemic curve (Ross and Hudson 1917, 219). This was an "important result" because, as the authors point out, the curve is similar in shape to those found by Brownlee (1907).

Part III of the published version was rather different to Part II in its genesis since none of it, bar the section headings, was prepared before Hudson's arrival. Indeed, Ross made a point of mentioning that he was particularly obliged to Hudson for help with Part III (Ross and Hudson 1917, 213). While keeping to Ross's structure, it seems that Hudson generated most of the content for this part, instructing Ross on mathematical techniques in the process. Having continued with the exploration of changes in the infection rate, taking into account seasonal variation, sudden change and continuous variation, graphical methods are employed to ascertain the variation over time of the two groups. Modifications are then made to Ross's original equations to allow for the assumption that the number of deaths from disease is proportional to the number of new cases. Significantly, the theory is extended to allow for affected individuals to recover into an immune state, i.e., there are now three groups to be considered rather than two. This leads to an early version of what is now known as the SIR model (Ross and Hudson 1917, 236–238).[45]

In their conclusion, Ross and Hudson were careful to spell out the limitations of their results. What they had produced was "an apparatus which can be used in a variety of special cases" (Ross and Hudson 1917, 238). The theory involved a large number of dependent variables, a fluctuation in any one of which could be accounted for in many different ways by supposing a suitably adjusted variation in almost any of the parameters. Nevertheless, the shape of the curves obtained from the different cases they examined gave them enough confidence to be able to state that "the rise and fall of epidemics as far as we see at present can be explained by the general laws of happenings, as studied in this paper" (Ross and Hudson 1917, 239), although with some caveats regarding the different assumptions they had had to make (such as, for example, regarding case-mortality and infectivity as constants).

In October 1916, Ross submitted their work as a single paper to the *Proceedings of the Royal Society*. Rather curiously, but maybe because it was wartime, the paper was refereed by Arthur Eddington.[46] He described it as "a guide in interpreting the significance of epidemic statistics" and recommended publication in full without modification, although he found the theory "rather dull" and the mathematics

[44] The reversion rate is the proportion of affected individuals who revert to being unaffected in a given element of time.

[45] The SIR model, where S is the number of susceptible individuals, I the number of infectious individuals and R the number of removed (recovered/immune or deceased) individuals, can be used to determine the critical number of susceptible individuals necessary for the disease to take hold in a given population. The SIR model was developed further in Kermack and McKendrick (1927) which is discussed in Sect. 15.7 of this paper.

[46] Eddington, who was the Cambridge Plumian Professor of Astronomy and Experimental Philosophy, and Director of the Cambridge Observatory, was principally working as an astrophysicist.

"elementary".[47] The paper appeared in May the following year although, because of its length, it was published in two consecutive parts with the second part following seamlessly on from the first, Ross having advised on the positioning of the split between the two parts (Ross and Hudson 1917).[48] As with the first part of the paper, its publication produced almost no reaction, from either the medical or the mathematical communities. There was a brief notice by George Pólya in the *Jahrbuch über die Fortschritte der Mathematik* but, unsurprisingly, Pólya found it unnecessary to reproduce the details because the emphasis was not mathematical (Pólya 1917).

At the beginning of Part II, Ross had announced that "records of epidemics are now being examined in order to find out how far the theoretical results which we have reached may be applied to them; but these studies must be reserved entirely for future discussion." (Ross and Hudson 1917, 213).[49] This was the work for which Hudson had originally been employed, and for which a knowledge of statistics was required. But although they had started to discuss it, it never materialised.[50]

In the middle of December, Hudson wrote to Ross to let him know that the Admiralty had offered her a post of temporary Technical Assistant in the Aircraft Construction Department, and in view of the "national importance and urgency of this work", she felt obliged to accept it and resign from her research work with him.[51] It was clear she had enjoyed their collaboration, and Ross in return was sorry to lose her, considering her services "invaluable", and hoped she would return to the work at a later date.[52]

Immediately after receiving Hudson's letter, Ross asked Major Greenwood if he could spare some time to continue the work. With the theoretical part completed, he was anxious to see if it could be fitted to actual epidemics, particularly since he had already begun to discuss this in a preliminary way with Hudson. But as he told Greenwood, for this he needed a "trained epidemiological statistician".[53] But nothing came of it, Greenwood presumably being too busy with the war effort or not feeling fit for the statistics. Ross tried to tempt Hudson back in 1917 but to no

[47] Referee's report by Arthur Stanley Eddington, November 1916, RSA RR/23/84.

[48] Letter from Ross to Secretary of the Royal Society, 17 October 1916, RCPSG 9/M/8/5/66.

[49] See also the letter from Hudson to Ross, 3 October 1916, RCPSG 9/M/8/5/52.

[50] In their correspondence Ross and Hudson discuss information about "cases" sent to them from "Dr Hay" and the need to ask him for further analysis according to "date of attack" but with no other details, see the correspondence from Ross to Hudson 16 September 1926, RCPSG 9/M/8/5/49, and Hudson to Ross, 15 September, 19 September and 3 October 1916, RCPSG 9/M/8/5/48, 9/M/8/5/50 and 9/M/8/5/52 respectively. It is possible that this correspondence refers to data from a 1904 study of "Two hundred cases of acute lobar pneumonia" by John Hay of Liverpool Medical Institution (Hay 1904).

[51] Letter from Hudson to Ross, 18 December 1916, RCPSG 9/M/8/5/57.

[52] Letter from Ross to Hudson, 19 December 1916, RCPSG 9/M/8/5/58.

[53] Letter from Ross to Greenwood, 19 December 1916, LSHTM Ross/163/11/01. For Greenwood's biography, see Farewell and Johnson (2016).

avail.[54] However, they kept in contact. In the spring of 1918, Ross asked Hudson for her opinion on whether a paper by Lotka was suitable for publication by the Royal Society. She was pithy in her assessment:[55]

> He [Lotka] has taken some of our equations, specialised them by drastic assumptions, borrowed some analysis from a German [Hertz], and got out an approximate result—not I think of any practical importance.

However, since Lotka claimed priority over some ideas, she cautioned Ross about holding the paper back in case his actions could be misinterpreted. In the event, the paper was published in the *Journal of the Washington Academy of Sciences* (Lotka 1919).

After the War, Hudson decided to return to geometry, and in early 1919 informed Ross of her plans. Once she had finished with the aeronautical work, she would go to Cambridge, where she had access to the library, to complete her book on Cremona transformations (which would appear in 1927).[56] She admitted to Ross that she was "sad to rule out the epidemic work" but she was clear in her ambition. To soften the blow, she suggested to him that he employ one of her female assistants at the Admiralty who were soon to be demobbed. A couple of days later she sent him a formal application from Annie Trout, a London university graduate with a first-class degree in mathematics. However, despite Hudson's recommendation and Trout's suitability, Ross told Trout that he was unable to employ a mathematical assistant for epidemiological work.[57] Although Ross was keen to continue the work—he had asked Hudson for the names and qualifications of suitable women—it seems he wasn't prepared to go on the stump for funds to support it. Had Hudson been available it might well have been a different story.

15.7 The Ross-Hudson Collaboration 1928–1931

There was then a hiatus for over a decade during which time Ross did no further work on the theory of epidemics. As he explained in his *Memoirs*:

> For years I had been toiling at the attempt to fix mathematics on the general theory of epidemics, and in 1918 [sic] the Royal Society published my paper and gave me the capable

[54] Letter from Ross to Hudson, 4 June 1917, RCPSG 9/M/8/5/61.

[55] Letter from Hudson to Ross, 30 March 1918, RCPSG 9/M/8/5/38. Hudson may have been irked by the fact that in his paper Lotka mentions Ross several times, including with reference to the Ross and Hudson paper, but her name is completely absent. In the summer of 1931 Ross invited her to lunch to meet Lotka but she declined being "too much of an invalid" due to her rheumatism. Letter from Hudson to Miss Lafford (Ross's secretary), 17 June 1931, RCPSG 9/M/9/1/26.

[56] Letter from Hudson to Ross, 12 February 1919, LSHTM Ross/173/20/42.

[57] Letter from Hudson to Ross, 14 February 1919, LSHTM Ross/173/20/43. Letter from Ross to Trout, 24 February 1919, LSHTM Ross/173/20/46. For more information on Trout, see Barrow-Green and Royle (2022, 554–555).

> assistance of Miss Hilda P Hudson. After a second paper, the war interrupted our studies; but so little interest was taken in them by the "health authorities", that I have thought it useless to continue them since then. (Ross 1923, 515).

But then something happened that reignited his interest. And that something was the publication in 1927 of the first of a (now famous) series of papers by William Kermack and Anderson McKendrick (Kermack and McKendrick 1927). By taking the work of Ross and Hudson as their starting point and generalising, Kermack and McKendrick had formulated a deterministic model for the spread of an infectious disease in a closed population, in terms of both its magnitude and its termination, and this is the model which is the basis for today's SIR model.[58] Contrary to what had been previously believed, their results showed that the termination of an epidemic was determined by the relationship between the population density and the rates of infectivity, recovery and death.[59] Although Kermack and McKendrick were working with a slightly different set of assumptions and simplifications to that of Ross and Hudson, their results were very similar. Both accounted for the usual course of an epidemic by the decrease in the number of those susceptible without having to assume any change in infectivity, and both led to functions with bell-shaped graphs in which the number of cases rises more quickly than it falls.

Ross had known McKendrick from time spent together in Sierra Leone in 1901 (Ross 1923, 439–449), (Heesterbeek 2005, 93–95), and it was Ross who had "impressed on the young physician lieutenant, not yet trained in mathematics, the tremendous power of mathematical methods in medical research" (Hirsch 2004). Ross and McKendrick had remained in touch over the years, so it is little surprise, even though Ross was now past the age of 70 and not in good health, that the paper galvanised him into action.[60] In July 1928, he got in touch with Hudson, and she agreed to spend one day a week working for him—indeed she found it "very gratifying" that he should still think of her in connection with this work.[61]

To start, Ross set Hudson the task of familiarising herself with his 'Theory of Happenings', the mathematical addendum in Ross (1911a). Although it had formed the basis for Part I of the Royal Society papers, she did not find it easy to read due to the muddled form of the algebra—too many related symbols, too many versions of each relation, inconsistent notation—and the number of errors and misprints

[58] For a discussion of Kermack and McKendrick (1927), see Anderson (1991).

[59] It was commonly thought that epidemics terminated either because the supply of those susceptible had been exhausted or because the virulence of the cause of the epidemic had waned.

[60] In 1911 McKendrick told Ross that he thought *The Prevention of Malaria* was a "capital book", and that he was "trying to reach the same conclusions from differential equations, but it is a very elusive business, and I am having to extend mathematics in new directions. I doubt whether I shall get what I want, but 'a man's reach must extend his grasp'". Letter from McKendrick to Ross, 6 August 1911, LSHTM Ross/106/28/112. Quoted in Kucharski (2020, 21–22).

[61] Letter from Ross to Hudson, 13 July 1928, RCPSG M 9/1/5; letter from Hudson to Ross, 21 July 1928, RCPSG 9/M/9/1/6; letter from Hudson to Ross, 28 December 1928, RCPSG 9/M/17/1/8.

Hudson did not wish to receive any payment for the work beyond out-of-pocket expenses. Letter from Ross to the Secretaries of the Royal Society, 6 November 1929, RCPSG 9/M/9/1/25.

(she noted 33).[62] It is little wonder then that it did not make much of an impact when it was published. Ross's plan was for them to write a new paper, 'Dependent Happenings', based on this material and including material from the Royal Society papers that he thought could be understood by the medical profession, with the aim of it being a preliminary to practical application, i.e., returning to the work they had just begun before Hudson was called away by the Admiralty.

At the same time, Ross was seeking funds from the Royal Society to get the three papers reprinted and published together in book form. Buoyed by the Kermack and McKendrick paper, he was more than ever convinced of the importance of the work, telling Hudson that he felt it should be read by "all doctors and health officers".[63] Hudson supported the republishing but felt that they should improve on the papers first. In particular, she felt the work would be incomplete without a numerical comparison using statistics from an actual epidemic, i.e., she wanted to do the work she had originally been employed to do. If this could be done, she said, they should then be able to deduce better values for the constants in their formula (Ross and Hudson 1917, 203), and may even be able to improve on Kermack and McKendrick's estimate for the threshold density (Kermack and McKendrick 1927, 715).[64] But, as she acknowledged, it would take some effort to obtain suitable statistics since those they needed would have to run to several thousands of cases.

Meanwhile Hudson herself had been busy writing a review of the Kermack and McKendrick paper for *Science Progress*, the review journal of which Ross was the editor.[65] It appeared the following year with the rather enticing title 'Contagion and calculus' which was probably due to Ross (Hudson 1929). Hudson thought well of the paper, noting that the results it contained accorded well with those obtained by her and Ross, although she did have a "few captious criticisms of details" (Hudson 1929, 522). With her mathematician's eye, she noted the lack of discussion about the relative order of quantities assumed to be small, and she found the use of $\sqrt{-q}$

[62] Letter from Hudson to Ross, 6 March 1929, RCPSG 9/M/17/1/19; letter from Ross to Hudson 19 March 1929, RCPSG 9/M/17/1/21. Hudson was not the only one to have been confused by Ross's notation. In 1916 Percy MacMahon, who was a supporter of Ross as a mathematician—he felt they shared common ground as "amateurs [not holding[mathematical chairs"—told Ross that he found his notation "somewhat difficult to get accustomed to" and that he hoped Ross would express any new mathematical results "in a language and notation that can be understood by the mathematical multitude." LSHTM Ross/163/34/15.

[63] Letter from Ross to Hudson, 11 November 1930, RCPSG 9/M/9/1/17.

[64] Letter from Hudson to Ross, 8 December 1928, RCPSG 9/M/17/1/7. The threshold density of a population is the population density above which an epidemic can be sustained.

[65] It was not Hudson's first review for *Science Progress*. In 1917, at Ross's request, she reviewed H.S. Carslaw's text on non-Euclidean geometry (Hudson 1917).

for a real constant (where q is negative), "a minor exasperation". It was a detailed review and Ross was happy with her efforts:[66]

> You have sent me an excellent review on the book [sic] by Kermack and McKendrick, and I am glad that you like the paper. ... You are now quite beyond me for the reasons which I mentioned that I have got no brain at all.

Their work together continued but by January 1929 Hudson was having doubts, telling Ross she felt "somewhat averse to publishing a new Royal Society paper unless it contains some distinctly new idea," although she was happy about a reprint of the original papers.[67] But Ross was not to be deterred, and by April 1929 he felt sufficiently confident to announce to readers of the *BMJ* the resumption of the collaboration, having outlined its earlier history:

> We are now beginning to consolidate and summarize our previous studies and to develop them by means of the finite calculus, the infinitesimal calculus, and by the use of integral equations. (Ross 1929, 674).

Notably, he made no mention of statistics. It seems that he had given up on the idea of seeing how well the theory worked in practice.

Within a couple of months, Ross had redrafted 'The Theory of Happenings' into the first part of a new article which he entitled 'Two-Party Aggregates'—a two-party aggregate being a set of individuals that can be divided into two groups depending on whether the individuals possess (or receive) a particular quality. Initially only his name was on the draft but after it had been typed, Hudson's name was added by hand.[68] But by September he had got distracted by other mathematical interests and put the article on hold never to return to it, despite telling Hudson that he would.[69] Since the article was essentially the contents of the 1911 mathematical addendum augmented by Hudson's clarifications and corrections, Hudson herself would have seen no value in pushing for publication, especially given Ross's poor health.

Ross did, however, persist with the reprinting of the Royal Society papers in book form, although with no new material apart from a short introduction which he alone authored (Ross and Hudson 1928). The book was eventually published in early 1931, Hudson having helped Ross with the proofs. Also in 1931, believing fallacious experimental demonstrations of the possibility or impossibility of mosquito-control were being presented by people ignorant of his early work, he republished his paper from the St. Louis Congress (Ross 1905a) in the *Journal of Tropical Medicine and Hygiene* under the title 'A Mathematical Justification of Mosquito-Control'.[70]

Although Ross did no more epidemiological research, he did produce a further manuscript. Entitled 'A priori epidemiology', it was written between 31 October and 16 November 1931, less than a year before he died, and was never published.

[66] Letter from Ross to Hudson, 30 October 1928, RCPSG 9/M/9/1/13.

[67] Letter from Hudson to Ross, 26 January 1929, RCPSG 9/M/17/1/15.

[68] Two-Party Aggregates, manuscript, 20 May 1929–5 June 1929, RCPSG 9/M/17/1/2-3. Letter from Ross to Hudson, 7 June 1929, RCPSG 9/M/17/1/25.

[69] Letter from Ross to Hudson, 29 September 1929, RCPSG 9/M/18/1/2.

[70] See A.B. Hill, *Journal of Tropical Medicine and Hygiene* 34 (1931), p.177.

Over the course of four pages, he recounted the history of his own work, and that of others, explaining the significance of working a priori, eventually reducing it to "a method of trial and error". Just how much the epidemiological work meant to him is clear from the opening paragraph:[71]

> My name has long been recognised even among the general British public as that of the man who largely verified Manson's mosquito theory of malaria and who endeavoured to have it employed for saving human life on a large scale in most tropical countries; but in my own opinion my principal work has been to establish the general laws of epidemics.

Even allowing for when it was written, it is remarkable to see that he valued the epidemiological work more highly than anything else he had achieved. Later in the manuscript, in the context of the Royal Society papers, he described Hudson as "an accomplished mathematician." It is an apt description even though Ross's mathematical interaction with Hudson had been nowhere near her true capability. Over the several years of their collaboration, Hudson had more than proved her worth to him, both as a mathematician and as a trusted colleague. As well as writing reviews for *Science Progress*, she also advised him on the suitability of mathematical articles submitted to the journal.[72] On a more personal level, in May 1916, he had asked her for advice on one of his papers, 'The iteration of certain functions', which, with her approval, he submitted to the Royal Society although it was ultimately rejected.[73] And the following year, when she was working in the Air Department, he asked her if she could assist the Cambridge biologist, G.F.H. Nuttall, who needed help with calculating the reproductive power of body lice.[74] On all occasions she willingly obliged.

15.8 Conclusion

Although the mathematical approach pioneered by Ross would prove very influential,[75] the actual mathematics required to complete Ross's Royal Society paper of 1916 was neither deep nor novel, as pointed out by both Eddington and Pólya. The fact that Ross, despite his enthusiasm for mathematics, required the assistance of Hudson points sharply to his lack of mathematical training and

---

[71] Ross, 'A Priori Epidemiology', unpublished manuscript, 31 October-16 November 1931, RCPSG 9/M/9/1/48.

[72] Ross to Hudson, 1 September 1916, RCPSG 9/M/8/15/46.

[73] The referees for Ross's paper were William Burnside (reject), Andrew Russell Forsyth (accept with modifications), Percy MacMahon (accept). For the referee reports, see RSA RR/23/81-83.

[74] Nuttall had originally asked Ross for help with the calculation. Ross was unable to do it, so he passed it on to Hudson, and Hudson, although very busy with War work, managed it while on a train. LSHTM Ross/173/27/16. One rather alarming result from her calculations was that the offspring of the daughters of one female louse would number 112,778 (Nuttall 1923, 163)!

[75] For remarks on the influence of Ross's approach, see the introduction to Heesterbeek and Roberts (2015) and to Dietz and Schenzle (2022).

general mathematical knowledge, and one might even be tempted to say lack of mathematical discipline. This lack showed in his earlier work too as, for example, in the jumble of algebra in Ross (1911a). The fact that Hudson, a Cambridge-trained mathematician but one who now specialised in geometry, not only had no difficulty in dealing with the mathematics he presented to her but also endeavoured to instruct him, serves to underscore the point.

Ross clearly felt at ease working with Hudson, and it was an ease that was quickly established. Although there is formality in the openings and closings of their correspondence, the body is often humorous and Ross often pokes fun at himself, as for example in this letter to Hudson of 28 January 1929:[76]

> I have also been rereading that dreadful paper by Ross & Hudson, and have almost understood my first paper, and have greatly admired the concluding part III by Hudson. It is very well done, and some of the figures which you produced ought to have shown [Clifford Allchin] Gill who wrote his medical work on the Genesis of Epidemics, that we have already included his curious lucubrations in our conjoint paper — but of course doctors seldom know a word of mathematics, so that they can write any rubbish they please.

Ross was not always the easiest person to get along with—he was involved in several notable polemics[77]—but he hit it off with Hudson. He clearly respected her mathematical ability and she in turn seemed able to get the best out of him, not losing patience when at times his responses (or lack of them) must surely have left her exasperated. As the relationship developed, the tone in his letters becomes more familiar, and the terms he uses to describe her in his publications and manuscripts become increasingly complimentary. Having begun as his assistant, by 1928 she had become a "very expert mathematician" (Ross 1928, 158). The two met often to discuss their work, and, in the later years of their collaboration, when Ross had become wheelchair bound, Hudson would stay over at Ross's house.[78]

That at the end of 1916 Hudson gave up working with Ross in favour of working for the Aircraft Construction Department is no surprise. While she could see the benefit of their collaboration, the War brought a different set of priorities, and the work was not optimizing her mathematically. Similarly, after the War, it was an obvious choice for her to return to geometry, the subject in which she excelled. But once her book was published in 1927, she had time to spare, and it is an indication of the pleasure she had previously found in working with Ross that she did not hesitate to pick up the reins with him again. However, by this time, he was beginning to suffer from ill health and, as time marched on, it became clear that little further progress would be made, the main outcome being the republication of the Royal Society papers, but with no new material, and that not until 1931. A year later Ross would be dead.

[76] Ross to Hudson, 28 January 1929, RCPSG 9/M/7/1/16.

[77] As well as falling out with Manson (Footnote 1), Ross was involved in a bitter priority dispute with Giovanni Grassi over the discovery of malarial transmission, see Capanna (2006). See also Chernin (1988a) which describes Ross's attack on Paul De Kruif's *Microbe Hunters* (1926).

[78] Letter from Hudson to Miss Lafford (Ross's secretary), 20 July 1929, RCPSG 9/M/9/1/4.

As well as pointing to Hudson's versatility as a mathematician, the epidemiological work shows her as an excellent collaborator, both of which characteristics would be important in her work for the Air Construction Department. That she was employed at all by Ross was a consequence of the War, and she accepted the position because she thought it was war-work, albeit of a different kind. It is hard to imagine any other circumstances under which Ross would have specified to the Royal Society that he required funds to employ "a lady mathematical worker". That Hudson applied for the post could hardly have been more fortunate for Ross. Not only was Hudson one of the best English women mathematicians of the day, but she was also a confident, careful worker with an instinct for successful collaboration, and with a family connection that put Ross at ease from the start.

Acknowledgments I first learnt about Hilda Hudson from Jeremy Gray when he and I were writing an article together on geometry in Cambridge (Barrow-Green and Gray 2006). Jeremy has been my constant guide and mentor since I began research as his doctoral student in 1989, and it gives me great pleasure to dedicate this article to him.

I owe special thanks to Clare Harrison, the archivist at the Royal College of Physicians and Surgeons, Glasgow, and Claire Frankland, the archivist at the London School of Hygiene and Tropical Medicine, both of whom went out of their way to assist me in the challenging conditions of the pandemic. I am also very grateful to Reinhard Siegmund-Schultze for many helpful remarks and suggestions made during the writing of this article.

Archival Sources

LSHTM: Sir Ronald Ross Collection, London School of Hygiene and Tropical Medicine, London WC1E 7HT.
RCPSG: Sir Ronald Ross Collection, Royal College of Physicians and Surgeons, Glasgow G2 5RJ.
RSA: The Royal Society of London, London SW1Y 5AG.

References

Anderson, R. M. 1991. Discussion: The Kermack-McKendrick epidemic threshold theorem. *Bulletin of Mathematical Biology* 53: 3–32.

Anon. 1916. Pathometry. *The British Medical Journal* Issue 2891: 766, 27 May 1916.

Barrow-Green, J. E. 2001. "The advantage of proceeding from an author of some scientific reputation": Isaac Todhunter and his mathematical textbooks. In *Teaching and Learning in Nineteenth-Century Cambridge*, ed. J. Smith and C. Stray, 177–203. Cambridge: Boydell Press.

Barrow-Green, J. E., and J. J. Gray. 2006. Geometry at Cambridge, 1863–1940. *Historia Mathematica* 33: 315–356.

Barrow-Green, J. E., and T. Royle. 2022. The work of British women mathematicians during the First World War. In *The Palgrave Handbook of Women and Science since 1660*, ed. C. G. Jones, A. E. Martin, and A. Wolf, 549–572. London: Palgrave Macmillan.

Brownlee, J. 1907. Statistical studies in immunity. *Proceedings of the Royal Society of Edinburgh* 26: 484–521.

Brownlee, J. 1915. On the curve of the epidemic. *The British Medical Journal* Issue 2836: 799–800, 8 May 1915.

Capanna, E. 2006. Grassi versus Ross: who solved the riddle of malaria? *International Microbiology* 9: 69–74.

Chernin, E. 1988a. Paul De Kruif's *Microbe Hunters* and an outraged Ronald Ross. *Reviews of Infectious Diseases* 410: 661–667.

Chernin, E. 1988b. Sir Ronald Ross vs. Sir Patrick Manson: A matter of libel. *Journal of the History of Medicine and Allied Sciences* 43: 262–274.

Cranna, V. 2018. Sir Ronald Ross and malaria in the First World War. https://blogs.lshtm.ac.uk/library/2018/11/07/sir-ronald-ross-and-malaria-in-the-first-world-war/.

Creese, M. R. S. 2004. Hilda Phoebe Hudson. *Oxford Dictionary of National Biography*.

Dietz, K., and J. A. P. Heesterbeek. 2002. Daniel Bernoulli's epidemiological model revisited. *Mathematical Biosciences* 180: 1–21.

Dietz, K., and D. Schenzle. 2022. Mathematical models for infectious disease statistics. In *A Celebration of Statistics. The ISI Centenary Volume*, ed. A. Atkinson and S. Fienberg, 549–572. New York: Springer-Verlag.

Farewell, V., and T. Johnson. 2016. Major Greenwood (1880-1949): a biographical and bibliographical study. *Statistics in Medicine* 35: 645–670.

Fine, P. E. M. 1975. Ross's a priori pathometry—a perspective. *Proceedings of the Royal Society of Medicine* 68: 547–551.

Gray, J. J. 2008. *Plato's Ghost. The Modernist Transformation of Mathematics*. Princeton: Princeton University Press.

Gray, J. J. 2013. *Henri Poincaré. A Scientific Biography*. Princeton: Princeton University Press.

Greenwood, M. 1916. The application of mathematics to epidemiology. *Nature* 97: 243–244.

Hay, J. 1904. Two hundred cases of acute lobar pneumonia. *The Lancet* 1: 1423.

Heesterbeek, J. A. P. 2005. The law of mass-action in epidemiology: A historical perspective. In *Ecological Paradigms Lost: Routes of Theory Change*, ed. K. Cuddington and B. E. Beisner, 81–104. Burlington: Elsevier.

Heesterbeek, J. A. P., and M. G. Roberts. 2015. How mathematical epidemiology became a field of biology: a commentary on Anderson and May (1981) 'The population dynamics of microparasites and invertebrate hosts'. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370: 1666:20140307.

Hirsch, W. M. 2004. Anderson Gray McKendrick. *Oxford Dictionary of National Biography*.

Hudson, H. P. 1917. The Elements of non-Euclidean Plane Geometry and Trigonometry. By H. S. Carslaw. *Science Progress* 11: 518–519.

Hudson, H. P. 1927. *Cremona Transformations in Plane and Space*. Cambridge: Cambridge University Press.

Hudson, H. P. 1929. Contagion and calculus. *Science Progress* 23: 521–523.

Kermack, W. O., and A. G. McKendrick. 1927. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A 115* (772): 700–721.

Kucharski, A. 2020. *The Rules of Contagion*. London: Profile Books.

Lorenat, J. 2020. Certain modern ideas and methods: "Geometric reality" in the mathematics of Charlotte Angas Scott. *The Review of Symbolic Logic* 13: 681–719.

Lotka, A. J. 1912. Quantitative studies in epidemiology. *Nature* 88: 497–498.

Lotka, A. J. 1919. A contribution to quantitative epidemiology. *Journal of the Washington Academy of Sciences* 9: 73–77.

Lotka, A. J. 1923. Contribution to the analysis of malaria epidemiology. I. General part. *American Journal of Epidemiology* 3 (Supplement 1): 1–36. Reprinted in *Lectures Notes in Biomathematics* 22 (1978): 302–347.

Nuttall, G. F. H. 1923. The biology of pediculus humanus. *Parasitology* 10: 180–185.

Parkes, S. 2007. Steamboat Ladies. *Oxford Dictionary of National Biography*.

Pearson, K. 1905a. The problem of the random walk. *Nature* 72: 294.

Pearson, K. 1905b. The problem of the random walk. *Nature* 72: 342.
Pearson, K., and J. Blakeman. 1906. A mathematical theory of random migration. *Drapers' Company Research Memoirs Biometric Series XV*.
Pólya, G. 1917. R. Ross and H. P. Hudson. The application of the theory of probabilities to the study of *a priori* pathometry. Part II, III. *Jahrbuch über die Fortschritte der Mathematik*, 46.0789.02.
Ross, R. 1901. *The Algebra of Space*. Liverpool: George Philip and Son.
Ross, R. 1905a. On the logical basis of the sanitary policy of mosquito reduction. *The British Medical Journal* Issue 2315: 1025–1029, 13 May 1905. Reprinted in *St Louis Congress of Arts and Science, Universal Exposition, St. Louis* Volume VI (1906), Boston and New York: Houghton, Mifflin and Company, 89–102; *Journal of Tropical Medicine and Hygiene* 34 (1931): 177–183 (under the title 'A Mathematical Justification of Mosquito-Control').
Ross, R. 1905b. The possibility of reducing mosquitoes. *Nature* 72: 151.
Ross, R. 1908. *Report on the Prevention of Malaria*. London: Waterlow.
Ross, R. 1910. *The Prevention of Malaria*. London: John Murray.
Ross, R. 1911a. *The Prevention of Malaria*. London: John Murray. Second edition with an Addendum on the 'The Theory of Happenings'.
Ross, R. 1911b. Some quantitative studies in epidemiology. *Nature* 87: 466–467.
Ross, R. 1915. Some a priori pathometric equations. *The British Medical Journal* Issue 2830: 546–547, 27 March 1915.
Ross, R. 1916. An application of the theory of probabilities to the study of *a priori* pathometry—Part I. *Proceedings of the Royal Society of London, Series A 92*: 204–230.
Ross, R. 1918. *Solid Space-Algebra, the Systems of Hamilton and Grassmann combined*. London: Harrison and Sons, Ltd. Republished 1930.
Ross, R. 1923. *Memoirs with a Full Account of the Great Malaria Problem and Its Solution*. London: John Murray.
Ross, R. 1928. *Studies on Malaria*. London: John Murray.
Ross, R. 1929. Constructive epidemiology. *The British Medical Journal* Issue 3562: 673–674, 13 April 1929.
Ross, R., and H. P. Hudson. 1917. An application of the theory of probabilities to the study of *a priori* pathometry—Part II, Part III. *Proceedings of the Royal Society of London, Series A* 93: 212–225, 225–240.
Ross, R., and H. P. Hudson. 1928. *A Priori Pathometry*. London: Harrison and Sons Ltd.
Royle, T. 2017. The impact of the women of the Technical Section of the Admiralty Air Department on the structural integrity of aircraft during World War One. *Historia Mathematica* 44: 342–366.
Semple, J. G. 1969. Hilda Phoebe Hudson. *Bulletin of the London Mathematical Society* 3: 357–359.
Smith, D. L., K. E. Battle, S. I. Hay, C. M. Barker, T. W. Scott, and F. E. McKenzie. 2012. Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *PLOS Pathogens* 8 (4): e1002588. https://doi.org/10.1371/journal.ppat.1002588.
Snyder, V. 1927. 'Cremona Transformations in Plane and Space' by Hilda Hudson. *The American Mathematical Monthly* 34 (9): 487–488.
Waite, H. 1910. Mosquitoes and malaria. *Biometrika* 7: 421–436.
White, F. P. 1928. 'Cremona Transformations in Plane and Space' by H. P. Hudson. *The Mathematical Gazette* 14: 148–149.

Part IV
Modernism

Chapter 16
How Useful Is the Term 'Modernism' for Understanding the History of Early Twentieth-Century Mathematics?

Leo Corry

Abstract The present article is intended as a critical assessment of some basic assumptions underlying the analysis of modernism in mathematics in its relationship with the broader aspects of cultural modernism, especially in the period 1890–1930. It discusses the potential historiographical gains of approaching the history of mathematics in the periods under such a perspective and suggests that a fruitful analysis of the phenomenon of modernism in mathematics must focus not on the *common features* of mathematics and other contemporary cultural trends, but rather on the *common historical processes* that led to the dominant approaches in all fields.

16.1 Introduction

It is widely acknowledged that the period roughly delimited by 1890 and 1930 was marked by deep change in mathematics. It was also a time of thoroughgoing transformations that impacted the visual arts, music, architecture, and literature. The latter has often been explained in terms of artistic responses to the sweeping processes of modernization affecting Western societies. The term "modernism" has typically been used to refer to such trends and the ways in which they implied highly innovative—sometimes *avant-garde*—aesthetic conceptions characterized by unprecedented radical breaks with long-standing traditions in each area of cultural expression. A question naturally arising in these circumstances is whether the development of mathematics during said period can be seen as part of the phenomenon of "modernism" considered in its broader context, and whether adopting such a perspective leads to important historical insights.

Herbert Mehrtens' pioneering study, *Moderne-Sprache-Mathematik* (Mehrtens 1990), opened the way to serious discussions on this issue. Following on his footsteps, Jeremy Gray published his well-known book, *Plato's Ghost. The Mod-*

L. Corry (✉)
Cohn Institute for History and Philosophy of Science, Tel Aviv University, Tel Aviv, Israel

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_16

ernist Transformation of Mathematics in 2018.[1] The present article is intended as a critical assessment of some basic assumptions underlying the possible discussions of modernism in mathematics and the potential historiographical gains of pursuing such discussions.

The processes of modernization that affected the content of mathematics during the said period concern the development of new methodologies, the rise of newly investigated mathematical entities and of new sub-disciplines, as well as the reshaping of the internal organization of mathematical knowledge, the transformation relationships between mathematics and its neighboring disciplines, the demise or total abandonment of activity in areas of research that were very important in the previous century, and the adoption of either implicit or formulated new philosophical attitudes. At the institutional level, the meteoric rise of the Göttingen school came to epitomize the substantial changes undergone by the discipline in this period, as other centers, both in the German-speaking world (such as Berlin, Munich, Vienna, Hamburg) and outside (Paris, Cambridge), also underwent a significant transformation. In terms of scientific leadership, the achievements of David Hilbert and his circle embodied, both symbolically and contents-wise, the personal dimension of the spirit of the period, alongside other prominent names, such as Emmy Noether, Giuseppe Peano, and Felix Hausdorff.

The suggestive idea of possible parallel developments and similar sources underlying both the broader cultural manifestations and mathematics arises in a comparable way when examining the dramatic changes that affected the neighboring discipline of physics. The classical theories of mechanics and electromagnetism had reached a climax towards the end of the nineteenth century yet now, its foundational assumptions had been put into question, and thoroughly new directions were leading the physicists' understanding of phenomena at both the microscopic and the astronomic level. In his classical 1971 study on the rise of the new quantum mechanics, Paul Forman postulated an organic association between the contemporary adoption of non-deterministic types of causality in physics and some leading cultural motifs which he associated with the modernist spirit of Weimar Germany such as irrationality, anti-scientism, and acausality (Forman 1971). In the epigraph of the article, Forman cited the German physicist Gustav Mie who, in his 1925 inaugural lecture in Freiburg, very explicitly expressed the kind of attitude that attracted Forman's curiosity, as he indicated that even physics, "a discipline rigorously bound to the results of experiment," evolves in ways that parallel those of the intellectual movements in other areas of modern life.

Forman's article has been widely read and cited, sometimes severely criticized, and also intriguingly reappraised by (Forman 2007; Carson et al. 2011). To the extent that one wants to either accept or reject the thrust of Forman's argumentation, what kind of lesson does it teach us about the issue of "mathematics and modernism", if at all? A similar question can be asked of works dealing with the development of other fields of knowledge and culture at the time, including areas

[1] Some of the main ideas were sketched earlier in (Gray 2004, 2006).

distant from mathematics. The present article suggests ways of addressing those questions and indicates some possible, specific directions in which this analysis might be profitably undertaken. The main pitfall against which I want to call attention is that of shooting the arrow and then tracing a bull's eye around it. Indeed, one of the main difficulties affecting discussions of "modernism" in general (not just concerning the history of mathematics) is that of finding the proper definition of the concept, to begin with. One might easily start by finding a general definition that can then be made to fit the developments of mathematics in the relevant period just to be able to put together all that we have learnt from historical research and thus affirm that, yes, "modernism" characterizes mathematics as it characterizes other contemporary cultural manifestations. Although this approach has some interest, it does not seem to be very illuminating, and indeed it runs the risk of being misleading since, by its very nature, it may force us to be unnecessarily "flexible" in our approach to making historical facts fit the desired definition.

The article opens with an overview of some prominent ways in which the term "modernism" has been used in the historiography of the arts, and calls attention to certain debates surrounding its relevance in that context. This is followed by a discussion of three concrete examples of works that investigate the relationship between modernism in general and the modern exact sciences: on the one hand, an investigation of the influence of scientific ideas on modern visual arts (in the writings of Linda Henderson), and, on the other hand, two books (by Herbert Mehrtens and Jeremy Gray) that explore the connections of modern mathematics with more general, modernist cultural trends. In the following two sections, I consider two examples of authors who discuss the roots and developments of modernist ideas in specific contexts (modernist painting in the writings of Clement Greenberg and Viennese modernism by Allan Janik and Stephen Toulmin) and examine the possible convenience of using their perspective in discussing modernism and mathematics.

Besides the critical examination of existing debates, on the positive side, a main claim raised discussed in this article is that a fruitful analysis of the phenomenon of modernism in mathematics must focus not on the *common features* of mathematics and other contemporary cultural trends (including other scientific disciplines – mainly physics), but rather on the *common historical processes* that led to the dominant approaches in all fields in the period of time we are investigating. To the extent that the existence of what is described as common, modernist *features* in the sciences and in the arts has been explained in the existing literature, this has been typically done in terms of a mysterious "Zeitgeist" or even "common cultural values" (as suggested, e.g., in (Miller 2000, 480; Yourgrau 2005, 3)). Though useful at first sight, such an approach is, in my view, far from satisfactory because the "Zeitgeist", if it indeed exists, is what needs to be explained. In contrast, a clearer understanding of the *processes* that led to the rise of modernism in other intellectual fields, may help us look for similar historical processes in mathematics.

It is pertinent to mention that Mehrtens pointed to this direction, as he stressed the difference between "*Moderne*", referring to the intellectual trend itself, and "*Modernisierung*", which refers to the historical *processes* leading to changes in the discipline of mathematics, within its broader social and cultural context. However, there seems to be much room for further exploration in this direction, which could lead to additional insights thus far overlooked by historians. If properly pursued, this might amount, I suggest, to a significant contribution to the historiography of the discipline. Likewise, and no less interestingly, a clearer understanding of the historical processes that led to putative modernist mathematics could shed new light on the essence and origins of modernism in general.

16.2 Modernism: A Useful Historiographical Category?

Despite its ubiquity, the fruitfulness of the concept of "modernism" as an analytic category in the context of general cultural history is far from being self-evident or settled. Indeed, its very meaning and the time span that it covers remains the subject of debate. Ranging from the seminal anthology edited by Malcolm Bradbury and James McFarlane (1976), and the more recent, two-volume collection edited by Astradur Eysteinsson and Vivian Liska (2007), to the volumes of the journal *Modernism/modernity*, by the Modernist Studies Association, the body of research literature is enormous. From this abundance of sources, I want to focus here on Ulrich Weisstein's article, "How Useful is the Term 'Modernism' for the Interdisciplinary Study of Twentieth-Century Art?" (1995), (whose title I have obviously appropriated). Based on the assumption that the idea of "Modernism" has indeed been used in fruitful ways in his own field of research, comparative literature, Weisstein wondered about its possible usefulness in researching other domains, including the visual arts and music. In doing so, he characterized the various kinds of modernist aesthetics in terms of their emphasis on the formal, as opposed to concrete subject matters and intentions, together with a consistent inclination to undertake radical breaks with the accepted norms of the field, by way of rites of passage and inspirational manifestos meant to embody avant-garde attitudes.

To be sure, Weisstein's characterization has both merits and drawbacks as an adequate prism from which to approach modernism in general. Yet the same can be said of many such checklists proposed by other authors pursuing the same task, most partially overlapping with and differing from Weisstein's, as well as from each other.[2] Thus, assessing the extent to which modern mathematics is properly defined as a modernist phenomenon by reference to any specific proposal of this kind—by checking whether or not, and to what extent, the suggested features are manifested—may end up being an unilluminating historiographical exercise. It runs the risk of providing a Procrustean bed into which the historical facts are forced, while shedding little new light on our understanding of the historical processes. "Modernism" may only become a truly useful historiographical category for our topic if it helps interpreting the known historical evidence in innovative ways, or if it

[2] See, e.g., (Calinescu 1987; Childs 2000; Eysteinsson 1990, 2021; Gay 2007).

would lead us to consider new kinds of materials thus far ignored, or underestimated, as part of historical research on the development of mathematics.

The question whether "modernism" can be used as a useful category to study the history of mathematics, moreover, is best understood when seen as part of a broader trend noticeable over the last 30 years, that involved attempts to take advantage and inspiration of historiographic conceptions, originating in neighboring fields, mainly in the historiography of other scientific disciplines (Barany 2020; Remmert et al. 2016). The trend arose, in the first place, in relation with the Kuhnian concepts of "revolution" and "paradigm" (Gillies 1992), Lakatos' "scientific research programs" (Hallett 1979a, b), and with ideas taken from the sociology of knowledge (MacKenzie 1993), which in an extreme version led to David Bloor's "strong program" (Bloor 1991). More recently, it has comprised reliance on ideas such as "research schools" (Parshall 2004), "traditions" (Rowe 2004), "images of science" (Corry 1989; Bottazzini and Dahan-Dalmédico 2001), "epistemic configurations" (Epple 2004), "material culture of science" (Galison 2003),[3] quantitative analyses (Goldstein 1999; Wagner-Döbler and Berg 1993), and various others.

When discussing mathematics in association with literature, art, or music, on the other hand, it is important to stress the obvious, namely, that in fields like art, literature and music, considerations of objectivity, universality, testability, and the like, if appearing at all as part of the aims of the artists or of the audiences, emerge in ways that differ sensibly from those of science (see, e.g., Corry 2007a, b, c; Engelhardt and Tubbs 2021). No less important is to keep in mind the different relationship each of these domains entertain with its own past and history. Many definitions of modernism put at their focus the idea of a "radical break with the past", and such definitions will necessarily apply in sensibly different ways to either the arts or to mathematics. Being guided, above all, by the need to solve problems and to develop mathematical theories, the kinds, and breadth of choices available to a mathematician (and, in particular, choices that may lead to "breaks with the past") are much more reduced and clearly constrained than those available to the artist. In shaping their artistic self-identity and in defining their creative agenda, modernist artists can choose to ignore, and even oppose any aspect of traditional aesthetics and craftsmanship. This implies taking professional risks, of course, especially when it comes to artists at the beginning of their careers, but it can certainly be done and, in fact, has been successfully achieved by prominent modernists.

The choices open before aspiring mathematicians intent on making "a radical break with the past" while remaining part of the mathematical community are much more reduced. Artists might decide to develop their work and career by innovating within the field to the extent that cuts all connection with the contemporary mainstream in the relevant community. The aspiring mathematicians, in contrast, must fully assume the central values of the professional ethos to become part of the guild. They will abide by the rules of classical logic and gain complete control

[3] Although more naturally seen as dealing with the history of physics, Galison's book devotes considerable attention to Poincaré's mathematics as well.

of the accepted "mathematical craftsmanship" in their field of choice. They must publish in the mainstream mathematical journals and will typically strive to do so in those broadly considered to be the leading ones. Moreover, in very few cases will an already established mathematician come up with radical proposals for changes in the standards of the field.[4]

The kind of radical changes that have affected mathematics, especially modern mathematics, touch upon the images of knowledge, and particular to innovative ways of organizing knowledge into sub-disciplines (as in the case of "Modern Algebra" (Corry 2007b)) or developing new methodologies over older ones (as in the case of computer-assisted number theory (Corry 2007c)).

When it comes to the relationship between mathematics and other scientific disciplines, particularly physics, however, there are important points to stress. Thus, given their radical new approach to the basic concepts of physics—time, space, matter, causality—it seems natural that historians undertaking the question of modernism in science and the arts, turned to the theory of relativity and quantum mechanics as a fundamental bridge across domains during the period in question. The aforementioned work of Forman is a foremost example of this trend. Indeed, Forman stressed, while focusing his account specifically on the impact of Oswald Spengler's ideas, that attempts at drawing such bridges were at the very heart of Weimar culture. Spengler's account of Western culture draws fundamental parallels between art, mathematics, science, culture, and society, and the main contribution of Forman's analysis is found in the detailed description of the strong impression this perspective on history caused both on scientists and mathematicians.

Additionally, there are more or less successful attempts at understanding these bridges that could be mentioned here (Miller 2000; Vargish and Mook 1999).[5] And yet, in spite of the disciplinary closeness between physics and mathematics, there are some important differences that affect our discussion here, particularly concerning

[4] The most prominent example that would come to mind is that of Luitzen J.E. Brouwer, whose doctoral advisor urged him to delete the more philosophical and controversial parts of the dissertation and to focus on the more mainstream aspects of mathematics that it contained. It was only somewhat later, as he became a respected practitioner of a mainstream mathematical domain, that he started publishing and promoting his philosophical ideas, and to devote his time and energies to developing new kinds of radical, intuitionistic mathematics. Brouwer promoted a kind of logic, later called "intuitionistic logic", deviating from the mainstream but not implying a call to abandon classical logic, but rather to revert logic to a previous stage in its evolution, where no considerations of the actual infinite had (wrongfully and dangerously, from his perspective) made deep headway into mainstream mathematics. See (van Dalen 1999, 89–99). Another interesting case is that of Doron Zeilberger's call, after a distinguished carrer in classical disciplines, for an abandonment of "Human-Supremacist", "human-generated, and human-centrist 'conceptual' pure math mathematics" in favor of computer-generated, "experimental mathematics".

[5] In an illuminating article about the use of the terms "classical" and "modern" by physicists in the early twentieth century, Staley (2005) addresses this difference from an interesting perspective. In his opinion, whereas in physics discussions about "classical" theories and their status were more significant for the consolidation and propagation of new theories and approaches than any invocation of "modernity", in mathematics, different views about "modernity" were central to many debates within the mathematical community.

what I have elsewhere called the “reflexive character of mathematical knowledge,” on which I want to comment briefly as part of this preliminary discussion.

Mathematics is in a unique position among the sciences to allow an investigation of aspects of the discipline with tools offered by mathematics itself (Corry 1989). Entire mathematical disciplines that arose in the early twentieth century are devoted to this quest: proof theory, complexity theory, category theory, etc. These analyze specific aspects of mathematical practice and mathematical theories, and do so with the help of tools provided by the discipline and with the same degree of precision and clarity that is typical, and indeed unique, to mathematics. Gödel’s theorems, for instance, involve results about the limitations of deduction mathematical theories defined by systems of axioms. The way that new methods were explicitly introduced to prove them does not differ from the way this is done in other mathematical situations. Biology, for example, cannot self-analyze the discipline with tools taken from the discipline itself, as biological theories are not biological entities.

On the other hand, literature can become the subject matter of a literary piece; painting can become the subject matter of painting, and so on, for other artistic endeavors. But whatever these domains can express about themselves, they will do so differently from what mathematics can say about itself. This unique feature of mathematics is not remarkable in itself and is also specifically relevant to the discussion of modernism, given the dual fact that (1) the reflexive study of the language and methods of specific cultural fields has very often been taken to be a hallmark of modernism in the arts, as I will stress below, and that (2) this reflexive character of mathematics became so prominently developed in the period that interests us here.

The differences and tensions arising in this complex, triangular relationship between mathematics, the natural sciences and the arts must be considered and stressed explicitly in any serious analysis of mathematics and modernism. This relationship, moreover, is subject to ongoing changes and conditioned by historical circumstances. Hence, a proper examination of the historical processes under which the three realms evolved in the period that interests us here, and their possible interactions, is necessary for such an analysis and to shed new light on the history of modern mathematics. Whatever one may want to say about modernism in mathematics and its relationship with modernism in other fields, one must remember that the changing relationship among the fields must be taken to be part of this historical phenomenon.[6]

[6] An even broader and more comprehensive such analysis should also pay attention to philosophy and the social sciences with their own specificities, but for reasons of space I will leave them outside the scope of the present discussion. See, e.g., (Ross 1994; Vrahimis 2012).

16.3 Modern Mathematics and Modernist Art

I move now on to examine some existing works explicitly addressing the connections between mathematics and the arts in the period between1890–1930 and to comment on them against the background of the ideas discussed in the previous section. First, I focus on an analysis of the possible influences of mathematics and the sciences on the arts. Then, I move to consider the opposite direction, which includes the important contribution of Jeremy Gray.

An outstanding example of analysis of the influence of physics and mathematics on modern art in the early twentieth century appears in the work of Linda Henderson (Henderson 1983, 2004, 2005, 2007). Henderson explored the ways in which certain scientific ideas dominating the public imagination at the turn of the century, provided "the armature of the cultural matrix that stimulated the imaginations of modern artists and writers" (Henderson 2004, 458). Artists who felt the inadequacy of current artistic language to express the complexity of new realities newly uncovered by science (or increasingly perceived by public imagination) were pushed into pursuing new directions of expression, and hence contributing to the creation of a new artistic language; the modernist language of art. But in showing this, Henderson also studiously undermined the all-too-easy, and often repeated image of a putative convergence of modern art and modern science at the turn of twentieth century in the emblematic personae and personalities of Picasso and Einstein.[7] Contrary to a conception first broadly and famously promoted in Sigfried Giedion's *Space, Time and Architecture* (Giedion 1941), Einstein's early ideas on relativity were not at all known to Picasso at the time of consolidating his new cubist conceptions. More generally, it was not before 1919, when in the wake of the famous Eddington eclipse expedition, Einstein was catapulted to world fame, as the popularizations of relativity theory captured public and artistic imagination (Levenson 2003, 218–37). It was only then that ideas of space and time related to relativity did offer new metaphors and opened new avenues of expression that some prominent artists undertook to follow. As Henderson's work illustrates, it was not relativity but central ideas stemming from classical physics in the late nineteenth century that underlie the ways in which science contributed to creating new artistic directions in the early modernist period. These ideas were related above all with the *ether*, but also with other concepts and theories that stressed the existence of suprasensible, invisible physical phenomena. These "invisible phenomena" comprised the discovery of X-rays, radioactivity, the discourse around the fourth dimension (especially as popularized through the works of the British hyperspace philosopher Howard Hinton (1853–1907)), and the idea of the cosmic consciousness introduced by the Russian esoteric philosopher Pyotr Demianovich Ouspenskii (1878–1947).

Henderson offers a superb example of how, by looking into the development of science, we can gain new insights into the issue of modernism in art. The main

[7] A typical version of which appears in (Miller 2002).

thread of her account emerges from within the internal development in the arts and focuses on some crucial historical crossroads where substantial questions about the most fundamental assumptions of art and of its language arose at the turn of the twentieth century. Faced with these pressing questions, certain artists sought to come to terms with these by looking for new ideas and directions of thought. Henderson then separately focuses on contemporary developments in science, developments that, in themselves, had nothing to do with modernism or with modernist *Zeitgeist*, and shows how these developments afforded new concepts, a new imagery, and new perspectives that the artists could take as starting point for the new ways they were attempting to develop in their own artistic quest. Thus, in Henderson's narrative there is no assumption of common ideas or common trends simultaneously arising in both realms for some unknown reason. In fact, whether the main scientific ideas were properly understood by the artists in question is not a truly relevant point in her account. She shows, in this way, how public perception of scientific ideas—not necessarily the truly important or more revolutionary ones at the time—played a central role in the consolidation of major trends and personal styles in modernist art (Cubism, Futurism, Duchamp, Boccioni, Kupka, etc.). Science appears here as offering a broadened world of ideas, metaphors, and images from which the artists could pick according to their needs, tastes, and inclinations.[8]

Henderson's work thus offers a remarkable example of an approach that, the direction of the impact were to be turned around, has the potential to lead to a truly illuminating attempt at making sense of modern mathematics as part of the broader cultural phenomena of modernism. Such an approach would ideally involve two steps:

1. The historian should first take a fresh look at the overall developments in the discipline of mathematics—its results, its language, its foundational conceptions, its relationships with neighboring disciplines, its institutions and values—to trace those places where the discipline and its practitioners face in this period of time *an inadequacy to address, in terms of the existing disciplinary tools, the complexity of a new reality*. This inadequacy may well manifest itself in terms of a deep crisis or anxiety systematically surfacing in the disciplinary discourse, that historians should identify and articulate.
2. In a second crucial step, the historian should provide an account of the ways in which this inadequacy was addressed by mathematicians following new paths. In this account, external inputs from the arts, music, architecture or philosophy would become instrumental in helping shape the course of events that transformed the discipline at the turn of the twentieth century.

Whether or not such an approach may successfully be applied to understanding in new ways specific situations in the development of modern mathematics is yet to be seen. At this point, I would like to take a brief look at two seminal books that

[8] Similar in this respect, with an emphasis on mathematics, are the account presented in (Gamwell 2015).

undertook the most thoroughgoing analysis to date of modern mathematics as part of the more general cultural phenomenon of modernism and to analyze, relying on the scheme suggested above, the scope and impact of their undertaking. One of them is, as already stated, Jeremy Gray's *Plato's Ghost,* but I start with Herbert Mehrtens' *Moderne-Sprache-Mathematik* (Mehrtens 1990), which pioneered the trend.

Mehrtens explicitly connected some of the basic features commonly associated with modern mathematics to modernization processes and their manifestations in various fields of culture and society. He examined the impact of the rise of new types of industries and professions (e.g., in the insurance area), and of trends in higher education. In his analysis, he incorporated—among other things—semiotic concepts and philosophical insights drawn from authors like Foucault or Lacan. He accorded prime importance to an examination of mathematical language while stressing a three-fold distinction between different aspects of the latter: (1) mathematics *as a* language (*Sprache Mathematik*), (2) the language used in mathematical texts (comprising systems of terms and symbols that are combined according to formal rules stipulated in advance), and (3) the language of mathematicians (*Sprechen der Mathematiker*), which comprises a combination of language used in fully formalized mathematical texts and texts written in natural language.

In these terms, Mehrtens discussed modernism in mathematics by referring to the main kinds of reactions elicited by the development of mathematics by the end of the nineteenth century, notably as they manifested themselves through debates about the source of its meaning in mathematics and about the autonomy of the discipline and its discourse. These reactions comprised a break with more traditional disciplines and a search for disciplinary autonomy of a kind and degree theretofore unknown in the field. In these terms, he identified two groups of mathematicians espousing diverging views. On the one hand, there was a "modern" camp represented by the likes of Georg Cantor (1845–1914), David Hilbert (1862–1943), Felix Hausdorff (1868–1942) and Ernst Zermelo (1871–1953). An increasing estrangement from the classical conception of mathematics was characteristic of their attitude as an attempt to explore some naturally or transcendentally given mathematical entities (such as numbers, geometrical spaces, or functions). They conceived the essence of mathematics to be the analysis of a man-made symbolic language, and the exploration of the logical possibilities spanned by the application of the rules that control this language. Mathematics, in this view, was a free, creative enterprise constrained only by fruitfulness and internal coherence. Hilbert was, in Mehrtens account, leading figure of this camp.

Concurrently, a second camp developed, denominated by Mehrtens as "counter-modern", led by mathematicians such as Felix Klein (1849–1925), Henri Poincaré (1854–1912), and Luitzen J. E. Brouwer (1881–1966). For them, the investigation of spatial and arithmetic intuition (in the classical sense of *Anschauung*) continued to be the primary thrust of mathematics. He also included mathematicians in this camp who lay their stress on real-world applications in physics, technology, economics, etc. The rhetoric of "freedom" of ideas as the basis of mathematics, initiated by Richard Dedekind (1831–1916) and enthusiastically followed by the modernist mathematicians (Corry 2017), was rejected, in Mehrtens' account, by the mathematicians of the counter-modernist camp, who priced above all finiteness, *Anschauung* and "construction." Brouwer appears here as the arch-counter-modernist. His idiosyncratic positions in both mathematical and political matters (as well as the affinities between Brouwer and the national-socialist Berlin mathematician (1886–1982)) allowed Mehrtens to identify what he saw as the common, counter-modernist traits underlying both levels (Mehrtens 1996).

An important and original point underlying Mehrtens' analysis is the emphasis on the simultaneous existence of these two camps and the focus on the ongoing critical dialogue between them, as the main feature of the history of early twentieth-century mathematics. This critical dialogue was, inter alia, at the root of a crisis of meaning that affected the discipline in the 1920s (the so-called "foundational crisis" (pp. 289–330)) and led to a redefinition of its self-identity. Moreover, by contrasting the attitudes of the two camps, Mehrtens implicitly presented the modernist attitude in mathematics *as a matter of choice* rather than one of necessity.

Mehrtens' book has been consistently praised for its pioneering status in the debate on modernism in mathematics and for the original approach, it has put forward. However, its limitations have also been consistently pointed out. Mehrtens' analysis focuses mainly on the programmatic declarations of those mathematicians he discusses and on their institutional activities. These are matters of real interest as sources of historical analysis, and it is worth stressing that the contents of mathematics are influenced by ideological considerations and institutional constraints. But as Moritz Epple remarks, in the final account, "Mehrtens does not attempt to analyze some of the more advanced productions of modernist or counter-modernist mathematicians, and makes, in fact, no claims about the internal construction of modern mathematics" (Epple 1997, 191).[9]

Thus, Mehrtens left many fundamental questions unanswered, and his argumentation was somewhat misleading. For one thing, the critical debate among "moderns" and "countermoderns" would appear to be, in Mehrtens' account, one that referred only to the external or meta-mathematical aspects, while being alien to questions of actual research programs, newly emerging mathematical results, techniques, or disciplines. In addition, the classification of mathematicians into these two camps, and the criteria of belonging to either of them, seems too coarse to stand the test of close historical scrutiny, and in the final account was too strongly circumscribed to the Göttingen mathematical culture. In this sense, Mehrtens' book, for all its virtues, falls short of giving a satisfactory account of "modern mathematics" as a "modernist" undertaking.

Having said that, I think that two fundamental elements of Mehrtens' analysis are highly relevant to any prospective, insightful analysis of modernism in mathematics. First is the possible, simultaneous existence of alternative approaches to mathematics that are open to choose, according to considerations that do not strictly derive from the body of mathematics itself. Some of the elements that Mehrtens identifies in distinguishing between moderns and counter-moderns seem to me

[9] See also (Epple 1996).

highly relevant, but I think they could be more fruitfully used by historians if approached in a less schematic way, namely, by realizing that in the work of one of the same mathematician (or, alternatively, in the works of several mathematicians associated with one and the same school or tradition) we can find elements of both the modern and the counter-modern trend. These various elements may interact and continuously change their relative weight along the historical process.

The second point refers to the historical processes that Mehrtens indicated as leading to the rise of modernist approaches in mathematics, namely the rapid growth of the discipline (together with other branches of sciences) by the late nineteenth century, and the enormous diversity and heterogeneity that suddenly appeared at various levels of mathematical activity (technical, language-related, philosophical, institutional). In this sense, Mehrtens follows the lead of those accounts of the rise of modernism in the arts that have presented it as a reaction to certain sociological and historical processes (such as urbanization, industrialization, or mechanization), and that in my view, if identified within the history of mathematics, may lead to new insights about the development of the discipline.

The second book to be mentioned here is Jeremy Gray's more recent *Plato's Ghost. The Modernist Transformation of Mathematics.* Gray's book provides a thoroughgoing account of the main transformations mathematics underwent in the period of our discussion, while comparing the main traits of these developments with the conceptions that previously dominated the discipline and that he schematically summarizes as "the consensus in 1880". Gray claims that the developments so described are best understood as a "modernist transformation". This concerns not just the changes that affected the contents of the leading mathematical branches but also additional aspects related to the discipline, such as its foundational conceptions, its language, its disciplinary relationship with physics, or even the ways in which the history of mathematics was now written or in which mathematics was popularized. Thus, for instance, Gray provides an illuminating survey of works written at the turn of the century, many of them by leading mathematicians, aimed at educated audiences of teachers, philosophers, psychologists, lawyers, and members of the Church. Such audiences, Gray suggests, reflected a new kind of growing interest of audiences that were ambiguous about science but wanted to hear about current developments (Gray 2008, 346–65). On the other hand, Gray remains less sure about the connection between modernist trends and the renewed interest in historical writing about mathematics (Gray 2008, 365–372).

Naturally, Gray is well-aware that "if the idea of mathematical modernism is to be worth entertaining, it must be clear, it must be useful, and it must merit the analogy it implies with contemporary cultural modernism." In addition, "there should be mathematical developments that do not fit at the very least those from earlier periods, and one might presume some contemporary ones as well." Accordingly, Gray's book opens with a characterization of modernism meant to provide the underlying thread of his analysis. In his own words:

> Here modernism is defined as an autonomous body of ideas, having little or no outward reference, placing considerable emphasis on formal aspects of the work and maintaining a complicated – indeed anxious – rather than a naïve relationship with the day-to-day world,

> which is the de facto view of a coherent group of people, such as a professional or discipline-based group that has a high sense of the seriousness and value of what it is trying to achieve. (Gray 2008, 1)

Gray intends this definition, not as a straitjacket determined by a strict party line but rather as an idea of a broad cultural field providing a perspective that may help historians integrate issues traditionally treated separately (including both technical aspects of certain sub-disciplines and prevailing philosophical conceptions about mathematics), or stressing new historical insights on previously unnoticed developments. Thus, for instance, the interactions with ideas of artificial languages, the importance of certain philosophers hitherto marginalized in the history of mathematics, the role of popularization, or the interest in the history of mathematics which had a resurgence in the said period.

One issue of particular interest raised by Gray in this context is that of "anxiety" (pp. 266–277). The development of mathematics in the nineteenth century is usually presented as a great success story, which certainly it is, and Gray does not dispute it. But at the same time, a growing sense of anxiety of a new kind, about the reliability of mathematics, the nature of proof, or the pervasiveness of error, was a recurrent theme in many discussions about mathematics, and this is an aspect that has received much less attention. Gray raises the point in direct connection with the anxiety that is often associated with modernism as a general cultural trait of the turn of the century. As an example of this anxiety, he calls attention to certain texts with such a manifest concern that historians previously overlooked or just regarded as isolated texts. Gray makes a clear and explicit connection between these texts with one another and with the broader topics of modernism.

Gray's book complements Mehrtens' in presenting a much broader and nuanced characterization of the discipline of mathematics in the period 1890–1930. On the other hand, in comparison with Mehrtens, Gray devotes much more attention to describing these characteristic features than to explaining the motivations and causes of the processes that ultimately led mathematics to become the kind of discipline that he aptly describes.[10]

Mehrtens' and Gray's books are, then, two significant attempts to approach the history of modern mathematics while relying on the idea of modernism as a historiographical category with significant explanatory added value. Against the background of my brief account and the many additional reviews of the books cited above, I return to my claim that for such attempts to be successful, it is necessary to focus more compellingly on showing (if possible) that the external processes that led to modernism in general and modern mathematics are similar and have common cultural roots.[11] One should not rule out the possibility that such kinds of external processes indeed took place and were meaningful in shaping the history

[10] For additional discussions on Gray's book, see (Feferman 2009; Rowe 2013; Schappacher 2012; Scholz 2010).

[11] An alternative, but not very convincing, way to connect mathematics with the general phenomenon of modernism appears in (Everdell 1997), where Cantor and Dedekind are presented

of mathematics. But in terms of existing research, little evidence of anything of the sort has been put forward by historians in rigorous detail thus far. The question, therefore, arises whether it is possible and illuminating to do so.

16.4 Greenberg's Modernist Painting and Modernist Mathematics

I proceed to discuss in this and in the next section two specific kinds of analysis of modernism that, while being completely unrelated to mathematics, do suggest directions that might be followed to turn the type of general directives delineated in the previous sections into concrete historical research. First, I discuss some ideas found in the writings of the celebrated and highly controversial art critic Clement Greenberg (1909–1994). For some historians of art, I should stress at the outset, Greenberg is total anathema and the foremost example of how the history of modern art should *not* be written and understood. Art historian Caroline A. Jones, for instance, described his views on modernism as "extraordinarily narrow" and as not proving "capacious enough for much painting of the modern period (even much "great painting", *pace* Greenberg)" (Jones 2000, 494). Jones published the most comprehensive account to date of Greenberg's writings and influence (Jones 2005). The reader willing to take the challenge of her ambitious book will get the direct taste of the kind of passionate opposition (and attraction) that the "Greenberg effect" has aroused among its critics.

Still, I find it pertinent to call attention to some of Greenberg's texts for their high suggestivity for the main aim of this article. Being an outsider to the world of art criticism, I can bypass the question of whether Greenberg's characterization of modernism in art is comprehensive enough. Likewise, I can certainly ignore the ways in which he allegedly turned his view from *descriptive* to *normative*, i.e., that he did not limit himself to providing a historical explanation of the process that led to the creation and predominance of certain styles in twentieth century art, but he also wanted to determine, along the same train of ideas, what good art is and should be.[12] Greenberg was certainly not just a detached commentator but a main figure, strongly involved in the art scene in New York who had the power and the tools to build and destroy at will the careers of many an artist. His support of Jackson Pollock is a well-known chapter of his achievements in this regard, and so is his very negative attitude towards Marcel Duchamp and Ad Reinhardt.

as the true (unaware) initiators of modernism because the way in which they treated the continuum in their mathematical work. See also (Pollack-Milgate 2021).

[12] One can find in Greenberg's own texts support for such a view, but in other places he emphatically denied that his analysis was ever intended as anything beyond pure description. See, e.g., (Greenberg 1983): "I wrote a piece called 'Modernist Painting' that got taken as a program when it was only a description."

Good examples of Greenberg's insights that I deem valuable for the present discussion are found in a famous article of 1960, "Modernist Painting", where he characterized the essence of modernism in terms that, if unaware of the context, one could easily take to be a description of modern mathematics. He thus wrote:

> The essence of Modernism lies, as I see it, in the use of characteristic methods of a discipline to criticize the discipline itself, not in order to subvert it but in order to entrench it more firmly in its area of competence. ... The self-criticism of Modernism grows out of, but is not the same thing as, the criticism of the Enlightenment. The Enlightenment criticized from the outside, the way criticism in its accepted sense does; Modernism criticizes from the inside, through the procedures themselves of that which is being criticized. (Greenberg 1995, 85)

Indeed, the reflexive character of mathematics (discussed above) reached a distinctive peak at the turn of the twentieth century and became the main tool for discussing and indeed criticizing the discipline. Think of the foundational works of Frege, Russell, Hilbert, Brouwer, Weyl or Gödel. As in Greenberg's description, this "criticism" worked from within, using the tools of the discipline, meant not to subvert it, but rather to entrench its status.[13]

For Greenberg, the source of this new kind of criticism coming from within could be traced back to Kant. It would seem natural that, given the essentially critical nature of the discipline, philosophy would engage in this kind to self-criticism, and Kant took it to new heights in his critical philosophy exploring the conditions of production of philosophy itself. However, Greenberg raised an interesting historical point here, relevant to our account. As the eighteenth century wore on, more rational justifications started to emerge in other disciplines as well, eventually reaching the arts. The latter, according to Greenberg, had been denied by the Enlightenment a serious task and the arts were thus gradually reduced to "pure and simple" entertainment. A type of Kantian self-criticism that would explore the conditions of production of art from within art itself (and here he meant mainly the visual arts) appeared as a possible way to redefine the kind of experience that would stress what is valuable in art in its own right and, particularly, what could not be obtained from any other kind of activity. Herein lies Greenberg's explanation of the origin, the essence, and indeed the justification of modernist art:

> Each art had to determine, through its own operations and works, the effects exclusive to itself. ... It quickly emerged that the unique and proper area of competence of each art coincided with all that was unique in the nature of its medium. (Greenberg 1995, 86)

And in the case of painting this led Greenberg to characterize modernism in terms of a preoccupation with two main dimensions of this artistic activity, namely, (1)

[13] It is worth stressing, however, that the issue of self-criticism and the ability of an individual (or a collective for that matter) to effectively distance himself from the normative framework in which he functions in order to be self-critical and innovative is a truly complex one, when considered from a broader philosophical point of view. For a through discussion that examines the views of philosophers like Brandom, Friedman, Davidson, Habermas, Rorty, and others, see (Fisch and Benbaji 2011).

the intrinsic fact of painting's *flatness* and the inherent physical delimitation of this flatness and (2) the gradual tendency of painting (recognizable since the last third of the nineteenth century) to estrange itself from the classical task of representation while occupying itself increasingly with questions of its nature. Thus, these two main characteristic features—painting's preoccupation with the question of flatness and its limitations—appear here as a direct consequence of the self-critical processes that Greenberg described above:

> It was the stressing of the ineluctable flatness of the surface that remained, however, more fundamental than anything else to the processes by which pictorial art criticized and defined itself under Modernism. For flatness alone was unique and exclusive to pictorial art. The enclosing shape of the picture was a limiting condition, or norm, that was shared with the art of the theater; color was a norm and a means shared not only with the theater, but also with sculpture. Because flatness was the only condition painting shared with no other art, Modernist painting oriented itself to flatness as it did to nothing else. (Greenberg 1995, 86)

Greenberg's focus exclusively on the question of flatness as *the* defining feature of modernist art has been one of the main points of criticism directed against him. We need not enter a debate about that here. What I do learn from Greenberg's analysis, however, is a possible underlying explanation of the historical conditions for the rise to pre-eminence of what Greenberg sees as Kantian-like self-criticism (art analyzing art with the tools of art alone) and which appears as a primary characteristic trait of modernist art. Since as already indicated, this kind of critical approach is also strongly distinctive of modern mathematics (and especially of the foundational quests typical of the turn of the twentieth century: mathematics analyzing the foundations and the limitations of mathematics with the tools of mathematics alone, and without the help of external, philosophical and metaphysical arguments) we are led to wonder about a possible new focal point of analysis airising from Greenberg's approach to the question: was the rise of a new kind of modern mathematics related to a search for what was unique and exclusive to mathematics and the peculiar nature of its medium? And if so: why did mathematicians engage in this search? What happened in, say, the last part of the nineteenth century, and not before that, that prompted at that time this kind of search, and what were the consequences of it?

We may then ask these questions for mathematics in general and not just for those places that are typically associated with modernist trends, namely the new kind of foundational research that appeared in the works of Frege, Russell, Hilbert, and others at the turn of the twentieth century. I will return briefly to these questions in the concluding section. At this point, I want to stress that an analogy with Greenberg's analysis might, in principle, help us understand the origins and causes of the processes (social, institutional, disciplinary, philosophical, internal, etc.) behind the rise of modern mathematics and not just to check against a checklist of features characteristic of modernism in art.

It is enlightening to consider some additional points raised by Greenberg, which are relevant to our discussion. Thus, for instance, strongly connected with the previous issue, Greenberg stressed the centrality of the quest for the autonomy of art. The impact of the process of self-criticism was translated, in Greenberg's analysis, to a focused search for "purity" in art as the guarantee for the preservation of the necessary standards,[14] and consequently, the status of the medium of art was transformed. In Greenberg's words:

> Realistic, naturalistic art had dissembled the medium, using art to conceal art; Modernism used art to call attention to art. The limitations that constitute the medium of painting – the flat surface, the shape of the support, the properties of the pigment – were treated by the Old Masters as negative factors that could be acknowledged only implicitly or indirectly. Under Modernism these same limitations came to be regarded as positive factors and were acknowledged openly. Manet's became the first Modernist pictures by virtue of the frankness with which they declared the flat surfaces on which they were painted. The Impressionists, in Manet's wake, abjured underpainting and glazes, to leave the eye under no doubt as to the fact that the colors they used were made of paint that came from tubes or pots. Cézanne sacrificed verisimilitude, or correctness, in order to fit his drawing and design more explicitly to the rectangular shape of the canvas. (Greenberg 1995, 86)

Again, the analogy with mathematics seems to me highly suggestive, but we need to analyze its validity very carefully. The search for autonomy, and eventually even segregation, is an acknowledged characteristic of at least certain essential parts of modern mathematics. In this sense, the analogy with modern art is evident and has often been mentioned. But what were the reasons for this? We are well aware of important internal, purely mathematical dynamics of ideas leading to the rise of a new kind of approach and practice that stressed the need for the autonomy of mathematical discourse and mathematical methods. Here perhaps with the help of a perspective similar to that suggested by Greenberg for art, we may look for some other, more external kinds of causes in the case of mathematics. The increased search for purity in mathematics can be related to a specific attempt to "guarantee its standards of quality". But what about "limitations that constitute the medium" of mathematics, that were treated by the Old Masters as negative factors that could be acknowledged only implicitly or indirectly", and that in modern mathematics could come "to be regarded as positive factors" and to be "acknowledged openly"? This appears as a remarkable, and far from self-evident characterization of modern art that Greenberg's analysis brings to the fore. But given the already mentioned, essentially inevitable, need to rely on historical continuity in mathematics (as opposed to the arts), a transposition of this kind of argumentation to mathematics is far from straightforward and requires additional care.

In further exploring this point, however, one might try to bring to bear ideas from sociologists of science such as Rudolf Stichweh, who has highlighted the systemic, interrelated character of discipline formation by the end of nineteenth century. Stichweh's analysis meant to show how the emergence and consolidation of an autonomous self-understanding of the various academic disciplines depended always on similar processes taking place in the neighboring disciplines at the same time (Stichweh 1984). Stichweh's perspective might open interesting avenues of research also for our topic, but at this point, I leave this as an open thread calling for further thought concerning the question of modernism in mathematics.

[14] A discussion of "purity" and its centrality in modernism, from a different perspective appears in (Cheetham 1991).

In referring to the "necessity of formalism" as the "essential, defining side" of modernism (at least in the case of painting and sculpture), Greenberg added another interesting explanation that seems very suggestive for mathematics as well:

> Modernism defines itself in the long run not as a "movement", much less a program, but rather as a kind of bias or tropism: towards aesthetic value, aesthetic value as such and as an ultimate. ... This more conscious, this almost exacerbated concern with aesthetic value emerges in the mid-nineteenth century in response to an emergency. The emergency is perceived in a growing relaxation of aesthetic standards at the top of Western society, and in the threat this offers to serious practice of art and literature. (Greenberg 1971, 171)

Keeping in mind that terms such as "formal", "abstract" or "aesthetic" have significantly different meanings and elicit different contexts in mathematics and in the arts, one can still ask whether the idea of associating the entrenchment of formalist approaches as part of the consolidation of modern mathematics with a reaction to an emergency, as described here by Greenberg for the arts, may bring with it new insights. Moreover, we can also ask if the "emergency" in question was not only similar but perhaps even the same one in both cases. I already mentioned the issue of "anxiety" discussed by Gray concerning the development of mathematics at the turn of the nineteenth century, which he related to what some mathematicians conceived as a relaxation of standards. There is no doubt that formalism in mathematics can be associated with a possible reaction to such a relaxation. Thus, formalism may appear here not just as a common trait perceived in mathematics and art but also as motivated by similar concerns in both cases. More on this I will say in the next section.

Finally, I would like to mention yet another suggestion of Greenberg that may be relevant for historians of mathematics in their field, as it touches upon the supposed radical break with the past that appears in so many characterizations of modernism. In an article entitled "Modern and Postmodern," Greenberg wrote:

> Contrary to the common notion, Modernism or the avant-garde didn't make its entrance by breaking with the past. Far from it. Nor did it have such a thing as a program, nor has it really ever had one. It's been in the nature, rather, of an attitude and an orientation: an attitude and orientation to standards and levels: standards and levels of aesthetic quality in the first and also the last place. ...
>
> And where did the Modernists get their standards and levels from? From the past, that is, the best of the past. But not so much from particular models in the past – though from these too – as from a generalized feeling and apprehending, a kind of distilling and extracting of aesthetic quality as shown by the best of the past. (Greenberg 1980)

I find it remarkable that Greenberg would stress this point in opposition to what so many considered an unavoidable trait of modernism. As I said above, truly radical breaks with the past seem rather unlikely in mathematics. As Greenberg stresses here, modernism may arise not from a radical break but from a conscious process of distilling and extracting quality from what proved to be the best practice in the past. I think that in laying the central elements of modern mathematics, some of the most influential mathematicians of the turn of the century acted precisely in this way. This was undoubtedly the case, as I have discussed in detail elsewhere, with

Dedekind’s early introduction of structuralist concerns in the algebra (Corry 2017) and with Hilbert’s introduction of the modern axiomatic approach (Corry 2004).

As already stated, however, there are also good reasons to react to Greenberg’s views with great care. It is not only that they are very much debated among art historians, but also that Greenberg did not write systematic, scholarly texts with all due footnotes and references. Most of his writings appeared as scattered articles, conferences, etc., and they sometimes follow a somewhat associative style. Thus, one must not be surprised to find deep changes and possibly conflicting views in them throughout the years.

And yet, even if the criticism directed at him is well taken, especially when one tries to apply his view to analyzing in detail the works of specific individual artists, this does not mean that the essential structure of the processes he describes cannot be reconstructed for the purposes I am pursuing here, and then followed in a more scholarly solid fashion. If one were able to develop explanations of these kinds for mathematics, then it may turn out that it is not only justified and valuable to use the term modernism in the context of the history of mathematics but also that it is not just a coincidence that modernism appears in mathematics as well as in the arts nearly contemporarily, and that this coincidence can be made sense of in more or less tangible terms.[15]

16.5 Wittgenstein’s Vienna and Modernist Mathematics

The second source I want to refer to in the search for ideas relevant to a possible fruitful discussion of modernism in mathematics is the book *Wittgenstein’s Vienna* by Allan Janik and Stephen Toulmin. The main topic of this book is an interpretation of the roots and meaning of Wittgenstein’s *Tractatus*. Contrary to accepted views—according to which the fundamental questions underlying the treatise were epistemology, philosophy of science, and logic taken for their own sake—the authors aimed at presenting Wittgenstein as a thinker deeply rooted in the intellectual life of Vienna at the turn of the twentieth century, for whom the question of language and its limitations was mainly an *ethical* concern and not merely a linguistic-analytic one. These ethical concerns, they contended, can

[15] Greenberg, of course, is not the only one to discuss modernism in terms of the processes that led to its rise, rather than by just providing a checklist of characteristic features. Also worthy of mention here is the work of Dan Albright (Albright 1997, 2000), who stresses the crisis of values in art that led to modernism. In his view, if in previous centuries, artists, writers, and musicians could be inherently confident about the validity of the delight and edification they provided to their audiences, during the twentieth century art found itself in a new and odd situation, plagued with insecurity. Faced with the crisis, radical claims about the locus of value in art were advanced in various realms at nearly the same time. The various radical modernist manifestoes thus produced reflect the need of the artist not only to create, as was always the case in the past, but also to promote new standards of value and to provide some new kind of justification to the very existence of art.

only be fully understood against the background of Viennese modernism in its manifold manifestations. Like in the case of Greenberg, I do not intend to come up here with an appraisal of Janik and Toulmin's analysis as a contribution to the Wittgenstein scholarship, but rather to focus on ideas potentially relevant to the topic of mathematics and modernism.

The basic question that Janik and Toulmin pursued thorough the book appears right at the beginning, phrased in the following terms:

> Was it an absolute coincidence that the beginnings of twelve-tone music, 'modern' architecture, legal and logical positivism, nonrepresentational painting and psychoanalysis – not to mention the revival of interest in Schopenhauer and Kierkegaard – were all taking place simultaneously and were largely concentrated in Vienna? (Janik and Toulmin 1973, 18)

The central hypothesis of the book is that "to be a *fin-de-siècle* Viennese artist or intellectual conscious of the social realities of Kakania [a term coined by Robert Musil to describe Austro-Hungarian society disparagingly (L.C.)] one had to face the problem of the nature and limits of language, expression, and communication. (p. 117)" Accordingly, they offered an account of the deep changes that affected art, philosophy, and other aspects of cultural life around 1900 in Vienna, as interrelated attempts to meet the challenges posed by questions of communication (language), authenticity, and symbolic expression. These challenges, in turn, derived from the deep social changes that affected the capital city of the Habsburgs: a medley of interacting tongues, the tension between the central imperial rule and the local and national aspirations, liberalization alongside decentralization of the traditional centers of power, changes in the production processes and social structures. And in this context, the most crucial instance of the philosophical side of this sweeping cultural phenomenon arose in the person whose writings, in their view, embodied the crucial influence on Wittgenstein, Fritz Mauthner (1849–1923), who developed a unique doctrine of "Critique of Language" (*Sprachkritik*) in several interesting books and is one of the few persons mentioned by name in the *Tractatus*.

In the received interpretation of Wittgenstein, the importance of this reference to Mauthner is often downplayed, but Janik and Toulmin made it the centerpiece of their analysis. Their alternative (and, as I see it, enlightening) approach to Wittgenstein affords a useful perspective for our discussion since the authors did not limit themselves to indicating general analogies between various fields of activity or a common, putative underlying ethereal *Zeitgeist* but instead emphasized concrete historical processes that were motivated by similar concerns stemming from the specific historical circumstances of turn-of-the-century Vienna.

Incidentally, an important focus of attention for Janik and Toulmin is found in contemporary science, and in the works of Ernst Mach (1838–1916), Heinrich Hertz (1857–1894), and Ludwig Boltzmann (1844–1906). For these three scientists, as it is well known, metaphysics had no place in science, and they devoted conscious and systematic efforts at finding those places where metaphysics had subtly but mistakenly been incorporated. This task, however, was not pursued in the same way by the three of them. Janik and Toulmin describe them as representing significantly different stages in a continuous process. Mach represents a first stage where the limits of physics were set “externally”, as it were, employing a more philosophical analysis. Hertz and Boltzmann, on the contrary, by following an approach that can retrospectively be described as “axiomatic”, pursued the same task “from within.” Hertz and Boltzmann sought to set the correct limits of physical science through an introspective analysis using the tools of science (and here, of course, we find a remarkable similarity with Greenberg’s stress on “criticism from within,” as explained above).

The interesting point in their analysis, however, is that they embedded this two-stage process in the more general, broad historical processes that underlie all other manifestations of Viennese modernism. First and foremost, among these manifestations were, for them, the processes leading from Mauthner to Wittgenstein. The philosophical critique of language undertaken by Mauthner as a response to the need mentioned above to establish the “limits of language, expression, and communication” starts from a point that is similar to that of Mach’s attempt in physics. And very much like Hertz and Boltzmann had further pursued Mach’s quest, but by way of an alternative, more internally focused path, so did Wittgenstein in relation to Mauthner. Hertz and Boltzmann, according to Janik and Toulmin, “had shown how the logical articulation and empirical application of systematic theories in physical science give one a direct *bildliche Darstellung* of the world, namely, a mathematical model which, when suitably applied, can yield true and certain knowledge of the world. And they had done so, furthermore, in a way that satisfied Kant’s fundamentally antimetaphysical demands – namely, by mapping the limits of the language of physical theory entirely “from within” (p.166). In similar terms, Janik and Toulmin presented the philosophical work of Wittgenstein as a continuation of Mauthner’s, in which the limits of language, in general, were mapped from within. They also examined and laid all the necessary stress on the ethical outlook, which in their interpretation, was so central to Wittgenstein’s undertaking and arose from the writings of Kierkegaard and Tolstoy. (Of course, this main element played no role in the story about Mach, Hertz and Boltzmann.)

The sociocultural elements underlying both aspects of the story, as described above, are expanded subtly to cover other fields of activity along the same lines: music, architecture, journalism, law, painting, and literature. And in all these fields Janik and Toulmin also added a third stage that was produced along the lines of commonly characterized historical processes. Thus, for instance, the three stages in music are represented by Gustav Mahler, then Arnold Schönberg, then Paul Hindemith. In the case of architecture, it is Otto Wagner, then Adolf Loos, and then Bauhaus.[16] And in the case of philosophy, the stage after Wittgenstein (who came

[16] (Galison 1990) presents an analysis that complements this view and locates the Bauhaus movement in relation with logical positivism, as part of Viennese modernism.

after Mauthner) is that of logical positivism.[17] The process which is common to all these threads can be briefly described as follows:

> In architecture as in music, then, the technical innovations worked out before 1914 by the 'critical' generation of Schönberg and Loos were formalized in the 1920s and 1930s, so becoming the basis for a compulsory antidecorative style which eventually became as conventional as the overdecorative style which it displaced. And we might pursue these parallels still further if we pleased – into poetry and literature, painting and sculpture, and even into physics and pure mathematics. In each case, novel techniques of axiomatization or prung rhythm, operationalism or nonrepresentational art, were first introduced in order to deal with artistic or intellectual problems left over from the late nineteenth century – so having the status of interesting and legitimate new *means* – only to acquire after a few years the status of *ends*, through becoming the stock in trade of a newly professionalized school of modern poets, abstract artists or philosophical analysts. (p. 254)

Not *Zeitgeist*-like arguments or superficial analogies, then, as part of the explanation, but a common ground to all these processes, namely, "a consistent attempt to evade the social and political problems of Austria by the debasement of language."

Can the insights of Janik and Toulmin be imported into the historiography of mathematics fruitfully? It is curious that the passage quoted mentioned pure mathematics as having been affected by the same circumstances as other cultural manifestations. They do not give details about what they may have had in mind when saying this. Indeed, at a different place, they did state that "in a very few self-contained theoretical disciplines—for example, the purest parts of mathematics—one can perhaps detach concepts and arguments from the historic-cultural milieus in which they were introduced and used and consider their merits or defects in isolation from that milieu" (p. 27). But my point is not what was done in the book in relation to mathematics, but what *could be done* by analyzing Viennese mathematics at the turn of the century from the perspective afforded by the book.

I briefly indicate here specific parameters that might be considered in an attempted answer. In the relevant period, Vienna did have an interesting, original, and very productive mathematical community. Its more prominent names included Wilhelm Wirtinger (1865–1945), Philipp Furtwängler (1869–1940), Eduard Helly (1884–1943), Kurt Gödel (1906–1978), Kurt Reidemeister (1893–1971), Witold Hurewicz (1904–1956), Walther Mayer (1887–1948), Johann Radon (1887–1946), Alfred Tauber (1866–1942), Olga Taussky (1906–1995), Heinrich Tietze (1880–1964) and Leopold Vietoris (1891–2002). Each of these mathematicians arrived in Vienna from different places at different times, bringing their baggage drawn from the mathematical traditions from which they stemmed. Of course, even before we start to consider the question that occupies us here, one should be able to come forward with a more articulate understanding of the Vienna mathematical community than we now have what the main mathematical fields were pursued, what kinds of interactions existed with the local scientific communities and with neighboring mathematical institutions, what were the internal mechanisms

[17] (Janik 2001, 147–69) discusses the somewhat different relation between Hertz's famous Introduction and the late Wittgenstein.

of production, training, and transmission of mathematical knowledge, etc. Such questions have been pursued, sometimes in detail, for Göttingen and Berlin, for the various USA centers of mathematics, and some British and Italian contexts, but, unfortunately, much less so for Vienna.

One existing work of relevant historical research does indicate, however, that there might be some room for pursuing this question along the broader conception of modernism, as suggested above. Moritz Epple has investigated the mathematical contributions of Kurt Reidemeister on Knot Theory in the 1920s while comparing it with work conducted simultaneously in the same field at Yale, on the one hand, and at Princeton, on the other hand. Epple discussed the intellectual atmosphere of the city as part of the relevant intellectual background to Reidemeister, and his discussion on the rise of modern topology is framed in the broader context of modernism in mathematics (Epple 1999, 299–322; 2004).

As part of his account, Epple stressed the existence of different paths into modernity that led to other varieties of mathematical modernism, even within the same institutional context, i.e., that of mathematics at Vienna, where Reidemeister worked alongside Wilhelm Wirtinger, producing different brands of modern mathematics (Epple 1999, 236–26). Reidemeister had strong intellectual interactions with Hans Hahn, Otto Neurath, Otto Schreier, and Karl Menger, all of them engaged in the activities of the Vienna Circle, which is of obvious relevance for any discussion on modernism. In addition, the most prominent members of the literary milieu in the city at the time had formal training in mathematics and strong connections with the local mathematicians. The three most famous examples of this are Robert Musil (1880–1942), Herman Broch (1886–1951) and Leo Perutz (1882–1957). The latter continued to be actively involved in mathematics throughout his life (Sigmund 1999; Engelhardt 2018, 2021).

What kinds of mathematics were done at the time in Vienna? To what extent can such kinds of mathematics be properly called modernist? Is this somehow connected with the work and the person of their Viennese neighbor Boltzmann? Inspired by Mehrtens' kind of analysis, Mitchell Ash has recently discussed the linkages between "modern ways of thinking about science" and the radical development of the visual arts in Vienna at the time. In his view, the significance of technological modernism "presupposes a concept of knowledge-based less on self-referential abstraction than on what can be done with, or to, nature as well as other human beings" (Ash 2018, 27). In an attempt to bring out basic features that link science and the arts in that specific cultural context, he illustrated the plurality of modernisms manifest in the sciences and culture of 'Vienna 1900′ by discussing the work of Ernst Mach and Ludwig Boltzmann, on the one hand, and the music of Arnold Schoenberg, on the other. For all the merits of Ash's analysis, the question remains open, whether the kind of mathematics practiced in Vienna was peculiar and different to what preceded it, and, more importantly, if the processes leading to the changes that brought about this possibly new conception are similar or similarly motivated as all the other complex processes described in Janik and Toulmin's book concerning Viennese culture in general.

It is relevant to stress that Epple's methodological proposals include reliance on Weber's idea of "patterns of rationality" as a way to contextualize the mathematical practice of a specific culture. But at the same time, his comparative analysis is based on the idea of an "epistemic thing", originally introduced by historians of experimental sciences (Rheinberger 1997). Epple uses this concept to explain in what senses Reidemeister's topological research differed from other, contemporary ones. In so doing, he suggested the plausibility that the specificity of Reidemeister's work was tied to Viennese intellectual modernism, even though, the nature of the tie remains to be explained.[18] In an ideal study of the mutual relationship between modernism and mathematics one might also be led to go the opposite direction, namely by understanding the specifics of Reidemeister's (and Hahn's and Menger's), and uncover new historical mechanisms behind the development of Viennese modernism.

A different, and perhaps highly relevant direction in which the analysis of Janik and Toulmin can provide illuminating hints to historians of mathematics has to do with the development of the modern axiomatic approach. I have devoted considerable attention in my research to the work of David Hilbert, to the centrality of the axiomatic approach for his work, and to the significant impact that this aspect of his work had on mathematics and physics in the early twentieth century, precisely at the time under discussion here. In my analysis, I have shown how the work of Hertz and Boltzmann had a direct influence on Hilbert and on the consolidation of the axiomatic approach and its application to both geometry and physics. I have also stressed the pervasive presence of Mach's ideas and his empiricist-oriented criticism in the background of Hilbert's work (Corry 2004, Chps. 2–3). Considering this, it is remarkable that in the three-stage model of Janik and Toulmin, precisely this thread, leading from Mach to Hertz and Boltzmann, which the authors single out as highly important, is not completed with its third stage. My suggestion here is that one might look at the process leading to, and at the consolidation of, Hilbert's new axiomatic approach precisely as that third stage. A historical analysis of the kind provided for Hilbert may plausibly be complemented with an eye on the types of processes described by Janik and Toulmin. In this way, the mathematics embodied in and promoted by Hilbert's approach could be seen as an aspect of mathematical modernism, not just because of a series of characteristic features associated with it, but rather because it might be seen as the outcome of a process with specific historical-cultural roots that gave rise to modernism in so many fields of culture at the time.

[18] Epple's description of the intellectual background to *fin-de-siècle* Vienna also strongly relies on the classical study (Schorske 1980).

16.6 Prospective Remarks

An emphasis on the formal, as opposed to thematic values; a rite of passage through avant-garde; a radical break with tradition (or even a "desire to offend tradition"); the wish to explore subjective experience as opposed to representing "outward experience"; a high degree of self-consciousness; a criticism of the basic principles of the discipline and its limits using the tools of the discipline. These are some characteristic features typically associated with modernism in its various cultural manifestations at the turn of the twentieth century. Some of them are mentioned and analyzed by the authors referred to above, and I take them to be an illustrative rather than an exhaustive sample of scholarly discussions on the topic of modernism. We may find such basic attitudes also in the mathematics of the period in question. Historians of modern mathematics might debate the degree to which such traits are central and pervasive and hence the extent to which it may be appropriate to describe, based on our current historical knowledge, modern mathematics as a modernist endeavor. My tentative proposal in this survey is, in contrast, that rather than exploring the topic in this straightforward way, we should ask ourselves if the perspective of modernism may lead us to look for new insights into making sense of the history of modern mathematics.

Considering, for instance, Jeremy Gray's emphasis on the sense of anxiety that arose at the end of the nineteenth century side-by-side with the enormous successes of the discipline. Talk about this success is standard in any historical account concerning this period, but the concomitant anxiety indicated by Gray has been much less discussed (if at all). By situating it in a modernist context, Gray draws our attention to the possibility that this is a more significant issue than we have realized thus far. He gives the example of an inaugural lecture delivered in 1910 at Tübingen by Oskar Perron (1880–1975). Perron was a proficient mathematician with acknowledged contributions to various fields, but his prominence was far from the high profile of a Hilbert or a Noether. Thus, one will not find his name often mentioned in discussions about mathematical modernism. But as Gray indicates, it is essential to hear what a mathematician like him had to share about his discipline at the turn of the twentieth century. In his lecture, Perron addressed mainly questions related to the gap between the public perception of mathematics and the actual practices in the discipline, particularly concerning the question of the certainty and exactness of its methods (Perron 1911). One should then ask whether this is an isolated phenomenon or a manifestation of a more generalized concern of the practitioners of mathematics at the time and whether, by looking at the kind of considerations discussed in the preceding sections, we can gain some innovative historical types of insights on this question.

Well, if we follow the lead opened by Gray, we do find instances that give us further food for thought. Thus, for instance, an interesting text by Alfred Pringsheim (1850–1941), who, like Perron, was a well-known mathematician and not one whose work is typically discussed in relation to modernism. In addition to his mathematical activities, Pringsheim was deeply immersed in the broad cultural trends of his time, and that to an unusual degree. He came from a wealthy Jewish family in Berlin who used his wealth to support art, and Alfred became a well-known art collector. He had a strong, well-cultivated musical background and became one of the earliest Wagner supporters. His daughter Katia, one of the first active women university students at Munich, married Thomas Mann. The family house in Berlin and his one in Munich (both known as "Palais Pringsheim") were prominent architectural icons (though they were far from any clue of modernist taste) (Perron 1952).[19] In 1904, on the occasion of the 145th anniversary of the Munich Academy of Sciences, Pringsheim gave a lecture entitled "On the Value and Alleged Lack of Value of Mathematics" (Pringsheim 1904). Without going into the details of the talk, I will say that it reflects the kinds of concerns addressed by Perron very closely. Incidentally, Pringsheim had been one of Perron's most influential teachers in Munich (Frank 1982).

One may mention additional texts that go in the same direction (von Mises 1922), and, more importantly, one is motivated now to look for more. Still, the question remains open whether we can *find* not just additional texts that serve as evidence for these kinds of concerns (which is quite likely), but rather if we can *understand their roots* and the processes leading to their rise and consolidation. And more specifically: whether these roots may be found to be directly connected, or at least closely related, to those found at the basis of modernism as a broad cultural phenomenon (and this, I think, is less likely, though still plausible).

Think, for instance, of Greenberg's explanation of the rise of modern painting in terms of the need existing in each art by the late nineteenth century to determine—purely with the help of its means—what was unique and exclusive to itself, according to the nature of its medium. The intense foundational activity, one of the acknowledged characteristic features of mathematics at the turn of the twentieth century, can be easily seen as a similar manifestation—purely from within the discipline—of the phenomenon indicated by Greenberg in the case of painting. Indeed, this is a point typically stressed in the debates about modernism in mathematics. But can we, in addition, explain the timing and the main thrusts of this foundational activity on the same grounds that Greenberg adduced for the arts? For Greenberg, in the wake of the Enlightenment, the arts were gradually assimilated into entertainment, pure and simple. This was a primary trigger that led to the kind of internally pursued self-criticism laying at the basis of modernism. Can we come up with a similar explanation in the case of mathematics? Dan Albright, to take another example, sees the roots of modernism as related to the new and odd situation for art, plagued with insecurity, as opposed to the confidence in the validity of the delight and edification it had provided to their audiences in previous times (Albright 2000). Perhaps this could be a fruitful lead to follow in connection with the topic of anxiety just mentioned above. Can we trace a direct relationship

[19] In this context it is natural to stress that also Wittgenstein was born to a privileged and immensely wealthy Viennese family, who generously supported the likes of Gustav Klimt and Alfred Loos as well as the poets Georg Trakl and Rainer Maria Rilke. The circle of friends of the Wittgenstein family included many distinguished figures of the Viennese musical milieu, such as Johannes Brahms (Monk 1991).

between the changes of status in the arts, with related changes of position in mathematics, with questions about certainty and the unity of mathematics, and with the increasing trend of foundational research at the turn of the twentieth century? Can the explanations of Janik and Toulmin about the centrality of the problem of language in the modernist culture of Vienna, and its social and ideological roots, be of any help in consolidating such an explanation?

Answering these questions would require, I believe, additional historical research taking into consideration that modernism is a historical phenomenon with an internal evolution and geographical specificities that are often overlooked. Thus, with the Great War in Europe precisely in the middle of the period that frames our discussion (1890–1930) and its profound social and cultural impact it is obvious that a single idea of "modernism" is too coarse to account for all the developments typically related with the term without further historicizing it. What is less obvious, but no less significant, are the differences among modernist cultures across the continent and in the USA. As mathematics is the quintessential universal endeavor, these geographical differences would seem irrelevant for the discussion. Still, I suggest that they are not and that the right way to consider modernism in mathematics would be, if at all, at the local level: modernist Paris mathematics, modernist Viennese mathematics, etc. Moreover, this approach would inherently emphasize the need to analyze not just the pronouncements of the Hilberts and the Weyls but also the pronouncements and the mathematical deeds of the Pringsheims the Perrons, and the Reidemeisters.

How useful is, then, the term 'modernism' for understanding the history of early twentieth-century mathematics? I hope to have shown that while the answer to this question may potentially be positive, there is a long way to go before this potentiality can be translated into reality. In particular, a plain characterization via *checklists of putatively defining features of what modernism is* (which is a much-debated question anyway) will not suffice. What may be of use for gaining new insights into the history of mathematics from the perspective of this question is a deeper understanding of the *historical processes leading to modernism* in its various cultural manifestations.

Acknowledgments An earlier version of this paper was written about ten years ago and remained unpublished. Nevertheless, it was posted on my website, and it was read in its preliminary format and even cited in several places. I am glad and proud to be able to publish a fully revised and updated version in this volume dedicated to Jeremy Gray. I thank Jeremy for our interesting conversations on the topic of modernism and beyond, and, above all, for his decisive and lasting contribution to our discipline.

For enlightening discussions and comments on the topic of the present text, I would also like to thank Moritz Epple, Gal Hertz and Menachem Fisch. For important editorial comments that led to an improved final version of this text, I am thankful to Karine Chemla, Hongxing Zhang, and two anonymous reviewers.

References

Albright, Daniel. 1997. *Quantum Poetics: Yeats, Pound, Eliot, and the Science of Modernism*. Cambridge: Cambridge University Press.

———. 2000. *Untwisting the Serpent: Modernism in Music, Literature, and Other Arts*. Chicago: University of Chicago Press.

Ash, Mitchell G. 2018. Multiple Modernisms in Concert: The Sciences, Technology and Culture in Vienna around 1900. In *Being Modern: The Cultural Impact of Science in the Early Twentieth Century*, ed. Robert Bud, Paul Greenhalgh, Frank James, and Morag Shiach, 23–39. London: University College London.

Barany, Michael J. 2020. Histories of Mathematical Practice: Reconstruction, Genealogy, and the Unruly Pasts of Ruly Knowledge. *ZDM Mathematics Education* 52 (6): 1075–1086. https://doi.org/10.1007/s11858-020-01175-5.

Bloor, David. 1991. *Knowledge and Social Imagery*. University of Chicago Press.

Bottazzini, U., and Amy Dahan-Dalmédico, eds. 2001. *Changing Images in Mathematics: From the French Revolution to the New Millennium*, Studies in the History of Science, Technology, and Medicine. Vol. 13. London: Taylor & Francis.

Bradbury, Malcolm, and James Walter McFarlane. 1976. *Modernism: A Guide to European Literature, 1890–1930*. New York: Penguin Books.

Calinescu, Matei. 1987. *Five Faces of Modernity: Modernism, Avant-Garde, Decadence, Kitsch, Postmodernism*. 2nd ed. Durham: Duke University Press.

Carson, Cathryn, Alexei Kojevnikov, and Helmuth Trischler, eds. 2011. *Weimar Culture and Quantum Mechanics: Selected papers by Paul Forman and Contemporary Perspectives on the Forman Thesis*. London: Imperial College Press. https://www.twirpx.com/file/3325749/.

Cheetham, Mark A. 1991. *The Rhetoric of Purity: Essentialist Theory and the Advent of Abstract Painting*. Cambridge/New York: Cambridge University Press.

Childs, Peter. 2000. *Modernism*. London/New York: Routledge.

Corry, Leo. 1989. Linearity and Reflexivity in the Growth of Mathematical Knowledge. *Science in Context* 3 (02): 409–440. https://doi.org/10.1017/S0269889700000880.

———. 2004. *David Hilbert and the Axiomatization of Physics (1898–1918)*, Archimedes. Vol. 10. Dordrecht: Springer. http://link.springer.com/10.1007/978-1-4020-2778-9.

———. 2007a. Calculating the Limits of Poetic License: Fictional Narrative and the History of Mathematics. *Configurations* 15 (3): 195–226. https://doi.org/10.1353/con.0.0034.

———. 2007b. From Algebra (1895) to Moderne Algebra (1930): Changing Conceptions of a Discipline. A Guided Tour Using the Jahrbuch Über Die Fortschritte Der Mathematik. In *Episodes in the History of Modern Algebra (1800–1950)*, ed. Jeremy J. Gray and Karen H. Parshall, 221–244. Providence: American Mathematical Society.

———. 2007c. Number Crunching vs. Number Theory: Computers and FLT, from Kummer to SWAC (1850–1960), and Beyond. *Archive for History of Exact Sciences* 62 (4): 393–455. https://doi.org/10.1007/s00407-007-0018-2.

———. 2017. Steht Es Alles Wirklich Schon Bei Dedekind? Ideals and Factorization between Dedekind and Noether. In *In Memoriam Richard Dedekind (1831–1916)*, ed. Katrin Scheel, Thomas Sonar, and Peter Ullrich, 134–159. Münster: Verlag für wissenschaftliche Texte und Medien.

Engelhardt, Nina. 2018. *Modernism, Fiction and Mathematics*. Edinburgh: Edinburgh University Press.

———. 2021. Mathematics and Modernism. In *The Palgrave Handbook of Literature and Mathematics*, ed. Robert Tubbs, Alice Jenkins, and Nina Engelhardt, 281–298. Cham: Springer. https://doi.org/10.1007/978-3-030-55478-1_16.

Engelhardt, Nina, and Robert Tubbs. 2021. Introduction: Relationships and Connections Between Literature and Mathematics. In *The Palgrave Handbook of Literature and Mathematics*, ed. Robert Tubbs, Alice Jenkins, and Nina Engelhardt, 1–20. Cham: Springer. https://doi.org/10.1007/978-3-030-55478-1_1.

Epple, Moritz. 1996. Die Mathematische Moderne Und Die Herrschaft Der Zeichen: Über Herbert Mehrtens, Moderne –Sprache – Mathematik. *NTM Zeitschrift für Geschichte der Wissenschaften, Technik und Medizin* 4: 173–180.

———. 1997. Styles of Argumentation in Late 19th-Century Geometry and the Structure of Mathematical Modernity. In *Analysis and Synthesis in Mathematics: History and Philosophy*, ed. Marco Panza and Michael Otte, 177–198. Dordrecht: Kluwer.

———. 1999. *Die Entstehung Der Knotentheorie: Kontexte Und Konstruktionen Einer Modernen Mathematischen Theorie*. Wiesbaden: Vieweg +Teubner.

———. 2004. Knot Invariants in Vienna and Princeton During the 1920s: Epistemic Configurations of Mathematical Research. *Science in Context* 17 (1–2): 131–164. https://doi.org/10.1017/S0269889704000079.

Everdell, William R. 1997. *The First Moderns: Profiles in the Origins of Twentieth-Century Thought*. Chicago: University of Chicago Press.

Eysteinsson, Astradur. 1990. *The Concept of Modernism*. Ithaca: Cornell University Press.

———. 2021. Modernism—Borders and Crises. *Humanities* 10 (2): 76. https://doi.org/10.3390/h10020076.

Eysteinsson, Astradur, and Vivian Liska, eds. 2007. *Modernism. Chlel.Xxi*. 2 vols. Amsterdam: John Benjamins. https://benjamins.com/catalog/chlel.xxi.

Feferman, Solomon. 2009. Modernism in Mathematics, Review of Plato's Ghost by Jeremy Gray. *American Scientist* 97 (5): 417.

Fisch, Menachem, and Yitzhak Benbaji. 2011. *The View from Within: Normativity and the Limits of Self-Criticism*. Notre Dame: University of Notre Dame Press. https://muse.jhu.edu/book/13384.

Forman, Paul. 1971. Weimar Culture, Causality, and Quantum Theory, 1918–1927: Adaptation by German Physicists and Mathematicians to a Hostile Intellectual Environment. *Historical Studies in the Physical Sciences* 3 (January): 1–115. https://doi.org/10.2307/27757315.

———. 2007. The Primacy of Science in Modernity, of Technology in Postmodernity, and of Ideology in the History of Technology. *History and Technology* 23 (1–2): 1–152. https://doi.org/10.1080/07341510601092191.

Frank, E. 1982. In Memoriam Oskar Perron. *Journal of Number Theory* 14: 281–291.

Galison, Peter. 1990. Aufbau/Bauhaus: Logical Positivism and Architectural Modernism. *Critical Inquiry* 16 (4): 709–752.

———. 2003. *Einstein's Clocks, Poincaré's Maps: Empires of Time*. 1st ed. New York: W.W. Norton.

Gamwell, Lynn. 2015. *Mathematics and Art: A Cultural History*. Princeton University Press.

Gay, Peter. 2007. *Modernism: The Lure of Heresy*. New York: Norton.

Giedion, Sigfried. 1941. *Space, Time and Architecture*. Cambridge, MA: Harvard University Press.

Gillies, Donald, ed. 1992. *Revolutions in Mathematics*, Oxford Science Publications. Oxford/New York: Oxford University Press.

Goldstein, Catherine. 1999. Sur La Question Des Méthodes Quantitatives En Histoire Des Mathématiques: Le Cas de La Théorie Des Nombres En France (1870–1914). *Acta Historiae Rerum Naturalium Nec Non Technicarum*, New Series, 3 (28): 187–214.

Gray, Jeremy J. 2004. Anxiety and Abstraction in Nineteenth-Century Mathematics. *Science in Context* 17 (1–2): 23–47. https://doi.org/10.1017/S0269889704000043.

———. 2006. Modern Mathematics as a Cultural Phenomenon. In *The Architecture of Modern Mathematics*, ed. Jeremy J. Gray and José Ferreirós, 371–396. New York: Oxford University Press.

———. 2008. *Plato's Ghost: The Modernist Transformation of Mathematics*. Princeton: Princeton University Press.

Greenberg, Clement. 1971. Necessity of 'Formalism'. *New Literary History* 3 (1): 171–175. https://doi.org/10.2307/468386.

———. 1980. Modern and Postmodern. *Art* 54 (6) http://www.sharecom.ca/greenberg/postmodernism.html.

———. 1983. *Taste*. January 18. http://www.sharecom.ca/greenberg/taste.html.

———. 1995. Modernist Painting. In *The Collected Essays and Criticism, Volume 4: Modernism with a Vengeance, 1957–1969 (The Collected Essays and Criticism, Vol 4)*, ed. John O'Brien, 85–93. Chicago: University of Chicago Press.

Hallett, M. 1979a. Towards a Theory of Mathematical Research Programmes (i). *The British Journal for the Philosophy of Science* 30 (1): 1–25.

———. 1979b. Towards a Theory of Mathematical Research Programmes (II). *The British Journal for the Philosophy of Science* 30 (2): 135–159.

Henderson, Linda Dalrymple. 1983. *The Fourth Dimension and Non-Euclidean Geometry in Modern Art*. Princeton: Princeton University Press.

———. 2004. Editor's Introduction: I. Writing Modern Art and Science– An Overview; II. Cubism, Futurism, and Ether Physics in the Early Twentieth Century. *Science in Context* 17 (4): 423–466.

———. 2005. *Duchamp in Context: Science and Technology in the "Large Glass" and Related Works*. Princeton: Princeton University Press.

———. 2007. Modernism and Science. In *Modernism*, ed. Vivian Liska and Astradur Eysteinsson, 383–401.

Janik, Allan. 2001. *Wittgenstein's Vienna Revisited*. New Brunswick: Transaction Publishers.

Janik, Allan, and Stephen Edelston Toulmin. 1973. *Wittgenstein's Vienna*. New York: Simon and Schuster.

Jones, Caroline A. 2000. The Modernist Paradigm: The Artworld and Thomas Kuhn. *Critical Inquiry* 26 (3): 488–528.

———. 2005. *Eyesight Alone: Clement Greenberg's Modernism and the Bureaucratization of the Senses*. Chicago: University of Chicago Press.

Levenson, Thomas. 2003. *Einstein in Berlin*. New York: Bantam Books.

MacKenzie, Donald. 1993. Negotiating Arithmetic, Constructing Proof: The Sociology of Mathematics and Information Technology. *Social Studies of Science* 23 (1): 37–65.

Mehrtens, Herbert. 1990. *Moderne Sprache, Mathematik: Eine Geschichte Des Streits Um Die Grundlagen Der Disziplin Und Des Subjekts Formaler Systeme*, 1. Aufl. Frankfurt: Suhrkamp.

———. 1996. Modernism vs. Counter-Modernism, Nationalism vs. Internationalism: Style and Politics in Mathematics 1900–1950. In *L'Europe Mathématique/Mathematical Europe. Histoires, Mythes, Identités/History, Myth, Identity*, ed. Catherine Goldstein, Jim Ritter, and Jeremy J. Gray, 518–529. Paris: Editions de la Maison des Sciences de l'Homme.

Miller, Arthur I. 2000. *Insights of Genius: Imagery and Creativity in Science and Art*. Cambridge, MA: MIT Press.

———. 2002. *Einstein, Picasso: Space, Time and The Beauty That Causes Havoc*. New York: Basic Books.

Monk, Ray. 1991. *Ludwig Wittgenstein: The Duty of Genius*. Later Printing. edition. London: Vintage.

Parshall, Karen Hunger. 2004. Defining a Mathematical Research School: The Case of Algebra at the University of Chicago, 1892–1945. *Historia Mathematica* 31 (3): 263–278. https://doi.org/10.1016/S0315-0860(03)00048-X.

Perron, Oskar. 1911. Über Wahrheit und Irrtum in der Mathematik. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 20: 196–211.

———. 1952. Alfred Pringsheim. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 56 (1): 1–6.

Pollack-Milgate, Howard. 2021. Mathematics in German Literature: Paradoxes of infinity. In *The Palgrave Handbook of Literature and Mathematics*, ed. Robert Tubbs, Alice Jenkins, and Nina Engelhardt, 299–318. Cham: Springer. https://doi.org/10.1007/978-3-030-55478-1_17.

Pringsheim, Alfred. 1904. Über Wert und angeblichen Unwert der Mathematik. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 13: 357–382.

Remmert, Volker R., Martina R. Schneider, and Henrik Kragh Sørensen, eds. 2016. *Historiography of Mathematics in the 19th and 20th Centuries*. New York: Birkhäuser.

Rheinberger, Hans-Jörg. 1997. *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Stanford: Stanford University Press.

Ross, Dorothy, ed. 1994. *Modernist Impulses in the Human Sciences, 1870–1930*. Baltimore: Johns Hopkins University Press.

Rowe, David E. 2004. Making Mathematics in an Oral Culture: Göttingen in the Era of Klein and Hilbert. *Science in Context* 17: 85–129. https://doi.org/10.1017/S0269889704000067.

———. 2013. Review of Jeremy Gray's Plato's Ghost: The Modernist Transformation of Mathematics (2008). *Bulletin of the American Mathematical Society* 50 (3): 513–521.

Schappacher, Norbert. 2012. Panorama Eines Umbruchs. *NTM Zeitschrift Für Geschichte Der Wissenschaften, Technik Und Medizin* 20 (September). https://doi.org/10.1007/s00048-012-0072-y

Scholz, Erhard. 2010. The Tale of Modernist Mathematics. *Metascience* 19: 213–216.

Schorske, Carl E. 1980. *Fin-De-Siecle Vienna: Politics and Culture*. New York: Vintage.

Sigmund, Karl. 1999. Musil, Perutz, Broch – Mathematik Und Die Wiener Literaten. *Mitteilungen DMV*, 1999.

Staley, R. 2005. On the Co-Creation of Classical and Modern Physics. *Isis* 96 (4): 530–558.

Stichweh, Rudolf. 1984. *Zur Entstehung Des Modernen Systems Wissenschaftlicher Disziplinen: Physik in Deutschland, 1740–1890*. Frankfurt: Suhrkamp.

van Dalen, Dirk. 1999. *Mystic, Geometer, and Intuitionist: The Life of L. E. J. Brouwer Volume 1: The Dawning Revolution*. New York: Oxford University Press.

Vargish, Thomas, and Delo E. Mook. 1999. *Inside Modernism: Relativity Theory, Cubism, Narrative*. New Haven: Yale University Press.

von Mises, Richard. 1922. Über Die Gegenwärtige Krise Der Mechanik. *Die Naturwissenschaften* 10: 25–29. https://doi.org/10.1007/BF01590406.

Vrahimis, Andreas. 2012. Modernism and the Vienna Circle's critique of Heidegger. *Critical Quarterly* 54 (3): 61–83.

Wagner-Döbler, Roland, and Jan Berg. 1993. *Mathematische Logik von 1847 Bis Zur Gegenwart: Eine Bibliometrische Untersuchung*. Berlin: Walter de Gruyter.

Weisstein, Ulrich. 1995. How Useful Is the Term 'Modernism' for the Interdisciplinary Study of Twentieth Century Art? In *The Turn of the Century. Modernism and Modernity in Literature and the Arts*, ed. Christian Berg, Frank Durieux, and Geert Lernout, 409–441. Berlin/New York: De Gruyter.

Yourgrau, Palle. 2005. *A World Without Time: The Forgotten Legacy of Godel and Einstein*. New York: Basic Books.

Chapter 17
What Is the Right Way to Be Modern? Examples from Integration Theory in the Twentieth Century

Tom Archibald

Abstract In this paper we attempt to identify some features of modernist agendas in twentieth-century integration theory. This entails looking at ways in which abstraction is used as a strategy, both in defining mathematical objects and in axiomatizing the theories involved. Choosing examples from 1910 to 1950, we consider also explicit or implicit remarks about such matters as a virtue and function of these modern tools. Work of H. Lebesgue, C. Carathéodory, F. Riesz, O. Nikodym, and N. Bourbaki are considered.

17.1 The Historiography of Modernity

Jeremy Gray's book *Plato's Ghost* presents the rich results of a long reflection on what is modern about modern mathematics, as well as an account of the history of that subject. "Modern" is a term that eludes strict definition, as so many different writers use it with no fixed meaning. Indeed, it is in part a label that emerged to describe a period, the features of which vary from time to time and place to place, and which are taken up differentially by different writers as characteristic. In this way it resembles the Enlightenment and the Romantic period. But while the label is elusive, it doesn't mean it is useless. Nonetheless, what we mean by it needs to be thought about and specified to some degree. The associated term "modernism" likewise needs to be explicated somewhat, both in a general cultural context, to some degree, and with in mathematics.

Work on this paper has been supported by the SSHRC (Canada), which support is gratefully acknowledged. I also thank 3 anonymous referees for their careful reading, which was sometimes more careful than the writing, and for thoughtful and helpful comments.

T. Archibald (✉)
Department of Mathematics, Simon Fraser University, Burnaby, BC, Canada
e-mail: tarchi@sfu.ca

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_17

One way to handle this is to look at a well-established locus or body of work that studies modernism. Consider Yale's Modernism Lab (https://campuspress.yale.edu/modernismlab/). Yale University, long a leading centre for studies of literature, can't provide an authoritative view, but it provides the scholar of mathematics with examples of how the identification of modernist features can be achieved. Here we note that the approach begins with a periodization, with early (literary) modernism being put between 1914 and 1926, though in fact the site proposes no strict definitions and reaches from 1890 (with work of the dramatist Ibsen seen as a precursor and major influence on uncontestedly modern figures) and stopping in 1940, around the deaths of the writers James Joyce and Virginia Woolf. A casual glance through these pages reveals some features of the literary work studied that are broadly shared. These include a concern with sex and sexuality, with Freud, Proust, and Woolf as examples. This is probably not so useful for looking at modern mathematics. On the other hand, there is a group of ideas about going beyond the conventions of earlier periods, for example by diminishing the role of plot, shedding ideas about how to expose a character in a realistic way, being strongly innovative in the use of language, and making the reader work hard to understand. These can be summed up in the credo of the poet and critic Ezra Pound: "Make it new." While Pound's view of modernism is certainly not entirely shared among many writers and artists who are customarily grouped under that label, the abandonment of conventional representation and of naturalism, and even the creation of works that do not claim to represent anything outside the artist's interior creative nexus, are largely common both to cultural modernism and to "the modern". There are limits here: buildings can be modern, but they need to function as buildings. Paintings and poems also need to function in some ways, so that they can be hung on the wall or spoken or printed, so not everything old is abandoned. The idea of taking good features and problems from older work and remaking it with a "modern" eye is one way to achieve this, as Picasso did with Velasquez, or Pound with Sextus Propertius.

A key feature of this kind of programmatic innovation is to understand various structural aspects of work in the preceding period and innovate by playing with them. That is, in literature, plot and character, but also the forms of narrative and description, the very ways of writing things down (for example in chapters and paragraphs), the conventions of vocabulary and dialogue, the dramatic or narrative arc, and the position of the author with respect to the characters and the reader. In Joyce's *Ulysses*, for example, we find an entire chapter written in the form of tabloid newspaper articles, complete with headlines. We also have the detailed depiction of the thought processes of characters as though these are accessible to the author, true also of Woolf. How these changes to tradition work, and exactly what they consist of, varies with the author. This is the subject of much of the critical writing about the period, raising questions for example about how the intent of the author is related to the reader's experience. This contrasts with modern mathematics, which—however it may resemble these literary approaches—still seeks the premodern aim of more or less univocal understanding of the meaning of mathematical objects and results. On the other hand, mathematical modernism shares many features with the work of moderns in other areas of cultural production. By transforming the treatment of older areas using new tools, mathematics likewise engages in an esthetically-motivated renewal of its procedures. One feature of this is abstraction, a frequent if not necessary accompaniment to much modern mathematics.

Another arena for the investigation of modernism lies in its institutions. To focus on painting, we choose two different types of institution to investigate what is modernism: the Museum of Modern Art (MoMA) in New York, and H. Arnason's *History of Modern Art*, widely and long used as the basic survey in courses introducing modern art to US undergraduates. MoMA was founded in 1929, itself thus a demarcation point for a kind of ascendancy of the modern in the visual arts in New York. The location, in the heart of cultural Manhattan, provided access to a collection of modern art that was soon to become central in the West, not least due to WW2. The first director of this museum was Alfred H. Barr, a Picasso specialist and enthusiast, and the work of Picasso and Matisse are central, augmented by the more strictly abstract work such as that of the Russian W. Kandinsky. The museum not only serves to define what should be included as modern; it also comments on what is major, by virtue of its collections policy, which feeds back with the community of collectors and the value of their work. This is strikingly different from what happens in either literature or mathematics. Indeed, Arnason's book largely encodes and extends the MoMA canon of artists, globalizing it to artistic collections around the world.

In both MoMA's collections and Arnason's selections we find a picture of modernism that is broadly consistent with what happens in literature. The emphasis is still on "making it new", again by playing with identifiable features of older work. Here the distinction is made between "representational" work, that is, work in which there are recognizable objects or even a scene, and "non-representational" in which nothing in particular can be identified. Picasso and Matisse usually have figures: figures and scenes of interiors, where shapes and colours are abstracted from actual likeness but heads and legs, tables and flowers are easy to see. More "purely" abstract work may not represent anything at all: just blotches with colour, or even black on white or white on white. Again, there is a heritage here pushing back into the nineteenth century, in this case going back somewhat further than in the literary case; and the end of the period is vague, but doubtless should at least be placed to include the Abstract Expressionist work of the 1950s. The label "Abstract Expressionist" reminds us that in this period, let us say going from about 1900 to at least 1960, there is a tendency to label groups of artists and works according to artistic approach: cubism, expressionism, surrealism, futurism, and so on were often accompanied by manifestos. The identifying aspects of these are sometimes expressed by protagonists or critics, and sometimes are merely inferred by shared features. And again, the function of these innovations may mean one thing for the artists, another thing for critics and audience. Again, this contrasts with mathematics.

In mathematics, "making it new" remains a feature of modernism. However, there are both similarities and differences between the use of the term in the broader cultural world and that in mathematics. In the mathematical context, there is certainly an emphasis on the analysis of underlying structures, the removal of certain features and the retention of others, to create new mathematical objects that abstract from the older ones. And indeed there are also broad similarities in what the mathematician is seeking to achieve and what the artist is trying for. The simple idea is that by the action of abstraction, one gives a new look at old structures that both addresses old problems and creates new ones. In mathematics, however, it is often the case that the new structure supplies a solution to an old problem that is completely recognizable as a solution in the older context. It also can be the case that it is demonstrated that no solution is possible: this is not a solution that would seem to be possible to articulate in the artistic context. And of course, the entire meaning of the phrase "solution of a problem" is completely different in these contexts. Solving the problem of depicting the interior mental processes of a woman conflicted by her adultery, or of expressing the visual essence of a bowl of fruit, is not so easily comparable to showing that the characteristic function of the rationals has a Lebesgue integral.

Modernism in mathematics has many features. In what follows, we will consider mostly the role of abstraction in making certain aspects of mathematics modern. The extent to which this can be regarded as programmatically intended to encapsulate something called "modernism" I will not address. Instead, I'll think of abstraction as a not-unusual feature of a modern approach that is often undertaken self-consciously and has ultimately the aim of formulating and solving mathematical problems. It is clearly not the only feature of modern mathematics worth considering: foundationalism, often pursued through rather abstract aims, is another related feature, for example.

This leads us to consider what problem or problems abstraction is intended to solve, and then look at how it does so. In the case of mathematics, there are several purposes that can be identified with hindsight. Abstraction seeks to elucidate classical or traditional problems—think of Galois theory, or ideal numbers. It seeks to clarify by the clear identification of assumptions, and the investigation of what happens without them: non-Euclidean geometry is an example here. It also has the salutary feature of renewing the field of problems, or changing from one set of problems to another, one which concerns the abstract objects themselves. But if we talk of its function in the transition to modern mathematics, one key aspect is precisely to give the mathematician alternative strategies to those of brute force computation.

One master of this older form was C. G. J. Jacobi:

> Mathematics clearly has the property that one can come to a discovery by calculation, quite in opposition to Goethe's verse:
>
> > Das ist eine von den alten Sünden,
> > Sie meinen: Rechnen, das sei Erfinden.
>
> For when we strike out initially on an incorrect path, calculation at once amplifies the error. Since we calculate with literal expressions, that is, with expressions which carry the nature of their origins within them, the result always gives the shortest path that we should take. However this method of discovery by calculation now is definitely no longer applicable in connection with the Abelian transcendents, since if we go off the correct path by the smallest amount, we get no result at all due to the huge complications of the calculation. Thus it

seems that the conduct of this research would lift mathematics up to a higher viewpoint. *C. G. J. Jacobi, Winter 1840.*[1]

The remarks above by Jacobi clearly indicate that, for him at around 1840, the limitations of the calculational method as a source of discovery were becoming apparent. His own work on elliptic functions, and their expression in terms of theta functions, had been extremely formula-heavy, and the *Fundamenta Nova* is essentially a collection of derivations of formulas. His early discoveries about modular functions provide a nice example of the kind of thing that is very difficult to extend, since the calculations simply get out of hand. His comment appears to mean that such research is in his view unsustainable, and that he sees Abelian functions as an area where a "higher level" of mathematics needs to be applied. Reading with hindsight, the remarks seem prophetic, since his student Riemann was to revolutionize this field and many others by inventing and introducing topological notions into complex analysis. The notion of a Riemann surface had a complicated reception and development, but it would seem reasonable to describe this as a move to a more "modern", conceptual viewpoint in which new structures or devices are introduced into mathematics in order to make progress. This shift has been described by several writers, notably Kragh Sørensen and Laugwitz, in terms of a move from a formula-based to a concept-based mathematics.

However, there is more than one kind of formula, and formulas have a variety of purposes. Different formulas for the same thing may, for example, give different representations of the same object. Closed form versus series expressions for functions give an obvious example here. Formulas may be used to classify objects, as in Legendre's classification of elliptic integrals, through the production of canonical forms. Formulas also sometimes serve the function of representing generic objects—here one may think of Gauss' hypergeometric series, for example, used to write large classes of functions, or Weierstrass's later product representation for entire functions.

Jacobi's comment above raises the prospect of a kind of "Jacobi limit" for the possibility of working with mathematical formulas directly in order to draw general conclusions. Jacobi's concern, expressed in the context of elliptic functions, was slow to be realized, if indeed it ever has fully been realized. We need only think of the multi-page expansions of late nineteenth-century perturbation theory, for example, or the arcane estimation procedures involved in Bernstein's work on Hilbert's twentieth problem, to see that some mathematicians in some contexts have continued to find concrete formulas extremely valuable and have retained them both as tools and as objects of study.

In some domains, however, we see specific cases where more abstract tools were brought to bear in ways that sought exactly to get around a variety of problems. The effectiveness of this approach came to be coupled with axiomatics, notably by Hilbert and his school, to give a mathematics that can be described as modern. This

[1] Transcription by C. Borchhardt quoted by Königsberger, p. 261.

brings us to our case, in which we look at some aspects of the question of how to make an abstract integration theory.

17.2 Integration and Abstraction

That the twentieth century is a period in which mathematics became increasingly abstract, and in which the generality of mathematical results came to be highly prized, goes without saying. Historical studies (Corry 2004, for example) have discussed the ways in which various abstract structures have taken centre stage in mathematical research, and the related value of generality in mathematics is examined in several chapters of the recent collective volume by Chemla et al. (2016).

But how and why is value assigned or attributed to abstract entities? When does an abstraction have the kind of explanatory power that inspires mathematicians to adopt it or emulate it? How do users of various abstract methods and entities justify those uses? We explore this question in the context of the theory of integration in the twentieth century. The centrality of the integral to mathematics since its invention was cemented in the nineteenth century, and the attempts to render it theoretically solid have been led by Cauchy and Weierstrass, in both the real and the complex domains. Gray has written extensively about this classical period, as has Umberto Bottazzini.

Many theories of the integral, with differing foundations and roles, were developed in the first half of the twentieth century, and this paper is not an attempt to survey them. What we will instead examine is the way in which, in a particular subset of cases, arguments were presented in favour of one approach or another. Such arguments turn, on the one hand, on the convenience, suitability, or intuitive character of one approach over another, for example with respect to its applications; and on the other hand, on the question of key properties or results being retained and extended in the newly-constructed, more abstract setting. The precise range of application is naturally implicated, and the question of generality enters here as a matter of course. Actual direct discussions about these meta-questions, if that is the correct term, occur only infrequently, and usually not in the papers where the innovations and results are presented. However, both specific and implicit statements show that these questions are not uncontested, and the accompanying dynamic provides a thread for understanding important features of the mathematics of the first half of the twentieth century.

In what follows, we proceed as follows. First, we look at Lebesgue's introduction of measure and his description of two approaches, where the axiomatic one was unrealized. In the immediate reception of Lebesgue's work Vitali's formulation of the notion of absolute continuity later led Lebesgue to his decomposition theorem and the discussion of singularities, which was to be a fruitful field for testing the virtues of various approaches to integration. Meanwhile, work of Carathéodory provided an incipient axiomatic approach, one that was to remain durable. At about the same time (during WWI), a more abstract version of measure and integration on more general sets was provided, among others by Fréchet and Daniell. We then proceed with a discussion of the virtues of abstraction as they were seen in the 1920s and 1930, focussing on the examples of F. Riesz and O. Nikodym, important both for Riesz's work in integration theory and for Nikodym's work on Lebesgue decomposition. The elaboration of Riesz's work eventually encompassed a functional approach involving a weak version of lattice theory. This sets the stage for a brief discussion of Bourbaki's version of the functional approach, notably due to Dieudonné. We conclude with a discussion of differing views about measure theory versus the functional approach in the 1950s. The whole would be capable of a much more thorough treatment, and we have left out many important issues, not least the question of existence of measures.

17.2.1 Lebesgue, Measure, and Axioms

Henri Lebesgue introduced his generalization of the Riemann integral in 1901, incorporating this and the associated theory of measurable subsets of $\mathbb{R}$ and real-valued functions of a real variable (Hawkins 1970). The work became his thesis and was published in 1902 in the *Annali*. This built on earlier ideas, notably of Borel, including the notion of outer content. To develop the notion and properties of the integral, Lebesgue defined what we now call Lebesgue measure. The resulting integral possessed the following properties, derived from this idea of measure.

(1) For any [real numbers] a, b, h, we have $\int_a^b f(x)dx = \int_{a+h}^{b+h} f(x-h)dx$
(2) For any a, b, c, we have $\int_a^b f(x)dx + \int_b^c f(x)dx + \int_c^a f(x)dx = 0$
(3) $\int_a^b [f(x) + \phi(x)]dx = \int_a^b f(x)dx + \int_a^b \phi(x)dx$
(4) If $f \geq 0$ et $b > a$, then $\int_a^b f(x)dx \geq 0$.
(5) $\int_0^1 1 \times dx = 1$.
(6) If $f_n(x)$ increases to $f(x)$, the integral of $f_n(x)$ has that of $f(x)$ as a limit.

In fact Lebesgue, having given these six conditions, makes a methodological comment.

> By stating the six conditions of the problem of integration we define the integral. This definition belongs to the class of definitions that we may call *descriptive*. In such definitions we state the characteristic properties of the entity we wish to define. in *constructive* definitions we state what operations we need to carry out in order to obtain the entity we want to define. Constructive definitions are the ones that are mostly used in analysis. However, we sometimes use descriptive definitions ... (Lebesgue 1903, 106–107)

A footnote here is particularly illuminating. "The use of descriptive definitions is indispensable for the first terms [les premiers termes] of a science when we wish to construct this science on a purely logical and abstract basis," he says referring both to the 1898 thesis of Jules Drach and to Hilbert's memoir on the foundations of geometry that appeared in the Annales of the ENS in 1900. He continues,

> The definition is thus called axiomatic because it lists the necessary axioms. It is therefore self-contained and forms a complete whole. (Lebesgue 1903, 107).

The simultaneous citation of the notoriously faulty thesis of Drach and the fundamental work of Hilbert is somewhat surprising. It is Hilbert who seems to be the main influence, as the later discussion of consistency proofs underlines.

Lebesgue's work quickly reached a certain public, partly through his published lectures given at the Collège de France (Lebesgue 1903) and partly through a subsequent work on infinite series that used his integral. As one would expect, besides simply defining the integral Lebesgue dealt with many other issues, including indefinite integration.

17.2.2 Absolute Continuity and Decomposition

Lebesgue's 1910 result that one can write any function of bounded variation as the sum of an absolutely continuous function and a "function of singularities" was made possible by intervening work due largely to his Italian reader, Guiseppe Vitali, who introduced the idea of absolute continuity. This notion was identified by several authors independently. The first of these was Vitali, in 1905 (Borgato 2012). Vitali was closely associated with Dini, Arzelà, and Fubini among others, and at the time of this research was a high school teacher. Vitali had worked for several years on an approach similar to Lebesgue's for defining integrals of measurable functions (in his own sense), apparently learning of Lebesgue's work in 1904 from Pincherle. Vitali is thus one of the earliest to have seen the interest of Lebesgue's work, along with F. Riesz.

Lebesgue had shown that an indefinite integral of a Lebesgue-integrable function (he uses the term *sommable*) may be differentiated to obtain this function, except on a set of measure zero. Vitali sharpened this by introducing the notion of absolute continuity (Vitali 1904–1905, 1021), defined below. Vitali showed that this property was necessary and sufficient for a function to be the indefinite integral of a summable (i.e. Lebesgue integrable) function, explicitly citing Lebesgue (1903). Vitali's discussion includes the relationship between absolute continuity of a function and boundedness of its variation, noting in detail that absolutely continuous functions are of bounded variation (evidently) but that the converse need not be so. In the 1905 paper he describes such a function in a generic way using the Cantor set; the standard "devil's staircase" example that makes this concrete was to come later.

This definition of absolute continuity was easily generalized and remained useful. The notion of absolute continuity, and the fact that absolutely continuous functions are antiderivatives and vice-versa, was key to Lebesgue's 1910 decomposition theorem.

17.2.2.1 Lebesgue's Decomposition Theorem, 1910/1927

To get an idea of the purpose and effects of the several layers of generalization that took place, we look at a succession of versions of a result due to Lebesgue, usually known as the Lebesgue-Nikodym theorem.

In Chapter VIII of his *Leçons*, Lebesgue took on the subject of indefinite integrals. The chapter starts with some basic results: the indefinite integrals of summable functions are continuous, and of bounded variation. We can think of the indefinite integral of a function f in a variety of ways. Defining it to begin with as

$$F(x) = \int_{\alpha}^{x} f(x)dx + C$$

where α is fixed and C is a constant, Lebesgue points out that in fact we can also think of

$$\Phi(\alpha, \beta) = \int_{\alpha}^{\beta} f(x)dx$$

as indefinite when the bounds are not fixed, but rather define some interval of integration. In the same vein, we can extend this concept to measurable sets E and consider the indefinite integral as a set function

$$\Phi(E) = \int_{E} f(x)dx.$$

Throughout Lebesgue thinks of f itself as continuous on some interval.

Such functions $\Phi(E)$ have all their important properties, and their relation to the more usual indefinite integral, determined by two features: countable additivity ("complete" additivity in Lebesgue's usage) and absolute continuity. This is worked out by Lebesgue in 1910, but stated more clearly in the 1927 version, where set functions are employed. By 1927, Φ is a linear functional on E. The result in question then states that the integral of a summable function $f(x)$ over a set E tends to 0 as $m(E)$ approaches zero. This is a nice way to grasp the idea of absolute continuity, in fact.

Definition A set function $\Phi(E)$ is absolutely continuous if, for all $\varepsilon > 0$, there exists η such that $m(E) \leq \eta \implies |\Phi(E)| \leq \varepsilon$.

(Here m is the measure.) Lebesgue shows that this is true for indefinite integrals, seen as set functions.

If the set function is not absolutely continuous, Lebesgue shows that one can write any function of bounded variation as the sum of an absolutely continuous function and a "function of singularities", which he defines in detail.

For clarity, not so evident in Lebesgue's version, I cite this in later language, due to Dieudonné:

Theorem *If E is the set of all continuous real-valued functions $x(t)$ on* $[0, 1]$*, every linear functional on E may be written uniquely as*

$$L(x) = \int_0^1 y(t)x(t)dt + S(x)$$

where y is measurable, positive, and well-defined on $[0, 1]$ *except on a set of measure 0, and $S(x)$ is a positive singular linear functional.*

This result of Lebesgue sets the stage for an elaboration of the functional approach, at first due to F. Riesz. It is also the launching ground for a more abstract version of the decomposition theorem due to Fréchet and published first in 1916. But we are getting slightly ahead of ourselves. Let us return to the teens of the twentieth century and consider another aspect of the reception of Lebesgue.

17.2.3 Carathéodory and Formal Measure Theory

The response to Lebesgue's work was initially rather mixed, but soon the theory attracted wide interest. In addition to Lebesgue's own exposition of 1903, accounts of the theory appeared in the widely-used textbook of De la Vallée Poussin from the the second edition of volume 2 in 1912 and the third edition of volume 1 (1914). It may be here that it was encountered by Constantine Carathéodory, whose first paper on the subject appeared in 1914 (Carathéodory 1914).

Carathéodory (1873–1950) had recently succeeded Klein in Göttingen, joining Hilbert and Landau as *Ordinarius*. He had also joined the editorial board of *Mathematische Annalen*, at the invitation of Klein (Georgiadou 2004, 90–95). As Georgiadou argues, in his career of about 10 years, he had been much influenced by Hilbert (in the calculus of variations especially) and by members of the Hilbert school (notably Erhard Schmidt and Ernst Zermelo). In this context his decision to enter the field of measure theory by creating an abstract, that is to say axiomatic, approach, is unsurprising.

Carathéodory begins his paper, then, with what he terms a "Formal Theory of Measure" (Carathéodory 1914, 405). Noting that outer measure μ^* is the key concept, he uses three of its properties as formal axioms:

I. To any set of points A in an q-dimensional space R_q is assigned a unique number $\mu(A)$ which can be zero, positive, or $+\infty$, called the outer measure of A.

II. If B is a subset of A then

$$\mu^* B \leq \mu^* A.$$

III. If A is the union of a sequence of finite or countably infinite point sets $A_1, A_2, \ldots,$ then

$$\mu^* A \leq \mu^* A_1 + \mu^* A_1 + \ldots.$$

(Carathéodory 1914, 405)

These axioms were immediately joined by the definition of measurable set:

A point set A is called measurable, if for any arbitrary point set W of finite outer measure the following relation holds:

$$\mu^* W = \mu^* AW + \mu^*(W - AW);$$

the measure μA will then be defined by the equation

$$\mu A = \mu^* A.$$

(Carathéodory 1914, 406)

(Note that he uses AW to mean $A \cap W$.) This then leads to a chain of theorems about measurability, for example of intersections and unions of measurable sets. Two further axioms are then adduced.

IV. If A_1 and A_2 are two point sets, whose distance $\delta \neq 0$ then

$$\mu^*(A_1 + A_2) = \mu^* A_1 + \mu^* A_2$$

V. The outer measure $\mu^* A$ is the lower bound of the measures μB of all measurable sets that contain A as a subset.

The fourth one allows us to prove, for example, that intervals are measurable. The fifth permits a definition of inner measure, and allows us to recover the basic notion that a point set with finite outer measure is measurable if and only if its outer and inner measures coincide.

The advantage of this axiomatic approach is that anything that satisfies these five axioms can be used to create a theory of measure that functions analogously to Lebesgue's. The second part of the paper consists of developing exactly such a theory of so-called linear measures.

Carathéodory lectured on this material in the summer semester of 1914. By December 1917 he had completed a book that originated in these lectures, *Vorlesungen über reelle Funktionen*, a scholarly and comprehensive treatise that was to serve as a basic reference in the field of real analysis for many years (Carathéodory 1918). This book had a second edition in 1927, little altered, and was reissued by the Chelsea Publishing Company in New York in 1948, its copyright having been vested in the U. S. Alien Property Custodian as of 1946. It remains in print, dedicated to Schmidt and Zermelo as the author's friends. The Chelsea edition enjoyed very wide circulation in the expanding postwar mathematical universe in the US, which likely accounts in part for the continued interest in this older work.

The approach of Carathédory rests on an axiomatic formulation of the properties of outer measure, specifying which sets with outer measure are measurable. Other axiomatic formulations that were to follow dispensed with this approach, instead axiomatizing the concepts of content and measure as such. At mid-century, at least some writers regarded the Carathédory approach as the "usual" one, Mayrhofer being one example (Mayrhofer 1952, III). In his case, he derived the properties of these usual outer measures. The nature of the algebraic assumptions on the collection of subsets of the space—set of sets, ring, field, Boolean ring, Boolean lattice—also is a locus of variability with implications for how the definitions get made and how the material is ordered.

17.2.4 An Aside: Hausdorff and Measurability

General problems associated with the existence of measures for certain kind of sets began to appear quite soon after Lebesgue's work began to be taken seriously. The questions tie very closely to set-theoretic issues (such as Zermelo's Axiom of Choice) and to point-set topology. Already in 1914, Felix Hausdorff, in his *Grundzüge der Mengenlehre* (469–472) posed the problem of measure "au sens large". Banach, in his 1923 paper "Sur le problème de la mesure" stated Hausdorff's problem this way:

> In his book "Grundzüge der Mengenlehre" (Leipzig 1914) Monsieur Hausdorff treats the following problem: Can we attach to each bounded set E of an m-dimensional space a number $m(E)$ satisfying the following conditions:
>
> (1) $m(E) \geq 0$,
> (2) $m(E_0) = 1$ for a set E_0 in the given space,
> (3) $m(E_1 + E_2) = m(E_1) + m(E_2)$, if $E_1 E_2 = 0$,
> (4) $m(E_1) = m(E_2)$ if the sets E_1 et E_2 may be superposed.
>
> He proves that this problem is impossible for spaces of three dimensions or more. (Banach 1923)

In fact Banach's quotation is a little misleading. Hausdorff sees the Lebesgue idea in continuity with the older work of Hankel, Cantor, Jordan, and Peano. These older theories considered "measures" that were additive over pairwise disjoint sets, while the work of Borel and Lebesgue extended this to countable additivity. This may seem small, but for Hausdorff "the passage from the finite to the countable in the newer content and integral theory may be designated as one of the greatest advances of mathematics (Hausdorff 1914, 400)." Retaining the Lebesgue description in terms of constructive versus descriptive definition, Hausdorff pointed out that Lebesgue had wished to provide an axiomatic treatment, but that he had not been successful.

17.3 Virtues of Abstraction: Riesz and Nikodym

If abstraction can have various forms, it can also be possessed of various mathematical virtues: clarity, elementariness, economy, simplicity, and so on. These terms are used almost always without much precision, the intent being that the reader will understand from the outcome that the suggested virtue is present. These terms function quite a lot as do the various critical terms in art and literature. They can suggest formal characteristics of the work, or emphasize the response of the reader. Sometimes these virtues are relative, and the advantages of one theory over another may be stated explicitly or left implicit. The examples that follow are not intended as exhaustive.

17.3.1 Riesz and the Functional Approach, 1912

We have already mentioned that Frigyes Riesz was an early user of Lebesgue, something on which Riesz prided himself later. According to a reminiscence of 1949, it was the 1906 book on trigonometric series that first drew his attention to the Lebesgue integral, but then he turned to the thesis and the book on integration. Indeed, this was the immediate background to Riesz's first major triumph, now known as the Riesz-Fischer theorem, the idea for which was sparked (again according to Riesz himself) by the reading of Fatou's thesis. Not so modestly, Riesz noted of his theorem, "it is possibly the first application of the Lebesgue theory." (Riesz 1949, 30).

Riesz was soon led to consider whether a more general approach to the Lebesgue theory would be possible. By 1912, he proposed an approach that he terms "élémentaire", in the sense that it does not require a detailed study of measurable sets first, making it easier to learn (Riesz 1912, 1–2). Apparently he had been considering this for several years at that point, but was spurred to it by remarks of Borel, who proposed to do something similar, though Riesz's approach was his own. The paper was programmatic rather than detailed, and, in the 3-page format required by the *Comptes rendus*, it was not possible to include proofs.

Riesz discussed the idea of "elementariness" in a justificatory passage:

> What we want is an elementary theory. It would be difficult to define what we mean by elementary, but one point of comparison is provided by the speed with which theories become familiar. Now, if the theory of M. Lebesgue is not yet familiar to all those who work on Analysis, it's because **it is preceded by an in-depth study of measurable sets**. Can we get rid of this general notion? (Riesz 1912).

The programme was one that is familiar today in its broad outlines. First, Riesz defined "simple functions" as those taking on constant values on subintervals of a given interval, except at a finite number of discontinuities. Restricting himself to bounded functions, the only notion used from measure theory was that of a set of measure 0 (one with an arbitrarily small countable or finite cover). Taking the integrals of simple functions as obvious in the Riemannian sense, general integrals are defined as the limits of integrals of simple functions f_n, where $f_n \to f$ for all x in the interval of integration, except at a set of measure zero. With this, he states, one can recover as integrable the same set of bounded functions as in the case of Lebesgue; and one can even go ahead and define (Lebesgue) measure of sets by considering the characteristic functions of the sets.

Hence, in the Riesz approach, one can define the integral without the niceties of measure theory, except for the rather simple notion of a set of measure zero. Notions about measurable sets and measurable functions can then be recovered. The key virtue cited here is one of simplicity, long thought of as a virtue of a theory, even if the meaning is imprecise.

Spurred by discussions with Mittag-Leffler in 1916, Riesz wrote out the details of this approach in 1917, and the result appeared in *Acta Mathematica* in 1920. In the time interval between the programmatic statement and the publication, several key mathematical developments occurred that were to spur Riesz to further activity concerning integration theory. We turn to a quick discussion of a few of these in order to provide continuity to our account.

While the Riesz-Fischer theorem made use of the Lebesgue theory, the 1909 representation theorem for linear operators required the use of a Stieltjes integral. The problem of retooling Stieltjes integration in a manner compatible with the Lebesgue approach was taken on by Johann Radon in Vienna, and his 1913 Habilitationsschrift presented the theory of what is sometimes called the Radon Integral, sometimes the Lebesgue-Stieltjes integral, for subsets of Euclidean space. This required a generalized notion of measure. The details are nicely discussed in Hawkins (1970). Fréchet, in 1916, reacted to the Radon paper by noting that the method employed by Radon could be extended beyond $\mathbb{R}^n$ to arbitrary abstract spaces. This does not seem to have had a rapid reception. Fréchet also presented a version of the decomposition theorem for such situations.

17.3.2 Otton Nikodym

The move away from a naturalistic view of mathematics is gradual, tentative on the part of some, and partial both in individual practice and across the communities of mathematicians. But as the view that mathematics is a human contrivance gained momentum, various ways of presenting mathematical results also developed in which "naturalness" took on a different meaning. Indeed, there is a broadening and diversification of what constitute "mathematical reasons" for presenting a set of results in a certain way, a way involving the choice of definitions, the precise setting, the basic concepts, the axioms. These are largely familiar to us, since we hear mathematicians use them all the time: this method is "efficient", that one is "natural", this one allows us to see clearly the "essential" relationships between different objects or results, and so on.

But these ways of speaking and presenting, even proving results obviously did develop over time. Different choices here reflect values and influences.

Radon had already developed a discussion of the situation under which a function is an antiderivative in his more general setting, building directly on the work of Lebesgue. Carrying this to a more abstract setting was the work of Otton Nikodym, leading to the well-known "Radon-Nikodym" theorem. In what follows I'll try to look at the presentation of Nikodym embedded in the literature to which it was responding, with particular attention to what we might term his "mathematical reasons" for doing things as he did. By examining these we will get some image of one approach to the way in which research mathematics is to be presented, in the view of one practitioner, but one whose approach was derived from influences that ranged across the practitioners of the nascent theories of measure, functional analysis, and topology.

Nikodym was Galician, born in 1887 in what was then Austria-Hungary. With high school qualifications in both mathematics and classics, he went to the university in Lviv; at the time the principal language of instruction was his native Polish. He completed his studies in mathematics there in 1911; Sierpiński had joined the faculty there in 1908, was his teacher, and remained a mentor. From 1911 to 1924 Nikodym taught in a high school in Kraków, and he was one of the founding members, in 1919, of the Polish Mathematical Society. In 1924, apparently at the urging of Sierpiński, he obtained a doctorate, after which he joined the faculty of the Jagellonian University in Kraków. From 1930 to 1945 he was in Warsaw, though the university was closed after the German invasion in 1939. After the war he moved to the US, where he had a long career.

Here we concentrate on (Nikodym 1930). By examining Nikodym's strategy in that paper, we can get some idea about what he wants from a "theoretical configuration." Nikodym's explicit aim is to generalize the work of Radon to arbitrary sets (not simply collections of subsets of euclidean space). As he knew he had been preceded by Fréchet, but here he uses a different approach. His reasons for doing so have to do with *generality, economy/parsimony, and efficiency*.

Indeed, Nikodym contrasts his approach with that of Fréchet, who had generalized Radon's integral using a family of sets that is assumed closed under countable union and set difference (so-called "additive families"), with an arbitrary real-valued set function as measure. Nikodym prefers a different foundation, namely the *corps d'ensembles* (a field of sets: closed under countable union and complement in a specified "universe" or *variété*) and a restriction to *non-negative measures*. Now every *corps* is an additive family, but not conversely. Nonetheless (his first stated reason for this choice) the more restricted collection of sets provides a notion of integral that is no less general.

One general remark (justified in the text of the paper) is that his point of view "provides more methods of proof". For example, the restriction to non-negative measures permits an extension of the absolute continuity of the Lebesgue integral to this more general context (that is, the retention of the Vitali property). As he puts it,"if my $E_n \to 0$, we have $\lim_{n\to\infty} \int_{E_n} f = 0$." This fact combined with countable additivity for the integral expresses a property that renders the theory of Lebesgue "so harmonious and important."

Further, absolute continuity can be shown by using the non-negative measure on the field to define a distance function which renders the field a complete metric space. This, then, furnishes the additional means for proof to which he refers: he has the whole arsenal of metric space theory, by then well-developed. Finally, in this context, which functions are measurable does not depend on the measure. (The integrability, or *sommabilité* does of course.)

Nikodym also draws to our attention that his definitions are different from those of Lebesgue, “in order to arrive more quickly at more interesting theorems.” An itemized list of the differences is not so interesting, but it is true that there is a real economy of presentation in this context compared to Lebesgue and Radon. Finally in the introduction he gives a theorem that is “a little more interesting:”

> The necessary and sufficient condition that a function $\mathscr{F}(E)$ be perfectly additive and μ-continuous is that there exists a μ-summable function $f(x)$ such that $\mathscr{F}(E) = \int_E f d\mu$ for each $E \in \mathscr{K}$.

Thus countably additive set functions may be generally represented as integrals with respect to non-negative measures. This is the “abstract” version of the Radon-Nikodym theorem. As we shall see below, this work facilitated a revised view of the decomposition theorem, which was later to be termed the Lebesgue-Nikodym theorem.

We have a certain amount of additional information about the importance that Nikodym attached to abstract approaches. For example, Nikodym gave a series of radio talks, published in 1946 as “Spojrzmy w głębiny myśli: cykl wykładów popularnych z dziedziny nauk ścisłych” (Let’s look deep into ideas: a series of popular lectures from the field of scientific research). Here the role of logic, and mathematics as a kind of avatar of logic in scientific research, is compared to the role of intuition. Nikodym stresses, using a variety of examples from mathematics and theoretical physics, that logically-based mathematics has both a formalizing role with respect to intuition, where it serves as a corrective; and a creative, innovative role where good formulations may lead to the development of guiding intuition.

> Abstract thought grew in deep minds, in silence and concentration; although it does not make a fuss, it weighs heavily on the fate of entire nations, guiding the progress of culture, imperceptibly but reliably.

17.4 Riesz’s 1928 Program and Its 1936/40 Realization

In 1928, the year before Nikodym’s work, F. Riesz spoke at the Bologna ICM on some aspects of his recent work. Here he recalled his 1909 representation theorem, which showed that any continuous linear functional can be represented as a Stieltjes integral with respect to a function of bounded variation; because of this, he noted in 1928, the study of functions of bounded variation becomes a study of linear operations (as he then phrased it). Thus, for example, the result of Jordan that a function of bounded variation can be written as the difference of two monotone functions translates into the fact that such a continuous linear functional can be decomposed into two linear non-negative operators of which it is the difference. This raised, for Riesz, general questions about decomposition of functions and functionals, which he was to explore intermittently over the next decade and more. Here he was particularly interested in the relationship between indefinite integration and the existence of derivatives, a location where the relation between measure and integration becomes particularly tricky.

Riesz in particular called attention to a result of Fréchet, coming from work of Lebesgue, De la Vallée Poussin, and Beppo Levi, which stated that a function of bounded variation α could be written as a sum

$$\alpha(x) = \alpha_1(x) + \alpha_2(x) + \alpha_3(x)$$

where $\alpha_1(x)$ is an antiderivative for $\alpha'(x)$, and is absolutely continuous; $\alpha_2(x)$, the "singularities function" is continuous, of bounded variation, with derivative zero almost everywhere; and $\alpha_3(x)$, the rest, is the "jump function". These results were extracted from the analytic representation of the function. Riesz now sought to turn the tables and approach the same material, as he put it, synthetically. The aim is to investigate such questions with linear operations (as he then termed operators) directly, and in particular operations on functions defined on abstract sets, explicitly following the lead of Percy Daniell's integration studies from around 1920 (we omit discussion of Daniell's work here).

In order to do this, Riesz introduced a notion of majorant and minorant operators; because there is a greatest minorizing operator, for example, these operate analogously to the sup and inf of sets of real numbers. Operators are called *disjoint* if min(A, B)=0, that is, if the only (non-negative) operator that minorizes both is the one that is identically 0.

If I is then the Lebesgue integral over an interval, an operator (*opération positive*) is defined as *singular* if it is disjoint from I. It is *regular* if it is disjoint from all singular operators; a singular operator is *continuous* if it is disjoint from all operators of the form $A_f = f(x_0)$ (evaluation at a single point), and *purely discontinuous* if disjoint from all operators that are continuous. These three categories of operators furnish the decomposition corresponding to that of Fréchet's version of Lebesgue decomposition:

> ...to decompose a non-negative operation A into three parts ...we have only to use a procedure that is almost automatic. Consider all the elements of the respective class that minorize A ...and take the least majorizer. Proceeding thus for each of the three classes we get three operations [i.e. operators]. These are precisely the same that Fréchet arrived at in decomposing the function of bounded variation that figures in the expression of A. (Riesz 1930).

Riesz then turns to a clarifying discussion of the relationship of these ideas with Lebesgue integration. This provides much of the intuitive content of the work, particularly the 1940 detailed version that was to prove influential. What Riesz terms a regular operator in this context is one that can be written as the Stieltjes integral (he is restricting himself here to continuous real-valued functions f) of some Lebesgue-integrable function ϕ:

$$A(f) = \int_a^b \phi(x) f(x) dx.$$

If for the moment we consider positive functions ϕ, the ordering of the operators is then straightforward, and the integral $I = \int_a^b f(x)dx$ is a kind of "unit.". Since the ϕ is determined (up to a set of measure 0), there is a one-to-one correspondence between regular operators and positive Lebesgue-integrable functions. That is, the Lebesgue theory of integration can be built as a theory of linear operators.

The discussion in this paper is largely programmatic: there are not many theorems. But this paper (Riesz 1930) prefigures much of his work over the next years in this area, and feeds directly into the work of Dieudonné, among others.

As for the question of justification, Riesz wrote later about this paper that by restricting himself to linear operators on continuous functions, the parts of the decomposition could be obtained without recourse to the explicit expression as a Stieltjes integral:

> Such a method has the advantage of being completely general, so that it can be used for linear operators of functions defined on abstract sets, as M. Daniell has studied in generalizing the notion of integral. (Riesz 1940, 174)

Thus the power of the Daniell integral is called upon to validate the level of generality: linear operations on functions defined on abstract sets.

Riesz was keenly aware of the technical problems that this generalization presented, and he investigated this question in front of the Bologna audience. He noted that, in this case, the regular operators are just those that can be written as

$$A(f) = \int_a^b \phi(x) f(x)dx$$

with respect to some non-negative summable function ϕ. In other words, up to sets of measure zero, there is a one-to-one correspondence between regular operators and summable functions. Thus, as he put it, starting with linear operators on continuous functions, without leaving "the sacred wood of classical analysis", one gets Lebesgue summability. Riesz clearly saw that by altering the classes of objects to which this applies and generalizing to arbitrary elements, he could unify the treatment of a broad class of theories of integral. On the whole, it's interesting that Riesz chose to present this work-in-progress to the ICM. Of course, he could be sure that it would be of broad interest to those working in analysis, and he could draw on a recognized achievement—his 1909 paper, explicitly cited—to pique the audience's interest. But, in fact, he may also have thought that this was one of the really interesting things he was working on. His 1940 paper, a very substantial piece of work, brought this to fruition, and also to the attention of Dieudonné, for whom it was a point of departure for recasting integration theory in a way consistent with the Bourbaki priorities.

17.4.1 The Riesz Generalization of 1936/1940

The paper (Riesz 1940) which elaborates on the statements at the ICM of 1928, restates results that Riesz published in Hungarian in 1936. We can hardly improve, for a brief account of the content and to illustrate the level of generality, on what von Neumann had to say about this paper in the first volume of the new *Mathematical Reviews*. After noting that the 1928 paper had inspired work by Freudenthal, Garrett Birkhoff, Kantorovich and others, von Neumann continued:

> The investigation deals with partially ordered linear spaces ... All other authors postulated further the lattice property, that is, the existence of a least upper bound and of a greatest lower bound for any two elements f, g. The author requires instead: Given any four elements $f_1, f_2, g_1, g_2 \geq 0$ with $f_1 + f_2 = g_1 + g_2$, there exist four elements $h_{11}, h_{12}, h_{21}, h_{22} \geq 0$ with $h_{11}+h_{12} = f_1, h_{21}+h_{22} = f_2, h_{11}+h_{21} = g_1, h_{12}+h_{22} = g_2$. This remarkable decomposition condition is weaker than the lattice postulate. The author then investigates those linear functions $L(f)$ in such a space, for which $L(f)$ is a real number not less than 0 when $f \geq 0$. They and their differences form together a space satisfying the same conditions, the dual of the former. These dual spaces are shown to be lattices, and even to possess the stronger property of lattice completeness. The author establishes this, as well as a detailed spectral theory of the dual spaces, by very elegant direct methods and discusses numerous applications. These include various integrals of Stieltjes, Hellinger and Daniell.

This review emphasizes the importance to von Neumann, and doubtless other readers, of the level of generality. This plays out in the use of algebraic structures (Boolean algebras and lattices, abelian groups) to generalize the algebra of sets of real numbers. Dieudonné later made the observation that such spaces—and even the more general case of topological vector spaces—are not markedly different from sets of functions, as far as integration and its relation to differentiation are concerned (Dieudonné 1948).

This, then, was about linear functionals (by the Riesz representation theorem, then, integrals) on abstract sets. But in the 1940 paper the level of generality is greater. And as Riesz noted, his weaker hypothesis here leads to a lattice structure for the dual space. This permitted Riesz the use of this formal tool to grasp the relations between the different results, as he put it. But besides this, the abstraction itself is presented as an achievement:

> Let us observe what is obvious anyhow, that in the following consideration one could have replaced numerical operations by others with values in any complete linear lattice. (Riesz 1940, 175)

To grasp what is going on here, and the usefulness of the generality, it's illuminating to look at the example of non-negative harmonic functions, defined in some open region in the plane or space. These may be decomposed as stated, but we can't use simply the min or the max of a pair of functions, since this need not be harmonic. However, we can define $\inf(f_1, g_1) = \phi$ where ϕ is the greatest harmonic function bounded above by both f_1 and g_1. The sup may be defined analogously. (Riesz observed that these may be constructed by balayage). But this method is general: one uses for the inf the greatest element below a given pair.

In the Riesz approach, the decomposition of operators into the sum of *disjoint* operators turns into a fundamental tool that permits, say, the definition of integrals or measure of sets in a given "fundamental domain" or base set Ω. The fact that the set of positive linear functionals becomes a complete lattice, and yields the smallest such complete lattice, provides the structure that allows the extension of the basic model for continuous functions, sketched in 1928, to sets of abstract elements.

The generalization works as follows. Given a "unit" positive operator E on a base set Ω and another operator in the same family A, we find the smallest complete family containing E by looking at the set E' of all operators disjoint from E, and then taking $[E]$ to be the set of all operators disjoint from all members of E'. This is a complete lattice, that is, it contains all infs and all existing sups of the original set. It's useful to keep in mind the 1928 example, in which E corresponds to $\int_a^b f(x)dx$ and A would then have a representation $\int_a^b \phi(x) f(x)dx$.

We then simultaneously decompose E and A. Define

$$P_\lambda = \text{positive part of } (\lambda E - A), \lambda \in \mathbb{R},$$

with a corresponding negative part. The meaning here is that for any λ and f, $P_\lambda f$ is the upper envelope of the positive part of all functions $\lambda Eg - Ag$, $g \leq f$. This is convex, hence has a right derivative, and for any λ we can decompose E into disjoint operators E_λ and $E - E_\lambda$, with

$$E_\lambda = P'_\lambda,$$

yielding

$$A_\lambda = \lambda P'_\lambda - P_\lambda.$$

This in turn permits the definition of an operator which corresponds to an integral with respect to E_λ, and Riesz concludes the paper with theorems laying the foundation for analysis in this context.

Riesz's account is clear and economical, and it's hard to abbreviate it with clarity. Let's look at an example, provided by Riesz himself, to get the import of the greater generality. Recall that when f is continuous, for any operator A we can write

$$Af = \int_a^b f(x)d\alpha(x)$$

and we may think of A as generated by $\alpha(x)$, a function of bounded variation, non-decreasing if A is positive. In fact if E is a positive operator generated by $\epsilon(x)$, and if A is also positive and bounded with respect to $E(0 \leq a \leq cE)$, it is already a theorem of Lebesgue that

$$Af = \int_a^b a(x) f(x) d\epsilon(x)$$

where $a(x)$ is the derivative of α with respect to ϵ and this integral is a Lebesgue-Stieltjes integral. The more general approach in this case permits treatment of the case where A is not restricted to being positive or bounded, using the parametrization and the fact that the positive part of $A = \lim_{\lambda\to\infty} A_\lambda$, among other things.

While features of this work interested many writers in the period, we will restrict ourselves in what follows to Bourbaki, and mostly Dieudonné.

17.5 The Decomposition Theorem as a Test of Good Abstraction: Bourbaki and Dieudonné

As is well-known, the Bourbaki group formed in 1934 for the specific purpose of writing a textbook, initially termed a *Traité d'analyse*, that would serve as a modern and unified treatment of the basics in all areas of mathematics. The name, Nicolas Bourbaki, was that of a French officer in the Franco-Prussian war. The original members were René Possel, André Weil, Jean Delsarte, Jean Dieudonné, and Claude Chevalley. Within a few weeks they were joined by others including Jean Leray and Szolem Mandelbrojt. In initial discussions they decided that their integration theory would be based on the Lebesgue integral.

However the interests in research of Weil and Dieudonné, as well as contemporary developments, soon dictated that the version of integration would be more modern than that of Lebesgue, and indeed the minutes don't imply that this was not always clear. In fact, Bourbaki is famous for the textbooks, but he also published a small number of research papers, the first of which appeared in the *Comptes rendus* of 1935 and concerns measure.

In the interval between this work and the 1940 work of Riesz, Bourbaki mathematicians undertook various projects related to integration. One rather singular piece on this emerged from the pen of André Weil. Weil's article "Calcul des probabilités" (Weil 1940) appeared in the *Revue Rose*, one of a pair of journals that brought developments in the scholarly world to the attention of a broad educated public. Weil used the article not only to explain the correspondence between ideas in the theory of measure and those in probability, due to Kolmogorov. He also used it to advance a particular view of mathematical activity, and to propose that measure should be a derived concept, with integration being treated as more fundamental. This called for a reform of the Lebesgue theory:

> But M. Lebesgue having given the leading role in his theory to measure ... all his successors have thought it important to imitate him in this decisive point. We believe that the moment has come, via a tighter analysis, to decompose his discoveries into their constitutive elements, to distinguish what is essential to the treatment of an integral from things that

> have to do with the particular properties of the sets on which we are usually operating. Such an analysis is in fact possible (1), ... and here is a sketch of the first results to which it leads. (Weil 1940, 206)

The footnote (1) refers us to the fact that the theory of integration will be taken up in a a forthcoming volume of Bourbaki, "a number that will also contain the first principles of the probability calculus" (Weil 1940, ibid). This promise to deal with probability turned out not to be kept.

We note that this statement seems to establish that Weil certainly did not know at that point of the earlier efforts by Riesz, dating back to 1912. The discussion that begins the paper explains the idea of a formal theory to the reader of the *Revue Rose*, in a way that basically outlines the Hilbertian formalist point of view.

17.5.1 Dieudonné and the Lebesgue-Nikodym Theorem

As mentioned earlier, Bourbaki prepared work on integration in this period, that remained unpublished. However, the contents are to be gleaned at least partially from use made of them by Jean Dieudonné. In a sequence of related articles titled "Sur le théorème de Lebesgue-Nikodym" I-V, published between 1941 and 1951, Dieudonné explored the Lebesgue decomposition result as generalized by Nikodym, first in the context of Riesz spaces (that is, ordered vector spaces), then in the context of topological vector spaces, the focus of much of his activity over this period. It's interesting that these papers were published in global dispersion: the US and France, but also India, Canada, and related work in Brazil.

We note in passing that the Lebesgue-Nikodym decomposition also was examined by von Neumann at about the same time. In his 1940 paper "On Rings of Operators III" (v. Neumann 1940, 126–130), he formulates this theorem in the context of rings of operators on Banach spaces. It functions as a tool, for example in results about the commutativity of functionals with functions in the base space, and the results rest on square-integrability.

Dieudonné's paper of 1941 builds directly on the insight of Riesz that the decomposition theorem holds equally well, not only for function spaces, but for "Riesz spaces", as Dieudonné was to baptize them in 1944: sets whose elements are ordered and possess an abelian group structure (Dieudonné 1941). The theorem Dieudonné sought to generalize is stated as follows:

Theorem *If E is the set of all continuous real-valued functions $x(t)$ on* $[0, 1]$, *every linear functional on E may be written uniquely as*

$$L(x) = \int_0^1 y(t)x(t)dt + S(x)$$

where y is measurable, positive, and well-defined on $[0, 1]$ *except on a set of measure 0, and $S(x)$ is a positive singular linear functional.*

Dieudonné observes that this result has been generalized to the Radon-Stieltjes integral by Nikodym, as we have discussed, and indeed terms it the Lebesgue-Nikodym decomposition. His assessment here of the achievement of Riesz follows the account of Riesz himself, and emphasizes the fact that the decomposition theorem is retained in the more abstract setting (Dieudonné 1941, 567–568).

Dieudonné immediately noticed one desideratum following Riesz's paper:

> However, the theory of F. Riesz does not permit us to give a sufficiently precise form to the "absolutely continuous" part of the decomposition as in the formula. This is primarily because the product of two elements of E [which can be any set] need not be defined. (Dieudonné 1941, 548)

To supply this lack, Dieudonné called upon the recent introduction of limits in a "filtering set" by Bourbaki. By this means, Bourbaki had shown how to define a locally convex topology on an appropriately restricted set of linear functionals on E (they need to be relatively bounded). This in turn permits a new definition of the "smallest complete family" containing the given U (which corresponds to the integral in the decomposition) in the sense of Riesz. From this standpoint, it was possible for Bourbaki to give specific form to the functional corresponding to the absolutely continous portion of the decomposition, and hence to obtain what he termed "a perfect generalization of the Lebesgue-Nikodym theorem." (Dieudonné 1941, 549)

Dieudonné notes that the idea of Bourbaki he discusses is from an unpublished 1939 "premier projet" of the portion of the treatise dedicated to the theory of integration. Who in Bourbaki authored this is unclear, though Weil was probably the first to publish about the Bourbaki integration theory.

The introduction of a topological approach to this set of questions might be termed a return to such an approach. While the setup is somewhat technical, I'll describe it here succinctly, since without it it is not easy to understand the relationship between the work of Riesz and that of Dieudonné. A *filter* $\mathscr{F}$ on a set E (a set of functions for Bourbaki) is a collection of subsets that is non-empty, does not contain the empty set, and is closed under intersection and "upward inclusion": every set containing an element of $\mathscr{F}$ is also in $\mathscr{F}$. The idea is due to another Bourbaki member, Henri Cartan, who had introduced it in 1937 to generalize the notion of limit of a sequence (Cartan 1937). Instead of the ordering by decomposition that Riesz had defined, Bourbaki used the idea of limits with respect to a filtering set. The latter is achieved by considering, on the one hand, finite partitions of a function, where if $x \in E$ is a function a partition $\mathscr{P}$ is a set of functions x_i such that $x = \sum x_i$. The partitions are ordered in the obvious way, so that $\mathscr{P}$ is a filter, or filtering set.

If we then seek a limit of a set of linear functionals $I_1, \ldots, I_n$ this is done with respect to a Lipschitz, positively homogeneous function ϕ on $\mathbb{R}^n$:

$$\phi(I_1, \ldots, I_n) = \lim_{\mathscr{P}} \sum_i \phi(I_1(x_i), \ldots, I_n(x_i)). \tag{17.1}$$

As Dieudonné noted. "F. Riesz also defines these "functions of linear functions" but in an entirely different, and more convoluted [détournée] way," (Dieudonné 1941, 548) a value judgement that would seem to privilege the resemblance of his own viewpoint to that of "ordinary" integration, and its topological content. For this provides the possibility of defining a locally convex topology on a suitably restricted space of functions E, those that can be expressed as the difference of positive functions (a property harking back to Jordan). This further permits a redefinition of the "smallest complete family" via a completion operation, but in this case, the object corresponding to the absolutely continuous part of a operator can be explicitly represented. We omit the details, as does Dieudonné.

Despite Dieudonné's preference for this topological approach, he notes that Riesz's method made it possible to avoid the assumption that the members of E were functions:

> This last part of the work of N. Bourbaki (of which the capital point is the proof of the identity of the part of F defined by the topological route, and the corresponding "smallest complete family") assumes as an essential point that the elements of E are functions. It remained to examine the possibility of arriving at analogous results while staying on the path followed by Riesz, that is, without supposing the functional character of the elements of the set E.

This is exactly what Dieudonné does in the following pages. Without the assumption on the elements, he proposed a *lattice* structure on E, which also needs a product, that is, it should be a commutative ring. This permits the use of filters, since the formula 17.1 does not require the arguments of the operators to be functions. This allows the definition of the topological structure, in a way that will not be elaborated here. We note that the ring and lattice structure brings us closer also to von Neumann. The delicate part rests on actually reproducing the theory of Riesz. As Dieudonné noted, the method consists of lifting an idea of von Neumann—one used in his proof of the Lebesgue-Nikodym theorem.

In fact, as Dieudonné doubtless suspected early on, it's possible to characterize exactly where the Lebesgue-Nikodym decomposition holds. In Dieudonné (1944), he gave a necessary and sufficient condition for the abstract result of 1941 to hold, and further showed that the only "Riesz rings" for which the theorem holds are isomorphic to rings of functions integrable (summable) with respect to a completely additive measure on a ground set (functions may differ on a set of measure zero). In other words, the abstract study of the L-N result will not in itself lead to applications beyond the classical ones. In the words of Riesz, we never leave the "sacred wood of classical analysis." Further investigation led him to the result, in fact, that even when the theorem fails, the corresponding Riesz space will be isomorphic to a space of integrable functions under fairly broad hypotheses.

In fact, this premonition, in Dieudonné's view, underlay also the work of Riesz himself, as well as other work of Kakutani, Freudenthal and others. In other words, when one considers an abstract linear space with order that is a lattice—a Riesz space, when some conditions are added—our intuitions that come from spaces for which the points are functions are reliable.

17.6 Concluding Remarks

17.6.1 Abstract Integration: Two Views from the Fifties

In 1953, The Bulletin of the AMS published reviews of two works that suggest a significant divergence of opinion about the correct way to present the theory of measure and integration. The two works in question here were Bourbaki's *Intégration*, reviewed by Paul Halmos, and K. Mayrhofer's *Inhalt und Mass*, reviewed by Jean Dieudonné. Each reviewer had sharp criticisms of the book he had been given. In Dieudonné's review of Mayrhofer, for example, he advances specific criticisms about content, and in general he feels that the work is out of date:

> Finally, the reviewer wants to take exception to the author's statement that measure theory (as understood in this book) is the foundation of the theory of integration. This was undoubtedly true some years ago, but is fortunately no longer so, as more and more mathematicians are shifting to the "functional approach" to integration. It is always rash to make predictions, but the reviewer cannot help thinking that, despite its intrinsic merits, this book, as well as its brethren of the same tendency, will in a few years have joined many an other obsolete theory on the shelves of the Old Curiosity Shop of mathematics. (Dieudonné 1953)

Dieudonné remained true to the iconoclastic role associated with his membership in the Bourbaki group, and his lack of appreciation of Mayrhofer's stately march through the recent, already classical theory originating with Carathéodory's "abstract" measures of 1914 is unsurprising given intervening developments. These included his own role as a member of the Bourbaki team that had produced the first four chapters of the *Intégration* volumes, in a sense a competing textbook, that had appeared in the previous year. It also reflects research work of his own, work immediately following on interests of Riesz and of the Bourbaki group.

The Bourbaki volume was also reviewed in the Bulletin—by another author of a textbook on measure theory, Paul Halmos, whose book on the subject had appeared in 1950. Halmos in turn was unappreciative of several important aspects of the Bourbaki approach:

> Putting ourselves in the place of a student, we must ask: "Is the subject important, is the book clearly written, and is the material well organized?" Putting ourselves in the place of the supervisor of a Ph.D. thesis in one of the applications of integration (e.g., ergodic theory, probability theory, length and area, Boolean algebra, or integral geometry), we must ask: "Is this the point of view that will help a student to understand and to extend his field of interest ?" I say that the answer to the student's question is yes and the answer to the professor's question is no (Halmos 1953).

Halmos's objections centre on the fact that the version of measure theory employed by Bourbaki is not well-adapted for use in areas he considers central, most notably probability and ergodic theory.

But this is not the only problem:

> Owing, no doubt, to the authors' predilection for using as definiens what for most mathematicians is the definiendum, there are many spots at which the treatment appears artificial. (Halmos 1953).

This statement is supported by several telling examples. These illustrate not only a presentation strategy for a complex collection of related theoretical results, but a set of values about what it is important to do in the mathematics of the period.

The differences between the two points of view signal different ways in which abstraction is justified and what the approach provides. These revolve to some degree around both subjects and methodology, both in regard to the role of axiomatization, and in the way in which definitions are made and employed. These are fundamental aspects of theory construction, and lie at the heart of the very abstract world of pure analysis in the first half of the twentieth century.

Halmos felt the approach in which measure is a derived notion was artificial, and ill-suited to the applications he thought most important: probability and ergodic theory. We have already seen that Weil came to a different conclusion, though of course Weil's programmatic statement did not contain a completely worked out version of probability theory.

The appeal of the measure-free approach to Bourbaki was surely the hope of a unified approach to integration in a large variety of contexts. The Riesz program immediately provided this hope, since the space of elements on which the operators acted was so general, and Riesz himself extracted the Lebesgue integral and the Daniell integral. Work of Dieudonné and others on integration of vector-valued functions also illustrates this hope. This was a era in which hopes of very general, very powerful approaches carried a lot of weight in directing research programs: we can think of category theory, and of the theory of topological vector spaces. Indeed, the latter theory, which Dieudonné worked hard to promote, is a place where one required the locally convex topology worked out in the 1941 paper.

The purpose of representation and decomposition theorems may be seen to be somewhat different in the papers of these different writers. Nevertheless, Riesz's 1909 insight that every linear operator was somehow an integral, and Dieudonné's hard-won account of every Riesz space being a function space, show a shared idea about what abstraction should be doing. The resistance to the functional approach to integration, which we have not taken up here, shows that not all agreed.

Despite the fact that the power of the generality was less universally agreed upon than some had hoped, the value of the functional approach was more than validated. The direct approach via measure theory likewise continued to be taught and used. The rather friendly contest between the two has no ultimate resolution, and the process shows that mathematicians resist discarding good approaches for the sake of unification of approach.

17.6.2 Abstraction as a Modernist Strategy

At the beginning of this paper we proposed a brief reflection on aspects of abstraction in modern art and literature, with the end in mind of comparing the way in which abstraction functioned there to the ways in which it functions in mathematics. We have illustrated its use in the work of various writers on analysis, largely confining ourselves to one related set of results in order to show both variations and commonalities between different authors.

In all the fields, previous problems are treated, not only by new methods, but by taking certain features out of the picture while retaining others. Of course, that is just to say that the word abstraction can be applied meaningfully in these cases. More than that, though, in mathematics as in the arts we see a movement away from naturalistic representation to a standpoint in which the work is determined by the maker, as well as by the "rules of the game," which however do not remain fixed and are to some degree up to the players. Riesz raised the bar and created a decomposition theorem for operators on arbitrary sets; Dieudonné " improves" it by naturalizing its methods via topology, but then sought to recover the more general viewpoint. The "sacred wood" of the classical remains a reference point. Likewise one does not "abstract off" the notion of correct proof. By contrast, literature and painting may seem more free. Perhaps there is nothing in mathematics that is equivalent to lying a canvas on the ground and pouring paint on it directly out of a bucket. But even there, the retention of recognizable elements of achievement need at least to be argued, sometimes by the artist, sometimes by critics—sometimes by collectors and dealers, as perhaps contemporary art shows. If the NFT can work for Jeff Koons and Emily Ratajkowski, perhaps it has a future in the mathematical world.

References

Banach, S. 1923. Sur le problème de la mesure. *Fundamenta Mathematicae* 4 (1): 7–33.

Borgato, M. T. 2012. Giuseppe vitali: Real and complex analysis and differential geometry. In *Mathematicians in Bologna, 1861-1960*, ed. S. Coen, 31–55. Basel: Birkhäuser.

Carathéodory, C. 1914. Über das lineare Mass von Punktmengen – eine Verallgemeinerung des Längenbegriffs. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1914: 404–426.

Carathéodory, C. 1918. *Vorlesungen über reelle Funktionen.* Leipzig: B. G. Teubner.

Cartan, H. 1937. Filtres et ultrafiltres. *Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences, Paris* 205: 777–779.

Chemla, K., R. Chorlay, and D. Rabouin. 2016. Prologue: generality as a component of an epistemological culture. In *The Oxford Handbook of Generality in Mathematics and the Sciences*, 1–41. Oxford: Oxford University Press.

Corry, L. 2004. *Modern Algebra and the Rise of Mathematical Structures* (2nd ed.). Basel: Birkhäuser Verlag.

Dieudonné, J. 1941. Sur le théorème de Lebesgue-Nikodym. *Annals of Mathematics (2)* 42: 547–555.

Dieudonné, J. 1944. Sur le théorème de Lebesgue-Nikodym. II. *Bulletin de la Société Mathématique de France* 72: 193–239.
Dieudonné, J. 1948. Sur le théorème de Lebesgue-Nikodym. III. *Annales de l'université de Grenoble. Nouvelle série. Section sciences mathématiques et physiques* 23: 25–53.
Dieudonné, J. 1953. Book review: Inhalt und Mass. *Bulletin of the American Mathematical Society* 59 (5): 479–480.
Georgiadou, M. 2004. *Constantin Carathéodory*. Springer.
Halmos, P. R. 1953. Book review: Intégration. *Bulletin of the American Mathematical Society* 59 (3): 249–255.
Hausdorff, F. 1914. *Grundzüge der Mengenlehre*. Leipzig: Veit.
Hawkins, T. 1970. *Lebesgue's Theory of Integration: Its Origins and Development*. University of Wisconsin Press.
Lebesgue, H. 1903. *Leçons sur l'intégration et la recherche des fonctions primitives*. Gauthier-Villars.
Mayrhofer, K. 1952. *Inhalt und Mass*. Springer.
Nikodym, O. 1930. Sur une généralisation des intégrales de M. J. Radon. *Fundamenta* 15: 131–179.
Riesz, F. 1912. Sur quelques points de la théorie des fonctions sommables. *Comptes rendus* 154: 641–643.
Riesz, F. 1930. Sur la décomposition des opérations fonctionnelles linéaires. In *Atti Congresso Bologna*, vol. 3, 143–148.
Riesz, F. 1940. Sur quelques notions fondamentales dans la théorie générale des opérations linéaires. *Annals of Mathematics (2)* 41: 174–206.
Riesz, F. 1949. L'évolution de la notion d'intégrale depuis Lebesgue. *Annales de l'institut Fourier Grenoble* 1: 29–42.
v. Neumann, J. 1940. On rings of operators. III. *Annals of Mathematics (2)* 41: 94–161.
Vitali, G. 1904–1905. Sulle funzioni integrali. . *Atti dell' Accademia delle Scienze di Torino 40*, 1021–1034.
Weil, A. 1940. Calcul des probabilités, méthode axiomatique, intégration. *Revue Scientifique (Revue Rose Illus.)* 78: 201–208.

Chapter 18
On Set Theories and Modernism

José Ferreirós

Abstract "Classical" set theory in the style of Cantor, Dedekind and Zermelo enjoyed dominance during the early twentieth century, playing a prominent role in many of the "modern" mathematical developments, in analysis, algebra, topology, and so on. Yet the theory was polemical, and one finds a characteristic pattern of second thoughts about set theory, after an initial enthusiasm; examples we briefly discuss are Borel, Weyl, and Rey Pastor, which contrast with Hilbert's optimism. As a result, one can speak of a proliferation of set theories – in the plural – during the period 1910–1950. Classic examples are the (more or less coherent) *predicative* proposals of Russell and Weyl, the "intuitionistic set theory" developed by Brouwer, and several other systems which we cannot discuss here. We consider in this paper whether such criticism of set theory was a symptom of traditionalism, which leads to an analysis of the notion of *modernism*, paying especial attention to the case of L.E.J. Brouwer. I shall argue that modernism is a somewhat ambiguous notion, and that Brouwer (like Weyl) can indeed be regarded as prototypically "modern" in a sense that was characteristic of the Inter War period 1918–1939.

The rise of "modern mathematics" in the first half of the twentieth century featured axiomatics and the idea of structures as key elements, both of them sustained by set-theoretic concepts. In 1910, David Hilbert wrote that set theory "occupies today an

Thanks are due to Moritz Epple, Erhard Scholz, Leo Corry, Javier Ordóñez, and very especially to Mark van Atten, who greatly improved my discussion of Brouwer, and Jeremy Gray for the many conversations that have helped shape my views on the topic. No doubt other historians have also influenced me; may they forgive my lack of memory. This work had a complicated story: it is partly based on a chapter meant for a collective book, which was ready in 2008, but I have eliminated some sections and combined it with new material for the present volume. Work on it was supported financially by research grants FFI2009-10224, P07-HUM-02594 and FFI2017-84524-P.

J. Ferreirós (✉)
Departamento de Filosofía y Lógica, Universidad de Sevilla, Sevilla, Spain
e-mail: josef@us.es

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_18

outstanding role in our science, and radiates [*ausströmt*] its powerful influence into all branches of mathematics".[1] Already in 1900, his choice of the first two among his celebrated 'Mathematical Problems' was signalling that set theory would be central in the future.[2] Also in 1910, his disciple Hermann Weyl stated that "set theory seems to us today the proper foundation, in the logical sense, of the mathematical sciences" (Weyl 1910, 302 and 304). On the basis of ε (the membership relation) and those relations that can be based upon it, one can establish mathematical definitions and axioms, the result being a structuralist conception of the discipline, with "completely isomorphic systems" playing a central role (Weyl 1910, 300–301).

Set theory articulated a new conception of mathematical ontology, deeply at variance with the traditional views of the seventeenth and eighteenth centuries focused on magnitude; set-theoretical methods not only supplemented more traditional ones, but they became central to a new form of axiomatics and to the mathematical study of "structures."[3] To remain in the "belle époque" 1890–1914, one can mention the influence of set theory on:

- the "modern theory of [algebraic] numbers" developed by Dedekind and Hilbert;[4]
- algebra and algebraic geometry (the paradigmatic expression "modern algebra" is found in a later book);[5]
- modern analysis, theory of integration, measure theory (linked with the names of French mathematicians Borel and Lebesgue);
- general topology (through contributions due to Cantor, Hausdorff, Brouwer and others);
- new researches on geometry due to Hilbert, the Italian school of Peano, Pieri, etc.

The mathematicians involved in these developments, around 1900, were mainly centred in Göttingen, Torino and Paris, but it was especially the Germans (seconded by the Italians) who forcefully promoted the new viewpoint. The influential group of Göttingen mathematicians around Hilbert became the main promoters of this *new mathematics*, which they paradigmatically identified with set-theoretic methods.

Yet, the rise of set-theoretic "modern" mathematics was polemical. Many relevant authors had second thoughts about set theory: after an initial period of enthusiasm with the new concepts about the infinite, they evolved into milder or

[1] Hilbert 1932, 466, from the obituary of his also famous friend Minkowski, originally published in 1910. Zermelo said that it was only "the influence of D. Hilbert" which made him realise the importance and deep significance of the fundamental problems of set theory.

[2] First Cantor's continuum problem, second the consistency of the set-theoretic definition of the real numbers.

[3] For a masterful exposition of these novelties, insisting on matters of ontology, epistemology and methodology, the reader is referred to (Gray 2008). The paradigmatic definition of a structure around 1950 was as a kind of set-theoretic construct (Bourbaki), but more recently a lively debate has been developing about other possible approaches, notably the one based on category theory.

[4] The expression is literal from (Hilbert 1897, iii) and it became a ready-made phrase later on.

[5] B. L. Van der Waedern, *Moderne Algebra* (Berlin: Julius Springer, 1930). See (Corry 1996).

stronger critique. Weyl himself was a prominent example. He explained that, as a student in Göttingen, he had been raised in the set-theoretic tradition and even "confined" into "that complex of notions which today enjoys absolute domination in mathematics," connected above all with the names Dedekind and Cantor, *but* he was able to find his "own way out of this circle" of ideas (Weyl 1918, 35–36). In his opinion, set theory and classical analysis involved a vicious circle and had to be abandoned:

> I started initially with an examination of Zermelo's axioms for set theory, which constitute an exact and complete formulation of the foundations of the Dedekind-Cantor theory. My intention was to fix more precisely the concept of 'definite set-theoretic property', which Zermelo employs in the crucial Axiom III of 'Subsets', since his explanation appeared unsatisfactory to me. And so I was led to the principles of definition of § 2. My attempt to formulate these principles as axioms of set formation and to express the requirement that there exist no more sets than those formed by finitely many applications of the principles of construction embodied in the axioms – and indeed, to do this *without presupposing the concept of the natural numbers* – drove me to a vast and ever more complicated formulation, but unfortunately not to any satisfactory result. Only in connection with certain general philosophical insights, which I could only find by renouncing conventionalism, did I realise clearly that I was wrestling with a scholastic pseudo-problem. And I became firmly convinced (in agreement with Poincaré, however little I share his other philosophical ideas) that *the conception of iteration*, *of the sequence of natural numbers, is an ultimate foundation of mathematical thought* (despite Dedekind's 'theory of chains', which seeks to give [it] a logical foundation ...)." (Weyl 1918, 36-37)

Weyl's case was by no means an exception. Already in 1910, Paris was the main center of criticism of set theory, especially as a result of the rejection of the Axiom of Choice by such influential mathematicians as Borel, Baire and Lebesgue. Their "five letters" discussing the matter in 1905 became well known, and Borel gave them publicity from 1914, when he included them as appendixes to his treatise on the theory of functions. The *vicious circle principle* was formulated by Poincaré and Russell, leading them (and Weyl) to the rejection of impredicative forms of definition that can be found in analysis and set theory.

Was such criticism of set theory a symptom of traditionalism? A lack of sight, a methodological blindness linked with anti-modernism? What is modernism, after all, and to what extent did it determine the new form of mathematics? These are the questions I would like to consider in this chapter, paying especial attention to the case of L.E.J. Brouwer. I will argue that *modernism* is a somewhat ambiguous notion, and that Brouwer (like Weyl) can indeed be regarded as prototypically "modern" in a sense that was characteristic of the Inter War period 1918–1939.

18.1 Introductory Remarks on Modernism

Jeremy Gray (2008) has offered a remarkable exposition of the twentieth-century transformation of mathematics into so-called modern mathematics, presented for a broad audience interested in history of science and history of mathematics. In an effort to make clear and graspable the links to cultural modernism, Gray emphasized traits such as the following four: the new vision of mathematics as an *autonomous* body of ideas; the view that it has little or no outward reference; renewed and "considerable" emphasis on *formal* aspects of the work; and the "complicated – indeed anxious – ... relationship with the day-to-day world".[6] This new self-conception was increasingly shared by a coherent community of professionals, with a high sense of the seriousness and value of their discipline (Gray 2008, 1).

The topic of modernism and its connection with the "foundational crisis" was also studied in a pioneering work, *Moderne – Sprache – Mathematik* (1990), by Herbert Mehrtens. This book was written in an interesting postmodernist spirit, its analysis taking inspiration from cultural history and semiotics. It was Mehrtens' thesis that one can better understand the early twentieth-century foundational crisis, not as having to do with problems of rigor and foundations alone, but as an expression of social readjustements within the discipline, of conflicting cultural definitions of the figure of the mathematician herself. Whence his subtitle, "a history of the fight for the foundations of the discipline and the subject of formal systems." According to him, different attitudes expressed in the famous "foundational crisis" of the 1920s correlate with the modernism or counter-modernism of the personalities involved.

Mehrtens presented Hilbert as the prototypical modernist, the "managing director" of the new business,[7] while his opponents are (almost automatically) counter-moderns, being critics of the modern methods. The catchwords for the new attitudes of the modern mathematicians, according to Mehrtens (1990, chap. 2), were *freedom*, *productivity*, *fruitfulness*, and it was also frequent to talk about abstraction and about formal language. The counter-moderns (*op. cit.*, chap. 3) expressed concerns about the *integrity* of mathematics, and favoured instead the *concrete*, the given, or *intuition*; they explicitly promoted restrictions on the way mathematics was to be conducted, having to do with meaning and truth. "Both sides are into formulating the limits of disciplinary discourse, and thus concerned with the «I» and the «us» of mathematicians. Here lies the dilemma of modernism, not merely mathematical, in the Self that is «ineffable» (Ernst Mach) and «no longer master in his own house» (Sigmund Freud)."[8] As a matter of fact, Mehrtens employs not

[6] Of course there are parallelisms in art history; an intriguing example of parallel traits in a very different domain is given e.g. by Gray's discussion of "Catholic modernism" (2008, 141–142). See also (Ferreirós & Gray 2006).

[7] Mehrtens (1990), 108. The colourful expression was employed by Hilbert's friend and colleague Hermann Minkowski in a letter commenting on Hilbert's celebrated lecture of 1900 on 'Mathematical problems.'

[8] Mehrtens 1990, 9–10. A not very careful reader of Mehrtens will thus extract the idea that there were merely two opposite sides – the modernists or progressives, and the conservatives calling for reactionary reform. Such a simplistic scheme would be untenable, but Mehrtens is more sophisticated and faithful to the events. His essay is about "the emergence of mathematical modernism around the turn of the century and about its nemesis [*Abwehr*], the «countermodernism,» that set itself as an opposition, as a shadow that would not lose it despite all the successes" (Mehrtens 1990, 7). He intimates that there is an internal link between both, so that the modernists could not exist without their shadow – some kind of inner dialectics.

only the two main categories of modernism (*Moderne*) and counter-modernism (*Gegenmoderne*), but on occasion also speaks of anti-modernism (*Antimoderne*). Only the anti-moderns are depicted as reactionary traditionalists.[9]

This is in line with what historian Herf wrote in his book *Reactionary Modernism* (1984, 12), i.e., that modernism was not a movement exclusively of the political Left or Right; we may add that it was not simply a movement of cultural optimism or pessimism, nor a movement of technological enthusiasm or aversion. Even in the characteristically optimistic *belle époque* we find some modernists who are cultural pessimists and reject the technological "modern world." This is well known with respect to writers and artists (consider the contrasting attitudes toward technology and the new industrial world in H. Broch and R. Musil), but the same tensions find natural expression inside the mathematics community (e.g., Brouwer and Hilbert, in that same order; see below).

Europe witnessed in the late nineteenth century the full impact of two great Revolutions of the previous century: industrialisation and democratisation – or reactions thereto, as in Germany. Indeed, a second industrial revolution was under way, related to technological, science-based industries exemplified by the electrical and chemical companies; and it might not be inappropriate to say that a second socio-political revolution was also happening: extension of voting to all of society, the rights of women, emergence of the welfare state. The most obvious understanding of modernism is that it consisted in different forms of *explicit and high-strung cultural reaction* to those historical experiences. Implicit in this perspective is the idea that we must *differentiate between modernity and modernism* – a central point in my attempt to rethink the issue.[10] Obviously, there is a broad gamut of possible cultural reactions to modernity – an intensive transformation of life and society – ranging from fully optimistic embracement to an equally extreme and pessimistic rejection. What is most characteristic of the modernists is their emphasis on "the new," but this can be coloured in different ways.

The first four decades of the twentieth century was a time when many were expecting full arrival of the new, if not fighting to promote the revolution. It could be a "new world" of technological marvels, but also new social ways, new kinds of human relations; or a "new art" in music, literature, painting, or a "new science" in radical break with the past; or else – most important – it could be a "new man" (or woman) having the traits expected by the Marxists, or those proposed by the Fascists, or by other groups.

Gray (2008) acknowledges explicitly that cultural modernism was a broad movement with strong tensions between diverging tendencies and interpretations, and this has as a consequence that the links between cultural and mathematical

[9] In fact this category seems to be reserved for those who developed political orientations akin to National-socialism (Mehrtens 1990, 14–15, 308*ff*).

[10] Judging from the reactions of some colleagues and referees, I come to think that perhaps there is a cultural element in this perspective of mine: perhaps there is a reflection of my cultural background in Spain and its complex cultural and political history in the fact that, from the beginning, it seemed obvious to me that modernism and modernity are to be distinguished.

modernism become rather elusive.[11] In fact, Gray chooses to downplay the issue (2008, 14), focusing instead on the "independent" modernist transformation of mathematics understood as a phenomenon internal to the discipline (*op. cit.*, 37). In this contribution it will be my aim to insist on the *plurality* of modernisms and the ensuing difficulty of employing the category in such a way that mathematical modernism and "modern mathematics" are simply identified.

Thus, in what follows my principal aim is to emphasize the complexity of the picture, warning against simplistic interpretations of the period. Generally speaking, cultural movements are *never* homogeneous and monovalent, rather they are marked by tensions. This should be a truism for historians, but historians of science (and here I think e.g. of recent debates concerning the Scientific Revolution) often tend to the monovalent picture. Of course, that complicates any attempt to relate scientific developments with the culture that made them possible. We find partisans of the modern methods in mathematics that show clear signs of cultural traditionalism, and opponents of the "new maths" that may be counted among the best examples of modernist mathematicians. Perhaps my most important argument will be a vindication of the figure of L.E.J. Brouwer as a modernist – so that, in my view, the identification of modernism with so-called "modern mathematics" becomes dubious.

18.2 Hilbert's Optimism

David Hilbert became in the 1900s one of the most reputed mathematicians in the whole world. He was a leader within the institution that led international mathematics at the time, the University of Göttingen, contributing significantly to the modernisation of the mathematicians' work. Work at Göttingen became much more collective and oral than was usual in the past, when mathematicians laboured in isolation on the basis of books and journals. There was a great number of students around, and the weekly meetings of the Göttingen Mathematical Society played a decisive role, with many visitors coming by.[12] Even more important, the puristic values of German mathematics were tempered thanks to the efforts of Felix Klein, leading to the forging of new links with engineering and the natural sciences. It is worth noticing that Hilbert continued with all his energies this tendency, which took the discipline mathematics away from the more extreme trend toward full autonomy.[13]

[11] The point is acknowledged both by Mehrtens (1990, chap. 7) and Gray (2008, chap. 1).

[12] For a thorough discussion of this topic, see Rowe (1989) and (2004).

[13] On this topic see Corry (2004). With R. Courant, Hilbert authored a key book for mathematical physics, *Methoden der Mathematischen Physik* (Berlin: Julius Springer, 1931), in the tradition of Riemann and Klein.

With his 1900 talk on mathematical problems, and especially with the choice of the first two,[14] Hilbert made – I surmise – a conscious attempt to *influence* the direction of mathematics. It was his bet for the future dominance of the set-theoretic orientation, in contrast to the severe criticisms voiced by powerful members of the older generation in Germany. Indeed, he was willing to invest all of the influence and respect he had accumulated over the years, in the attempt to preserve the new set-theoretic methods. This is what Hilbert did during the famous foundational crisis of the 1920s, and in this way he became a scientific icon, the name most directly associated to "the modern" in mathematics. Another reason why Hilbert became an icon of modern mathematics was the strong link established between his name and the "mathematics of axioms".[15]

In the 1920s, already more than 60 years old, Hilbert was still inventing methods based on axiomatisation for grounding mathematics anew, namely proof theory and metamathematics. But he was also using highly charged rhetoric: he depicted Cantor as the prophet of the new mathematical paradise, Hilbert himself as his vindicator; the intuitionist Brouwer was presented as a leader of the reactionary party, a follower to that villain figure of Leopold Kronecker (leading mathematician at Berlin from the 1860s to the 1890s, and ardent critic of the modern infinitary methods).[16] The following sentence is well known, but deserves to be quoted again:

> We shall carefully investigate those ways of forming notions and those modes of inference that are fruitful; we shall nurse them, support them, and make them usable, whenever there is the slightest promise of success. No one shall be able to drive us from the paradise that Cantor created for us. (Hilbert 1926, 375–76)

Hilbert's rhetoric aligned Brouwer with Kronecker, yet it should be clarified that Brouwer accepted infinitary methods, unlike Kronecker. A few years before, when Hilbert's disciple Weyl saluted Brouwer as "the Revolution" (Weyl 1921), Hilbert answered that he was not a revolution at all, but rather an attempted *Putsch* repeating the failed attempt of Kronecker. Clearly it was a fight between Progress and Reaction; one could even say, without abandoning the biblical language that Hilbert liked to employ, between Good and Evil.

Both the Promethean connotations of such rhetoric, and the abstractness of the axiomatic and set-theoretic mathematical style, suggest the possibility of finding links between the "new maths" and cultural modernism. Set theory helped to establish the new mathematical developments on foundations that not only seemed to guarantee methodological rigor, but especially a freedom of thought that was

---

[14] Even though his public presentation was limited to 10 of the 23 problems in his list, those two were among the chosen ones. On this topic, see Ferreirós (2007), chap. IX.

[15] In the present context it is important to notice that his understanding of axiomatics was profoundly linked with set-theoretic methods. On this topic, see Kanamori and Dreben (1997), Gray (2000), Ferreirós (2009).

[16] In the 1920s, Arthur Schoenflies did historical work that established Kronecker's reputation as a malevolent enemy of Cantor, and even a major cause of his mental illness. Much earlier, Hilbert had obtained an impression of Kronecker's ways of promoting his enmity to the "new mathematics" from his friends Minkowski and Hurwitz, who knew well the Berlin master.

strongly emphasized by Cantor, Dedekind and Hilbert. Cantor went as far as saying that "freedom" is "the essence" of mathematics, while Dedekind liked to repeat that numbers are "free creations of the human mind." Such is the ultimate referent for Hilbert's defiant cry of 1926, the portrait of himself as a rebellious Adam fighting to remain in the Garden.

According to Herf (1984, 12), the "central legend" of modernism "was the free creative spirit ... who refuses to accept any limits and who advocates what Daniel Bell has called the «megalomania of self-infinitization»." In this central inspiration a romantic motif lingers on. Both Cantor and Dedekind seem to have received the cultural impact of romanticism early in their lives, but of course ideas and trends associate rather freely in different minds: flexibility is the rule. Dedekind preserved an emphasis on the "free creative spirit" and on the absence of limits for mathematical "creation" (concept formation, introduction of structures), but in a classicist reading that reinforced more and more the idea of the logical laws at play. Cantor, by contrast, remained closer to Herf's description, much more of a romantic throughout his life, and even combated explicitly the customary depiction of the human mind as finite.[17]

All of this was part and parcel of a clear move towards the autonomy and self-containment of mathematics. Hilbert's talks (especially his famous conference of 1900 on 'Mathematical problems') are examples of *belle époque* optimism, transmitting the message that the future is ours, full of freedom, and the received concepts and methods are basically all right, perhaps requiring minor corrections. Hilbert's later *proof theory* was self-containment at its utmost, as he tried to certify the foundations of mathematics by purely mathematical means (and it is tempting to see its failure as one of the many failures of extreme modernist tendencies shortly before the Second World War).

There is little doubt that Hilbert can be dubbed a "modern man," to pause for a moment on this aspect of the question.[18] He was far from the old habits of German University mandarins, to the point of being criticised for his careless way of meddling with students. There are clear signs of his progressive stance in matters of culture and society, like his promotion of social democrat philosopher L. Nelson or his way of defending that Emmy Noether should be appointed a University lecturer ("*Meine Herren*, the faculty is certainly not a public bath", see Reid 1972, 143). He had the attitudes of an enterprising man of science, his life fully devoted to his specialised business, which he conceived as an autonomous enterprise. In fact, one of Hilbert's contributions in connection with set theory and foundational studies

[17] Interestingly, Brouwer thought that mental proofs are, in general, infinite objects (see his *Collected Works*, vol. 1, p. 394); I thank van Atten for referring me to this passage.

[18] This is an aspect of the figures under review that we shall take into account. Considering *modern* as an actor's category, we avoid any need to define it more properly, but it might be worthwhile to ask whether there is a prototypic "modern persona" (in the sense of *scientific personae*, see Daston and Sibum 2003).

was to free their discussion from the philosophical and metaphysical elements that figured prominently in Cantor and other members of the older generation.[19]

But perhaps one should ask for more when talking about modernism. I have mentioned the need to distinguish between modern*ism* and modern*ity*. The mere fact that the institutional conditions of mathematical work were significantly modernised in Göttingen does not imply that the relevant actors were modernists. Indeed, historical evidence pointing to the distinction is very close at hand. Felix Klein was as important as Hilbert for the modernisation of the enterprise of mathematics at Göttingen, but the ambivalence of Klein's attitudes makes it difficult to classify him (Mehrtens counts him among the *counter*-moderns, see 1990, 206*ff*). Historically Klein was a central figure in the promotion of the modern methods, associated with the names of Riemann, Jordan, Cantor, Lie, Hilbert, through his original work (e.g. the Erlangen Programme), his editorial activities in *Mathematische Annalen*, his activity as leader of a mathematical school, and not least his promotion of the rising young star Hilbert during the 1890s.[20] Now, if we have to differentiate between modernism and modernity in the case of Klein, the same must apply to Hilbert and others.

Of course, applying such criteria strictly, not many names will be left. The best examples of culturally modernist mathematicians that I know are Felix Hausdorff (b. 1868), L.E.J. Brouwer (b. 1881), Hermann Weyl (b. 1885), and Alfred Tarski (b. 1902).[21] But of those four names, two are strongly linked with the critique of modern mathematics and the proposal of alternative methodologies: Brouwer and Weyl were "counter-moderns," to use Mehrtens' rather unsatisfactory label.[22] I shall only discuss in detail one of those names.

18.3 Disenchantment

Authors who were as optimistic as Hilbert, concerning set theory, did not abound: one might mention Hausdorff, Fraenkel, von Neumann, Mahlo or Tarski as examples. Many others were exploring the pros and cons of set theory in an (apparently)

[19] A detailed analysis of his foundational views would show this in full clarity. But here it may suffice to indicate a simple symptom: in the 1910 quotation given above, first page, Hilbert wrote *Mengentheorie* – and not *Mengenlehre*. In doing so he was avoiding the traditionalistic overtones so frequent in Cantor's work.

[20] See (Rowe 1989) on the "intellectual alliance" between Klein and Hilbert.

[21] Perhaps Betrand Russell and Alfred N. Whitehead might be other good candidates. See on Hausdorff (Mehrtens 1990) or (Epple 2006), on Brouwer see van Dalen's biography (1999) and (van Stigt 1990) (1996), on Weyl the book edited by E. Scholz (2001) and (Scholz 2006), and about Tarski (Feferman and Feferman 2004).

[22] Although Weyl's allegiance to Brouwer's ideas only lasted for 4–5 years, from his 1918 book *Das Kontinuum* to the end of his career he remained a critic of set theory and preferred alternative restricted methods; it has been said that, throughout his career, he consciously avoided employing the most significant and polemical axiom of set theory, the Axiom of Choice (Ferreirós 2007, 339, quoting Dieudonné).

noncommittal way: this is the case of such important experts as Sierpiński and Luzin, and probably A. Church too.[23] This fits the 'atmosphere of insecurity' that, as I have argued (Ferreirós 2007, chap. X), reigned during the 1920s and 30s. But one can also find many relevant mathematicians who entertained serious doubts about set theory: Poincaré, Borel and Skolem of course, but also Lebesgue, Baire, Weyl, van der Waerden, ... And the list continues even to this day.[24]

A key figure in this respect was Émile Borel, highly influential in the period 1910–1940. His *Leçons sur la théorie des fonctions* (Borel 1898) were not only a pioneering work introducing set theory to the French community, but also –in subsequent editions – a work promoting skepticism towards Cantorian ideas. The second edition appeared in 1914, the third in 1928,[25] and he kept adding new appendixes. Borel comments that he was a "dangerous revolutionary" when the first edition came out, but later became a "reactionary" in the eyes of his students, because of his attitude towards AC and his requirements about "definitions" (Borel 1950, x). He recognizes the "freedom" of mathematicians to introduce concepts and axioms, modulo consistency, but he also remarks that "some branches of this science, having had their hour of fame, have been abandoned" due to sterility (1950, xi).

Especially important was Note IV (2nd edn 1914, p. 135) entitled "Les polémiques sur le transfini". This included the well-known *five letters* exchanged in 1905 between Borel, Lebesgue, Baire and Hadamard (pp. 150–158). By publicizing widely these letters, Borel contributed to promoting critical stances. Let me quote in particular an interesting passage due to René Baire in 1905:

> The expression *given a set* is employed all the time, but does it have a sense? Not always, in my view. Once we speak about the infinite (even countable, and here I am tempted to be more radical than Borel), the assimilation, *conscious or unconscious*, with a bag of balls that one passes from hand to hand, must disappear completely – we are, in my view, in the *virtual*, which is to say that we establish conventions which allow us later, once an object is defined *by a new convention*, to affirm certain properties of this object. But to believe that one has gone farther than this, that does not seem legitimate. In particular, from a set being given (we may be in agreement to say, for instance, that we give ourselves the set of [all] sequences of positive integers), *it is false – I think – to consider the parts of this set as given* [i.e., the powerset]. A fortiori I refuse to attach a sense to the fact of conceiving a choice within each part of a set [as Zermelo does, JF].

[23] In papers such as 'Auswahlaxiom und Kontinuumshypothese' (1938), Sierpiński offered a most competent exposition of the uses of and paradoxical consequences of the axiom of choice (AC) and the continuum hypothesis (CH) in mathematics, but he declared himself to be neither for nor against AC.

[24] Fields medal W. P. Thurston e.g. calls the axioms of set theory 'polite fictions' (Thurston 1994). Many first-rate mathematicians would opt for minimalism as regards set theory and the higher infinite: the Bourbaki themselves, Quine or Wang proposed restrictions on set theory around 1950; but also much later we find influential examples such as Martin-Löf or Feferman from the 1970s, Voevodsky in the 2000s.

[25] The fourth in 1950. First edn was a mere 134 pages, the fourth featured 150 more of appendixes.

Also noteworthy is Borel's talk given at the Rome ICM 1908 (p. 159*ff*), where he asserts that the concept of uncountable set is "*purely* negative". The idea of a countable infinity of choices is quite dubious, but the idea of an uncountable infinity of choices is "entirely devoid of sense" (Borel 1950, 160). In another paper of 1912 (quoted by Lusin 1929, 22), he is forcefully against admission of the full second number class:

> wherever *all* of the transfinite numbers of the second number class must effectively take part (and not merely those that are less than one of them, fixed in advance), it seems to me that we are abandoning the domain of Mathematics.

Thus the notion of $\aleph_1$ is illusory.

Let me provide another example, the leading mathematician in Spain and Argentina at the time, Julio Rey Pastor (1888–1962). His book *La matemática superior* is interesting for several reasons, among them the contrast between the first edition (1916, written under the influence of his stay at Göttingen in 1913) and the second in 1951. Rey Pastor speaks of his "youthful enthusiasm" for set theory, later abandoned. Early on, he explained that mathematics would be defined in the future as "the *Science of Sets*". But in the later edition he changed the first chapter of the book, and wrote:

> At the beginning of the 20th century, ephemeral golden age of Cantorism, one considered the transfinite Arithmetic and Geometry of Cantor as the indispensable basis of Analysis; the new treatises devoted a long initial chapter to them. The successes of the French school (Borel, Baire, Lebesgue, Fréchet) confirmed this prestige, while the criticism of Borel and Poincaré moderated the excessive enthusiasm, making clear that the whole of classical Analysis can be rigorously built without the actual infinite but with some Cantorian concepts that are indispensable to promote the theory of real functions ... (accumulation points, neighborhoods, ...). The rise of the unexpected paradoxes, early in the century, cooled down the enthusiasm of which this author was infected, giving the transfinite an undue prominence in these conferences. (Rey Pastor 1951, 20–21)

According to him, mathematics in the nineteenth century, while going from the concrete to the abstract, had never lost sight of reality and the natural sciences. Meanwhile, it is "the characteristic of our times: arbitrary abstract constructions, arbitrary postulates and arbitrary functions; a sum of arbitrariness that horrified Poincaré" (Rey Pastor 1951, 7).

As a result of such criticisms, different mathematicians proposed different approaches to set theory (and related systems), some of them very wide and Cantorian (e.g. von Neumann's), some much more restrictive. The restrictions were sometimes proposed in naïve or informal terms, as happened with Borel, Lebesgue, Baire or Luzin, sometimes they were duly formalized – an example being *simple* type theory, which may be regarded as a variant of set theory (Ferreirós 2007, 348–356). And one could even find an "intuitionistic theory of sets" proposed by Brouwer (1919) right after the War. The complexity of this confusing situation in the Inter-War period was even greater, due to the unclarity and divergences concerning the scope and limits of logic (see Ferreirós 2007, 345–348, 357*ff*).

18.4 Brouwer: An "Entirely Modern Man"

Luitzen Egbertus Jan Brouwer (1881–1966) belonged to a generation which falls squarely within the high time of modernism: he was 19 years younger than Hilbert. One of the very best mathematicians in his generation, regarded as a key founder of topology, he is remembered above all for his deviant ideas about the foundations and methods of mathematics. But that was not the only major theme in Brouwer's life: he had very strong philosophical interests and an eccentric lifestyle; 2 years before his dissertation in 1907, he published a monograph entitled *Life, Art, and Mysticism*. During the next few years, however, he devoted his efforts to purely mathematical questions of high contemporary relevance, which established his great reputation and won him a professorship in 1913.[26]

His high achievements in the field of topology made him internationally famous, as he introduced new methods, rigorising the field, and proved central results like the fixed point theorems and the invariance of dimension under bicontinuous mappings. This is perhaps his most important result of the period. Cantor had established in 1878, to everyone's astonishment, that two spaces with different number of dimensions (e.g., a simple line and all of three-dimensional space) can be bijected, so that to *every* point in space there is one corresponding point on the line! Dedekind conjectured that such bijections must be highly discontinuous, and that dimension can be proven invariant if the one-to-one correspondence is supposed to be bicontinuous. The result was only proven rigorously by Brouwer in 1912, establishing the topological nature of the concept of dimension (interestingly, Cantor's attempted proof of 1879 was the only major error in his publications).

Brouwer's reputation brought him offers of chairs at Berlin and Göttingen in 1919 (remember that Göttingen was the world centre for mathematics). The Berlin professors, on that occasion, stressed that Brouwer had provided "the firm foundation" for topology, "long sought in vain," and they opined that he "is equalled in the originality of his methods by none of the mathematicians of the younger generation" (Van Dalen 1999, 300). Hilbert too seems to have regarded Brouwer as the most brilliant among his generation, but the young Dutch was by this time fully absorbed into developing his unorthodox views on mathematics. This led to the "foundational crisis" in the 1920s, a time when some perceived the danger of a "schism" within the mathematical community, which also led to estrangement of the once cordial relations between both leaders.

There was an enormous difference in scientific and personal outlook between Hilbert and Brouwer, evident already in the 1900s. Hilbert was full of optimism, while the young Brouwer's philosophical works are pessimistic, and his dissertation of 1907 is a call for deep rethinking and reform of mathematics. In 1908 he published in Dutch severe criticism of one of the ideas presented by Hilbert in

[26] The main biography of Brouwer is (Van Dalen 1999); a solid briefer presentation is (Van Atten 2003). See also the introduction to an English edition of *Life, Art, and Mysticism*, (Van Stigt 1996); for the reception of intuitionistic mathematics, and more, see (Hesseling 2003).

1900 (the view that all mathematical problems are solvable) under the title 'The unreliability of the logical principles'. The principle of excluded middle, basic for traditional mathematical proofs by *reductio ad absurdum*, is characteristically rejected in intuitionistic logic and mathematics.

Initially the relation between the two was cordial, and young Brouwer clearly admired Hilbert: he even compared him to "a prophet."[27] Yet the fight between both in the 1920s has stamped their differences so much, that it seems almost impossible for most mathematicians to regard Brouwer as a "modern." After all, while concentration on foundational issues died out in the midst of the deep fog brought about by Gödel's incompleteness results, a stream of mathematical results continued to flow nonetheless, and mathematicians remained faithful to the mathematical style that Hilbert had defended – i.e., the set-theoretic principles and methods; axiomatic, structural mathematics.

Meanwhile, in his Cambridge Lectures after World War II, Brouwer says things such as "in modern or intuitionistic mathematics" (1981, 92), or speaks of "the modern theorems" of intuitionism (1981, 94).

18.4.1 Foundations

Apart from mysticism, topology and intuitionism, a fourth great theme throughout Brouwer's intellectual life was the philosophy of language (Van Dalen 1999). The philosophical themes were densely connected to the intuitionistic mathematics he came to propose. The question of language is particularly relevant, for Brouwer would continue to emphasize that mathematics does *not* depend on language or logic, being prior to language and logic (Brouwer 1928). The point for the intuitionist is that mathematics is a mental construction erected freely by the mind: "Mathematics is inner architecture" (Brouwer 1948, 96), it is "an introvert science, directed at beauty" (Brouwer 1981, 90).

Language is merely an instrument of social domination, and it makes impossible a real knowledge: it is "absolutely a function of the social activity of humanity," a medium for the "transmission of the will in the cultural community" (Brouwer 1928, 179). It lacks exactness or certainty, and so "*for pure mathematics there exists no certain language*" (1928, 180). The efforts of the formalist school are therefore based on a false belief, of almost magical character (i.e., that the formalisation of language can bring certainty to mathematics); mathematical knowledge can only be based on mental constructions, on live thinking. The mistake springs from a much older and more consequential error, "namely, the reckless trust in *classical*

[27] Van Atten calls my attention to a 1909 letter to Van Scheltema: "This summer the first mathematician of the world was in Scheveningen; I was already in contact with him through my work, but now I have repeatedly made walks with him, and talked as a young apostle with a prophet. He was 46 years old, but with a young soul and body; he swam vigorously and climbed walls and barbed wired gates with pleasure. It was a beautiful new ray of light through my life."

logic" (1928, 181). Already in 1908, this line of thought derived into denouncement of axiomatic systems, which from his standpoint cannot be the real foundation of mathematics, and very especially of the principle of *excluded middle* when applied to infinite collections.

The mental constructions of mathematics start from the basic experience of temporal succession, which engenders the idea of two-ity when a life-moment falls apart into two different things (something new comes up, while memory retains the previous thing). There is change but also retention, there is unity in multitude:

> Mathematics is a free creation, independent of experience; it is developed from a single *a priori* ur-intuition, which can both be called *constancy in change* and *unity in multitude*. (summary of the 1908 Dissertation, van Dalen 1999, 117)

By abstraction and iteration, this "originary" intuition unfolds the notions of finiteness, and "the infinite as a conceptual reality" – the totality of natural numbers, the intuitionistic continuum, "and finally the whole of pure mathematics" (Brouwer 1928, 177). And the whole process is completely autonomous, independent of natural science or any other sphere of human activity:

> the mathematical state of mind is usually indifferent with respect to natural science and definitely unfavourably inclined towards the promotion of the exploitation of nature and towards technique (1948 report to Curators, in van Dalen 1999, vol. 2, 828).

Brouwer came to regard the ideas of his early period as the "first act" of intuitionism, with a "second act" after the Great War, when he devoted himself fully to a deep reform of pure mathematics (see Brouwer 1981). The outcome is, in my opinion at least, a monument to the stature of Brouwer as a thinker and mathematician[28]: he was able to erect a whole new building of mathematics, including very subtle theories of the continuum and of analysis, on the basis of principles and methods solidly based on his methodological assumptions, but deeply at deviance with regular mathematics (which caused the enterprise to be extremely difficult and original).

In 1919, precisely the year when he received calls to the Universities of Berlin and Göttingen, he published in the *Jahresbericht* of the German Association of Mathematicians (*DMV*) a paper entitled 'Intuitionistische Mengenlehre'. The year before he had started a series of papers on this same topic (in German, published in the transactions of the Dutch Academy of Sciences), with the goal of revising set theory thoroughly and developing it "independently of the logical principle of excluded middle." This was not a minor reform, it was actually a crucial trait of intuitionism, that makes it strongly deviant from "classical" mathematics. In his paper for the *DMV* Brouwer stressed that he had been elaborating these ideas since 1907, before his involvement with topology. He emphasized that the foundations

[28] Readers who may want to disagree will certainly be able to find good company: see e.g. the amusing remarks on Brouwer offered by Grattan-Guinness (2000, 480 *ff*). More insightful analyses can be found in Mehrtens (1990, 257–287) and Gray (2008, 413 *ff*) and van Atten (2003).

for set theory provided by Zermelo was to be rejected, and that a new form of "*constructive* set theory" was required (Brouwer 1919, 203–204).[29]

His proposals were the most radical in this age of radicalism: the set theory of Cantor and Dedekind was to be completely overturned, classical ideas about the real numbers had to be abandoned, and with them classical analysis, replaced by a very novel intuitionistic analysis. A technical point that is relatively easy to grasp is the following: the basic theorem that an infinite, bounded set of real numbers always has a least upper bound, has to be abandoned.[30] The world of mathematics was to be constructed anew, from scratch, in such a way that the meaningless features of traditional approaches would be erased, and a new pure edifice would be erected, full of sense. Applause came from unexpected sides, especially when Hilbert's most brilliant student, a fully mature mathematician among the very best in their time, saluted him with the words: "Brouwer – that is the Revolution" (Weyl 1921).

In his dissertation of 1907, Brouwer had actually explained how he could accept some of Cantor's ideas, including the transfinite numbers ω, $\omega + 1$, ... up to a certain point – as long as they are denumerable and in a certain sense constructible – but not the further concept of a "totality of all" such denumerable numbers. We have seen Borel defending a similar standpoint. Cantor introduced this totality in the 1883 *Grundlagen*, and called it the "second number class;" thesis 13 of Brouwer's dissertation states: "The second number class of Cantor does not exist" (van Dalen 1999, 113). It was not the set-theoretic paradoxes that caused his reaction; as he remarked in 1923,

> an incorrect theory, even if it cannot be checked by any contradiction that would refute it, is none the less incorrect, just as a criminal policy is none the less criminal even if it cannot be checked by any court that would curb it. (quoted in Hesseling 2003, 62)

It is simply an illusion to conceive of mathematics as dealing with independently existing objects, with an objective reality somehow external to the mind. But this is what 'modern' mathematics does: the objects of its theories are imagined to be elements of an actually given totality, a domain independent of the thinking subject. This feature is deeply embedded in the methods employed in mathematics, and following Bernays (a key collaborator of Hilbert) it is often called the "Platonism" of modern mathematics (Bernays 1935).[31] Meanwhile, the constructivists' treatment of mathematics – exemplified by intuitionism – is based on careful consideration of the processes by which numbers, functions, etc., are defined or constructed. Each and every thing that a mathematician can legitimately talk about must have been explicitly constructed in a mental activity.

[29] This is not the place for a detailed presentation of intuitionistic mathematics. Brouwer made an effort to be clear and readable in his lectures (1981), see also Heyting (1956).

[30] This is a feature not only of intuitionism, but more generally of constructivism; the development of analysis is severely affected by having to circumvent that simplifying principle.

[31] See also Ferreirós (2008). The point is discussed in any good textbook on philosophy of mathematics.

As time went by, Brouwer realized that it was better to avoid talking of "sets" at all, and he introduced new terminology ("species" and "spreads"). But the core of his intuitionism, the novel ideas that he was developing and making public among other places in the *Mathematische Annalen* (edited by Hilbert), are a direct continuation of the views on "intuitionistic set theory" proposed in 1919. On their basis, Brouwer submitted mathematical analysis to deep revision. This explains why von Neumann could write: "Today there exists no single general set theory – but a naive, an intuitionistic, as well as several formalistic-axiomatic systems" (von Neumann 1928, in his *Collected Works*, vol. 1, 321).

As Brouwer's reconstruction of mathematics developed in the 1920s, it became more and more clear that intuitionistic analysis was extremely subtle, complicated and foreign. Brouwer was not worried: "the spheres of truth are less transparent than those of illusion," as he would remark in 1933. But Hermann Weyl, even though he welcomed Brouwer's "Revolution" in 1921 and was convinced that he had delineated the domain of mathematical intuition in a completely satisfactory way, remarked: "the mathematician watches with pain the largest part of his towering theories dissolve into mist before his eyes" (Weyl 1925, 534). Soon Weyl abandoned intuitionism, although he remained a constructivist (see Mancosu 1998).

18.4.2 The Man

As for Brouwer the man, Hesseling (2003, 86) mentions two characteristic traits upon which both friends and foes would agree: that he was an outstanding mathematician, and also an impressive personality. A close friend of his youth, the Dutch poet and socialist thinker Adama van Scheltema, regarded Brouwer as "*in all respects* the paradigm of a man of exceptional genius – for being a *genius* he lacks the connection between his own mind and the world around him" (van Dalen 1999, 26).[32] That seems to have been said with the romantic, nineteenth-century idea of a genius in mind: a stormy personality, a soul capable of rising to heaven and reaching the depths of misery. Van Atten summarizes that Brouwer seems to have been an independent and brilliant man of high moral standards, but with an exaggerated sense of justice, often making him pugnacious. As a consequence, in his life he energetically fought many battles: "It still strikes me as curious that a person can get involved in so many disputes," writes van Dalen (1999, ix). To Mannoury's daughter, Brouwer once said: "Indeed, your father is one of the few people with

[32] That's something Scheltema sought to remedy in their student years, endeavouring "tirelessly . . . to make him come closer to the material world."

whom I have never had a quarrel. But he brought out the good in people, and I the bad."[33]

After visiting Amsterdam in 1928, Edmund Husserl reported to Heidegger in a letter of May 5, 1928:

> Among the most interesting things in Amsterdam were the long conversations with Brouwer, who made a great impression upon me, that of a wholly original, radically sincere, genuine, entirely modern man.[34]

That same year, another great philosopher, Wittgenstein, was impacted by listening to a talk that Brouwer gave in Vienna; the topic was 'Mathematics, science, and language' (Brouwer 1928), and the experience is said to have marked Wittgenstein's return to creative thinking (see Monk 1990).

Brouwer was deeply individualistic and felt deeply distrustful of established social views, bourgeois life and the like. These ideas were expressed in his 1905 monograph *Life, Art and Mysticism*. In the early days, he had strong qualms about becoming a mathematician, as he was deeply engaged with philosophy and mystical ideas. The thought of becoming an "expert" seemed to him a temptation of the will, a way of falling into attitudes that would separate him from the true perspective on life. During the 1900s he was strongly critical of science, technology, and modern civilisation; and much of this was retained throughout his life. Causal thinking, language, and technology were viewed as negative forces. An admirer regarded his 1905 monograph as the most formidable accusation against "our 'civilisation'."

Brouwer could be a high-nosed man, extremely demanding with himself and others, scornful towards most of his fellows in humankind, imbued of an elitist conception. In his youth he thought of himself as a "king," one of the best and noblest, who worked for the redemption of things and of their fellow men (van Dalen 1999, 37–38), i.e., to "put somewhat in the right position" the "grand totality" (*op. cit.*, 34). Unsurprisingly, he soon abandoned the socialist convictions of Scheltema: "I rarely think about politics, but my political sympathies are in the liberal, anti-democratic direction," he wrote in 1909 (van Dalen 1999, 35).

Our 'genius' was a high-strung and nervous person, described as uncompromising, and indeed he could act as a "justice fanatic" (in the words of Bieberbach). His 1905 monograph displays chapter titles as expressive as 'The sad world' and 'Man's downfall caused by the intellect.' He criticized severely the intellect (related with means-end rationality) for being the source of evil, and also social activities;

---

[33] I take these sentences verbatim from M. van Atten, "Luitzen Egbertus Jan Brouwer", *The Stanford Encyclopedia of Philosophy (Summer 2011 Edition)*, Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/sum2011/entries/brouwer/. Gerrit Mannoury, a polymath who among other things published mathematics and philosophy, was one of Brouwer's most important and influential teachers.

[34] I thank Mark van Atten for calling this passage to my attention, and also for the translation. From E. Husserl's *Briefwechsel* Vol. IV, (1994), p. 156.

he called for a return to nature and for placing confidence in the self and its will.[35] The book ends with a Schopenhauerian exaltation of nihilism.[36]

Recall that, according to Daniel Bell, modernist spirits had a definite tendency to the "megalomania of self-infinitization;" one may read from this viewpoint the words "you feel almighty," that Brouwer wrote in his 1905 booklet. Consequently with this worldview, he believed in a God and devoted himself to the tasks imposed by God – precisely the "utter powerlessness" experienced by the ego at times, reveals all around him the presence of a "higher power." But he also managed to find ways to reconcile his deep beliefs with the ways of life.

A life that was unconventional, far from the established manners of the bourgeois. It was centred on a hut in Blaricum, a village not too far from Amsterdam, from which Brouwer could pay regular, but scarce visits to the University. There he enjoyed life among the artists, the vegetarians, and the health seekers.[37] He led a most Spartan lifestyle, with peculiar diets and habits (van Dalen 1999, 62*ff*), and became notorious for his sexual freedom, never hidden from his wife. Surprising facts in the life of someone who was convinced – at least in his passionate youth – that "the illusion of woman" burdens your soul's karma, and that the ultimate goal should be to sacrifice everything, neglect everything, in order to reach the highest perfection: "the world of freedom, of painless contemplation, of – nothing."[38]

Brouwer could also be extremely perceptive, e.g. in his denunciation of the "human cancer" with its will for power and domination, and its way of turning the beautiful world into barren land (van Dalen 1999, 68). This happened at a time when it was not easy to figure out the destructive power of mankind, although it happened in the Netherlands, a man-made land. Science and technology were to him clear expressions of such a will, and so is mathematics most often. But the pure cultivation of mathematics can become independent from all that, it can be a pure activity worthy of the human mind. It was in this way that he reconciled himself with the idea of exploiting his great mathematical talents.

Brouwer's reconciliation with mathematics as a pure activity worthy of the human mind depended on the proviso that it be *kept apart* from its "applications," even those in pure science. The tendency to purification was already strong in

[35] Hesseling 2003, 30–34. Van Stigt 1996. Intuitive introspection and mystical views were given preference, with the names of Meister Eckehart, Jakob Böhme and the Bhagavad Gîta emerging as important references.

[36] Nihilism was a natural consequence of the philosophical solipsism of young Brouwer. As early as 1898, being only a boy of 17, he wrote: "the only truth is my own ego of this moment, surrounded by a wealth of representations in which the ego *believes*, and that makes it *live*" (van Dalen 1999, 18).

[37] As he described it himself around the turn of the century: people of both sexes in vest with bare black feet and blue nails, the sunbathing of bare backs, the gnawing of raw turnips and carrots (see van Dalen 1999, 28).

[38] As regards his wife, *Life, Art and Mysticism* makes clear the expectations of submission that Brouwer had at the time of choosing her (see van Dalen 1999, 73).

the nineteenth century, although not exactly for Brouwer's reasons.[39] It started in northern Germany, and with the great success of the German mathematicians it spread out all over Europe and into the United States, leaving its mark well into the twentieth century. Brouwer was not the only modernist who was interested in mathematics precisely because of its purity and autonomy from worldly concerns. An interesting example is the writer Robert Musil, who had been an engineer and admired mathematics greatly:

> It is understandable that an engineer is preoccupied by his specialty instead of coming into the freedom and space of the world of ideas, even if his machines are delivered to the furthest corners of the world; ... About mathematics one cannot say this; it is the new method itself, the spirit itself, in it lay the sources of time and the origin of an immense transformation. (*Der Mann ohne Eigenschaften*, 1930)

By contrast, an important part of the modernisation brought about by Klein and Hilbert in Göttingen consisted precisely in *moderating* the purist tendencies of German mathematicians, countering them with the creation of strong ties between mathematicians, physicists, and engineers.[40] Little wonder, then, that in 1919 Brouwer manifested a preference for Berlin instead of the leading center, Göttingen: his worldview was antagonistic to the Göttingen ways, much more in line with the purism and idealism represented by Berlin.

18.5 The Ambivalence of Modernism

All of those are, of course, reasons why people who identify modernism with modernity have no doubts: Brouwer can only be classified as a conservative, an inheritor of nineteenth century idealism, a man who turned against the modern world, perhaps even anti-modern. Thus, for instance, van Dalen writes that he "was basically a conservative" (1999, 72) and finds it hard to understand why others think of him as a revolutionary innovator. But this may be an ill-informed judgement, ignoring as it does how strongly modernism was linked with the late romantic legacy, how clearly represented are all of Brouwer's cultural themes among the more pessimistic and revolutionary modernists. Mehrtens is subtler and has catalogued Brouwer as the prototypical "counter-modern" (Mehrtens 1990, 188).

Let me suggest again that we better make a distinction, since otherwise – to speak now about people in the arts – we would have to reclassify many of the modernists,

[39] See e.g. Goldstein et al. (1996), Bottazzini and Dalmedico (2001).

[40] Proofs abound: the Göttingen Association for the Promotion of Applied Physics (1898), created by Klein in association with industrialists; his efforts to hire Ludwig Prandtl, who became head of the Institute for Technical Physics in 1905 and did pioneering work on aerodynamics; the professorship of applied mathematics created in 1904 for Carl Runge; the important contributions that Hilbert and Minkowski made to physics, their deep involvement with the subject in the 1900s and 1910s; the close links with the physicists, which Max Born for instance describes. See Rowe (1989).

people like the writer Hermann Broch, the painter Wassily Kandinsky, and so on. Kramer remarks:

> That the emergence of abstraction early in the second decade of this century represented for its pioneer creators a solution to a spiritual crisis; that the conception of this momentous artistic innovation entailed a categorical rejection of the materialism of modern life; and that abstraction was meant by its visionary inventors to play a role in redefining our relationship to the universe—all of this, were its implications even dimly grasped, would come as a shock to many people who now happily embrace the history of modern art as little more than a succession of styles, or art fashions, ...[41]

Properly revised, those are fitting words for Brouwer's intuitionism, its origins and its goals. Thinking about a mathematical counterpart to abstraction in painting, it seems to me that intuitionism is clearly the best. And if you consider again Herf's characterisation of modernism (quoted above), it should be obvious that Brouwer fits it better than any of the other figures we have mentioned so far. Modernism was not a movement of the political left or right, it could well be anti-democratic and elitistic, it was often against the myths of progress and modern civilisation.[42] Its "central legend," says Herf (1984, 12), was "the free creative spirit, at war with the bourgeoisie, who refuses to accept any limits" –how well does this fit Brouwer!

Brouwer's outlook may seem reactionary, and yet there is nothing in his proposals that resonates with the traditionalistic tendencies of a Cantor. It is true that Brouwer opposed technological progress, and he also rejected socialism because of his fierce individualism – is this not a modernist trait? One should also remember the aestheticizing tendencies of modernism, which again fit Brouwer's life and work so well. His truly rooted mathematics, based on intuition, is highly deviant from mainstream modern mathematics, but this can hardly constitute an argument for classifying him from the cultural and intellectual point of view. The great difference between mathematics and art, at this point, is that the former – unlike the latter – has seen a dominant orientation during the twentieth century. But if we let this fact influence too much our cultural judgements about the past, we shall commit an anachronism.

I should add that, if the reader is pondering the possibility of using the label "reactionary modernism" for Brouwer, there are some reasons why this runs into trouble. Herf's reactionary modernists were not only cultural traditionalists, but also nationalists, and they managed to combine some form of romanticism and cult of the soul with advanced technology. Brouwer was clearly not a traditionalist, he was against nationalism, and also against modern technology. The point concerning the latter has already been done, and as regards the former suffice it to say that, in the middle of the Great War, he wrote approvingly of anti-nationalism, denouncing

[41] Kramer (1995), p. 3. How that relationship was to be redefined was explored by Kandinsky not only in his paintings but in his influential treatise *Concerning the Spiritual in Art* (1911).

[42] Notice that Herf is not defining reactionary modernism here, but modernism in general. In this respect, van Atten comments that intuitionism was hailed by figures from both left and right, such as A. Khinchin in 1926 Soviet Russia (Verburgt and Hoppe-Kondrikova 2016) and O. Becker in 1933 Nazi Germany (van Atten 2003).

"words-of-power" such as 'Fatherland' as a "*means of defense of injustice*" (van Dalen 1999, 248–249). There have been attempts to align Brouwer with reactionary political ideas, largely because he was a great admirer of German culture and fought against the decision of the allied scientists to exclude German scientists from international meetings and scientific organisations during the 1920s. But the historical circumstances of these events were far more complex than a simplifying picture suggests (see vol. II of van Dalen's biography, but also Rowe and Felsch 2019).

Thus I think we have to admit Brouwer's modernism and with it the ambivalent nature of this cultural trend. Let me try to summarize the point in a simple and quick way, by adapting words of the modernist artist Ad Reinhardt[43]: The key point of Brouwer's modern mathematics was awareness of mathematics of itself – maths preoccupied with its own processes and means, with its own identity and distinction, maths preoccupied with its own unique statement, conscious of its own evolution and history and destiny – moving toward its own freedom, its own dignity, its own essence, its own reason, its own morality, and its own conscience.

The point concerning freedom may be the most delicate one, even polemical, in connection with Brouwer's work. His mathematics is a free creation of the human spirit, but it is certainly perceived as very *restrictive* by the modern mathematician. His tendency was against merely formal freedom, towards meaning and truth; it was a form of creative thinking free from transcendental assumptions like those of Platonism, and completely independent from questions of nature or reality (naturalism or realism). The greater freedom of modern mathematics would be regarded as based on a "high degree of arbitrariness" and the "belief in a transcendental world" that "taxes the strength of our faith hardly less than the doctrines of the early Fathers of the Church" (literal words of Weyl in 1946, many years after he abandoned intuitionism; quoted in Ferreirós 2007, 390).

In my opinion, it is not only this cultural trend that turns out to be ambivalent, but quite generally cultural movements of the kind are rather complex and ambivalent – think of the Enlightenment, of Romanticism, of post-modernism.[44] Historical epochs are marked by conflicts and tensions, because the new is not monovalent, and because the new and the old may combine in an endless variety of ways. This complicates significantly the task of historians who want to trace the links between cultural movements and the sciences, and makes it very difficult to formulate them in textbook-style summary form. But it is only by tackling that complexity, that we begin to confront the difficult question of the status of science as a form of culture.

[43] I thank the organisers of a Frankfurt meeting that took place in 2006, in particular M. Epple, for bringing the relevant quote to my attention.

[44] I have made this same point for Romanticism in Ferreirós (2003).

18.6 Concluding Remarks

The previous analysis has been based on two theses concerning modernisms: that their plurality has to be recognized explicitly, and that they have to be differentiated from modernity, a much broader movement. Let us now reflect these issues more explicitly.

Jeremy Gray has emphasized the *autonomy* of mathematics as a clear characteristic of its modern, twentieth-century form, and I could not agree more. However, in my view this was largely an effect of the modernisation of the discipline mathematics (alongside the other scientific disciplines) in the nineteenth and twentieth centuries.[45] It should not be identified too quickly with modern*ism*, albeit modernist mathematicians ought to be expected to emphasize that autonomy. Gray goes on to add that modern maths acknowledged "little or no outward reference," and although all of the figures we have discussed are good examples, the issue is subtle and deserves careful scrutiny. In their practice, Hilbert and followers took full advantage of the autonomy of the new modernised configuration of mathematics – indeed they produced quintessential examples thereof. But in their practice, they also preserved carefully the 'outward' links to physics, natural science, and technology: for Klein and Hilbert and Courant, the outer reference was of the essence.

Some authors stress the view that modern mathematics is a language, a purely syntactic, self-referential language (Mehrtens 1990, 8, 12). But this is an overstatement, quite independently from the fact that Brouwer opposed such a view frontally. That viewpoint is akin to strict formalism, a position adopted by some mathematicians and philosophers from about 1930; Hilbert, in particular, was far from it.[46] Often he emphasized the visual and the empirical origins of mathematics, and *always* its heavy load of conceptual contents. In a lecture course during the winter of 1919–1920 he explained:

> There is no talk of arbitrariness here. Mathematics is not like a game in which the problems are determined by rules invented arbitrarily – it is a conceptual system with inner necessity, that can only be this and not any other way. (Hilbert 1992, 14)

On this basis, one might even conclude that Hilbert was a counter-modernist, like his older colleague and powerful ally Klein; but of course an overall evaluation has to be more detailed and nuanced.

[45] On this topic there is abundant literature: see e.g. Bos et al. (1981), Rowe (1989), Goldstein et al. (1996), Bottazzini and Dalmedico (2001) – and of course Mehrtens (1990), chap. 5, and Gray (2008, 32 *ff*).

[46] See Ferreirós 2017, 68. One can find attributed to Hilbert this sentence: "Mathematics is a game played according to certain simple rules with meaningless marks on paper." But even if the phrase is repeated in a thousand web pages, it cannot be found anywhere in his work – it is just made up by putting together several things that, jointly, amount to a severe misrepresentation of Hilbert's thought. In effect, he often emphasised the meaningfulness of mathematical statements and the depth of conceptual content expressed in them.

Consider the question of outward reference in the works and thoughts of Hilbert and Brouwer. Here, Hilbert stroke very clearly and decidedly a middle ground, avoiding any kind of modernist excess[47]: not only was he fully in favour of the Platonist methodology, which accepts (sometimes as an *as-if*)[48] working on the assumption of an independent, ready-given realm of mathematical objects; he was crucial in the disciplinary effort of maintaining and promoting relations between physics and mathematics, and became a great propagandist of the central role of mathematics in all scientific theorizing.[49] Meanwhile, as we have seen, our modernist Brouwer was from the beginning a staunch defender of the need to separate mathematics from the sciences and technology, from any kind of reference to the so-called external world. Intuitionistic mathematics is inner architecture, based purely and solely on mental constructions; it is "an introvert science".

Another key feature of modernism in mathematics, according to Gray, is its emphasis on the formal. What this means is less clear than in the case of painting, because mathematics (like music) has been rather formal and abstract throughout its history. If anything, mathematics has undergone several transitions to more abstract and formal versions, and this has happened repeatedly in different cultural contexts. Certainly Hilbert emphasized the formal structures and introduced novel ways of playing with them, but he was never a strict formalist in the later sense. His metamathematics is a methodology: recourse to the syntactic play of signs in order to secure the consistency of the traditional theories attacked by intuitionism. Furthermore, he was not at all willing to abandon or relegate old contents – while Brouwer was precisely the one who favoured this move. So in my view neither of them fully complies with this characteristic trait of Gray's working definition of modernism.

The last crucial trait is the "anxious relation to the day-to-day world," and here things become more clear. I cannot find the least trace of such an anxiety in Hilbert, as his work and speeches were full of confidence in the future of mathematics, and its central role in relation to the sciences and the material world. Brouwer's work, on the other hand, generated a lot of anxiety because it put at high risk traditional analysis (even the calculus) and with it, as it seemed, the scientific applications of mathematics and its practical utility in the day-to-day world. Brouwer was not anxious, since philosophically these were for him welcome

[47] In questions like this it becomes obvious that Hilbert was indeed Klein's follower in his pragmatic and well-thought disciplinary politics. In my view, this is fully modern (in the sense of modernisation) but not modernist at all.

[48] Hilbert's views on this topic evolved, until he (together with Bernays) embraced a kind of as-if position in the context of his metamathematics of the 1920s; see Ferreirós (2009). Hans Vaihinger, who like Hilbert presented himself as a follower of Kant, published his book *The Philosophy of As If* in 1911.

[49] In a public lecture, 1930, he emphasised that mathematics is "the instrument that mediates between theory and practice, between thought and observation," without which "today's astronomy and physics would be impossible." This was recorded and is available on the internet: see http://math.sfsu.edu/smith/Documents/HilbertRadio/HilbertRadio.mp3 (accessed Sept. 2012).

traits. But Hilbert finally surrendered to anxiety and in 1928, thinking that the end of his life was coming, he dealt a severe blow to Brouwer's position and influence in the mathematical community, in what Einstein (disapprovingly) called the *war of the frog and the mice*.[50] This probably had some role in the practical finishing of the foundational debate, just like Gödel's results had a role in its (certainly deep) theoretical redefinition afterwards.

Summing up, it seems to me that two routes are possible. One of them is to deny one of my tenets, propounding that mathematical modernism must be expected to accompany mathematical professionalisation and modernisation. This is the route taken by Jeremy Gray, who for this reason must de-emphasize the links between mathematics and culture (Gray 2008, 14).[51] If this way of proceeding is taken as a model, it would be quite interesting to apply it to the arts: what are the crucial features of the modernisation of the profession of artist, that the modernist artist ought to respect? And, how would the ensuing picture of modernism in the twentieth century deviate from current views?

The other route, that I have tried to sketch, keeps a close eye on the connections between mathematics and the intellectual atmosphere – but to do so, it disassociates mathematical modernism from what is usually called "modern" mathematics.

References

Benacerraf, P., and H. Putnam, eds. 1983. *Philosophy of Mathematics: Selected Readings*. Cambridge: Cambridge University Press.

Bernays, P. 1935. Sur le platonisme dans les mathématiques, *L'Enseignement Mathématique* 34: 52–69. English in (Benacerraf & Putnam 1983).

Borel, E. 1898. *Leçons sur la théorie des fonctions*. 1st ed. Paris: Gauthier-Villars.

———. 1950. *Leçons sur la théorie des fonctions*. 4th ed. Paris: Gauthier-Villars.

Bos, H., H. Mehrtens, and I. Schneider, eds. 1981. *Social History of Nineteenth Century Mathematics*. Basel: Birkhäuser.

Bottazzini, U., and A.D. Dalmedico. 2001. *Changing Images in Mathematics*. London: Routledge.

Brouwer, L. E. J. 1905. *Leven, Kunst en Mystiek* [Life, Art and Mysticism]. Delft: Waltman. English translation in *Notre Dame Journal of Formal Logic* 37 (3).

———. 1907. *Over de grondslagen der wiskunde* [On the foundations of mathematics], Dissertation, Univ. Amsterdam. English translation in *Collected Works* (Amsterdam, North-Holland, 1970).

———. 1919. Intuitionistische Mengenlehre, *Jahresbericht der DMV* 28. English translation in *Collected Works* (Amsterdam, North-Holland, 1970).

———. 1928. Mathematics, Science, and Language; & The Structure of the Continuum. In (Mancosu 1998), 1175–185, 1186–197.

[50] See van Dalen (1990) or the short presentation in Hesseling (2003), 81–86, but notice that there were political issues at stake, tensions within the German mathematical community that led to a dramatic power struggle for control of the journal *Mathematische Annalen.* For this wider context, see Rowe and Felsch (2019).

[51] Perhaps he should have gone one step further, as modern mathematics has been so successful that anxiety is not one of its distinguishing traits – unlike the case of the arts.

———. 1948. Consciousness, philosophy and mathematics. In (Benacerraf & Putnam 1983), 90–96.
Brouwer, L.E.J. 1981. In *Cambridge Lectures on Intuitionism*, ed. D. Van Dalen. Cambridge University Press.
Cantor, G. 1883. *Grundlagen einer allgemeinen Mannichfaltigkeitslehre*, Leipzig: Teubner. English trans. in W. B. Ewald (ed.), *From Kant to Hilbert: A Source Book* (Oxford University Press 1996), vol. 2.
Corry, L. 1996. *Modern Algebra and the Rise of Mathematical Structures*. Basel: Birkhäuser. 2nd edn 2003.
———. 2004. *Hilbert and the Axiomatization of Physics (1898-1918): From "Grundlagen der Geometrie" to "Grundlagen der Physik"*. Dordrecht: Kluwer.
Daston, L. & H. Otto Sibum, eds. 2003. Scientific Personae [monographic]. *Science in Context* 16(1–2).
Epple, M. 2006. Felix Hausdorff's Considered Empiricism. In Ferreirós & Gray (2006), 263–289.
Feferman, A.B., and S. Feferman. 2004. *Alfred Tarski, Life and Logic*. Cambridge: Cambridge University Press.
Ferreirós, J. 2003. Del Neohumanismo al Organicismo: Gauss, Cantor y la matemática pura. In *Ciencia y Romanticismo*, ed. J. Montesinos, J. Ordóñez, and S. Toledo, 165–184. Tenerife: Fundación Orotava de Historia de la Ciencia.
———. [1999] 2007. *Labyrinth of Thought: A History of Set Theory and Its Role in Modern Mathematics*. Basel: Birkhäuser. 2nd ed..
———. 2008. The Crisis in the Foundations of Mathematics. In *Princeton Companion to Mathematics*, ed. T. Gowers, 142–156. Princeton: Princeton University Press.
———. 2009. Hilbert, Logicism, and Mathematical Existence. *Synthese* 170 (1): 33–70.
———. 2017. Wigner's "Unreasonable Effectiveness" in Context. *The Mathematical Intelligencer* 39: 64–71.
Ferreirós, J., and J.J. Gray, eds. 2006. *The Architecture of Modern Mathematics: Essays in History and Philosophy*. Oxford University Press.
Goldstein, C., J. Gray, and J. Ritter. 1996. *L'Europe mathematique/Mathematical Europe: Histories, Myths, Identities*. Paris: Editions Maison des Sciences de l'Homme.
Grattan-Guinness, I. 2000. *The Search for Mathematical Roots 1870-1940*. Princeton: Princeton University Press.
Gray, J.J. 2000. *The Hilbert Challenge*. Oxford: Oxford University Press.
———. 2008. *Plato's Ghost: The Modernist Transformation of Mathematics*. Princeton: Princeton University Press.
Herf, J. 1984. *Reactionary Modernism: Technology, Culture, and Politics in Weimar and the Third Reich*. Cambridge: Cambridge University Press.
Hesseling, D. 2003. *Gnomes in the Fog: The Reception of Brouwer's Intuitionism in the 1920s*. Basel: Birkhäuser.
Heyting, A. 1956. *Intuitionism, an Introduction*. Amsterdam: North Holland.
Hilbert, D. 1897. [*Zahlbericht*:] Die Theorie der algebraischen Zahlkörper. *Jahresbericht der DMV* 4, 175–546. In: *Gesammelte Abhandlungen* Vol. 1, 1932.
———. 1926. Über das Unendliche, *Math. Annalen* 95 (1): 161–190. Reprinted in *Grundlagen der Geometrie.* Berlin: Teubner, 1930. English translation in J. van Heijenoort (1967), 367–392.
———. 1932. *Gesammelte Abhandlungen*. Vol. 3. Berlin: Springer.
———. 1992. *Natur und mathematisches Erkennen*. Basel, Birkhäuser Verlag, ed. David E. Rowe.
Husserl, E. 1994. *Briefwechsel* IV. Eds. K. & E. Schuhmann. Dordrecht: Kluwer.
Kanamori, A., and B. Dreben. 1997. Hilbert and Set Theory. *Synthese* 110 (1): 77–125.
Kramer, H. 1995. Kandinsky & the Birth of Abstraction. *The New Criterion* 13: 3.
Lusin, N.N. 1929. *Leçons sur les ensembles analytiques*. Paris: Gauthier-Villars.
Mancosu, P., ed. 1998. *From Hilbert to Brouwer: The Debate on the Foundations of Mathematics in the 1920s*. Oxford: Oxford University Press.
Mehrtens, H. 1990. *Moderne – Sprache – Mathematik. Eine Geschichte des Streits um die Grundlagen der Disziplin und des Subjekts formaler Systeme*. Frankfurt: Suhrkamp.

Monk, R. 1990. *Ludwig Wittgenstein: The Duty of Genius*. London: Penguin.
Reid, C. 1972. *Hilbert*. New York: Springer.
Rey Pastor, J. 1951. *La matemática superior: Métodos y problemas del siglo XIX*. Buenos Aires, Iberoamericana (1st ed. Madrid, 1916).
Rowe, D. 1989. Klein, Hilbert, and the Gottingen Mathematical Tradition. *Osiris* 5: 186–213.
———. 2004. Making Mathematics in an Oral Culture: Göttingen in the Era of Klein and Hilbert. *Science in Context* 17: 85–129.
Rowe, D., and V. Felsch. 2019. *Otto Blumenthal: Ausgewählte Briefe und Schriften*. Vol. 2. Berlin: Springer.
Scholz, E., ed. 2001. *Hermann Weyl's Raum – Zeit – Materie, and a General Introduction to His Scientific Work*. Basel: Birkhäuser.
———. ed. 2006. Practice-Related Symbolic Realism in H. Weyl's Mature View of Mathematical Knowledge. In: Ferreirós and Gray (2006), 291–309.
Thurston, W.P. 1994. On Proof and Progress in Mathematics. *Bulletin of the American Mathematical Society* 30 (2): 161–177.
van Atten, M. 2003. *On Brouwer*. Belmont, CA: Wadsworth.
van Dalen, D. 1990. The War of the Frogs and the Mice, or the Crisis of the Mathematische Annalen. *Mathematical Intelligencer* 12 (4): 17–31.
———. 1999. *Mystic, Geometer, and Intuitionist: The life of L. E. J. Brouwer*. Oxford University Press. Vol. I: The Dawning Revolution [Vol. II, 2005: *Hope and Disillusion*.]
van Heijenoort, J., ed. 1967. *From Frege to Gödel*. Harvard: Harvard University Press.
van Stigt, W.P. 1990. *Brouwer's Intuitionism*. Amsterdam: North-Holland.
———. 1996. Introduction to Life, Art, and Mysticism. *Notre Dame Journal of Formal Logic* 37 (3): 381–387.
Verburgt, L.M., Hoppe-Kondrikova, O. 2016. On A. Ya. Khinchin's 'Ideas of intuitionism and the struggle for a subject matter in contemporary mathematics' (1926): a translation with introduction and commentary. *Historia Mathematica*, 43 (4): 369–398.
von Neumann, J. 1928. Die Axiomatisierung der Mengenlehre. *Mathematische Zeitschrift* 27: 669–752.
Weyl, H. 1910. Über die Definitionen der mathematischen Grundbegriffe. *Mathematisch-naturwissenschaftliche Blätter* (Leipzig), 7: 93–95, 109–113. Reprinted in *Gesammelte Abhandlungen.* Berlin: Springer, 1968, 298–304.
———. 1918. *Das Kontinuum*. Leipzig: Veit.
———. 1921. Über die neue Grundlagenkrise in der Mathematik, *Math. Zeitschrift* 10. English version in: Mancosu 1998.
———. 1925. Die heutige Erkenntnislage in der Mathematik, *Symposion* 1. English version in: Mancosu 1998.

Chapter 19
Mathematical Modernism, Goal or Problem? The Opposing Views of Felix Hausdorff and Hermann Weyl

Erhard Scholz

Abstract This chapter contains a case study of the work and self-understanding of two important mathematicians during the rise of modern mathematics: Felix Hausdorff (1868–1942) and Hermann Weyl (1885–1955). The two had strongly diverging positions with regard to basic questions of the methodology and nature of mathematics, reflected in the style and content of their research. Herbert Mehrtens (1990) describes them as protagonists of what he sees as the two opposing camps of "modernists" (Hilbert, Hausdorff et al.) and "countermodernists" (Brouwer, Weyl et al.). There is no doubt that Hausdorff may be described as a mathematical "modernist", while the qualification of Weyl as "countermodern" seems off the mark, once his work is taken into account.

19.1 Introduction

In the history of recent mathematics there is a wide consensus that mathematics underwent a deep transformation in its epistemic structure and its social system during, roughly, the last third of the nineteenth century and the first third of the following one. This led to *modern* mathematics in the sense of the twentieth century. Jeremy Gray called this phase a "modernist transformation of mathematics" (Gray 2008). His book presents a wide panorama of this period of shift in knowledge. The choice of the attribute "modernist" alludes to a wider cultural context of contemporary change in (visual, literary and sound) art and architecture. For good reasons Gray left it open in what way, or even whether, the transformative tendencies in these different branches of culture can be comprehended as different expressions of a common historical phenomenon. This question is still open and discussed in the contributions by L. Corry and J. Ferreirós to this book.

E. Scholz (✉)
School of Mathematics and Natural Sciences, University of Wuppertal, Wuppertal, Germany
e-mail: scholz@math.uni-wuppertal.de

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_19

By this choice of the word, and the question indicated by it, Gray took up a suggestion of Herbert Mehrtens made in *Moderne Sprache Mathematik* (sic! without punctuation) (Mehrtens 1990). For me the title of this book is difficult to translate, because it uses an ambiguity of the German language. It may be translated as "Modern Language Mathematics" or—adding punctuation—"Modernity, Language, Mathematics". Without the punctuation the first alternative would be the correct translation, but one may also understand it in the second way.[1] Mehrtens indicated the huge task of bringing together the historical understanding of the change in the practice of mathematics as a social (and institutional) system (chap. 5) and the knowledge style developed with it (chaps. 1–4). Apparently influenced by considerations from general history and history of art, and most importantly by Paul Forman's provocative thesis on the role of Weimar culture for the rise of quantum mechanics (Forman 1971, 1973), he proposed to highlight the radicality of the modern transformation of mathematics by establishing a narrative of two opposing camps,[2] the protagonists of *modernity* ("Moderne"), the "modernists" driving the modern transformation, and those opposing it, the "countermodernists", representing some (slightly mythical) entity called *countermodernity* ("Gegenmoderne").

For both camps Mehrtens found two main, or at least typical, protagonists. David Hilbert (1862–1943) and Felix Hausdorff (1868–1942) (et al.) for the modern camp versus Luitzen E.J. Brouwer (1881–1966) and Hermann Weyl (1885–1955) (et al.) for the countermodern camp. Mehrtens made the separation of the camps plausible by arguing essentially on the discourse level *about* mathematics, including the debate on foundational issues, with only marginal references to mathematical knowledge (the mathematical discourse itself, to state it in his terminology). The strict separation of the camps did not appear particularly convincing to many readers and was *not* taken up by Jeremy Gray. But in a weakened form it seems to remain a part of the debate on modernism in mathematics. The contributions of L. Corry and J. Ferreirós to the present book discuss this point from different perspectives; they also give good explanations of the terms " cultural modernism" and "modernism in mathematics", which need not be doubled here.

Hausdorff and Weyl, two protagonists of the opposing camps identified by Mehrtens, happen to have been subjects of historical studies of mine for some decades. The present book offers a good occasion for laying down my view of Mehrtens' presentation of these mathematicians as representatives of his opposing camps. The paper presents a simple case study trying to determine the adequacy and usefulness of the proposed categories for our historical understanding of the

[1] The first reading resonates with Mehrtens' way of presenting mathematics as a language, organized in two levels of *discourse*: the discourse *of* mathematics in the production, repectively documentation of knowledge, and the discourse *on* (about) mathematics, a meta-discourse which may include the foundational studies in the sense of Hilbert (Mehrtens 1990, chap. 6). A non-anonymous referee (N. Schappacher) of the present paper opts for the second alternative.

[2] No social, political, or cultural revolution ever happened without having to fight counter-revolutionary forces.

twentieth century.[3] Before drawing the conclusion in the last subsection of the paper, I try to avoid, as far as possible, the qualifications "modernist" or "countermodern", also of "modernism", at least as mathematics is concerned. On the other hand, I use the descriptive attribute "modern" for the deep transformation of mathematics between roughly 1860 and 1940 and the words "rise of modernity" for the social and cultural transformations in the late nineteenth century and the first half of the twentieth century.

The paper consists of three sections. It begins with informations about the life and work of our protagonists, in order to make the paper accessible for readers who are not so well acquainted with their place in of the history of mathematics in the twentieth century. Although this report on their mathematical work has to be extremely short it has major importance for judging how they contributed to and viewed the rise of mathematical modernity. This section ends by analysing their parallel work on Riemann surfaces in 1912. Both mathematicians attacked the problem of making the understanding of Riemann surfaces more precise than was usual at that time. The answers given to this challenge illustrate in a nutshell their individual predilections and different working styles (Sect. 19.2.3). The section following discusses differences between our authors with regard to three questions which were important for both: the understanding of the mathematical continuum (Sect. 19.3.1), the relation between axiomatics, construction and the foundations of mathematics (Sect. 19.3.2), and the question of which role mathematics can, or ought to, play in the wider enterprise of understanding nature (Sect. 19.2.3). The final section (Sect. 19.4) discusses how Hausdorff and Weyl saw themselves in the rise of modern society, before returning to the initial question of modernity and/or countermodernity.

19.2 Hausdorff and Weyl, Two Representatives of Twentieth Century Mathematics

19.2.1 Two Generations, Two Social Backgrounds

The main scientific work of our two authors took place in the first half of the twentieth century, although they belonged to two consecutive generations. Felix Hausdorff (1868–1942) was roughly Hilbert's generation although 7 years younger; Hermann Weyl (1885–1955) was a central figure of the next generation. Both were Germans, with Hausdorff coming from a Jewish-German family. During large parts of their life both worked in German speaking countries. Between 1913 and 1930 Weyl lived in Switzerland and moved back to Germany (Göttingen); but only for a

[3] For a discussion of Brouwer see the contribution of José Ferreirós to this volume.

short while. After the rise to power of the Nazis in 1933 he emigrated to the United States.

Hausdorff obtained his doctorate (1891) and habilitation (1895) at the University of Leipzig with mathematical studies of the refraction and absorption of light as part of the research program of the astronomer H. Bruns. He then turned towards Cantor's theory of sets and transfinite cardinal and ordinal numbers, the most abstract type of mathematics available at the time.[4] Like many Jewish mathematicians of this time he remained lecturer (*Privatdozent*) for a long time before he obtained his first associate professorship (*Extraordinarius*) at Bonn University in 1910. Three years later he accepted a call to Greifswald as full professor. In 1921 he returned to Bonn in the same position and stayed there for the rest of his life. Because of his Jewish origins Hausdorff was in more immediate danger than Weyl after the Nazis came to power, but he hesitated to emigrate in the early 1930s. When he finally tried to leave after 1939 he did not succeed. In 1942 when the anti-Jewish repression of the German Nazi regime reached its climax, he committed suicide with his wife and sister in law, in order to elude deportation and death in the concentration camps (Brieskorn and Purkert 2021; Siegmund-Schultze 2021).

Weyl had a later and quite different start into academic life. He obtained his dissertation (1908) and habilitation (1910) in Göttingen with research in real analysis (singular differential equations) under the guidance of D. Hilbert and the intellectual influence of F. Klein. Swiftly accepted as a promising young researcher he received a call as full professor at the Eidgenössische Technische Hochschule Zürich only 3 years after his habilitation. He was happy to live in a neutral country during the (first) Great War of the twentieth century (World War I). He had to serve in the German army for only about a year from May 1915 onward, before he was released and could go back to Switzerland in 1916. He enjoyed the liberal culture of Switzerland and learned to value it against the torn and crisis stricken German social life during the inter-war years of the twentieth century. In 1930 he hesitatingly accepted a call to Göttingen as successor of David Hilbert.[5] In 1933, after the rise of the Nazi movement to power, he emigrated to the USA following a call to the Institute of Advanced Studies where he was able to support other less privileged emigrants (Siegmund-Schultze 2009). He stayed there until his retirement in 1951 and shuttled between Princeton and Zürich during the last years of his life. Only at rare occasions did he visit post-war Germany.

Both men came from well-to-do families; they grew up in German life and culture of the late nineteenth and early twentieth centuries and shared its humanistic higher

[4] A careful scientific and intellectual biography of Hausdorff can be found in (Brieskorn and Purkert 2021) and in (Hausdorff 2002–2021, vol. 1B); the turn towards Cantorian transfinite numbers is described in great detail in chap. 7 of (Brieskorn and Purkert 2021).

[5] For Weyl there is nothing close to a full biography like the one for Hausdorff, written by Brieskorn and Purkert. An intellectual biography of his early years (until 1927) can be found in (Sigurdsson 1991) and (Rowe 2002), a documentation of his time in Zürich in (Frei and Stammbach 1992). For more extensive scientific biographies see (Coleman and Korté 2001; Chevalley and Weil 1957; Newman 1957; Atiyah 2002; Dieudonné 1976).

school education. Hausdorff's father was a successful textile merchant and owner of a small publishing house in Leipzig. As a traditional Jew he participated in the community on a national level and was active against the rising anti-semitism in late nineteenth century Germany. He stood in opposition to the Jewish reformers and contributed to the scholarly orthodox Talmud discussion (Brieskorn and Purkert 2021, chap. 1). Weyl's father was a director of a local bank and city councillor of Elmshorn, a medium sized town in Northern Germany. Already during his school time Weyl grew up deeply immersed in German philosophical thinking by reading Kant's *Critique of Pure Reason* at his parental home.[6] Although his initial enthusiasm for a naive version of Kantianism broke down in the early years of his mathematical studies at Göttingen after encountering Hilbert's axiomatic approach to geometry,[7] he remained attracted by German idealistic philosophy in different guises, in particular Husserlian phenomenology and Fichtean constructive idealism (Ryckman 2005; Sieroka 2010, 2019).

The young Hausdorff passed through a different intellectual development. In sharp opposition to his father's orthodox Judaism, he too was attracted by Kant's critical philosophy, though we was more fascinated by Schopenhauer's pessimistic philosophy of life and the young Nietzsche's radical cultural thoughts. In his later student's years he joined a circle of modernist intellectuals at Leipzig and became active as a freelance writer, essayist and philosopher under the pseudonym Paul Mongré. Unlike Weyl, he considered the liberation from any metaphysical bonds as a desirable goal of late nineteenth century thought (Brieskorn and Purkert 2021, chaps. 5, 6). These differences in general intellectual outlook between our protagonists would turn out to play a major role when it came to their predilections in mathematics and the way they reflected on their scientific work.

19.2.2 A Short Glance at Our Authors' Main Contributions to Mathematics

Before we discuss the attitudes of Weyl and Hausdorff on methods, role and goals of mathematical research we ought to recollect their main achievements in mathematics. We deal here with two "giants" of science and have to restrict to a selective survey on the most important topics of their scientific achievements. For

[6] Weyl (1955a, p. 632f.).

[7] The impression Hilbert's approach to he foundations of geometry made on the young student Weyl at Göttingen was described by himself in hindsight: Hilbert examines the independence of the axioms of geometry "…not only by drawing on the so-called non-Euclidena geometry, then nearly a century old, but by constructing, mostly on an arithmetical basis, a plethora of other strange geometries. Kant's bondage to Euclidean geometry now appeared naive. Under the overwhelming blow, the structure of Kantian philosophy, on which I had hung with faithful heart, crumbled into ruins" (Weyl 1955a, English, p. 206). Weyl was much more critical with respect to Hilbert's later extension of the axiomatic method to the natural sciences, beautifully described in Corry (2004).

more extended report on F. Hausdorff see (Brieskorn and Purkert 2021), for Weyl (Coleman and Korté 2001; Chevalley and Weil 1957; Atiyah 2002). Readers who are not so well acquainted with these topics may prefer to skip the following passages and jump directly to Sect. 19.2.3.

Felix Hausdorff is well known for his axiomatization of the concept of topological space in his opus magnum *Grundzüge der Mengenlehre* (Main Features of Set Theory) (Hausdorff 1914b). But this book presented far more than this. Its first 6 chapters contained the most comprehensive introduction to Cantorian set theory in the first decades of the twentieth century. Among other things, it included a detailed study of transfinite order structures to which Hausdorff had contributed himself in the preceding years (including the study of η_α sets which later became important in foundational studies of set theory). The second half of the book established a program for founding axiomatically basic fields of mathematics using the framework of set theory. As we now understand, this was a main trend of the modernization of mathematics in the first half of the twentieth century culminating in the work of "modern algebra" and Bourbaki's vision of mathematics around the middle of the century. Hausdorff himself exemplified the method for topological spaces (chap. 7), metrical spaces (chap. 8), functions (chap. 9), measure theory and integration (chap. 10). In these chapters he could in particular draw the consequences of developments of the first abstract (topological) space concepts about the turn of the centuries, due to M. Fréchet, F. Riesz, E.H. Moore and others.[8]

At the end of the book Hausdorff published a paradoxical disjoint decomposition of the 2-sphere, using the axiom of choice, and republished the argument separately in (Hausdorff 1914a).[9] According to his own description it showed that in Euclidean space "one third of the sphere" may be congruent to "one half" of it (ibid, p. 430). Hausdorff was a master of logical precise argumentation without using a formal system for logic itself. He was fond of counter-intuitive effects in the world of transfinite set theory. Some years later he generalized a measure theoretic approach initiated by C. Caratheodory and introduced a class of measures on subsets of metric spaces (*Hausdorff measures*) which enables to characterize the dimension p of point sets, where p may assume *fractional* values (Hausdorff 1919). The concept developed in this small paper on measure and dimension has been of enormous influence in mathematics (non-linear partial differential equations, dynamical systems, ergodic theory) and physics (potential theory, turbulent flows, meteorology) and has become widely known through with the rise of fractals and computer graphics at the end of the twentieth century.

[8] See Brieskorn and Purkert (2021, p. 353ff.); a more extended discussion of the early development of topological space concepts is given in the commentary (in German) on the historical background of Hausdorff's axioms in Hausdorff (2002–2021, vol. 2, pp. 675–708).

[9] Hausdorff was well aware that Zermelo's axiom of choice was an important contribution to the clarification of the foundations of transfinite set theory, but considered this as a step only in a larger enterprise which had not come to an end, and to which the study of transfinite numbers ought still to contribute; see Brieskorn and Purkert (2021, chap. 7.3) and the remark at the end of their sec. 7.2.1.

In his investigations of set theory he introduced important fundamental concepts like co-finality and co-initiality of ordered sets, or (Hausdorff) gaps in dense ordered structures. After establishing the distinction of regular and singular initial (cardinal) numbers he observed that regular cardinal numbers with limit index, should they exist at all, would be of "exorbitant" size (Plotkin 2005, 233). Later this became the starting point for the study of so-called "large cardinal numbers". Among his diverse contributions to set theory we also find the Hausdorff maximality principle of partially ordered sets, the general recursion formula of aleph-exponentiation, and the concept of so-called η_α sets which later became important in model theory.[10]

Other contributions of his relate to different fields of mathematics of the twentieth century, e.g., the Baker-Campbell-Hausdorff formula in the theory of Lie groups, or the Hausdorff-distance of compact subsets of a metrical space, which was later used by M. Gromov to measure the "distance" of metrical spaces from being isometric. The resulting Gromov-Hausdorff distance of metrical spaces became an important tool for differential topology (Hausdorff 2002–2021, vol. 1B, p. 779). In Hausdorff's lecture manuscripts we find many interesting insights, e.g. with regard to an axiomatic foundation of probability theory (Hausdorff 2002–2021, vol. *5*, 595–723) similar to the one in Kolmogorov's famous book of 1933, but 10 years earlier.

The mentioned topics show already the profile of Hausdorff's contribution to mathematics: set theory as the basis for work and as a framework of modern mathematics, order structures, point set topology, metric spaces, measure theory with particular attention to paradoxical or seemingly paradoxical (fractional dimension) results, and functional analysis. After his turn towards pure mathematics in the late 1890s, mentioned in the last section, the contributions to applied mathematics of his early phase were no longer of interest to him. On the other hand, his interests in pure mathematics show traces of his epistemological reflections around 1900 in which he assigned mathematical arguments an important role for the critique and decomposition of classical metaphysics (Epple 2021, sec. 5f.). Hausdorff's interest and active participation in philosophical reflection of mathematics faded away after his turn towards pure mathematics research. As Purkert/Brieskorn write, his alter ego Paul Mongré "bid farewell to the public" about 1910 (Brieskorn and Purkert 2021, p. 318).

Hermann Weyl, on the other side, was acknowledged as a leading figure of the post-Hilbert generation of mathematicians already during his life time. His research was as broad as the one of his academic teacher Hilbert; it comprised many fields inside mathematics and its foundations as well as long lasting contributions to mathematical physics. He was widely read in philosophy and did not hesitate to share his philosophical reflections on mathematics and science with the interested public. His most influential work in mathematics proper results from his studies in Lie theory starting in the mid-1920s (Weyl 1926). He combined E. Cartan's characterization of infinitesimal groups (Lie algebras) and their representations

[10] This list is a a selection of the survey of Hausdorff's main contributions to set theory in (Brieskorn and Purkert 2021, p. x). For more details see there.

with an integral approach used by I. Schur to the characters of certain groups (the special orthogonal ones). Weyl was able to generalize Schur's method to all the classical groups and to study their representations (Hawkins 2000; Eckes 2011). In his Princeton years he extended this approach, in cooperation with R. Brauer, to give a modern access to the invariants of the classical groups (Weyl 1939).

Although lying deep inside mathematics proper (i.e., "pure" mathematics), for Weyl this research topic was multiply intertwined with questions coming from theoretical physics and leading back to the latter. He had started to develop interest in infinitesimal symmetries in his thoughts about general relativity and generalized Riemannian geometry by introducing what he called a "length" gauge (today scale gauge). This led him to propose a geometrically unified field theory of gravity and electromagnetism in the framework of the first gauge theory of electrodynamics with local symmetries of geometric scale as the *gauge group* (Weyl 1918b,d).[11] In this context he contributed importantly to clarifying conceptual and mathematical questions in general relativity (Weyl 1918c) and differential geometry (Scholz 1999, 1995).

His proposal of a scale gauge theory of electromagnetism did not work out directly as a physical theory but could be "recycled" after the advent of the new quantum mechanics in form of a gauge theory for the phase of wave functions of charged particles (Vizgin 1994; Scholz 2004). Through the intermediation of W. Pauli (1933), (Pauli 1941), Weyl's idea of a gauge field approach to electromagnetism was generalized by C.N. Yang and L. Mills in 1954 to a more general gauge group of isotopic spin $SU(2)$ (O'Raifeartaigh and Straumann 2000). After a long interlude of laborious research in high energy physics it acquired a central role in the standard model of elementary particle physics in the 1970s (O'Raifeartaigh 1997). About the same time it entered also the research of differential topology and was used for defining new topological invariants (Kreck 1986). On a different, although connected route Weyl started to use group representation theory in the new quantum mechanics after 1926 (Weyl 1928). Together with E. Wigner he may be considered as a main actor for propagating symmetry considerations in the study of quantum systems which again became a main tool for particle physics in the second half of the twentieth century (Borrelli 2017, 2015).

A third field (aside from the representation theory of Lie groups and theoretical physics) in which Weyl intervened with long lasting consequences was the debate on the foundations of mathematics in the first third of the twentieth century. In spite of his high regard for Hilbert as a mathematician he was not at all convinced philosophically by his teachers proposal for a formalistic solution of the problems arising in transfinite set theory after 1905 and the consequences for analysis,

[11] In Jed Buchwald's contribution to this volume "gauging" is discussed in the pre-Weylian perspective of under-determination of the electromagnetic potential (decomposed in its scalar and its vector part) up to exact differentials as a history of "gauge" *ante letteram*. We learn from it that important physical questions of this under-determination have been posed and answered long before the explicit concept of *gauge* was introduced; for the later development see, among others, (O'Raifeartaigh 1997; O'Raifeartaigh and Straumann 2000).

arithmetic and mathematics in general. He did not share Hilbert's conviction that the consistency of a system of mathematical concepts and their postulated relations could justify it as mathematically valid theory (Hilbert spoke even of the "existence" of the conceptual objects, once their consistency was proved). He rather insisted on some symbolic/conceptual generation of the objects starting from the basic knowledge of natural numbers, perhaps added by elementary intuitions of continuity.

While the first problem, a constructive foundation of natural numbers, seemed relatively uncontroversial,[12] the second problem, a genetic foundation for the concept of the continuum, remained a lifelong problem for Weyl, and not only for him. Already in 1918 he sketched an alternative for the foundation of analysis, based on a reduced concept of predicatively defined real numbers (Weyl 1918a); but it remained unsatisfactory even for himself. He then fought for some years on the side of Brouwer for an intuitionistic program in the foundations of mathematics attacking Hilbert harshly (Weyl 1921). In the later 1920s he made attempts at founding the concept of an n-dimensional continuum by the methods of combinatorial topology. But also this approach remained open ended. This may have contributed to rethinking his earlier harsh critique of Hilbert's formalist program; in later writings he came to a more balanced view of Hilbert's foundational program (see below, Sect. 19.3.2).

Just as with the presentation of Hausdorff's work this survey is necessarily extremely selective: other fields of Weyl's work, e.g., in convex geometry, real and complex analysis have fallen completely through the cracks. The next subsections gives the chance for partially correcting this at least with regard to complex analysis.

19.2.3 Contrasting Trajectories: Riemann Surfaces as an Example

By an odd historical coincidence, Hausdorff and Weyl lectured on complex function theory at the same time without knowing each other's work. In the winter semester 1911/1912, Hausdorff gave an introduction to function theory at Bonn university, whereas Weyl lectured on Riemann's theory of Abelian integrals in Göttingen. Both had to struggle with the concept of a Riemann surface, which at that time was still only vaguely defined, and both made proposals about how to attack this question, with long range repercussions. Weyl's notes became a draft for his book on the idea of Riemann surface ("Die Idee der Riemannschen Fläche") published in the following year (Weyl 1913). This book is widely known for presenting the first definition of a manifold at least for the 2-dimensional case. For Hausdorff the lecture

[12] Hilbert hoped to be able to build a consistency proof for large parts of mathematics including full Peano arithmetic and real analysis on it between his talk at the International Congress of Mathematicians in 1904 until his lecture course on logic in 1920 (Sieg 2000).

led him to think about neighbourhood systems which turned into his axiomatics of topological spaces 2 years later.

Weyl drew upon Hilbert's sketch of an axiomatic characterization of the (real) plane, based on topological concepts (Hilbert 1903b).[13] Hilbert defined a *plane* as a "system of things" (set), with elements ("things") called *points*, which could be bijectively mapped as a whole on the "number plane". He then used Jordan domains of the latter for characterizing *neighbourhoods* ("Umgebungen") of points in the plane. Weyl could take up this idea but had to modify it. For Riemann surfaces, thought to arise semi-constructively from analytic function elements ("analytische Gebilde") in the sense of Weierstrass, he had to localize Hilbert's idea and could no longer presuppose a global bijection with the number plane. This led to the first definition of a manifold $\mathfrak{F}$ in Riemann's sense, although restricted to the 2-dimensional case, by establishing an axiomatics of neighbourhood systems in $\mathfrak{F}$, with bijective maps to open disks in the Euclidean plane (Weyl 1913, p. 17f.). This sufficed for defining continuity, differentiability and even analyticity of maps between such manifolds and of functions and to build Riemann's theory of Abelian differentials based on such a fundament. In particular the topological notions of triangulation, simple connectedness, covering surfaces, group of covering transformations and the topological genus of the surfaces etc. were thus put on an essentially clarified mathematical basis, if one kept the constructive context (analytic function elements and disks as coordinate images) in mind.

A (later) analysis from the more refined point of view of Hausdorff's topology would show that Weyl's axioms were not strong enough to provide a self-sufficient formal axiomatization: the Hausdorff separation property was not secured by his axioms although implicitly presupposed in the derivations. But this was not Weyl's concern during the next few decades. Only when he prepared the English translation from lectures given in 1954 at Harvard and Princeton, and in the third German edition did Weyl add Hausdorff separation axiom (Remmert 1997, p. xii).

The idea to talk about *neigbourhoods* of points in Weierstrassian function theory and even for characterizing more general spaces based on set theory was not an exclusive privilege of the Göttingen mathematicians. Weierstrass had used the terminology already, and also F. Riesz used it for generalized spaces (Riesz 1908; Rodriguez 2006). Hausdorff had earlier lectured on Cantorian set theory in Leipzig in summer semester 1901 and again in Bonn in 1910, but without taking up this idea in his discussions of topological aspects in general sets. In his lectures on function theory he was confronted with neighbourhoods of function elements in a natural way. In his lecture notes of winter semester 1911/12 we find clear evidence that he realized at this point that the study of neighbourhood systems was the clue for "ordering the system of points" which arose from the study of equivalence classes of analytic function elements, and also more generally. Moreover he became aware that the structural properties of such neighbourhood systems had to be analysed. This he did more extensively during the summer semester 1912, in which he gave

[13] Also in Hilbert (1903a, appendix IV).

his next course on set theory at Bonn. Here we find four structural properties of systems of neighbourhoods in metrical spaces which were essentially the axioms for a topological space, published in his book (Hausdorff 1914b). Moreover he already announced that these structural properties could be used as axioms for general spaces (Epple et al. 2002, p. 714ff.).

This small episode seems characteristic for the different thought styles for Hausdorff and Weyl. The former used the analysis of conceptual features of Weierstrassian function elements underlying the concept of Riemann surface as a stepping stone for a general characterization of topological spaces in the framework of sets. Weyl, on the other hand, took the same as a starting point to establish an axiomatic clarification of the intuitive concept of Riemann surface which had been in use already for several decades. He kept this framework closely linked to the construction of global objects from Weierstrassian function elements, aiming at concrete mathematical objects with multiple structures. This difference may seem a nuance of research orientation only; but we will see that it is characteristic for their contrasting positions with regard to the aim and character of mathematics. During the following years it would develop into an open opposition.

19.3 Mathematics in the Tension Between Formal Thought and Insight

19.3.1 Two Opposing Views of the Continuum: A Modified Classical Concept Versus a Set-Theoretic Perspective

Already in the period 1910–1914 Hausdorff and Weyl had developed quite different ideas about how to deal with the mathematical concept of *continuum*. As we know already Hausdorff was attracted by the epistemic perspective opened up by Cantor's transfinite set theory once he got to know of it. Weyl, by contrast, became increasingly sceptical with regard to any truth claim for the latter after discussing foundational problems in his Habilitation lecture (1910) (see below). W. Purkert was able to reconstruct from indirect evidence (remarks on the infinite in philosophical essays written under the pseudonym Mongré) that Hausdorff got to know Cantor's theory during the year 1897, the year of the First International Congress of Mathematicians at Zürich (Brieskorn and Purkert 2021, p. 262ff.). Hausdorff/Mongré was fascinated by the intellectual perspective of Cantor's treatment of the transfinite cardinal and ordinal numbers (although not yet clarified in sufficient detail, not to speak of its axiomatization). At this time he pursued a philosophical program in the footsteps of Kant, radicalized by Nietzsche, for "proving", more precisely by arguing with the use of mathematical metaphors, that no knowledge of the "thing in itself" is possible, and in particular no insight into the structure of "absolute time" or "absolute space" or even "cause" is possible (Stegmaier 2002; Epple 2021; Mongré 1898a, 1899). Wearing the hat of Mongré, he tried to convince his readers by an “apagogic proof” (a proof by contradiction) that absolute time or space, if assumed, cannot have any type of structure. For this goal Cantorian set theory seemed to him an ideal tool. Relative structures, i.e., not completely absolute ones, were of course possible also for him, i.e., order structures in the case of time and geometrical or even topological structures (before the advent of the word) in the case of space. According to Mongré/Hausdorff such non-absolute structures were “selected” by the mind, to make human action and survival possible; they could then form a rather individualistic “cosmos”. The individualistic exaggerations found in the early Mongré’s literary and philosophical utterances were, however, step by step moderated by what Hausdorff later called a *considerate empiricism*, which respected empirically founded scientific knowledge, including theoretical refinement and critique (Epple 2006).

As M. Epple and other authors have argued, Hausdorff’s mathematical research topics between 1900 and 1914 were still embossed by his interest in logically consistent, although intuitively surprising, perhaps even paradoxical insights into order structures ($\sim$ time) and/or topological, metrical and measure structures ($\sim$ space) (Epple 2006, sec. 5.6). Hausdorff’s first important works in transfinite set theory consisted in profound and technically demanding contributions to order structures (Hausdorff 1901). W. Purkert observed that an additional motivation for this work seems to have come from his interest in approaching a proof of Cantor’s famous *continuum hypothesis*; i.e. the assumption (at first a claim of Cantor) that the cardinal number of the subsets of the natural numbers $2^{\aleph_0} = \mathfrak{c}$ (which encodes the cardinal number of the “continuum”, i.e. the real numbers) is the first non-denumerable cardinal number

$$2^{\aleph_0} = \aleph_1$$

or, even more generally, $2^{\aleph_\nu} = \aleph_{\nu+1}$.

In Hausdorff’s view, the “continuum” itself would have to be understood by using all kinds of different types of topological and/or measure structures. Intuitive insight into the nature of the continuum seemed him of ephemeral value only, perhaps important for the imagination of the individual mathematician, but without any epistemic value with regard to truth claims. His great book *Grundzüge der Mengenlehre* was a splendid exemplification of this general view.

Weyl had a completely different view of the continuum, which was deeply influenced by the long tradition in mathematical and philosophical thought. Riemann’s concept of *manifold* appeared to him as the most promising modern clue to the topic. Its logical and formal foundations remained an open question for him until the end of his life, although he himself made at least three attempts to come to grips with it (Scholz 2000): a constructive approach in Weyl (1918a), influenced by E. Borel and H. Poincaré, which was designed to avoid the pitfall of impredicative definitions;[14] an intuitionistic one in Weyl (1921); and a combinatorial topological one (Weyl 1924) or in his lecture course on *Axiomatics* in Göttingen 1930/1931 (Weyl 1930/1931, §37).

In a paper written for the Lobachevsky anniversary in 1925, though published only posthumously, we find an explicit remark explaining why Weyl would not agree with Cantor's or Hausdorff's approach to the continuum, at least understood in the sense of a manifold describing physical space. In such a manifold the local descriptions by coordinates in a "number space" are "arbitrarily projected into the world" and everything else, in particular the metric structure is turned into a field on the space. This could well be reflected in a Kantian type of approach which shaped his understanding of the role of the spacetime concept in general relativity. A few years earlier he had contributed to a deeper understanding of the underlying concepts by posing the problem of space anew, in response to the changed situation after the rise of special and general relativity (Scholz 2016; Bernard and Lobo 2019). In the mid-1920s Weyl resumed a (relativized) Kantian perspective and sharpened his criticism of a set-theoretic substitute for it in the following way:

> Space thus emerges [by separating the topological manifold from the metrical and other fields on it, ES] even more clearly as the form of appearances in contrast to its real content: the content is measured once the form has been referred to coordinates. [Set theory, one may say, goes even further; it reduces the mf [manifold] to a set as such and considers already the continuous connection as a field on the latter. It should, however, be clear that in doing so it violates the essence of the continuum which by its nature cannot be smashed into a set of isolated elements. The analysis of the continuum should not be founded on the relation between element and set, but on the one between part and the whole. ...] (Weyl 1925/1988, p. 4f., second square brackets in original; translation ES)[15]

In other words, Weyl considered the notion of transfinite sets as an overstretched formal concept without substantial content, at least insofar as physical spacetime was concerned.[16] Below we shall see that his scepticism regarding the continuum as a concept of mathematical physics also pertained to its role in mathematical analysis and in the foundations of mathematics.

[14] For an appraisal of its mathematical long range import see Feferman (2000); a critical historical view is given in Schappacher (2010).

[15] "Deutlicher tritt dadurch der Raum als Form der Erscheinungen seinem realen Inhalt gegenüber: der Inhalt wird gemessen, nachdem die Form willkürlich auf Koordinaten bezogen ist. [Die Mengenlehre, kann man sagen, geht darin noch weiter; sie reduziert die Mf auf eine Menge schlechthin und betrachtet auch den stetigen Zusammenhang schon als ein in ihr bestehendes Feld. Es ist aber wohl sicher, daß sie dadurch gegen das Wesen des Kontinuums verstößt, als welches seiner Natur nach gar nicht in eine Menge einzelner Elemente zerschlagen werden kann. Nicht das Verhältnis von Element zur Menge, sondern dasjenige des Teiles zum Ganzen sollte der Analyse des Kontinuums zugrunde gelegt werden. Wir kommen darauf sogleich zurück.]" (Weyl 1925/1988, §37).

[16] Ferreirós (2016) calls this a "pointillist" view of the continuum, see in particular the discussion in chap. 10.4.

19.3.2 Axiomatics, Construction and the Problem of the Foundations of Mathematics

Weyl understood axiomatics as the defining basis of a conceptual framework on which a mathematical theory could be built. So far this was similar to Hilbert, but unlike the latter Weyl was not particularly convinced of the usefulness of trying to protect its internal consistency by a metatheoretical investigation, in particular if the metatheory was based on more or less strong transfinite methods the epistemic meaningfulness of which he doubted anyway. He pleaded for safeguarding the consistency of a mathematical theory in a different way, so to speak "on the job", constructing its object fields and clarifying their axiomatic principles in an integrated approach.[17] The task of an axiomatic formulation was to clarify the structure of some field of mathematical thought; its objects were to be constructed and dealt with symbolically. This should not depend on overly strong hypotheses about the infinite, in particular without making use of transfinite set theory and only if unavoidable with applying the principle of the excluded third without a constructive underpinning. Weyl's axiomatization of 2-dimensional manifolds and Riemann surfaces was an early example. And he stuck to this conception essentially all his life (i.e., with gradual modifications only). In the late 1930s he came into contact with members of the early Bourbaki group, in particular C. Chevalley, after which he began to develop more respect for how algebraists use axiomatics as a research tool in its own right, although still not with respect to foundational issues (which were not in the focus of Bourbaki anyhow).

Weyl was extremely sceptical with regard to Hilbert's program of founding mathematics by axiomatization and a formal analysis of proof structures with the aim of showing internal consistency. He considered such a justification of mathematical theories or even of mathematics as a whole as nothing but a formal showpiece. This might be impressive because of its acumen, but it would fail to meet the goal of justifying the substance of mathematics. In Weyl's view a meaningful justification would presuppose a clarification of the basic conceptual ingredients of a mathematical theory by *symbolic construction*, as he called it (Weyl 1927, 1949). A preliminary version of how a constructive approach to analysis might work was given in his famous book *Das Kontinuum* (Weyl 1918a); for its long-range impact see (Feferman 2000). But Weyl was discontent with his own achievements, in part because it justified only a restricted variant of analysis (without the general existence of a supremum of a bounded set of the reals). After he had constructed his reduced (denumerable) range of real numbers, he opened the discussion of the relation to geometry with a self-critical remark, deploring that the intuition of connectivity inherent in the geometrical concept *continuum* was not depicted in his constructive number continuum:

[17] This is nicely demonstrated in his lecture course on *Axiomatic* in the winter semester at Göttingen (Weyl 1930/1931).

> Once we have torn the continuum apart into isolated points, it is difficult to reconstruct ex post the connectivity between the single points, which is based on their non-independence, by some conceptual equivalent. (Weyl 1918a, 79, translation ES)[18]

So his constructive (denumerable) continuum of 1918 offended against the "essence" of the continuum at least as much as did a Hausdorffian set-theoretic approach (criticized in the quotation at the end the last subsection). Its only advantage was its (semi-finitist) constructive methodology rather than one in which transfinite sets were postulated axiomatically. Irrevocable connectivity between points by their inseparable infinitesimal neighbourhoods was what Weyl looked for. For a while he believed one could find this in the intuitionist approach proclaimed by L.E.J. Brouwer more or less at the same time (Brouwer 1919). So Weyl's attempts at laying the cornerstones of a constructive clarification for analysis shifted for some years (between 1919 and 1923) towards a strong support for Brouwer's more radical intuitionistic program, most decidedly expressed in his open polemics of (Weyl 1921). This most radical phase of his contributions to the foundations of mathematics has attracted much attention in the history and philosophy of mathematics (Rowe 2002, 2024; Hesseling 2003; Scholz 2000), (Mehrtens 1990, sec. 4.1), and with more technical details (Coleman and Korté 2001, sec. 6).

By the mid-1920s, however, he started to accept that Hilbert's formalist program was, after all, a defensible position. He remained sceptical, however, with regard to the epistemic value of such a formalist approach to the foundations of mathematics and to the concept of continuum, because it did not live up to his (undefined and probably undefinable) criteria of "insight" and "meaning". On different occasions he sketched how he would conceive of a constructive symbolic approach to the continuum, based on methods taken from combinatorial topology. He thus explored how far a symbolic representation of cell complexes with (denumerably) infinite sequences of barycentric subdivisions would carry (Weyl 1924, 1922, 1930/1931, 1940, 1985); but a purely combinatorial constructive characterization of topological manifolds, which he considered as the best mathematical approach to the "continuum", remained an unsolved problem.

By the middle of the century he accepted and appreciated the meanwhile widely spread axiomatic approach in mathematics:

> . . . the axiomatic attitude has ceased to be the pet subject of the methodologists [researchers in formal logic and foundations of mathematics, ES] its influence has spread from the roots to all branches of the mathematical tree (Weyl 1940).

But it remained important for him that axiomatic postulates were not dissolved from "symbolic construction". He was not satisfied with taking finitist methodology serious *only* at the level of metatheorical investigation (as Hilbert had proposed in his proof theoretic program for showing the consistency of axiomatic theories). He

[18] "Nachdem wir das Kontinuum in isolierte Punkte zerrissen haben, fällt es jetzt schwer, den auf der Unselbständigkeit der einzelnen Punkte beruhenden Zusammenhang nachträglich durch ein begriffliches Äquivalent wieder herzustellen" (Weyl 1918a, 79). The translation in (Weyl 1987, 103f.) suppresses the details "ex post" and the proxy character of the "conceptual equivalent".

demanded that on *all levels* of knowledge production and reflection mathematics ought to be a "...dexterous blending of constructive and axiomatic procedures" (Weyl 1985, 38). The foundations of mathematics, on the other hand, remained an open problem for him until the very end of his life.

In one respect Hausdorff's view of axiomatics was not very different from Weyl's (or most other twentieth century mathematicians): both saw axiomatic systems as providing the format for defining the basic concepts of a mathematical theory. But in other respect they differed drastically. Hausdorff was an excellent logically sharp thinker who did not see a need for formalizing logic, as we noticed already. He could never imaging giving up a principle like the law of the excluded middle. For Hausdorff this would unnecessarily reduce the range of mathematics, which was completely unacceptable to him. Symbolically supported *creation* combined with logical precision took the place occupied by symbolic *construction* for Weyl in the generation of mathematical knowledge. Hausdorff thus took up transfinite set theory as outlined by Dedekind and Cantor, a program adopted by Hilbert and his school, which established a language and thought milieu for symbolic creativity par excellence. He knew, of course, about the open questions in the foundations of set theory, in particular that the comprehension of infinite totalities had to be handled with care, but he saw no risk in building mathematical theories along these lines. Still in the late 1920, when he published the so-called second edition of his book on set theory (in fact, a new book) he emphasized the aspect of creativity in his lucid rhetoric:

> It is the eternal achievement of Georg Cantor to have dared this step into infinity, under interior and exterior struggles against seeming paradoxies, popular prejudices, philosophical statements of power (infinitum actu non datur), but also against reservations pronounced by the greatest mathematicians. By this he has become the creator of a new science, set theory, which today forms the grounding of the whole of mathematics. In our opinion, this triumph of Cantorian ideas is not altered by the fact that a certain antinomy arising from an excessively boundless freedom of forming sets still needs a complete elucidation and elimination. (Hausdorff 1927, 11, Werke *3*, 55)[19]

He knew that in the environment of Hilbert (Zermelo, Fraenkel, Bernays) the axiomatization of set theory was under way, but he saw no pressure to proceed along these lines, and was far from feeling any "anxiety" that something would go wrong with the foundations set theory and mathematics (Brieskorn and Purkert 2021, sec. 7.3).

[19] "Es ist das unsterbliche Verdienst Georg Cantors, diesen Schritt in die Unendlichkeit gewagt zu haben, unter inneren wie äußeren Kämpfen gegen scheinbare Paradoxien, populäre Vorurteile, philosophische Machtsprüche (infinitum actu non datur), aber auch gegen Bedenken, die selbst von den größten Mathematikern ausgesprochen waren. Er ist dadurch der Schöpfer einer neuen Wissenschaft, der Mengenlehre geworden, die heute das Fundament der gesamten Mathematik bildet. An diesem Triumph der Cantorschen Ideen ändert es nach unserer Ansicht nichts, daß noch eine bei allzu uferloser Freiheit der Mengenbildung auftretende Antinomie der vollständigen Aufklärung und Beseitigung bedarf." (Hausdorff 1927, 11, Werke *3*, 55).

From such a viewpoint a methodology which demanded a reduction of symbolic creation to procedures that would deliver only denumerable ranges of objects (Weyl's constructivism or Brouwer's intuitionism) appeared to him ridiculous, or even worse. He did not state such an opinion publicly, but made his views perfectly clear in a letter to Abraham Fraenkel written June 9, 1924, in response to Fraenkel's refinement of Zermelo's system of axioms (Fraenkel 1923). He thanked for this and for the discussion of the set theoretic antinomies, because this spared him from dealing with matters of less personal concern for him ("Dinge, die mir nicht liegen"). From now on he would be able just to refer to Fraenkel's book. He continued:

> You have even succeeded in making the oracle pronouncements of Brouwer and Weyl understandable – without making them appear to me any less nonsensical! You and Hilbert both treat intuitionism with too much respect; one must for once bring out heavier weapons against the senseless destructive anger of these mathematical Bolsheviks! (Hausdorff 2002–2021, vol. *9*, 293, translation (Rowe 2024))[20]

Hausdorff's surprisingly militant language has to be understood against the background of Weyl's polemical language in his paper propagating an intuitionistic "revolution"—and the turbulent conditions in post-war Germany during the early 1920s. It indicates a deep dividing line among early twentieth century mathematicians (in Germany) with regard to basic methodological convictions and the value of certain research programs. But can this dividing line be better understood by declaring our two protagonists as belonging to two separate camps of *modernists* (Hausdorff) and *counter-* or even *antimodernists* (Weyl) as proposed by (Mehrtens 1990)?—David Rowe calls Weyl, on the contrary, a "reluctant revolutionary" (Rowe 2002). This seems to me much more to the point; we will come back to this question in the final discussion.

19.3.3 Mathematics and the Material World

Although Hausdorff no longer contributed actively to the natural sciences after his disappointing experiences with his early works in astronomical optics, he held strong views on the usefulness of mathematics for acquiring knowledge about the physical world (Brieskorn and Purkert 2021, chap. 3). These views, however, were not widely known at the time. His contributions to probability theory remained relatively unnoticed (Brieskorn and Purkert 2021, sec. 4.2, 10.3); the lecture containing a set theoretic axiomatization of probability remained unpublished (Hausdorff 1923/2006).

[20] "Es ist Ihnen sogar geglückt, die Orakelsprüche der Herren Brouwer und Weyl verständlich zu machen – ohne dass sie mir nun weniger unsinnig ercheinen! Sowohl Sie als auch Hilbert behandeln den Intuitionismus zu achtungsvoll; man müsste gegen die sinnlose Zerstörungwuth dieser mathematischen Bolschewisten einmal gröberes Geschütz auffahren! . . . " (Hausdorff 2002–2021, vol. *9*, 293).

In the 1890s and early 1900s he was highly interested in the question of non-Euclidean geometry and in philosophico-mathematical question of space and time concepts (Epple 2021), (Brieskorn and Purkert 2021, sec. 5.6). In his radical thoughts on philosophical (epistemological and ontological) questions, published under the name Paul Mongré, mathematics played an important role for undermining the belief in fixed, perhaps even a priori, forms of knowledge of the external (material) world. The great variety of geometrical or, in nuce, even topological structures for spacelike thinking, and of order structures for timelike thinking became an important tool for him in putting established notions of mathematical physics, astronomy and cosmology into question. On the other hand, he made clear that the ordering of sense perceptions and scientific empirical knowledge needed mathematics for acquiring a well defined and intelligible form. He called such a methodology *considered empiricism* ("besonnener Empirismus"), in contrast to empiricism sans phrase and positivism on the one hand and neo-Kantianism, or any other rejuvenated version of German idealism, on the other (Epple 2006).

In later years Hausdorff did not completely lose interest in mathematical physics but this interest clearly moved toward the background of his attention. We know that he prepared talks, perhaps even an introductory publication for a wider public, on (special) relativity (Epple 2021, sec. 5.1), but he never took up questions from mathematical physics for his own research. From the beginning of the twentieth century onward his research profile became that of a "pure" mathematician who appreciated the role of mathematics for an open minded and critical understanding of the material world. In his early years he had formulated an aphorism expressing what he considered the role of mathematics in natural science:

> What we are missing is a self-critique of science; the verdicts of science given by art, religion and sentiment are just as numerous as useless. Perhaps this is the ultimate task [or even destination, ES] of *mathematics* (Mongré 1897, aphorism 401, transl. ES).[21]

Weyl, on the other hand, was a highly creative contributor to mathematical and theoretical physics, besides his great achievements in pure mathematics. As is well known, he made outstanding contributions to Einstein's theory of gravity and early cosmology (Giulini forthcoming; Goenner 2001; Rowe 2016), the generalization of Riemannian geometry as a scale covariant (conformal) framework for relativistic field theory (Vizgin 1994; Ryckman 2005; McCoy 2021), to the introduction of the gauge principle into the rising quantum mechanics (Straumann 1987; O'Raifeartaigh 1997), and finally he displayed, conjointly with B.L. van der Waerden and E. Wigner, the usability of group representations as a basic frame for studying symmetries in quantum physics (Eckes 2012; Schneider 2011; Scholz 2006). All of this turned out to be of long ranging influence on the course of physics during the twentieth century, and probably also beyond (Yang 1986; Mackey 1988; Borrelli 2015, 2009; Scholz 2018).

[21] "Uns fehlt eine Selbstkritik der Wissenschaft; Urtheile der Kunst, der Religion, des Gefühls über die Wissenschaft sind so zahlreich wie unnütz. Vielleicht ist dies die letzte Bestimmung der *Mathematik*." (Mongré 1897, Aph. 401).

In addition to his direct contributions to mathematical and theoretical physics, Weyl published (and proposed in talks) profound reflections on the epistemology and ontology of the physical world, and the role of mathematics in it, most notably (Weyl 1927, 1948/1949, 1955b). Transformations of mathematical structures played a great role in his reflections; but in stark contrast to Hausdorff he proposed to identify, as clearly as possible, what he considered the automorphisms (global and gauge) of "Nature" herself to which the transformation group of the descriptive symbol system ought to adapt as smoothly as possible. Weyl's objective-transcendental constructive mode has recently been taken up in the philosophy of physics in (Catren 2018). We have seen that Hausdorff used the method of transformation and the related structure groups with the opposite goal of undermining a belief (at least a naive one) in being able to discern such structures in the world, i.e., in a deconstructivist mode *ante letteram*.

For Weyl, philosophical reflections seemed important also for securing a cultural basis for mathematics, in particular the parts which were not amenable to what he accepted as constructive (i.e., essentially by denumerable procedures). From the mid-1920s onward he realized, at first hesitatingly, that the principle of the excluded third and axiomatically postulated transfinite mathematical objects of higher cardinality may be of importance and *acceptable* because of their role in making the difficult structures of modern physics intelligible, at least in an indirect symbolic way.

> From the formalist standpoint, the transfinite component of the axioms takes the place of complete induction and imprints its stamp upon mathematics. The latter does not consist here of evident truths but is *bold theoretical construction*, and as such the very opposite of analytical self-evidence. . . .
>
> In axiomatic formalism, finally, consciousness makes the attempt to 'jump over its own shadow'. to leave behind the stuff of the given, to represent the *transcendent*—but, how could it be otherwise?, only through the *symbol*. (Weyl 1949, 64ff.)

Hausdorff found joy in searching for logically consistent insight into transfinite constructions in the wide sense; he considered this a goal worth pursuing for its own sake, which carried an intrinsic value. Weyl, in contrast, considered such symbolical thought (dealing with a stronger transfinite than denumerable constructivism would accept) as meaningful only when it could be related to natural sciences directly or indirectly (Weyl 1949, 61). On this score, the differences between our protagonists could not have been larger. But does one of these opinions devaluate the other as a legitimate position for a twentieth century "modern" mathematician? It would seem more reasonable to consider both as understandable reactions of creative mathematicians to the challenge of the cultural and social modernization they lived in and contributed to.

19.4 Outlook: Hausdorff's and Weyl's Stance Towards Modernity

19.4.1 Hausdorff: Liberation, Rationalism and the "End of Metaphysics"

We are well informed about Hausdorff's perception and evaluation of cultural development in late nineteenth century Germany through the publication of his *alter ego* Mongré, in particular his time-critical essays in the *Neue Deutsche Rundschau*, a leading journal of liberal intellectuals in Germany (Brieskorn and Purkert 2021, Chap. 6), (Hausdorff 2002–2021, vol. *7*). As mentioned above, he came from a conservative Jewish parental home, the religious traditions and creed of which he did not share. He grew up in a German environment in rapidly modernizing change, which allowed for a slow and selective emancipation of Jewish people on the one hand, but on the other hand was also hatching a rising anti-Semitism in daily life. On this background Hausdorff developed a sharp-minded, critical, highly individualistic view of life and culture which at the turn to the twentieth century was characterized in Germany by a streaky mix of turbid tradition and cheered up modernism. Later in his life he characterized his own cultural and philosophical trajectory as having developed "from Wagner to Schopenhauer, from there back to Kant and forward to Nietzsche" (Hausdorff 2002–2021, vol. *9*, 503). With "Nietzsche" Hausdorff at this place referred to the *young* writer whom he emphatically talked about, at a different place, as the

> . . . affectionate, tempered, appreciative, freethinking Niezsche and the cool, dogma free, system-less sceptic Nietzsche and the (. . .) world blessing, all positive ecstatic Zarathustra (Brieskorn and Purkert 2021, 181).[22]

This picture of Nietzsche stands in stark contrast to the later "fanatic" Nietzsche who, in addition, was contorted for the worse by his the editors under the influence of his sister. The late, fanatic Niezsche preached a morality which, according to an observation made by Hausdorff as early as 1902, contained the potential for "turning into a world-historic scandal which might dwarf the inquisition and the witch trials, such that they would appear as harmless aberrances" (Brieskorn and Purkert 2021, 180).[23]

In short, the young Hausdorff/Mongré developed into an enlightened Nietzschean dissident in late nineteenth century Germany. He considered the cultural

[22] ". . . von dem gütigen, maßvollen, verstehenden Freigeist Nietzsche und von dem kühlen, dogmenfreien, systemlosen Skeptiker Nietzsche und von dem Triumphator des Ja-und Amenliedes, dem weltsegnenden, allbejahenden Ekstatiker Zarathustra" (Brieskorn and Purkert 2021, 181).

[23] "In Nietzsche glüht ein Fanatiker. Seine Moral der Züchtung, auf unserem heutigen Fundamente biologischen und physiologischen Wissens errichtet: das könnte ein weltgeschichtlicher Skandal weden, gegen den Inquisition und Hexenprozeß zu harmlosen Verirrungen verblassen" (Brieskorn and Purkert 2021, 180).

modernisation as an emancipatory chance, with the intellectual and social liberation of the individual as the cultural task of the time. Some of his writings as P. Mongré had the flavour of a drastically exaggerated emphasis on the role of individual perception of the world, and the fiction of the happiness of the "higher" persona standing above the happiness of the many and in contrast to any other kind of social bonds (Mongré 1898c), (Brieskorn and Purkert 2021, sec. 6.1.1). To him the motif of individual freedom seemed to fit well with Cantor's battle cry for set theory: "the essence of mathematics lies in its freedom". In contrast to Cantor himself, Hausdorff took set theory as a model and tool for dissolving metaphysical bonds not only within mathematics, but in general, with mathematics as a trailblazer. In his view, mathematical thought ought to be tied back to the social and outer material world only in the indirect and sceptical form of his "considerate empiricism" (see above).

In his book "Chaos in Cosmic Selection. An Epistemological Essay" Hausdorff/Mongré hoped to be able to do away with metaphysics once and for all. The book ended with the often cited, (all too) proud claim:

> Therewith the bridges have been torn down, which, in the imagination of all metaphysicians, connect the chaos [the transcendent world, ES] and the cosmos [the ordered sensible and intelligible world, ES] in both directions, and the *end of metaphysics* has been declared, the explicit one no less than the masked one, both of which the science of the coming century is obliged to scrap from its architecture (Mongré 1898a, 209; *7*, 803, emph. in the original).[24]

He broadened this argument in a more popular article "The unclean century" in the widely read journal *Neue Deutsche Rundschau* (Mongré 1898b). It contains a beautiful, in large parts satirical, general accounting with the cultural inconsistencies of the semi-modern culture in Wilhelmian Germany. Hausdorff/Mongré attacked, among others, the militaristic habitus among Germany's self-defined elites, who continued to regard duelling as the proper form of honour-saving conflict resolution. He also spoke out against certain aspects of Neo-Kantianism in German humanistic education as cultural hypocrisy, but he also criticised the rising neo-religiousness of diverse flavours as obscurantism, and the unacknowledged metaphysical elements in natural science (Brieskorn and Purkert 2021, sec. 6.1.2). All this he saw as a ballast from earlier times that had to be done away with by

> ...an act of cleanliness with which any retiring century should recommend itself to its successor (Mongré 1898b, 352).[25]

For the young Hausdorff (Mongré) some sort of purified modernity appeared as a desirable future state of the human world. Needless to say, this optimistic perspective was broken by the two Great Wars, the deep world crisis between them, and the rise of Nazism in Germany, with all the humiliations and cruelties against

[24] "Damit sind die Brücken abgebrochen, die in der Phantasie aller Metaphysiker vom Chaos zum Kosmos herüber und hinüber führen, und ist das *Ende der Metaphysik* erklärt, – der eingeständlichen nicht minder als jener verlarvten, die aus ihrem Gefüge auszuscheiden der Naturwissenschaft des nächsten Jahrhunderts nicht erspart bleibt" (Mongré 1898a, 209; *7*, 803).

[25] Man "...vollzieht einen Act der Reinlichkeit, mit dem jedes scheidende Jahrhundert sich seinem Nachfolger empfehlen sollte" (Mongré 1898b, Werke *7*, 352).

the Jewish population, that he himself had to go through. One of his last letters written in January 1941, about a year before his suicide he was forced into, ended with realistic resignation:

> Nietzsche always feared that Europe might perish because of a hysteria of pity: one cannot claim that this diagnosis was particularly realistic (Hausdorff 2002–2021, vol. *9*, 357).[26]

19.4.2 Weyl: Awareness of Crisis and the Search for Metaphysical Horizons

Weyl was among those who, while still at school, was strongly affected by Kant's critical philosophy. For him this by no means lead to a complacent and indolent attitude, so forthrightly attacked by Mongré in his essay about the "unclean century". In retrospect Weyl characterized the effect of Kant's teaching of the "ideality of space and time" quite differently:

> ... by one jerk I was awoken from the 'dogmatic slumber'; the world was most radically put into question for the mind of the adolescent (Weyl 1955a, *4*, 632).[27]

Thus, for Weyl, the reading of Kant had an effect usually ascribed to "modernity" or "modernism": a radical detachment of assuming simple bonds to reality. This detachment was even enhanced, when he entered Göttingen university and learned of Hilbert's studies of the foundations of geometry. The "multitude of different unfamiliar geometries" studied in the axiomatic approach destroyed his simplified picture of an "edifice" of Kantian philosophy, which he had erected in his mind (ibid. 633). This retrospective description indicate that Weyl, in contrast to Hausdorff/Mongré, sensed the confrontation with a "modern" view of the world, and the adoption of it for himself, as a deeply irritating experience. In much of his later writings we find an embarrassment about the basic detachment of mathematical knowledge from the link to the external world. Weyl would sometimes speak of a "transcendent" reality, apparently alluding also to the religious connotation of the word besides a vague reference to an outer nature beyond the one "given" to the senses and to phenomenal insight.

Many authors have argued that the experience of the breakdown of civil norms during the Great War and the following deep social crisis in Germany aggravated Weyl's, and others, sensitivity with regard to the stability also of scientific and even mathematical knowledge (Mehrtens 1990; Sigurdsson 2001; Schappacher 2003). The latter had been untightened already in the later nineteenth century by the loss of

---

[26] "Nietzsche hat immer befürchtet, dass Europa an einer Hysterie des Mitleids zugrunde gehen würde: man kann nicht behaupten, dass diese Diagnose sehr zutreffend war" (Hausdorff an J. Käfer, 2. Jan. 1941, (Hausdorff 2002–2021, vol. *9*, 357)).

[27] "... mit einem Ruck war ich aus dem dogmatischen Schlummer' erwacht, war dem Geist des Knaben auf radikale Weise die Welt in Frage gestellt" (Weyl 1955a, *4*, 632).

credibility of traditional metaphysics and an imputed direct reference to an external reality. This seems to have strongly influenced Weyl's sensitivity for crisis in the debate on the foundations of analysis and set theory.

This awareness of crisis was not only linked to the immediate post-war years and the early 1920s. Although in the second half of the 1920s he was willing to accept that Hilbert's proof theoretic ("finitist") program might even be successful with regard to a formal legitimation of the use of (strong) transfinite methods in mathematics, this would not solve, in his opinion, the problem of meaning of such parts of mathematics, which were based on transfinite axiomatic methods. In Weyl's view (as we know not in Hilbert's view) Gödel's incompleteness theorem for a sufficiently strong formal system embracing arithmetic and a formalized logic as strong as the one of Russell's *Principia Mathematica* dealt a "terrific blow" to Hilbert's program (Weyl 1946, *4*, 279). This was written after the second, even more devastating war than the one after which he had declared the new "crisis" in the foundations of mathematics. Weyl gave a short survey of the development of the research in the foundations of mathematics during the last few decades; then he repeated his diagnosis of the situation, given roughly 30 years ago:

> From this history one thing should be clear: we are less certain than ever about the ultimate foundations of (logic and) mathematics. Like everybody and everything in the world today, we have our 'crisis'. We have had it for nearly fifty years. (Weyl 1946, *4*, 279)

As we also know, this did not hinder him from participating in the enterprise of modern mathematics and physics, though it shaped his selection of research topics and methods. He continued:

> Outwardly it does not seem to hamper our daily work, and yet I for one confess that it has had a considerable practical influence on my mathematical life: it directed my interests to fields I considered relatively 'safe', and has been a constant drain on the enthusiasm and determination with which I pursued my research work. (ibid.)

Here Weyl speaks of 'safeness' in the sense of cultural meaning of mathematics, including its links to the clarification of knowledge in the natural science, in particular physics. In addition, this remark may also be read as a partial explanation for Weyl's never-ending efforts to find support in philosophical reflection of his work, an effort which did not stop short of explicit metaphysical considerations. This stands in the sharpest possible contrast to Hausdorff whose verdict of (classical) metaphysics we have seen above.

19.4.3 Final Remarks: Modern—Countermodern—Trans-Modern?

Neither of our two protagonists maintained a Platonist view of the objects of mathematical knowledge. Hausdorff rejected any claim of ideal order beyond the insights gained by logically precisely framed symbolic production in the realm of transfinite sets opened by Dedekind and Cantor. He was convinced that such highly elaborate argumentations of transfinite set-theory can be expanded without encountering contradictions, as long as this was carefully restrained by the comprehension principle. Although Hausdorff did not, to my knowledge, publish a short, conclusive verbal description of his view of mathematics, he may well be called a *symbolic formalist*. That is, he emphasized that mathematics deals with "objects of thought, symbols of undetermined meaning" which require no further constraint than that of logical consistency.[28] In this respect he had a completely different conception of set theory than Cantor who attached ontological meaning to transfinite sets; Hausdorff's view may rightly be called "modernist".

Weyl, as we have seen, had a rather different understanding of what mathematics is about, or at least ought to be about. His constructivist perspective or, over a time period, even intuitionist understanding (sui generis) of mathematical objects did not allow him to join the radical modernist attitudes of Hilbert, Hausdorff and, later, the young mathematicians of the Bourbaki generation. His decidedly constructivist perspective was, however, not at all "countermodern". It even had strong resemblances to certain features of modernist architecture (Bauhaus) or art (cubism). Also Weyl's most important philosophical inspirations received from Husserl's phenomenology and Fichte's "constructivism" (as he himself described it in (Weyl 1955a, *4*, 641) and the latter's contemporary liberal interpreter Fritz Medicus cannot be qualified as a "countermodern", in contrast to the conservative nationalist interpreters of the South-German Ficheans, or even as an "antimodern" influence on Weyl.

Finally, if one takes the corpus of their mathematical research work into account, surely the most important sources for the description of mathematicians, we see here two towering figures of mathematics in the "modern" period of the late nineteenth and the twentieth century. It would be misplaced to describe one of them, Weyl, as a "countermodern" mathematician and only the other one, Hausdorff, as modern. But, of course, the qualification of Weyl's and Brouwer's position in the foundations of mathematics as representatives of "countermodernism" ("Gegenmoderne") in Herbert Mehrtens' influential book (Mehrtens 1990, 289ff., 301) has an evident factual base. There were real differences between the two authors, which may be described in simple terms as follows: While Hausdorff, at least as a young man, welcomed the rising modernity/modernism in science and culture enthusiastically as a liberating movement, Weyl was irritated and suffered from the loss of security brought about by the intellectual and social turnovers of the late 19th and early 20th centuries. In consequence he tended to distance himself from modernist positions in the reflective *discourse* on mathematics, as Mehrtens calls it. This made him a modern scientist (not a modernist), who was critical of many aspects of modernity, not only with regard to epistemic questions but also with regard to the social destructions which were part of the rise of modernity.

[28] The closest approximation to such a short characterisation can be found in section 1, "Der Formalism", of an unpublished fragment (Hausdorff 1904/2021) written about 1904, in particular folio 4ff, vol. *6*, p. 474ff.

After the second Great War of the twentieth century Weyl was shocked by the destructive potential which had developed on the basis of scientific achievements. It seems that he experienced the rise of modern society, in the sense of late nineteenth- and twentieth-century high capitalism, including its scientific culture, as a challenge and a crisis set in permanence. In a manuscript written close to the end of his life and published posthumously by T. Tonietti, he deplored the state of things, dramatically clothed in his grave, humanistic style. He warned that modern science may be characterized by a kind of *hubris* (violent arrogance) and went on

> For who can close his eyes against the menace of our own self-destruction by science? The alarming fact is that the rapid progress of scientific knowledge is not paralleled by a corresponding growth of man's moral strength and responsibility, which have hardly changed in historical times (Weyl 1985, 12).

Weyl could not even follow Hardy's move after the first War for exculpating pure mathematics on the basis of its "uselessness" in practical matters, which, according to Hardy, would protect it against a participation in "exploitation of our fellow-men" and the destruction up to their "extermination" (Weyl's words). Weyl did not believe in such an escape route and emphasized:

> However the power of science rests on the combination of experiment, i.e., observation under freely chosen conditions, with symbolic construction, and the latter is its mathematical aspect. Thus if science is found guilty, mathematics cannot evade the verdict. (ibid.)

This sentiment echoes a similar, although less dramatically stated fright which was expressed in Hausdorff's downplayed remark of 1941 that Nietzsche's warning that modern history might suffer from too much pity for humanity or nature has turned out as not "particularly realistic".

Weyl was a sceptical modern actor all his life. As we know too well, the dangers of extermination of mankind by war and/or destruction of our natural habitat are now even more severe than in the 1950s. But science is not only an accomplice of the destructive sides of modernity; it also plays the role of collecting the warning signs and is necessary for exploring exit strategies from the ongoing destruction. From our vantage point of the early twenty-first century, Weyl may appear as a modern scientist who tried to dive through the wave of modernism towards some not yet clearly visible type of trans-modern culture. The latter would demand remaining true to the enlightened elements of modernity, while diminishing the destructive forces against nature and our fellow-men.

Acknowledgments I thank Walter Purkert, Norbert Schappacher, José Ferreirós, Lizhen Ji, Jinze Du, Matthias Kreck and David Rowe for their often detailed and valuable comments to preliminary versions of this chapter.

References

Atiyah, M. 2002. *Hermann Weyl, November 9, 1885 – December 9, 1955*, Volume 82 of *Biographical Memoirs National Academy of Sciences*. Washington, D.C.: National Academy Press.

Baeumler, A., and M. Schroeter. 1927. *Handbuch der Philosophie. Bd. II. Natur, Geist, Gott*. München: Oldenbourg.

Bernard, J., and C. Lobo. 2019. *Weyl and the Problem of Space. From Science to Philosophy*. Cham: Springer Nature Switzerland.

Borrelli, A. 2009. The emergence of selection rules and their encounter with group theory, 1913–1927. *Studies in History and Philosophy of Modern Physics* 40: 327–337.

Borrelli, A. 2015. The making of an intrinsic property: 'symmetry heuristics' in early particle physics. *Studies in History and Philosophy of Modern Physics A* 50: 59–70.

Borrelli, A. 2017. The uses of isospin in early nuclear and particle physics. *Studies in History and Philosophy of Modern Physics* 60: 81–94.

Brieskorn, E., and W. Purkert. 2021. *Felix Hausdorff. Mathematiker, Philosoph und Literat*. Berlin: Springer. English translation in preparation.

Brouwer, L. E. J. 1919. Intuitionistische Mengenlehre. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 28: 203–208. In (Mancosu 1998, 23–27).

Catren, G. 2018. Klein-Weyl's program and the ontology of gauge and quantum systems. *Studies in the History and Philosphy of Modern Physics* 61: 25–40.

Chevalley, C., and A. Weil. 1957. Hermann Weyl (1885–1955). *L'Enseignement mathématique* 3: 157–187.

Coleman, R., and H. Korté. 2001. Hermann Weyl: Mathematician, physicist, philosopher. In (Scholz 2001, 161–388).

Corry, L. 2004. *David Hilbert and the Axiomatization of Physics (1898-1918). From Grundlagen der Geometrie to Grundlagen der Physik*. Dordrecht: Kluwer.

Dieudonné, J. 1976. Weyl, Hermann. In *Dictionary of Scientific Biography*, ed. C. Gillispie, vol. 14, 281–285. Princeton University Press.

Eckes, C. 2011. *Groupes, invariants et géométrie dans l'oeuvre de Weyl. Une étude des écrits de Hermann Weyl en mathématiques, physique matématique et philosophie, 1910–1931*. Ph. D. thesis. http://math.univ-lyon1.fr/homes-www/remy/TheseChristopheEckes-26sept2011.pdf.

Eckes, C. 2012. Principes d'invariance et lois de la nature d'après Weyl et Wigner. *Philosophia Scientiae* 16(3): 153–176.

Epple, M. 2006. *Felix Hausdorff's Considered Empiricism*, In (Gray 1999, 263–289).

Epple, M. 2021. *Felix Hausdorffs Erkenntniskritik von Zeit und Raum*. In (Hausdorff 2002–2021, vol. *6*, 1–207).

Epple, M., H. Herrlich, M. Hušek, G. Preuß, W. Purkert, and E. Scholz. 2002. Zum Begriff des topologischen Raumes. In *F. Hausdorff, Gesammelte Werke, Bd. 2, Grundzüge der Mengenlehre*, 675–744. Berlin: Springer. http://www.aic.uni-wuppertal.de/fb7/hausdorff/Topologischer_Raum.pdf.

Feferman, S. 2000. The significance of Weyl's 'Das Kontinuum'. In *(Hendricks et al. 2000)*, 179–194.

Ferreirós, J. 2016. *Mathematical Knowledge and the Interplay of Practices*. Princeton University Press.

Forman, P. 1971. Weimar culture, causality, and quantum theory, 1918–1927: Adaptation by German physicists and mathematicians to a hostile intellectual environment. *Historical Studies in the Physical Sciences* 3: 1–116.

Forman, P. 1973. Scientific internationalism and the Weimar physicists: The ideology and its manipulation in Germany after World War I. *Isis* 64: 151–180.

Fraenkel, A. 1923. *Einleitung in die Mengenlehre, Zweite Auflage*. Berlin: Springer.

Frei, G., and U. Stammbach. 1992. *Hermann Weyl und die Mathematik an der ETH Zürich, 1923 – 1930*. Basel: Birkhäuser.

Giulini, D. forthcoming. Zu Kap. IV: ART und Verallgemeinerung. In (Weyl 2024).
Goenner, H. 2001. Weyl's contributions to cosmology. In *Hermann Weyl's Raum-Zeit-Materie and a General Introduction to His Scientific Work*, ed. E. Scholz, 105–138. Basel: Birkhäuser.
Gray, J. 1999. *The Symbolic Universe: Geometry and Physics 1890–1930*. Oxford: University Press.
Gray, J. 2008. *Plato's Ghost. The Modernist Transformation of Mathematics*. Princeton: University Press.
Hausdorff, F. 1901. Über eine gewisse Art geordneter Mengen, I–V. *Berichte Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften Leipzig, Math.-phys. Classe* 53: 400–475. In (Hausdorff 2002–2021, vol. *1B, 39–182*). English in J.M. Plotkin (ed.) *Hausdorff on Ordered Sets*. Providence (Rhode Island), AMS, 2005.
Hausdorff, F. 1904/2021. Raum und Zeit. Nachlass Hausdorff, Kapsel 49, Faz. 1067. In (Hausdorff 2002–2021, vol. *6*, 473–478).
Hausdorff, F. 1914a. Bemerkung über den Inhalt von Punktmengen. *Mathematische Annalen* 75: 428–433. In (Hausdorff 2002–2021, vol. *4*, 3–10).
Hausdorff, F. 1914b. *Grundzüge der Mengenlehre*. Leipzig: Veit & Co. Reprint Chelsea 1949ff. In (Hausdorff 2002–2021, vol. 2).
Hausdorff, F. 1919. Dimension und äußeres Maß. *Mathematische Annalen* 79: 157–179. In (Hausdorff 2002–2021, vol. *4*, 21–43).
Hausdorff, F. 1923/2006. Vorlesung Wahrscheinlichkeitstheorie (1923). In (Hausdorff 2002–2021, vol. *5*, 595 756–723).
Hausdorff, F. 1927. *Mengenlehre. Zweite neubearbeitete Auflage*. Leipzig: Walter de Gruyter. In (Hausdorff 2002–2021, vol. *3*, 42–451).
Hausdorff, F. 2002–2021. *Gesammelte Werke*, ed. E. Brieskorn, W. Purkert et al., 9 Bände. Belin: Springer.
Hawkins, T. 2000. *Emergence of the Theory of Lie Groups. An Essay in the History of Mathematics 1869–1926*. Berlin: Springer.
Hendricks, V., S. A. Pedersen, and K. F. Jørgensen (eds.). 2000. *Proof Theory. History and Philosophical Significance*. Dordrecht: Kluwer.
Hesseling, D. 2003. *Gnomes in the Fog. The Reception of Brouwer's Intuitionism in the 1920s*. Birkhäuser.
Hilbert, D. 1903a. *Grundlagen der Geometrie*, 2. Auflage. Teubner.
Hilbert, D. 1903b. Über die Grundlagen der Geometrie. *Mathematische Annalen* 56: 381–422.
Kolmogoroff, A. N. 1933. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer.
Kreck, M. 1986. Exotische Strukturen auf 4-Mannigfaltigkeiten. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 88: 124–145.
Mackey, G. 1988. Weyl's program and modern physics. In *Differential Geometric Methods in Theoretical Physics*, ed. K. B. M. Werner, 11–36. Dordrecht: Kluwer. Reprint in (Mackey 1992, 249–274).
Mackey, G. 1992. *The Scope and History of Commutative and Noncommutative Harmonic Analysis*. Providence, Rhode Island: American Mathematical Society/London Mathematical Society.
Mancosu, P. 1998. *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s*. Oxford: University Press.
McCoy, C. 2021. The constitution of Weyl's pure infinitesimal world geometry. Preprint http://philsci-archive.pitt.edu/19334/1/21.07.08Eprint.pdf.
Mehrtens, H. 1990. *Moderne - Sprache - Mathematik: eine Geschichte des Streits um die Grundlagen der Disziplin und des Subjekts formaler Systeme*. Frankfurt/Main: Suhrkamp.
Mongré, P. 1897. *Sant'Ilario – Gedanken aus der Landschaft Zarathustras*. Leipzig: C.G. Naumann. In (Hausdorff 2002–2021, vol. *7*, 85–473).
Mongré, P. 1898a. *Das Chaos in kosmischer Auslese*. Leipzig: C.G. Naumann. In (Hausdorff 2002–2021, vol. 7, 587–807).
Mongré, P. 1898b. Das unreinliche Jahrhundert. *Neue Deutsche Rundschau* 9 (5): 443–452. In (Hausdorff 2002–2021, vol. *8*, 341–352).

Mongré, P. 1898c. Massenglück und Einzelglück. *Neue Deutsche Rundschau* 9 (1): 64–75. In (Hausdorff 2002–2021, vol. *8*, 275–288).

Mongré, P. 1899. Das Chaos in kosmischer Auslese (*Selbstanzeige*). *Die Zukunft* 8 (5): 222–223. In (Hausdorff 2002–2021, vol. *7*, 811f.).

Newman, M. 1957. Hermann Weyl, 1885–1955. *Biographical Memoirs of Fellows of the Royal Society London* 3: 305–328.

O'Raifeartaigh, L. 1997. *The Dawning of Gauge Theory*. Princeton: University Press.

O'Raifeartaigh, L., and N. Straumann. 2000. Gauge theory: Historical origins and some modern developments. *Reviews of Modern Physics* 72: 1–23.

Pauli, W. 1933. Die allgemeinen Prinzipien der Wellenmechanik. In *Handbuch der Physik*, *Zweite Auflage, Band 24.1. Quantentheorie*, ed. H. Geiger, 83–272. Berlin: Springer.

Pauli, W. 1941. Relativistic field theory of elementary particles. *Reviews of Modern Physics* 13: 213–232.

Plotkin, J. 2005. *Hausdorff on Ordered Sets*, Volume 25 of *History of Mathematics*. AMS/LMS.

Remmert, R. 1997. Proömium. In *H. Weyl. Die Idee der Riemannschen Fläche*, ed. R. Remmert, ix–xii. Springer.

Riesz, F. 1908. Stetigkeitsbegriff und abstrakte Mengenlehre, 18–24. In (Riesz 1960, vol. 1).

Riesz, F. 1960. *Gesammelte Arbeiten, Bd. I.* Akademie der Wissenschaften.

Rodriguez, L. 2006. Friedrich Riesz' Beiträge zur Herausbildung des modernen mathematischen Konzepts abstrakter Räume. Synthese intellektueller Kulturen in Ungarn, Frankreich und Deutschland. Dissertationsschrift Universität Mainz.

Rowe, D. 2002. Hermann Weyl, the reluctant revolutionary. *Mathematical Intelligencer* 25 (1): 61–70. In (Rowe 2018, 331–341).

Rowe, D. 2016. A snapshot of debates on relativistic cosmology, 1917-1924 (part II). *Mathematical Intelligencer* 38 (3): 52–60.

Rowe, D. 2018. *A Richer Picture of Mathematics. The Göttingen Tradition and Beyond*. Springer.

Rowe, D. 2024. Brouwer and Hausdorff: On reassessing the foundations crisis during the 1920s. To appear in *Science in Context*: 35, no. 4.

Ryckman, T. 2005. *The Reign of Relativity. Philosophy in Physics 1915–1925*. Oxford: University Press.

Schappacher, N. 2003. Politisches in der Mathematik: Versuch einer Spurensicherung. *Mathematische Semesterberichte* 50 (1): 1–27.

Schappacher, N. 2010. Rewriting points. In *Proceedings of the International Congress of Mathematicians Hyderabad, India, 2010*, 3258–3291.

Schneider, M. 2011. *Zwischen zwei Disziplinen. B.L. van der Waerden und die Entwicklung der Quantenmechanik*. Berlin: Springer.

Scholz, E. 1995. Hermann Weyl's "purely infinitesimal geometry". In *Proceedings of the International Congress of Mathematicians, Zürich Switzerland 1994*, 1592–1603. Basel: Birkhäuser.

Scholz, E. 1999. Weyl and the theory of connections. In (Gray 1999, 260–284).

Scholz, E. 2000. Hermann Weyl on the concept of continuum. In (Hendricks et al. 2000, 195–220).

Scholz, E. (ed.). 2001. *Hermann Weyl's Raum - Zeit - Materie and a General Introduction to His Scientific Work*. Basel: Birkhäuser.

Scholz, E. 2004. Hermann Weyl's analysis of the "problem of space" and the origin of gauge structures. *Science in Context* 17: 165–197.

Scholz, E. 2006. Introducing groups into quantum theory (1926–1930). *Historia Mathematica* 33: 440–490.

Scholz, E. 2016. The problem of space in the light of relativity: the views of H. Weyl and E. Cartan. In *Eléments d'une biographie de l'Espace géométrique*, ed. L. Bioesmat-Martagon, 255–312. Nancy: Edition Universitaire de Lorraine. https://arxiv.org/abs/1310.7334.

Scholz, E. 2018. The unexpected resurgence of Weyl geometry in late 20th century physics. In *Beyond Einstein. Perspectives on Geometry, Gravitation, and Cosmology in the Twentieth Century*, ed. D. Rowe, T. Sauer, and S. Walter, 261–360. Heidelberg/Berlin: Springer/Birkhäuser.

Sieg, W. 2000. Toward finitist proof theory. In (Hendricks et al. 2000, 95–116).

Siegmund-Schultze, R. 2009. *Mathematicians Fleeing from Nazi-Germany. Individual Fates and Global Impact*. Princeton: University Press.

Siegmund-Schultze, R. 2021. Kein Überleben für einen älteren Mathematiker: Felix Hausdorffs gescheiterte Emigration und Tod. *Mitteilungen der Deutschen Mathematiker-Vereinigung* 29 (3): 132–136.

Sieroka, N. 2010. *Umgebungen. Symbolischer Konstruktivismus im Anschluss an Hermann Weyl und Fritz Medicus*. Zürich: Chronos.

Sieroka, N. 2019. Neighbourhoods and intersubjectivity: Analogies between Weyl's analyses of the continuum and transcendental-phenomenological theories of subjectivity. In *Weyl and the Problem of Space. From Science to Philosophy*, ed. J. Bernard, C. Lobo, 99–122. Cham: Springer Nature Switzerland.

Sigurdsson, S. 1991. *Hermann Weyl, mathematics and physics, 1900 – 1927*. Ph. D. thesis, Harvard University, Department of the History of 'Science.

Sigurdsson, S. 2001. Journeys in spacetime. In *Scholz (2001)*, 15–47.

Stegmaier, W. 2002. Ein Mathematiker in der Landschaft Zarathustras. Felix Hausdorf als Philosoph. *Nietzsche Studien* 31: 195–240. [Similarly in (Hausdorff 2002–2021, VII, 1–83)].

Straumann, N. 1987. Zum Ursprung der Eichtheorien. *Physikalische Blätter* 43: 414–421.

Vizgin, V. 1994. *Unified Field Theories in the First Third of the 20th Century. Translated from the Russian by J. B. Barbour*. Basel: Birkhäuser.

Weyl, H. 1913. *Die Idee der Riemannschen Fläche*. Leipzig: Teubner. ^{2}1923, ^{3}1955, ^{4}1997. English Weyl (1964).

Weyl, H. 1913/1964. *The Concept of a Riemann Surface*. Reading: Addison-Wesley. English translation of (Weyl 1913, 3rd ed.) by G. MacLane.

Weyl, H. 1918a. *Das Kontinuum. Kritische Untersuchungen über die Grundlagen der Analysis*. Leipzig: Veit. ^{2}1932 Berlin: de Gruyter. English Weyl (1987).

Weyl, H. 1918b. Gravitation und Elektrizität. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin*, 465–480. In (Weyl 1968, II, 29–42), English in (O'Raifeartaigh 1997, 24–37).

Weyl, H. 1918c. *Raum, - Zeit - Materie. Vorlesungen über allgemeine Relativitätstheorie*. Berlin: Springer. Weitere Auflagen: ^{2}1919, ^{3}1919, ^{4}1921, ^{5}1923, ^{6}1970, ^{7}1988, ^{8}1993.

Weyl, H. 1918d. Reine Infinitesimalgeometrie. *Mathematische Zeitschrift* 2: 384–411. In (Weyl 1968, II, 1–28).

Weyl, H. 1918/1987. *The Continuum*. English translation by S. Pollard and T. Bole. Kirksville, Missouri: Thomas Jefferson University Press. Corrected republication New York Dover 1994.

Weyl, H. 1921. Über die neue Grundlagenkrise der Mathematik. *Mathematische Zeitschrift* 10: 39–79. In (Weyl 1968, II, 143–180). English in (Mancosu 1998, 86–121).

Weyl, H. 1923/1924. Análisis situs combinatorio. *Revista Matematica Hispano-Americana* 5, 6, 5: 43ff, 6: 1–9. In (Weyl 1968, II, 390–415, 416–432) [142].

Weyl, H. 1923/2024. *Raum - Zeit -Materie. Vorlesungen über allgemeine Relativitätstheorie. 5. Auflage. Neu herausgegeben und kommentiert von D. Giulini und E. Scholz*. Berlin: Springer. In preparation, to appear in 2024.

Weyl, H. 1925/1926. Theorie der Darstellung kontinuierlicher halbeinfacher Gruppen durch lineare Transformationen. I, II, III und Nachtrag. *Mathematische Zeitschrift* 23, 24, 23: 271–309, 24: 328–395, 789–791. In (Weyl 1968, Bd. 2, 542–646).

Weyl, H. 1925/1988. *Riemanns geometrische Ideen, ihre Auswirkungen und ihre Verknüpfung mit der Gruppentheorie. Ed. K. Chandrasekharan*. Berlin: Springer.

Weyl, H. 1927. *Philosophie der Mathematik und Naturwissenschaft*. München: Oldenbourg. In (Baeumler and Schroeter 1927, Bd. II A); separat. Weitere Auflagen ^{2}1949, ^{3}1966. English with comments and appendices Weyl (1949), French Weyl (2017).

Weyl, H. 1928. *Gruppentheorie und Quantenmechanik.* Leipzig: Hirzel. ^{2}1931, English translation R.P. Robertson, New York: Dutten 1931.

Weyl, H. 1939. *The Classical Groups, Their Invariants and Representations.* Princeton: University Press. ^{2}1946.

Weyl, H. 1940. The mathematical way of thinking. *Science* 92: 437–446. In (Weyl 1968, vol. *3*, 710–718).

Weyl, H. 1946. Mathematics and logic. A brief survey serving as a preface to a review of "The Philosophy of Bertrand Russell". *American Mathematical Monthly* 53: 2–13. In (Weyl 1968, IV, 268–279).

Weyl, H. 1948/1949. Similarity and congruence: a chapter in the epistemology of science. ETH Bibliothek, Hochschularchiv Hs 91a:31, 23 Bl. Typoskript mit Korrekturen. In (Weyl 1955b, 3rd ed. 2016).

Weyl, H. 1949. *Philosophy of Mathematics and Natural Science*. Princeton: University Press. ^{2}1950, ^{3}2009.

Weyl, H. 1955a. Erkenntnis und Besinnung (Ein Lebensrückblick). *Studia Philosophica* 15: 153ff. In (Weyl 1968, IV, 631–649) [166]; English in (Weyl 2009, 204–221).

Weyl, H. 1955b. *Symmetrie. Ins Deutsche übersetzt von Lulu Bechtolsheim*. Basel/Berlin: Birkhäuser/Springer. ^{2}1981, 3. Auflage 2017: Ergänzt durch einen Text aus dem Nachlass 'Symmetry and congruence', und mit Kommentaren von D. Giulini, E. Scholz und K. Volkert.

Weyl, H. 1968. *Gesammelte Abhandlungen*, 4 Bde. Ed. K. Chandrasekharan. Berlin: Springer.

Weyl, H. 1985. Axiomatic versus constructive procedures in mathematics. Edited by T. Tonietti. *Mathematical Intelligencer* 7 (4): 12–17, 38.

Weyl, H. 2009. *Mind and Nature. Selected Writings on Philosophy, Mathematics, and Physics. Edited and with an Introduction by Peter Pesic*. Princeton: University Press.

Weyl, H. 2017. *Philosophie des mathématiques et des sciences de la nature. Traduit de l'anglais par Carlos Lobo*. Genève: MétisPresses.

Weyl, H. Ms 1922. Kombinatorische Analysis Situs. *Nachlaß Weyl, Zürich: ETH Bibliothek Handschrift 91a:3*. German text of Weyl (1924), dated 1922 with annotation "Zuerst vorgetragen in Zürich 1918". (Copy of the typoscript in the library Mathematisches Institut der Universität Göttingen).

Weyl, H. Ms 1930/1931. Axiomatik. *Bibliothek Mathematisches Institut Göttingen*. Vorlesung Göttingen, WS 1930/1931.

Yang, C. N. 1986. Hermann Weyl's contributions to physics. In *Hermann Weyl 1885–1955. Centenary Lectures*, ed. A. Borel, K. Chandrasekharan, R. Penrose, and C. Yang, 7–22. Berlin: Springer.

Part V
Mathematicians and Philosophy

Chapter 20
The Direction-Theory of Parallels: Geometry and Philosophy in the Age of Kant

Vincenzo De Risi

Abstract The direction-theory of parallels was a mathematical theory that gained enormous importance and popularity for about a century, from the 1770s to the 1870s. It was conceived for the purpose of proving the famous Parallel Postulate and establishing the foundations of Euclidean geometry. The development of this geometric theory was intertwined with that of mathematical epistemology. Proponents of the theory discussed at length such topics as the analyticity of mathematics, the role of intuition in geometry, mathematical constructivism, and the relationship between geometry and the structure of space. In the first few decades of its life, the direction-theory of parallels became the most important benchmark on which to test Kant's philosophy, and Kantians and anti-Kantians alike wrote articles and books on it. The direction-theory was later generally accepted by the leading post-Kantian philosophers of the nineteenth century. It was finally subjected to fatal criticism by Lewis Carroll and Gottlob Frege.

> [Enter:] The phantasm of Herr Niemand, carrying a pile of phantom-books, the works of Euclid's Modern Rivals, phantastically bound.
> (Lewis Carroll)

20.1 Euclid's Rivals on the Theory of Parallels

Back from Wonderland, Lewis Carroll published (under his worldly name of Charles Dodgson) a treatise on the foundations of geometry written in "dramatic form" and "lighter style". *Euclid and His Modern Rivals* appeared in 1879, and lambasted

V. De Risi (✉)
Laboratoire SPHère, CNRS, Paris, France

Max-Planck-Institut für Wissenschaftsgeschichte, Berlin, Germany

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_20

recent treatments of elementary geometry in England, while extolling the logical virtues of the original *Elements* by Euclid. An entire act of this tragedy of geometry is dedicated to the modern theories of parallel lines, which are expounded by the German Herr Niemand (Mr. Nobody) and rebutted by the Greek (and infernal) judge Minos – as a champion of Euclid. The greatest part of this act deals with the so-called "direction-theory" of parallels.[1]

The direction-theory of parallels was, like Herr Niemand, German. It was first conceived in 1778, with the stated aim of reforming Euclid's theory of parallels and proving the Parallel Postulate. This was a famous challenge in geometry, and since Antiquity mathematicians had been busy proposing new demonstrations of the postulate and at disproving other mathematicians' proofs.[2] The direction-theory fared much better for a long time and acquired great momentum at the turn of the nineteenth century. The theory began to be taught in schools, was endorsed by the most important philosophers, and was generally accepted by the mathematical community. It soon crossed Germany's borders and spread all over Europe. In 1870, the British *Association for the Improvement of Geometrical Teaching* recommended that students in England should learn geometry from modern textbooks rather than from Euclid's *Elements* – as was also the long-held custom in Germany.[3] Many of these new English manuals embraced the direction-theory of parallels as a more suitable approach for students than Euclid's.

The discovery of non-Euclidean geometry by Lobachevsky and Bolyai did not slow down this booming phenomenon. The importance of these pioneering works was not recognized for several decades, and it was only in the course of the 1870s that non-Euclidean geometry began to gain wider acceptance – when Klein and Poincaré legitimized it even outside the field of foundational studies.[4] At the end of the decade, Dodgson's witty drama disposed of this German theory belatedly imported into Britain and defeated by modern mathematics. The direction-theory, that had flourished for 101 years in more than as many books, finally stepped through the looking-glass.[5]

[1] Dodgson (1879). The theory of parallels is discussed in Act Two, the direction-theory being dealt with in pp. 70–131 of the work.

[2] For an outstanding presentation of the history of non-Euclidean geometry, see Gray (1989).

[3] An exhaustive presentation of the English debate on the teaching of Euclid is found in Moktefi (2011). For its connection with non-Euclidean geometry, see Gray (2004), pp. 95 ff.

[4] On the acceptance of non-Euclidean geometry in the second half of the nineteenth century, see Voelke (2005) and Volkert (2013).

[5] An important historical discussion of various attempts to prove the Parallel Postulate through arguments from direction, composed just when the era of such attempts was drawing to a close, is to be found in Schotten (1890–1893). For further discussions on the *attardés* who were still hoping to prove the Postulate through the direction-theory at the end of the nineteenth century, see Pont (1986).

The gist of the direction-theory is easy to convey. Whereas Euclid had defined parallel lines as straight lines that do not meet, the direction-theorists defined them as *straight lines having the same direction*. Sameness of direction is commonly understood as to be a transitive relation: if *A* has the same direction as *B*, and *B* the same direction as *C*, then *A* has the same direction as *C*. This apparently harmless assumption, applied to parallel lines, entails however the transitivity of parallelism – something that fails in non-Euclidean geometry. As a consequence, Euclid's modern rivals surreptitiously introduced an assumption equivalent into the Parallel Postulate in the definition of parallelism, and triumphantly derived the former from the latter. Minos did not need too much effort to expose the blatant *petitio principii* of these proofs, and Dodgson's book could expound the criticism in painful detail.

Five years after Dodgson's rebuttal of the direction-theory, Gottlob Frege produced a logical analysis of its shortcomings in his *Grundlagen der Arithmetik*. Frege identified the main mistake in the theory in an incorrect epistemological assumption. According to him, geometry must begin with concrete objects, such as straight lines and circles. Abstract notions, such as direction, cannot be assumed beforehand and employed to define these basic geometric objects. Quite the opposite: abstract notions may only be defined by looking at the relations obtaining between concrete objects. *If* such a relation is an equivalence relation – and, therefore, if it is transitive – *then* it partitions the set of objects into equivalence classes to which an abstract notion may be attached. The transitivity of relations among concrete objects is a prerequisite for abstracting notions in the first place.

> The trouble is, that this is to reverse the true order of things. For surely everything geometrical must be given originally in intuition. But now I ask whether anyone has an intuition of the direction of a straight line. Of a straight line, certainly; but do we distinguish in our intuition between this straight line and something else, its direction? That is hardly plausible. The concept of direction is only discovered at all as a result of a process of intellectual activity which takes its start from intuition. On the other hand, we do have a representation of parallel straight lines.[6]

According to Frege, the only possible procedure to abstract the notion of direction is the following. We consider the set of (concrete) straight lines in a plane and Euclid's relation of parallelism (i.e. non-incidence). Since the Parallel Postulate is true, this relation is transitive and in fact an equivalence relation. Therefore, we may partition the set of straight lines into mutually exclusive equivalence classes of non-incident lines. Each of these classes defines a direction as an abstract notion. By means of this procedure, we transform the sentence "line *a* is parallel (non-incident) to line *b*", which is about concrete objects given in intuition, into the statement that "the direction of line *a* is identical to the direction of line *b*", revolving on abstract, non-intuitive notions. From here, we get to the concept of direction in general.[7]

[6] Frege (1884), § 64, p. 75. Transl. by Austin in (Frege 1960), p. 75, modified.

[7] This last passage is further belabored by Frege in great detail, but it does not concern us here. Note that Lobachevsky and Bolyai were able to define the direction of asymptotic parallel lines, that is to say, of one special kind of parallels in hyperbolic geometry that have the transitive property.

Following Frege's analysis, one cannot even formulate the notion of direction without presupposing the Euclidean transitivity of parallelism and therefore the truth of the Parallel Postulate. The mathematical mistake of the direction-theorists was rooted in a faulty epistemology, which upturned the priority between intuition and concept. Frege's analysis attempted to reach much deeper than Dodgson's, and pointed a finger at the philosophical assumptions of the direction-theorists. In doing so, he exposed his own not inconsiderable share of philosophical commitments.[8]

It is remarkable that neither Dodgson nor Frege rebutted the direction-theory by exploiting the consistency of non-Euclidean geometry. A decade after his book on Euclid, Dodgson himself was still attempting to prove the Parallel Postulate and show the impossibility of non-Euclidean geometry.[9] Frege, committed to making use of intuition in geometry, explicitly rejected non-Euclidean geometry as a delusional, un-scientific theory engendered by an ill-conceived philosophy of mathematics.[10] Dodgson's and Frege's criticisms to the direction-theory only concerned the possibility of defining parallel lines through the notion of direction and to prove the Parallel Postulate from such a definition. Dodgson took the Parallel Postulate as a statement provable by different means, Frege as an indemonstrable axiom given by pure intuition, but both of them firmly believed in its unconditional truth.

The cultural phenomenon of the direction-theory of parallels and its demise poses a double challenge to the historians of mathematics. On the one hand, it is astonishing that such a clearly-faulted theory was accepted and taught for a hundred years. The theory was based on a trivial mistake – assuming the truth of the postulate in the definition of parallels – and it seems that Europe must have been relinquished by all gods if no one noticed such an obvious blunder. On the other hand, the direction-theory was destroyed by the discovery of non-Euclidean geometry, and yet the most important – and very belated – criticisms of it came from two important conservative (even "countermodern") logicians who opposed such a discovery.

The century in which the direction-theory thrived and waned marked the passage from a pre-modern epistemology to a modern – or modernist – philosophy of

[8] Recently Mancosu (2015) has advanced the conjecture that Frege came to know the direction-theory particularly from his reading of Schlömilch's *Grundzüge der Geometrie* (Schlömilch 1849). Among Frege's sources, Karl Georg Christian von Staudt (1798–1867) correctly assumed the Parallel Postulate as an axiom (in the form of Lorenz, see below), and then defined direction starting from (transitive, Euclidean) parallelism: see Staudt (1847), §§ 35–36, pp. 14–15. Frege had already endorsed a similar procedure in his doctoral dissertation from 1873, *Über eine geometrische Darstellung der imaginären Gebilde in der Ebene*. See Frege (1967), pp. 3–4 and 49.

[9] Dodgson (1888). In the treatise, Dodgson provisionally accepted a further "quasi-axiom" in order to prove the Parallel Postulate, admitting to have failed to offer a completely unhypothetical proof of it. The 1888 book also offers a further discussion of the direction-theory of parallels in Appendix IV, § 5, and several important remarks on the Parallel Postulate in non-Archimedean planes.

[10] Frege's clearest statements on this topic are in his unpublished *Über Euklidische Geometrie*. In it, he states that no one can serve two masters (Euclidean and non-Euclidean geometry), and concludes that non-Euclidean geometry is to be listed among the non-scientific disciplines having only historical interest – like astrology and mummies. See Frege (1983) pp. 182–84.

mathematics.[11] It is from this vantage point, which also people like Dodgson and Frege contributed to shape, that we may look down at the direction-theory as an inexplicable failure of logic and common sense. The direction-theorists, however, were not naïve. Their mathematics was rather based on a quite developed pre-modern epistemology. The direction-theory was elaborated in some crucial decades between the end of the eighteenth century and the beginning of the nineteenth century, in which German mathematicians and philosophers were highly engaged with epistemological questions. The direction-theorists played an important role in this debate, and they were not unaware of the challenges later raised against them by Dodgson and Frege. Only, their pre-modern answers were different from the modern ones.

In the eighteenth century, for instance, the philosopher and mathematician Christian Wolff had argued that the axioms of mathematics should follow from the definitions of the terms involved. Consequently, it was *expected* that the Parallel Postulate could be drawn from a suitable definition of parallel lines. Dodgson's complaint of a *petitio principii* would have been met with puzzlement by many mathematicians endorsing Wolff's views. According to other Leibnizians, intuition should play no role in the foundations of geometry, and it was nonsense to claim that the abstract notion of direction should follow the concrete intuition of parallel straight lines. Quite the opposite: the concept of direction is *simpler* than, and should therefore precede, that of a straight line. Frege's objection would have appeared to these philosophers as a serious epistemological error.

These examples show that in order to understand why the direction-theory was so widely upheld for an entire century, we have to dive deeper into the philosophy of mathematics that engendered it. Similarly, in order to appreciate its demise, that did not happen for purely mathematical reasons (the discovery of non-Euclidean geometry), we should follow the development of the nineteenth-century epistemology of mathematics up to the point in which it could no longer back the assumptions of the theory. The philosophy of Kant played an important role in this debate, since the direction-theory was born in the age of Kant and was immediately drawn into the disputes on transcendentalism. Kant's shadow projected however much further into the following century, and Frege's and Dodgson's remarks on the syntheticity of geometry and the role of intuition still depended on their reading of the *Critique*.[12] In short, the history of the cultural phenomenon of the direction-theory may only be written together with a history of the philosophy of mathematics in the eighteenth and nineteenth centuries.

[11] The topic of modernism in mathematics is wonderfully addressed in one of Jeremy Gray's most fascinating books: Gray (2008). The notion of countermodernism in mathematics was famously introduced by Mehrtens (1990) and is discussed at length by Gray.

[12] Frege's commitment to (broadly) Kantian views on geometry in the *Grundlagen* are well-known; see the classic paper by Dummett (1991). Dodgson quoted with approval Kant's *Critique* in order to rebut the ideas of Herr Niemand: see Dodgson (1879, p. 55).

In the present essay, I will not attempt this much. I rather concentrate on the first period of the direction theory, roughly covering the fifty years from 1778 to the discovery of non-Euclidean geometry. In these decades the theory was still confined to Germany and the debate surrounding it was sensitive to the important transformations of the philosophical landscape. Historians have not yet explored this early history of the theory – in fact they have not even recognized it as a topic of investigation. The available narratives on the subject (including Dodgson's own history of Euclid's rivals) generally concentrate on the second half of the nineteenth century, when the direction-theory already had its current name and final shape. This may convey the false impression that the theory was originally conceived this late.

An enquiry into the sources of the German and British direction-theorists of the nineteenth century, however, discloses another story – of which the present essay offers a first sketch. Under the name of a theory of *Lage* or *situs*, the "direction-theory" (as it was later called) was born much earlier and had slowly transformed into its nineteenth-century counterpart. The appreciation of this fact is not only relevant in offering a more exact genealogy of the theory. Rather, it offers a rationale for its invention, since it can be shown that the direction-theory is deeply rooted in the logical and epistemological discussion that took place in Germany in the last three decades of the eighteenth century. Without this early history, the reasons for the success and demise of the direction-theory are destined, I claim, to remain unfathomable.

In Sect. 20.2 of the present essay, I detail the background of the direction-theory and its roots in Leibniz' program of an *analysis situs*. In Sect. 20.3, I deal with Karsten's invention (1778) of the direction-theory in the context of a broadly Wolffian epistemology. In Sect. 20.4, I mention the first reactions to Karsten's theory and in particular Hindenburg's own attempt and his philosophical qualms on intuition. In Sect. 20.5, I briefly discuss the impact of Kant's new philosophy of mathematics on the direction-theory, and the beginnings of the analyticity-syntheticity debate. In Sect. 20.6, I introduce Schwab's theory, that soon became the standard view of the subject. In Sect. 20.7, I briefly mention the reception of the direction-theory among later philosophers and its general acceptance in Germany.

20.2 Leibniz, Kästner, and the Analysis Situs

The direction-theory of parallel lines did not declare its name before its full development. The early theorists did not put any special emphasis on the notion of direction (*Richtung*) and rather defined parallels through the concept of *situs*

(German: *Lage*). Over the years, the term '*Richtung*' became more and more common – to the point that it gave the name to the whole theory in the 1810s.[13]

I am not aware of any important theory of parallel lines grounded on the notion of direction (or *situs*, or similar) predating the German attempts in the eighteenth century. The only exception is possibly offered by the German mathematician Nikolaus Kauffmann (Latin: Mercator, 1620–1687), who in 1678 published a reworking of Euclid's *Elements*. In it, Mercator took the lead from Euclid's definition of an angle as the *inclination* (κλίσις) between two lines, to define parallel lines as lines that *do not incline* the one towards the other. Following this definition, Mercator states an axiom to the effect that if two lines incline in the same way towards a third, they do not incline the one towards the other (i.e. are parallel). From this axiom he easily proves *Elements* I, 30 (the transitivity of parallelism) and from this all other standard properties of parallel lines in Euclidean geometry.[14] Mercator's attempt has many points in common with the further developments of the direction-theory but I have not been able to trace a direct filiation from these 1678 *Elementa* to the German theory of parallels presented a century later.[15]

By contrast, an explicit, and yet fully fabricated, filiation may be found between the direction-theories and the geometrical essays of Gottfried Wilhelm Leibniz (1646–1716). Leibniz worked for his entire life on a new mathematical theory, that he called *analysis situs*, aimed at grounding all geometry on the notion of "situation". Leibniz did not publish any of the many essays that he wrote on the subject. In the first half of the eighteenth century, however, a few letters that he had sent to Huygens and Johann Bernoulli were published, and the German world was informed of Leibniz' grand geometrical project – a lost science which left no traces. The imagination of several mathematicians was tickled, and in the course of the following two centuries many different mathematical endeavors (combinatorial geometry, vector calculus, projective geometry, topology) were developed under the name of Leibniz' mysterious *analysis situs*. The theory of parallels made no exception, and a long-lasting narrative was engendered, according to which Leibniz had in fact embraced a direction-theory of parallels.[16]

[13] The notion of direction was very much open to debate in the first half of the seventeenth century, and a definition of parallelism through direction was no simple matter, as it involved several commitments on the nature of space as an "affine" (rather than centered) structure. On the cosmological debate that brought the notion of direction to the core of the Copernican Revolution, see Miller (2014).

[14] Mercator (1678). The definition of parallelism is on p. 2. The axiom is the third one on the same page, and the easy proof of *Elements* I, 30 happens as Theorem 7 on p. 7.

[15] A few decades later, Edmund Scarburgh (1705) also talked about the "Tendency and Inclination towards one another" of non-parallel lines, even though he did not rely on transitivity to prove the Postulate. This could be a further source for the British direction-theorists of the nineteenth century.

[16] For a late assessment of Leibniz's theory of parallels as a theory of direction, see Killing (1893–1898), vol. 1, p. 5.

Unbeknownst to the direction-theorists, Leibniz had indeed strived to prove the Parallel Postulate through his *analysis situs*. Among his unpublished papers, now preserved in the *Leibniz-Archiv* in Hannover, we find many different attempts at establishing Euclidean geometry.[17] Several of these attempts are quite ingenious, and the notion of *situs* is employed in them in unexpected ways. None of them, however, ever attempted to define parallel lines as lines having the same situation, or to assume the transitivity of parallelism through the notion of situation (or direction). The standard references made by the direction-theorists to Leibniz' *analysis situs* were therefore entirely unwarranted.

While Leibniz' geometry had no authentic impact on the eighteenth-century theory of parallels, his epistemology exerted an important influence on the debate on the theory of direction. Several generations of German philosophers and mathematicians (and especially Christian Wolff, 1679–1754) shared Leibniz' views on philosophy of science. Leibniz' idea that all axioms of geometry could be proven starting from the definitions of the terms employed, had a pivotal role in orienting the German debate towards the search for a definition of parallel lines that could improve on Euclid's. Leibniz' ideas on the analyticity of truth brought mathematicians to disregard intuition in geometry. Leibniz' insistence that geometry is the science of space (rather than the science of the individual figures in it), and that space is a complex structure of situational relations, offered a completely new perspective on the meaning of the Parallel Postulate. In short, Leibniz remained a looming figure throughout all the discussions on the theory of parallels that took place in Germany in the crucial years 1770–1830. The great majority of the direction-theorists styled themselves as Leibnizians.

We can follow in detail how Leibniz' heritage inspired the birth of the direction-theory. In the second half of the eighteenth century, the most famous mathematician in Germany was a devoted Leibnizian: the Göttingen professor Abraham Gotthelf Kästner (1719–1800). After having attempted to prove the Parallel Postulate himself for many years, Kästner became disillusioned with obtaining a demonstration of it with standard mathematical tools. He collected a large number of treatises dealing with parallels, and turned with great hopes to Leibniz' analysis of situation.[18] One of his students, Georg Simon Klügel (1739–1812), defended a dissertation in which he took advantage of his professor's library, and expounded a good number of failed attempts at proving the Postulate – including several by living mathematicians. This dissertation, the *Conatuum praecipuorum theoriam parallelarum demonstrandi recensio*, was published in 1763 and attracted considerable attention. Kästner added a note at the end of his student's book:

[17] Leibniz' papers on the theory of parallels have been published in De Risi (2015).

[18] See the preface to Kästner (1758), as well as §§ 27–28, pp. 13–14, on the provability of axioms from definitions. On Kästner's involvement with the Parallel Postulate, see Peters (1962).

> I hardly hope that we will ever obtain the true demonstration [of the Parallel Postulate] – of which you, bringing the light of geometry, have vanished the specters – unless we cultivate more assiduously the theory of *situs*, the analysis of which perished with Leibniz.[19]

Kästner also remarked that since all axioms of mathematics are grounded on, and provable from, the definitions of the terms, the major challenge for a theory of parallels was to find a new and appropriate definition of parallel lines.[20]

The gauntlet had been thrown in the name of Leibniz' *analysis situs*.

20.3 Karsten and the Birth of the Direction-Theory

The two main living mathematicians criticized by Klügel were Wenceslaus Johann Gustav Karsten (1732–1787), a professor at Bützow; and János András Segner (1704–1777), a Hungarian scholar who had been the first professor of mathematics at the University of Göttingen (and Kästner's predecessor in that chair) and later moved to Halle.

Segner had employed the notion of *Lage* (or *situs*) in his theory of parallels, but had made no use of it in his proof of the Postulate in his 1756 *Cursus mathematicus*.[21] Segner's proof implicitly assumed that any straight line, passing through a point inside an angle, cuts this angle. This principle is in turn a reformulation of the transitivity of parallelism, itself equivalent to the Parallel Postulate, and in hyperbolic geometry a straight line may be entirely contained within an angle. Klügel was not able to pinpoint Segner's mistake, and concentrated on a marginal issue of no consequence for the demonstration.[22] Segner may have taken note of Klügel's criticism, since in 1764 he slightly revised his proof in the German translation of the *Cursus*; but since the criticism was incorrect, the

[19] Klügel (1763), p. 33: "Habituros nos aliquando, veram eam cuius admoto geometriae lumine spectra dissipasti demonstrationem, vix speraverim nisi diligentius exculta doctrina situs, cuius analysis cum Leibnitio interiit".

[20] Kästner, however, seems to have thought that the culprit here was Euclid's famously obscure definition of a straight line, rather than Euclid's definition of parallels: "Der Grund, warum man in diesem Axiome [the Parallel Postulate] nicht die Evidenz der übrigen findet, ist ... daß man von der geraden Linie nur einen klaren Begriff hat, nicht einen deutlichen" (Kästner, 1790, p. 414). Kästner himself did not attempt to define parallel lines through the notion of direction.

[21] The reference to *Lage* in relation to the theory of parallels occurred already in Segner (1747), p. 218; in this treatise there is a rather naïve proof of the Parallel Postulate. Segner's more mature proof of the postulate was expounded in his Latin treatise on mathematics. See Segner (1756): in § 11 of the section on *Geometria*, p. 144, is Segner's assumption on angles; the proof of the Parallel Postulate, depending on such an assumption, is in § 31, pp. 150–51.

[22] Klügel (1763), § 11, pp. 15–16.

new demonstration did not fare any better.[23] The first complete explanation of Segner's *petitio principii* was offered by Johann Friedrich Lorenz (1738–1808) only in 1791.[24]

Karsten, by contrast, was struck hard by Klügel's *Recensio*. In 1758, he had published a work in which he accepted Segner's proof of the Postulate. By 1760, however, he had also offered an alternative demonstration of it, that was loosely inspired by a standard (and faulty) proof offered by the Persian mathematician Nasīr al-Dīn al-Tusi in the thirteenth century, which had later become commonplace in the European literature on the subject.[25] Klügel criticized this latter proof with valid reasons, and Karsten realized that he needed a different demonstration.[26] In 1778, Karsten was called to Halle to take Segner's chair, and he took this momentous occasion to give an inaugural speech on the theory of parallels – which was published as a *Versuch einer völlig berichtigten Theorie von den Parallellinien*. In this important paper, Karsten publicly recognized that Klügel's "*bekannte Disputation*" has disproven his and Segner's proofs, and accepted Kästner's suggestion of developing Leibniz' theory of *Lage* in order to ground a novel approach to the theory of parallels.

Karsten proposed a new definition of parallel lines that he intended to supplant Euclid's. To this effect, Karsten introduced the notion of the *Lage* of a straight line. He claimed, with Leibniz, that a general definition of *Lage* cannot be given, as this is one of the most basic notions that we make use of in geometry: a *simple concept*, in fact, that admits of no conceptual analysis. This notwithstanding, it is possible to give an implicit definition of the *Lage* of a straight line by stating the conditions under which two straight lines have the same *Lage*. Needless to say, the relation of the sameness of *Lage* is called *parallelism*, and two lines having the same *Lage* may be called *parallel lines*. Karsten made a comparison with the simple notion of a magnitude, which cannot be explicitly defined either, but may be implicitly defined through congruence (i.e. sameness of magnitude).[27] We can note that Karsten's

[23] Segner (1764), § 259, pp. 198–200. The preface to this work is dated 1763, and it is not obvious that Segner had read Klügel's *Recensio* at the time. Note that Segner published in 1767 a second edition of his 1747 treatise without relevant changes to the simple proof expounded there.

[24] Lorenz' aim in this book was to vindicate Segner by making explicit the hidden assumptions in his work. Therefore, he assumed the principle on the straight line inside an angle as an *axiom* that he thought to be self-evident and much clearer than the original postulate of Euclid. See Lorenz (1791), where the new principle is stated in § 44 of the section on *Planimetrie* (pp. 102–103), and the Parallel Postulate is proven, as Prop. 10, a few pages later (§ 83, pp. 118–21).

[25] The first proof is in Karsten (1758); Segner's assumption on angles is here accepted in § 73 (pp. 34–35), and the Parallel Postulate is proven in § 76, p. 36. The second proof is in Karsten (1760), § 91, pp. 31–35. Al-Tusi's proof had been expounded by Clavius, Wallis, Arnauld, and others, and was well-known in the eighteenth century.

[26] Klügel (1763): §§ 11–12, pp. 15–17, on Karsten's first proof from 1758; and § 8, pp. 12–13, on Karsten's demonstration from 1760.

[27] Even though Karsten is not quoting Leibniz explicitly, he is clearly drawing on the point made at § 47 of Leibniz's *Fifth Letter* to Samuel Clarke (1716), in which Leibniz claims that one cannot

argument is not much different from Frege's introduction of the notion of direction by abstraction, even if their epistemological motivations were completely different.

The sameness of *Lage*, according to Karsten, may be rigorously constructed in geometry. This is the core of Karsten's theory, which is neither a purely analytic discussion based on definitions, nor a mere appeal to the intuition of parallel lines.

Karsten writes that the relative position (the *Lage*) of two intersecting straight lines is expressed by their angle of incidence (as in Mercator's theory, mentioned above). From this it follows that each of two straight lines which form equal angles with a common transversal has the same *Lage* with respect to this transversal. So far, so good; but then Karsten changed the relation of "having the same *Lage* with respect to a third line" into that of "having the same *Lage*" *überhaupt*. That is to say, he claimed that two straight lines which form equal angles with a transversal have the same *Lage* with respect to one another, irrespective of the specific transversal that has been used to establish this relation. Thanks to this unjustified assumption, Karsten claimed that two straight lines which have the same *Lage* with respect to a certain transversal, will also have the same *Lage* with respect to any other transversal. Given Karsten's definition, this amounts to saying that two straight lines forming equal angles with one transversal will also form equal angles with any other transversal. This is a false claim in hyperbolic geometry, and one from which one can deduce the Parallel Postulate – as Karsten himself did in the subsequent pages.[28]

It should be remarked that Karsten himself did raise a few doubts about the soundness of his proof and was not completely satisfied with it. He claimed that the unanalyzability of the notion of *Lage* forces us to accept several principles which immediately flow from the nature of this simple notion. In this respect, Karsten introduced some latitude into the rather severe epistemology developed by Leibniz and Wolff.[29]

Karsten restated his theory without relevant changes in the 1780 edition of his textbook of elementary mathematics, and again in 1786, in an extended and

---

define the notion of "place" but only the notion of "having the same place" (see Robinet 1957, pp. 142–45).

[28] Dodgson did not refer to Karsten (nor to other earlier German authors) in the course of his books, but in his own attempt at proving the Parallel Postulate (1888, p. 69) he was crystal clear that the notion of the sameness of direction is in fact equivalent to the property that two parallels make equal angles with any transversal whatsoever – which is in fact Karsten's unwarranted assumption. As we know today, the latter assumption is only equivalent to the Parallel Postulate if we also accept the Axiom of Archimedes.

[29] Karsten's epistemology had already been stated in abridged form in his 1778 essay, where he explicitly claimed that his first six propositions are to be considered as axioms (pp. 14 and 16). He returned to the question, however, in §§ 1–16 of his essay *Von den Parallellinien*, included in Karsten (1786), pp. 115–30, which is a short essay in Wolffian epistemology – with a twist. In the latter essay, it is pretty clear that Karsten was not thinking about assuming the propositions on the *Lage* of straight lines as authentically unprovable axioms, but rather as statements that accept some kind of proof (*Beweis*) or at least of an explanation or an exhibition in a figure. He had no concept of different axiom systems, and believed that everything about Euclidean geometry could be justified in one way or another.

(slightly) improved version, as an essay *Von den Parallellinien* included in his monumental *Mathematische Abhandlungen*.[30] In the following years, Karsten's books enjoyed a large diffusion, and the direction-theory became well-known in the German world. We find traces of Karsten's theory even in technical schools and in the teaching at military academies.[31]

Karsten did not call his own theory a "direction-theory" of parallels, even though he often equated the notion of *Lage* with that of the *Richtung* (direction) of the straight lines.[32] There can be no doubt, however, that in his inaugural lecture of 1778 he exploited for the first time a line of argument that in the following years transformed into the fully-fledged theory of direction later criticized by Dodgson and Frege.

20.4 Hindenburg and Transitivity

After Karsten, the most important step forward in the direction-theory of parallelism was made by the Leipzig professor Carl Friedrich Hindenburg (1741–1808), better known as the leading figure of the German group of mathematicians working on combinatorics.[33] In his 1781 *Neues System der Parallellinien*, Hindenburg made good use of a few of Karsten's ideas and attempted to prove the Parallel Postulate through a different route.[34]

[30] Karsten (1780), pp. 383–414; the 1768 first edition did not include a theory of parallels. Karsten (1786), pp. 113–202.

[31] For instance, the Meinert's military textbook (1790) briefly mentions many attempts to prove the Parallel Postulate but states that Karsten's theory is the best one, and "Die vorstehende Theorie der Parallellinien ist völlig die Karstensche [. . .]. Fast sollte man glauben, wenn diese noch nicht alle Schwierigkeiten gehoben hat, daß sie schwerlich durch Hülfe der Elementargeometrie gehoben werden können. Auf die ausübende Mathematik haben die bisher erregten Zweifeln gegen die völlige Nichtigkeit der euklideischen Theorie der Parallellinien keinen Einfluß" (pp. 59–60; I thank Thomas Morel for pointing out this passage to me). But even a more theoretical work (even though practically oriented) such as Schmidt (1797), still followed Karsten's proof (cf. pp. 132–34).

[32] This is clear already in Karsten's essay from 1778. It should be noted, however, that the notion of *Richtung* is especially used by him in the later presentation of his theory, to be found in his *Mathematische Abhandlungen* from 1786. Here he still distinguished between *Lage* and *Richtung*, in the weak sense that in lines with the same situation two different directions may be spelled out (say, toward the left or the right side); more often, however, he simply wrote "*Lage oder Richtung*" as synonyms.

[33] For a biographical sketch of Hindenburg and a detailed analysis of his involvement with German combinatorics (as well as his relation with Leibniz' thinking), see the recent Noble (2022).

[34] Hindenburg (1781). The first section, *Über die Schwürigkeit bey der Lehre von den Parallellinien*, is a collection of criticisms of previous proofs of the Parallel Postulate (Hindenburg knew Klügel's *Recensio*); the second section, *Neues System der Parallellinien*, contains Hindenburg's own proof; and the third section, *Anmerkungen über das neue System der Parallellinien*, published in a later issue of the journal (still in the same year) replies to a few criticisms that had been

Hindenburg's epistemology was also different from Karsten's. He was critical towards the Leibnizian tradition in logic, and rejected the idea that an axiom should be proven from definitions.[35] This notwithstanding, Hindenburg appreciated the theory of *Lage*, and agreed with Kästner and Karsten that a fully developed *analysis situs* was required to handle the theory of parallels. The latter theory, however, had to be freed from the constraints of Leibniz' epistemology and developed through constructive, ruler-and-compass constructions.

Hindenburg maintained that geometry has two parts: one, which had been developed since ancient times, dealing with magnitudes; and another, instantiated by his own studies on combinatorics, dealing with *Lage*. Karsten had failed to keep these two branches apart and committed a μετάβασις εἰς ἄλλο γένος of sorts by defining the *Lage* of a straight line through a reference to the *magnitude* of an angle. The geometry of *Lage*, and the theory of parallels in particular, should be developed without any recourse to the notion of quantity. Accordingly, Hindenburg put much emphasis on the transitivity of parallelism, that seemed to him a purely situational (non-metric) property.

In sum, Hindenburg's program in the theory of parallels aimed to offer a geometrical (rather than a merely logical) and non-metric demonstration of the transitivity of parallelism.

Hindenburg considers two straight lines, *a* and *b* both parallel to a third line *c*, and attempts to prove that *a* and *b* are also parallel to one another. The demonstration is articulated in two different cases, depending on the position of line *c*. In Case One, lines *a* and *b* lie on different sides of the common parallel *c*. Hindenburg disposes of this case quickly, stating that if *a* and *b* were not parallel, the straight line *c* would meet one of them in the direction of their intersection, against the hypothesis. In the more difficult Case Two, *a* and *b* are both on the same side of the common parallel *c*, and Hindenburg builds a complex network of logical implications to bring back this case to the previous one, thus proving the theorem.

The system of logical implications underpinning Case Two was seen as the most problematic part of Hindenburg's proof. Several German mathematicians levelled objections to the logical form of the demonstration, and Karsten himself was among its fiercest opponents.[36] Hindenburg got caught in a difficult logical controversy. He published an essay *Noch etwas über die Parallellinien* (1786), almost entirely

levelled against the *Neues System* in a review which had appeared in the *Königsbergsche Gelehrte und Politische Zeitungen*.

[35] It is possible that Hindenburg's epistemological ideas were indebted to Johann Heinrich Lambert (1728–1777), who had sharply criticized Wolff's abuse of definitions in mathematics. Hindenburg had been a correspondent of Lambert in the latter's last year of life; their letters are to be found in Lambert (1781–1787, vol. 5.1, pp. 137 ff.), and do not concern geometry. In the *System*, Hindenburg explicitly quotes both Lambert's *Briefwechsel* (whose first volume had been published by Johann III Bernoulli in the same year 1781) and the *Neues Organon*. Later on, in 1786, Hindenburg published for the first time Lambert's important *Theorie der Parallellinien*, written in 1766 and left by him in manuscript form. On Lambert, see Gray and Tilling (1978).

[36] See Karsten, *Von den Parallellinien*, §§ 33–45, in Karsten (1786), pp. 145–162.

dedicated to the matter, and in the course of the debate he became convinced that the peculiar logical reasoning employed in Case Two was a form of *consequentia mirabilis*, a correct inferential scheme that had been employed in the theory of parallels by Gerolamo Saccheri in 1733.[37]

This entire logical dispute eventually proved to have been in vain. Even though no mathematician recognized it for many years, the problem of Hindenburg's proof was not in the difficult Case Two, but rather in the "self-evident" Case One – that no one discussed. In it, Hindenburg tacitly assumed that the straight line *c* cannot be entirely contained within the angle formed by *a* and *b* (if they are not parallel), and this assumption is equivalent to the Parallel Postulate. In fact, this was just Segner's presupposition in his fauly proof from the 1750s.

While Hindenburg's proof was believed wrong for the wrong reason, his engagement in a logical dispute showed how the alleged superiority of his mathematical method over Karsten's more logical approach was an illusion. As a matter of fact, Karsten "analytic" proof of the Parallel Postulate was much more constructive and geometrical than Hindenburg's "synthetic" demonstration. The Karsten-Hindenburg dispute was the prelude to a much wider debate on the role of intuition in geometrical proofs.

20.5 Kant and the Analyticity Debate

In the same year, 1781, in which Hindenburg published his essay on parallel lines, Immanuel Kant (1724–1804) printed the first edition of his *Critique of Pure Reason*. In the following 10 years, the philosophical landscape of Germany was completely transformed, and the debate on the analyticity and the syntheticity of mathematics came to the forefront. Kant's statements on synthetic *a priori* judgments and space as a pure intuition engendered a number of reactions. Leibnizian philosophers restated that no recourse to intuition was needed to prove mathematical theorems, and attempted to offer purely analytic demonstrations on the theory of parallels. Geometry was taken as the benchmark of Critical philosophy, and the direction-theory as the crowning effort in the foundations of mathematics. At times, it seemed that Kant's whole philosophy had to stand or fall according to whether the direction-theory itself fell or stood.

Kant was aware of Karsten's attempts to develop a direction-theory of parallels, and in private correspondence and notes, he explicitly criticized his analytic approach to geometry. In particular, Kant complained that the notion of direction (*Richtung*) cannot be defined without the concept of a straight line, and that,

[37] Hindenburg's defense of his Case Two had already begun in his *Anmerkungen* from 1781, which are almost entirely devoted to this purpose, but the main efforts in this direction came in the second section of his essay *Noch etwas über die Parallellinien* from 1786. On the history of the *consequentia mirabilis* in geometry, see my edition of Saccheri (2014).

therefore, one cannot define straight lines (or parallel straight lines) through direction – a remark, which is not much different from Frege's later claim on the same subject.[38] Kant agreed that the notion of *Lage* and Leibniz' famous *analysis situs* may perhaps be employed to prove the Parallel Postulate, but was skeptical on a non-metric approach to the notion of parallelism.

In the 1780s, Kant himself attempted to prove the Parallel Postulate, and seems to have arrived at the conclusion that no synthetic *a priori* proof of it can be given. This was not, however, a statement of absolute indemonstrability, and much less an opening towards non-Euclidean geometries. Kant thought, on the contrary, that a purely analytic proof of the Postulate (that he called a "philosophical proof") could be obtained starting from a viable definition of parallel lines. This proof could be an example of Leibniz' lost *analysis situs*, here interpreted as a metric theory, but Kant's reflections on the topic are unfortunately too brief and obscure to offer a perspicuous mathematical meaning. In any case, given the epistemology professed in the *Critique*, and the criticisms that Kant moved to Karsten's approach, the recourse to a philosophical proof to prove a mathematical theorem is quite an astonishing claim. Kant never published his thoughts on parallels and his tentative demonstration remained buried among his private notes.[39]

But while Kant did not publicly express his views on the theory of parallels, this was repeatedly done by the mathematician Johann Friedrich Schultz (1739–1805), who was a friend and follower of Kant, as well as one of the most prolific authors of proofs of the Parallel Postulate. In 1784, Schultz published his first important treatise on the topic, the *Entdeckte Theorie der Parallelen*, in which he expounded a theory based on the manipulation of infinite magnitudes that Kant himself found dubious and untenable. In the same essay, however, Schultz also strongly criticized both Karsten's and Hindenburg's attempts on the basis of Kant's philosophy of mathematics. He claimed that their proofs were entirely grounded on the *analysis* of the notion of *Lage*, and merely deduced from a definition what they had themselves

---

[38] The context of Kant's claim is a reply to the philosopher Salomon Maimon (1753–1800). In his *Versuch über die Transzendentalphilosophie* from 1790 (pp. 65–68), Maimon endorsed a broadly Leibnizian epistemology of mathematics with several Kantian nuances. He attempted to show that the proposition stating that a straight line is the shortest line between two points, which Kant had famously claimed to be synthetic and thus irreducible to the definition of a straight line (*KrV*, B16), could in fact be proven by conceptual analysis alone. To this effect, Maimon defined a straight line as a line such that every part of it has the same direction or *Lage*. Maimon had the opportunity to send the draft of his philosophical essay to Kant himself, who replied by criticizing Maimon's definition (Kant to Herz, May 26th, 1789; in KgS 11, pp. 48–54). In the very same years of Dodgson's criticisms of the direction-theory, Kant's opinion on direction was endorsed by Helmholtz in his 1878 presidential speech at the University of Berlin on *Die Tatsachen in der Wahrnehmung*. See Helmholtz (1921), p. 182.

[39] Kant himself attempted to prove the Parallel Postulate, making use of the metrical notion of "equidistance" and was convinced that this was the only viable way to address the issue. In this connection, he mentioned the *Geometrie der Lage*. See the *Reflexionen* nn. 8–10 in KgS 14, pp. 33–51, probably dating from 1784 to 1790. Kant's standard distinction between philosophical proofs and mathematical demonstrations is to be found (among other places) in *KrV*, A734–35/B762–63. On the topic, see De Risi (2013) and Heis (2020).

introduced in it. By contrast, geometry should be grounded on synthetic *a priori* judgments, which are the only ones that can actually extend our knowledge.[40]

Karsten, who honestly believed that any genuine geometrical theory must be analytic, did not reply to Schultz' criticisms. He did state, however, that he could not accept Schultz' demonstration, since arguments employing the infinite can only provide probable conclusions. Any strict mathematical proof should be analytic and finitistic.[41]

Hindenburg, on the other hand, was outraged by Schultz' allegations, which were not very different from those he himself had raised against Karsten. This time, he defended Karsten against Schultz, claiming that Karsten's new principles on the theory of *Lage* were to be understood as true synthetic axioms rather than analytic consequences of the definition of *Lage*. Karsten had not proven these principles, Hindenburg added, but he himself had done so in a geometric way: thus synthetically proving the Parallel Postulate and securing the direction-theory. Hindenburg also retorted Schultz' accusations, and stated that a theory of parallels like the one Schultz was advocating – employing the notion of infinite magnitudes – was wholly philosophical and un-mathematical. *This* was an analytic theory if there ever was one.[42]

Finally, Hindenburg dragged Kästner into the dispute, by publishing a letter of his in which Kästner criticized a proof of the Parallel Postulate advanced by the Swiss mathematician Louis Bertrand (1731–1812), that was very similar to Schultz'. Pushed by the events, Kästner took up the pen himself and, as a good Leibnizian, strongly criticized Schultz' theory of parallels as a monstrosity and restated, against Schultz' master Kant, that geometry is wholly analytic.[43]

Kant was unhappy to have been drawn into the fight. He had avowed Schultz' theory of parallels and did not want to be attacked by a famous mathematician such as Kästner. In 1790, he wrote a dense and deep reply to him, that he transmitted to Schultz with the request of publishing it in his name. Shultz did publish Kant's text, but could not help adding a conclusion with a further endorsement of the theory of infinite magnitudes.

---

[40] See for instance the appreciation of Kant's mathematical epistemology in Schultz (1784), pp. 27–28. Schultz' criticisms of Karsten and Hindenburg are here in pp. 31–65. Discussing the analyticity of Karsten's proof, Schultz remarked "daß es aber überhaupt schlechterdings unmöglich sey, die Lehre von den Parallellinien *durch bloße Analysirung des Begrifs ihrer Lage* festzustellen" (p. 41, my emphasis).

[41] Karsten's treatment of infinity was intended as an answer to a *Preisaufgabe* of the Berlin Academy (probably suggested by Lagrange), which had asked, in 1784, for papers dealing with the mathematics of the infinite. Karsten also explicitly discussed Schultz' attempt in §§ 52–76 of his essay *Von den Parallellinien* in Karsten (1786), pp. 168–202.

[42] Hindenburg's reply can be read in Hindenburg (1786, pp. 392–97); Hindenburg stated that Schultz's proof was merely philosophical rather than mathematical on p. 368 of the same essay.

[43] Bertand's proof is in (Bertrand 1778), vol. 2, p. 20. For Kästner's letter on Bertrand, see Hindenburg (1786). Kästner's views on infinity had already been expounded in his essay *De vera infiniti notione*, to be found in Kästner (1771, pp. 35–38). Kästner's anti-Kantian essays are to be found in Kästner (1790). Kant is never explicitly mentioned in them, but Schultz is.

Kant was extremely disappointed by Schultz' insubordination, as he feared a further rejoinder by Kästner. When it did not arrive, and he and Kästner exchanged polite private letters, Kant hoped – for a moment – to have settled the matter.[44]

He was greatly mistaken.

20.6 Schwab and His Critics

Schultz' reply had awakened the most relentless adversary of Kant's philosophy of mathematics and the most prominent proponent of the direction-theory of parallels.

Johann Christoph Schwab (1743–1821) was one of the founders of the *Philosophisches Magazin*, a journal that soon became the anti-Kantian organ in Germany. In a 1791 issue of the journal, Schwab published a paper against Schultz, and offered an analytic proof of the fact that one side of a triangle is shorter than the sum of the other two. The real target of the paper was Kant, who however decided not to reply. Other Kantians took up the fight, and Schwab's proof was later rebutted by August Wilhelm Rehberg (1757–1836). Rehberg, who was an important correspondent for Kant in mathematical matters, exposed several hidden assumptions in Schwab's proof, and restated the position that, without an appeal to an *a priori* intuition, one cannot hope to prove such basic geometrical statements. Schwab counter-replied to Rehberg's criticisms, and wrote further papers against Schultz.[45]

A long debate arose in Germany concerning Kant's philosophy of mathematics, the role of intuition in geometry, and the analyticity of proofs. The philosophical positions of the protagonists had many nuances, which cannot be recounted in full here, and sometimes the terms of discussion were uncertain. For instance, Schwab claimed against Rehberg that a single analytic proof of a geometrical theorem was enough to ruin Kantianism, whereas Rehberg maintained that showing the non-analyticity of a single proof was sufficient to smite "Leibnizians" down. As a result, they played with *ad hoc* examples of analytic and synthetic proofs, without being able to settle the matter or find common ground for decision.

---

[44] Kant's reply to Kästner is now to be found in KgS 10, pp. 410–23. Kant complained to Schultz, in his letters from August 2nd and 16th, 1790, in KgS 11, pp. 184 and 200–201. The happy ending of Kant's controversy with Kästner is witnessed to by a very respectful exchange in the same months (Kant to Kästner, August 5th, 1790, in KgS 11, p. 186; Kästner to Kant, October 2nd, 1790, KgS 11, pp. 213–15).

[45] Schwab's first paper is *Über die geometrischen Beweise*, published in the *Philosophisches Magazin* from 1791. See above for Maimon's attempt at the same result. Rehberg's reply to Schwab is the paper *Über die Natur der geometrischen Evidenz*, published in the *Philosophisches Magazin*, from 1792. Schwab counter-replied to Rehberg in the same issue of the journal (pp. 461–69), but the discussion continued for many years. A further attack of Schwab against Schultz, for instance, is to be found in his *Über das Unendliche des Herrn Hofpredigers Schultz* in the *Philosophisches Archiv* from 1792. On the Schwab-Rehberg discussion, see Webb (1987).

A decade later, Schwab decided that the time was ripe for a final assessment, and began writing books presenting his views on the analyticity of mathematics. He set himself no less an aim than to resuscitate Leibniz' *analysis situs* and prove the Parallel Postulate.

Schwab's first work on the topic is a *Tentamen novae parallelarum theoriae*, published in 1801, in which he presented his epistemological views and a new theory of parallels. This was later supplemented by a longer essay in French, the *Essai sur la situation* from 1808, which expounded Schwab's more general views about Leibniz's project of an *analysis situs*. And finally by a *Commentatio* on the First Book of the *Elements* (1814), in which Schwab attempted to reform elementary geometry and prove all the axioms in Euclid.[46]

Schwab followed Hindenburg's idea that a theory of *situs* (i.e. *Lage*) should complement the geometry of magnitudes, and that the whole of mathematics should be grounded on the two independent pillars of quantity and situation. These are simple notions, as Karsten had said, and cannot be defined. Schwab did not provide a definition of them by abstraction, and instead conceded to his adversaries that they can only be given in intuition. This does not mean, however, that geometry is based upon intuition, and much less upon synthetic *a priori* judgments.[47] Schwab rather endorsed an epistemology of eidetic abstraction, and stated that we are able to develop, out of the intuitions of *situs* and magnitude, a purely intellectual science which needs no further reference to the senses. Geometry is thus sensible in its origins, but nonetheless purely intellectual in its developments, and should rely solely upon sound logical reasoning without any recourse to any intuitive or diagrammatic support. Intuition does not provide us with any propositional content, and all axioms of geometry are provable from definitions and logic alone.

Schwab supplemented this broadly Leibnizian and Aristotelian picture of science with a number of mathematical developments. He defined a straight line as a line "all the points of which have the same situation" (whatever this may mean), and parallel lines as straight lines having the same reciprocal situation.[48] Finally, Schwab claimed that the statement that two things identical to a third are also identical to one another (i.e. the transitivity of identity) is a general logical principle that enjoys no less validity than does the Principle of Contradiction itself. From its

[46] Schwab was a prolific writer, and published many more books on philosophy, often discussing Kant's views (sometimes critically, and sometimes agreeing with him); in a few of them he restated his geometrical examples and the theory of parallels. Nonetheless, he did not write those eight volumes on the theory of parallels that De Morgan found somewhere attributed to him: "*Eight* volumes on the theory of parallels? If there be such a work, I trust I and it never meet, though ever so far produced" (De Morgan, 1915, vol. 1, p. 230).

[47] On Kant, see for instance Schwab (1808), pp. 28–29; (1814), §§ 5–6, pp. 8–10.

[48] Schwab gave no definition of a straight line in the *Tentamen*, and the latter was only added in the *Essai*: "Une ligne droite est donc celle dont tous les points ont la même situation" (1808, p. 20). His definition of parallel lines in the *Tentamen*, "Duae lineae rectae in eodem plano jacentes sunt parallelae inter se, si eundem situm habent ad se invicem" (1801, p. 1), is, in turn, identical to that in the *Essai*: "Deux droites sont parallèles, lorsqu'elles ont la même situation entr'elles, ou lorsque la situation de l'une est identique avec celle de l'autre" (1808, § 28, p. 26).

application to the simple notions of magnitude and *situs,* there immediately follows Euclid's first Common Notion ("if two things are equal in magnitude to a third, they are equal in magnitude to one another") and the transitivity of situation ("two things that have the same situation with respect to the same thing, have the same situation with respect to one another"). The latter principle, applied to the definition of parallel lines, entails the transitivity of parallelism and thus (as Schwab correctly proved) the Parallel Postulate itself.[49]

Schwab's "demonstration" collected a few aspects of the previous studies on the direction-theory, and offered a purely philosophical (and mathematically trivial) proof of the Parallel Postulate embedded in a rich, if eclectic, epistemology. For this very reason, Schwab's construal of the theory of parallels represented an important standpoint in the German debate. His epistemological views were clearly stated and defended and he was considered for many years to be the Leibnizian champion against the new wave of Kantianism in mathematics.

The historical relevance of Schwab's proposal is also due to the criticisms that his theory elicited among mathematicians close to Gauss.

Karl Felix Seyffer (1762–1822), an important astronomer in Göttingen who was one of Gauss' teachers, reviewed Schwab's essay only to conclude that the only correct attitude toward the Parallel Postulate was to accept it as an unprovable principle, since there is no salvation outside the Church of Euclid ("*nulla salus extra Euclidem*").[50]

Ferdinand Carl Schweikart (1780–1857) is particularly remembered nowadays since he was among the first to propose, in a private letter from 1818, the consistency of an "astral geometry" (*astralische Grössenlehre*) in which the Parallel Postulate is false. He also realized that there are in fact as many different astral geometries as the real values of a *Constante* which, with quite some interpretative generosity, may be related to Gaussian negative curvature.[51] Schweikart's early treatise on parallel lines (1807), however, was very much concerned with Schwab's theory of direction, which he criticized harshly and at great length. On that occasion,

[49] In the *Tentamen*, Schwab simply stated the transitivity of parallelism as an axiom ("Axioma: Si duae rectae in plano eundem situm habent ad se invicem; habent etiam eundem situm ad rectam tertiam", p. 3), followed by a discussion about the grounds of its validity. In the following works, however, and especially in the *Commentatio*, it is very clear that every axiom has to be proven from purely logical principles. A Greek source of Schwab's proof may be Proclus, who mentioned in passing that some relations have the property of transitivity; in particular, "similarity" is considered by Proclus to be transitive in general, and parallelism is just a "similarity of position" (ὁμοιότης θέσεως); thus, parallelism is transitive and the Parallel Postulate true (Proclus, *In primum Euclidis* 373).

[50] Seyffer's review of Schwab's *Tentamen* appeared in Seyffer (1801).

[51] Schweikart's note was attached to Gerling's letter to Gauss from January 25th, 1819, and may now be found in Gauss (1863–1917), vol. 8, pp. 179–81, along with Gauss's first reaction to it (vol. 8, pp. 181–82).

Schweikart indulged in some Kantian or post-Kantian reflections on the need of mathematical intuition for any finite mind (*endlicher Geist*).[52]

Finally, Carl Friedrich Gauss (1777–1855) himself wrote a review of Schwab's *Commentatio* which proved in the end to be his sole (very limited) public expression of his thoughts on the Parallel Postulate. Against Schwab, Gauss claimed that definitions and logical principles

> are able to accomplish nothing by themselves, and that they put forth only sterile blossoms unless the fertilizing living intuition (*Anschauung*) of the object itself prevails everywhere.[53]

This statement appeared to many scholars to be an endorsement of that very Kantianism which Schwab was attempting to wreck. Gauss' appeal to intuition was in all probability addressed toward an *empirical* acquaintance with the structure of space rather than an *a priori* intuition, and elsewhere he decidedly opposed Kant's views on mathematics, advocating rather an empiricist stance on the matter. Still, in the absence of further evidence, Gauss' criticisms of Schwab's anti-Kantianism appeared to foster Kant's transcendental epistemology. In later times, when his views on non-Euclidean geometries were finally disclosed to the public, a few neo-Kantian philosophers were able to claim – thanks to this review of Schwab – that Gauss had admitted the compatibility between non-Euclidean space and *a priori* intuition.[54]

20.7 Post-Kantian Philosophy and the Direction-Theory

These few dissonant voices notwithstanding, Schwab acted as a pied-piper to the whole scholarly world, and in the subsequent decades the direction-theory of parallels became mainstream in Germany. Like all pieces of standard science, the direction-theory was provided with a history and a name. In 1807, Johann Ephraim Scheibel (1736–1809) published a history of the theory of parallels from Euclid to his own time, which was crowned by the modern theory of *Lage*. He traced

[52] Schweikart's criticisms of Schwab are to be found in Schweikart (1807), which also refers to Kant.

[53] Gauss's review of Schwab's *Commentatio* was published in 1816 in the *Göttingische gelehrte Anzeigen*, and can now be read in Gauss (1863–1917) vol. 4, pp. 364–68. The quoted sentence is translated into English in Ewald (1996), vol. 1, p. 300. See Gauss's (much later) letter to Farkas Bolyai from March 6th, 1832, on the fact that the new geometrical discoveries made by János Bolyai (and Gauss himself) show that Kant was definitely wrong in believing that space is only a form of our intuition ("... der klarste Beweis, dass Kant Unrecht hatte zu behaupten, der Raum sei *nur* Form unserer Anschauung"; Gauss, 1863–1917, vol. 8, pp. 220–224). On Gauss' role (or lack thereof) in the discovery of non-Euclidean geometry, see Gray (2003).

[54] The first attempts to claim that the Kantian conception of geometry is compatible with non-Euclidean geometries are probably those by Nelson (1905–1906); cf. in the same years also Meinecke (1906); and later Natorp (1910).

the origin of the notion back to Euclid himself, who had exploited the notion of position (θέσις) in his book on *Data*.[55] In 1816, Carl Christian Hermann Vermehren (1792–1858) wrote an essay on the direction-theory in which the notion of *Richtung* supplanted that of *Lage* (or *situs*).[56] It is probably at this time that the theory began to be widely known with its current name.

In the meantime, the mathematician Karl Christian Langsdorf (1757–1834) updated Wolff's textbook for students, the *Anfangsgründe aller mathematischen Wissenschaften*, adding to it the theory of direction.[57] Among the other important and early essays on the theory, we can mention at least those by Andreas Jacobi (1801–1875) from 1824 and by Joseph Knar (1800–1864) from 1827 and 1828. The success of the theory was so extensive that even the Kantian philosopher Carl Siegmund Ouvrier (1751–1819) wrote an essay on parallel lines and the faculties of the mind, in which he accepted pure intuition as the basis of mathematics and also a theory of direction as the foundation of the theory of straight and parallel lines.[58]

We will not follow in any detail the diffusion of the direction-theory in German textbooks on geometry, which was as pervasive as it is mathematically uninteresting. Rather, I would like to point out the lesser-known fact that the most important philosophers in Germany plainly endorsed Schwab's theory.

In 1805, Johann Gottlieb Fichte (1762–1814) rejected Kant's views about the givenness of spatial intuition and claimed that space itself is generated by the thinking subject. He stated that the first determination of space is given by tracing an infinite straight line which is the archetypal diameter of space and the beginning of any further construction of finite magnitudes and figures. This line determines a "primal direction" (*Ur-Richtung*) in space, and therefore defines a first sheaf of parallel lines as those straight lines sharing the same original direction. As a consequence, the Parallel Postulate is firmly established upon the basis of the

[55] See Scheibel (1807). Scheibel's work includes a long discussion of Kästner's three essays on space and geometry which had been published against Kant and Schultz.

[56] Vermehren (1816): the discussion on direction is to be found on pp. 19–26, while on p. 21 Vermehren assumed an axiom regarding the transitivity of sameness of direction.

[57] Langsdorf edited the geometrical section of Wolff's *Anfangsgründe* in 1797, that is, before Schwab's contribution to the theory of parallels. Already in this book, Langsdorf mentioned the new theories by Karsten and Hindenburg, which he did not see as fundamentally opposing Wolff's original views. In 1802, Langsdorf wrote his own *Anfangsgründe*. In them, he started with a definition of the *Richtung* of a straight line (section on *Geometrie*, § 14, pp. 129–30) so as to be able to give a definition of parallel lines (§ 15, pp. 131–32), which, albeit grounded upon it, recurred to the notion of equidistance; a little further on in the text, the Parallel Postulate is proven (Theorem 14, pp. 173–76). In this work, written a year after Schwab's first essay on parallels, the latter's influence is manifest.

[58] The main theorem of Jacobi (1824) is on p. 58: "Per punctum quoddam una tantum recta unius ejusdemque certae directionis duci potest", which is "proven" without much discussion. Andreas Jacobi was the brother of the more famous mathematician Karl Friedrich Jacobi. Knar (1827) discusses several earlier attempts and quotes Karsten as the initiator of the theory that Knar himself advocated and tried to perfect. See also Knar (1828). A Kantian direction-theory is expounded in Ouvrier (1808).

original construction of space. The direction-theory of parallels appears to be the metaphysical foundation of space itself.[59]

Jakob Friedrich Fries (1773–1843) offered his own "synthetic" proof of the Parallel Postulate in 1822. This proof is far more mathematical than Fichte's but is still entirely based upon the notion of the direction of a straight line. Fries defined the straight line as a line the parts of which have the same *Richtung*, then complemented this definition by three *Axiome der Richtung*, and finally proved from all these assumptions that the interior angle sum of a triangle is equal to two right angles. The Parallel Postulate is an easy consequence of the latter fact.[60] Fries' philosophical influence on mathematicians was remarkable: Gauss appreciated his work, and Schlömilch (Frege's source on the direction-theory) was a student of his. No one, however, seems to have followed his quite naïve proof of the postulate.

Johann Friedrich Herbart (1776–1841) conceived of space as a structure produced by a system of monads, and for this reason he is sometimes credited for a philosophy of space that could possibly accommodate non-Euclidean manifolds. Historians generally regard him as the main philosophical source for Riemann's thinking, and speculate on his role in establishing the possibility of a plurality of geometries. Yet, in 1829 Herbart also claimed that the Parallel Postulate can be analytically proven – without any recourse to pure or empirical intuition – from the very notion of *Richtung*.[61]

Georg Wilhelm Friedrich Hegel (1770–1831), in the posthumous second edition of the *Wissenschaft der Logik* (1832), discussed at length the epistemology of mathematics and stated (in perfect agreement with the previous tradition) that any axiom can be proven starting from the definitions of the terms. The Parallel Postulate, in particular, may be proven from the definition of direction, or better of the sameness thereof, *die Gleichheit der Richtung*, even though Euclid and the other ancient writers had done well in assuming it without proof given the *plastische Charakter* of the beautiful endeavour of Greek science, which did not waste time on such trifles.[62]

---

[59] Fichte's discussion of the theory of parallels was not published during his lifetime, and is to be found among his lecture notes from 1805 called the "*Erlangen Logik*", in which he also mentions Schultz's proof, judging it to be unpersuasive. See Fichte (1962–, vol. II 9, pp. 124–37). An English translation of this text, with a commentary, is to be found in Wood (2012).

[60] Fries' discussion on the Parallel Postulate is to be found in (Fries 1822), §§ 66–69, pp. 355–380; the axioms of direction are on p. 369, the definition of a straight line on p. 376, the proof of the interior angle sum on p. 379. On Fries and geometry, see Gregory (1983).

[61] Herbart's statement on the Parallel Postulate is to be found in § 257 of the second volume of his *Allgemeine Metaphysik* (1828–1829), vol. 2, pp. 240–42. The extent of the influence of Herbart's philosophical views on Riemann's foundational mathematical works continues to be a subject of great debate. For an informed assessment, see Scholz (1982). On Herbart's theory of intelligible space, see Banks (2005).

[62] In his young years, Hegel studied a proof of *Elements* I, 29 (a proposition equivalent to the Parallel Postulate) through the definition of parallel lines as equidistant straight lines – in a broadly Wolffian perspective. These private notes have now been published in Hoffmeister (1936), pp. 288–300. Hegel's first public mention of the Parallel Postulate was in the third volume of the

None of these philosophers ever doubted of the truth of the Parallel Postulate. The influence of their views on the generations that followed was enormous, and it may explain the general acceptance of the direction-theory even outside the mathematical community. Still many decades later, in 1879, the same year in which Dodgson published his rebuke of Euclid's modern rivals, the illustrious philosopher Hermann Lotze (1817–1881) wrote a book in which he rejected non-Euclidean geometry on the basis of the direction-theory of parallels.[63]

The direction-theory of parallels was one of the most prominent phenomena in the history of the philosophy of mathematics between the eighteenth and the nineteenth centuries. We regard it nowadays as a huge misunderstanding, a conceptual dead end, and a reactionary force in the development of mathematics towards non-Euclidean geometry. Yet, the debate it engendered revolved around the most important topics in the epistemology of mathematics – abstraction, intuition, infinity, analyticity – and shaped the history of thought for a century. These magmatic reflections on the nature of mathematics were the philosophical background in which Lobachevsky, Bolyai and Gauss first conceived non-Euclidean geometry as an alternative theory of parallels. These same reflections, continuing into the following generations, produced the rebuttals by Dodgson and Frege, who – no longer pre-modern and yet not modern, and even possibly countermodern – seriously engaged the epistemology of a waning age. Modernism itself, eventually, came out of this debate.

The philosophical interest of the century of the direction-theory, with its supporters and its opponents, lies precisely in its being a passageway to modernity.

Wissenschaft der Logik (1816), where he claimed (in the section dedicated to the notion of a "Theorem") that all propositions of mathematics can be derived from their definitions, and the Parallel Postulate no less than any other. In the second edition of the first volume (1832; first remark to the notion of "Number"), however, he provided a much longer discussion of the analytic status of mathematics, criticizing Kant for his claim that the latter science should rather be understood to be synthetic and grounded in intuition. Here, in particular, Hegel concentrated on the definition of a straight line, attempting to show that its property of being the shortest line between two points can in fact be proven by pure logic and without any appeal to intuition (Schwab's example). The reference to Greek science, the equality of directions and the Parallel Postulate follows (cf. Hegel 1984, p. 200).

[63] Lotze strenuously opposed non-Euclidean geometry on philosophical grounds and criticized the views of both Helmholtz and Riemann simply restating Karsten's old (1778) argument on the transitivity of direction in parallelism – as nothing had happened in this field in the previous century (Lotze 1874–1879, vol. 2, § 131, pp. 247–49). The notions of *Richtung* and parallelism were already at work in Lotze's first attempt in philosophy, even though at the time he had not dared to propose a proof of the Parallel Postulate: see Lotze (1841), pp. 184 ff.

References

Banks, Eric. 2005. Kant, Herbart and Riemann. *Kant-Studien* 96: 208–234.
Bertrand, Louis. 1778. *Développement nouveau de la partie élémentaire des mathématiques. Genève.*
De Morgan, Augustus. 1915. *A Budget of Paradoxes*, ed. D.E. Smith, 2nd ed. Chicago: Open Court
De Risi, Vincenzo. 2013. La dimostrazione kantiana del Quinto Postulato. In *Kant und die Philosophie in weltbürgerlicher Absicht*, ed. S. Bacin, A. Ferrarin, C. La Rocca, and M. Ruffing, vol. 5, 31–43. Berlin: de Gruyter.
———. 2015. *Leibniz on the Parallel Postulate and the Foundations of Geometry*. Basel/Boston: Birkhäuser.
Dodgson, Charles Lutwidge. 1879. *Euclid and his Modern Rivals*. London: MacMillan.
———. 1888. *A New Theory of Parallels*. London: MacMillan.
Dummett, Michael. 1991. Frege and Kant on Geometry. *Inquiry* 25: 233–254.
Ewald, William Bragg. 1996. *From Kant to Hilbert: a source book in the foundations of mathematics*. Oxford: Clarendon Press.
Fichte, Johann Gottlieb. 1962. *Gesamtausgabe*, ed. R. Lauther et al. Cannstatt: Frommann.
Frege, Gottlob. 1884. *Grundlagen der Arithmetik*. Breslau: Koebner.
———. 1960. *The Foundations of Arithmetic*, ed. J.L. Austin, 2nd ed. New York: Harper.
———. 1967. *Kleine Schriften*, ed. I. Angelelli. Hildesheim: Olms.
———. 1983. *Nachgelassene Schriften*, ed. H. Hermes, F. Kambartel, and F. Kaulbach, 2nd ed. Hamburg: Meiner.
Fries, Jakob Friedrich. 1822. *Die mathematische Naturphilosophie*. Heidelberg: Mohr.
Gauss, Carl Friedrich. 1863–1917. *Werke*. Hildesheim: Göttingen, Leipzig.
Gray, Jeremy. 1989. *Ideas of Space: Euclidean, Non-Euclidean, and Relativistic*. Oxford: Clarendon Press.
———. 2003. Gauss and Non-Euclidean Geometry. In *Non-Euclidean Geometries: János Bolyai memorial volume*, ed. A. Prékopa and E. Molnar, 61–80. New York: Springer.
———. 2004. *János Bolyai, Non-Euclidean Geometry and the Nature of Space*. Cambridge, MA: Burndy.
———. 2008. *Plato's Ghost. The Modernist Transformation of Mathematics*. Princeton: PUP.
Gray, Jeremy, and Laura Tilling. 1978. Johann Heinrich Lambert, mathematician and scientist, 1728–1777. *Historia Mathematica* 5: 13–41.
Gregory, Frederick. 1983. Neo-Kantian Foundations of Geometry in the German Romantic Period. *Historia Mathematica* 10: 184–201.
Hegel, Georg Wilhelm Friedrich. 1984. *Gesammelte Werke*, ed. F. Hogemann and W. Jaeschke, vol. 21. Hamburg: Meiner.
Heis, Jeremy. 2020. Kant on Parallel Lines. Definitions, Postulates, and Axioms. In *Kant's Philosophy of Mathematics*, ed. C. Posy and O. Rechter, 157–180. Cambridge: CUP.
Helmholtz, Hermann von. 1921. *Schriften zur Erkenntnistheorie*, ed. M. Schlick and P. Hertz. Berlin: Springer.
Herbart, Johann Friedrich. 1828–1829. *Allgemeine Metaphysik*. Königsberg: Unzer.
Hindenburg, Carl Friedrich. 1781. Über die Schwürigkeit bey der Lehre von den Parallellinien. Neues System der Parallellinien. Anmerkungen über das neue System der Parallellinien. *Leipziger Magazin zur Naturkunde, Mathematik und Oekonomie*, 145–168 and 342–371.
———. 1786. Noch etwas über die Parallellinien. *Leipziger Magazin für reine und angewandte Mathematik*: 359–404.
Hoffmeister, Johannes. 1936. *Dokumente zu Hegels Entwicklung*. Stuttgart: Frommann.
Jacobi, Andreas. 1824. *De undecimo Euclidis axioma iudicium*. Jena: Croeck.
Karsten, Wenceslaus Johann Gustav. 1758. *Praelectiones matheseos theoreticae elementaris*. Rostock: Berger.
———. 1760. *Mathesis theoretica elementaris atque sublimior*. Rostock: Röse.
———. 1780. *Anfangsgründe der mathematischen Wissenschaften*. 2nd ed. Greifswald: Röse.

———. 1786. *Mathematische Abhandlungen*. Halle: Renger.

Kästner, Abraham Gotthelf. 1758. *Anfangsgründe der Arithmetik, Geometrie, ebenen und sphärischen Trigonometrie und Perspectiv*. Göttingen: Vandenhoeck.

———. 1771. *Dissertationes mathematicae et physicae*. Oldenburg: Richter.

———. 1790. Was heißt in Euclids Geometrie moglich?, Über den mathematischen Begriff des Raums, Über die geometrischen Axiome. *Philosophisches Magazin* 2 (4): 391–430.

Killing, Wilhelm. 1893–1898. *Einführung in die Grundlagen der Geometrie*. Padeborn: Schöningh.

Klügel, Georg Simon. 1763. *Conatuum praecipuorum theoriam parallelarum demonstrandi recensio*. Göttingen: Schultz.

Knar, Joseph. 1827. Über die Theorie der Parallellinien. *Zeitschrift für Physik und Mathematik* 3: 414–439.

———. 1828. Berichtigung meiner Ansicht über die Theorie der Parallellinien. *Zeitschrift für Physik und Mathematik* 4: 427–436.

Lambert, Johann Heinrich. 1764. *Neues Organon*. Leipzig: Wendler.

———. 1786. Theorie der Parallellinien. *Magazin für reine und angewandte Mathematik*, 13–64, and 325–358.

———. 1781–1787. *Deutscher gelehrter Briefwechsel*, ed. J. Bernoulli. Berlin.

Langsdorf, Karl Christian. 1802. *Anfangsgründe der reinen elementar- und höheren Mathematik aus Revision der bisherigen Principien gegründet*. Erlangen: Palm.

Lorenz, Johann Friedrich. 1791. *Grundriß der reinen und angewandten Mathematik, oder der erste Cursus der gesamten Mathematik*. Helmstädt: Fleckeisen.

Lotze, Hermann. 1841. *Metaphysik*. Leipzig: Weidmann.

———. 1874–1879. *System der Philosophie*. Leipzig: Hirzel.

Maimon, Salomon 1790 *Versuch über die Transzendentalphilosophie*. Berlin: Voß.

Mancosu, Paolo. 2015. Grundlagen, Section 64: Frege's Discussion of Definitions by Abstraction in Historical Context. *History and Philosophy of Logic* 36: 62–89.

Mehrtens, Herbert. 1990. *Moderne Sprache, Mathematik: Eine Geschichte des Streits um die Grundlagen der Disziplin und des Subjects formaler Systeme*. Frankfurt: Suhrkamp.

Meinecke, Wilhelm. 1906. Die Bedeutung der Nicht-Euklidischen Geometrie in ihrem Verhältnis zu Kants Theorie der mathematischen Erkenntnis. *Kant-Studien* 11: 209–232.

Meinert, Friedrich. 1790. *Lehrbuch der gesammten Kriegswissenschaften für Officiere bei der Infanterie und Kavallerie: zweiter Theile, Gemeine Geometrie*. Halle: Hemmerde.

Mercator. 1678. *Euclidis elementa geometrica novo ordine ac methodo fere demonstrata*. London: Martyn.

Miller, David Marshall. 2014. *Representing Space in the Scientific Revolution*. Cambridge: CUP.

Moktefi, Amirouche. 2011. Geometry: The Euclid Debate. In *Mathematics in Victorian Britain*, ed. R. Flood, A.C. Rice, and R. Wilson, 320–336. Oxford: OUP.

Natorp, Paul. 1910. *Die logischen Grundlagen der exacten Wissenschaften*. Leipzig: Teubner.

Nelson, Leonard. 1905–1906. Bemerkungen über die nicht-Euklidische Geometrie und den Ursprung der mathematischen Gewißheit, *Abhandlungen der Fries'schen Schule* 1–3. Now in L. Nelson, *Gesammelte Schriften* 3: 9–52. Hamburg: Meiner 1974.

Noble, Eduardo. 2022. *The Rise and Fall of the German Combinatorial Analysis*. Cham: Birkhäuser.

Ouvrier, Carl Siegmund. 1808. *Theorie der Parallelen als Ankündigung eines neuen Versuchs über das Erkenntnißvermögen*. Leipzig: Schiegg.

Peters, Wilhelm Servatius. 1962. Das Parallelenproblem bei A.G. Kästner. *Archive for History of Exact Sciences* 1: 480–487.

Pont, Jean-Claude. 1986. *L'aventure des parallèles. Histoire de la géométrie non euclidienne: précurseurs et attardés*. Berne: Lang.

Rehberg, August Wilhelm. 1792. Über die Natur der geometrischen Evidenz. *Philosophisches Magazin* 4 (4): 447–460.

Robinet, André. 1957. *Correspondance Leibniz-Clarke*. Paris: PUF.

Saccheri, Gerolamo 2014. *Euclid Vindicated from Every Blemish*, ed. V. De Risi. Basel/Boston: Birkhäuser.

Scarburgh, Edmund. 1705. *The English Euclide*. Oxford: Theatrum Sheldonianum.

Scheibel, Johann Ephraim. 1807. *Zwey mathematische Abhandlungen: I. Vertheidigung der Theorie der Parallellinien nach dem Euclides*. Breslau: Korn.

Schlömilch, Oskar. 1849. *Grundzüge einer wissenschaftlichen Darstellung der Geometrie des Masses*. Eisenach: Baerecke.

Schmidt, Georg Gottlieb. 1797. *Anfangsgründe der Mathematik zum Gebrauch auf Schulen und Universitäten*. Frankfurt: Barrentrapp.

Scholz, Erhard. 1982. Herbart's Influence on Bernhard Riemann. *Historia Mathematica* 9: 413–440.

Schotten, Heinrich. 1890–1893. *Inhalt und Methode des planimetrischen Unterrichts*. Leipzig: Teubner.

Schultz, Johann. 1784. *Entdeckte Theorie der Parallelen, nebst einer Untersuchung über den Ursprung ihrer bisherigen Schwierigkeit*. Königsberg: Kanter.

Schwab, Johann Christoph. 1791. Über die geometrischen Beweise, aus Gelegenheit einer Stelle in der Allgemeinen Litteratur-Zeitung. *Philosophisches Magazin* 3 (4): 397–407.

———. 1792. Über das Unendliche des Herrn Hofpredigers Schultz. *Philosophisches Archiv* 1 (3): 70–75.

———. 1801. *Tentamen novae parallelarum theoriae notione situs fundatae*. Stuttgart: Erhard.

———. 1808. *Essai sur la situation, pour servir de supplément aux principes de la géométrie*. Stuttgart: Cotta.

———. 1814. *Commentatio in primum elementorum Euclidis librum*. Stuttgart: Steinkopf.

Schweikart, Ferdinand Carl. 1807. *Die Theorie der Parallellinien nebst dem Vorschlage ihrer Verbannung aus der Geometrie*. Jena: Gabler.

Segner, János András. 1747. *Vorlesungen über die Rechenkunst und Geometrie*. Lemgo: Meyer.

———. 1756. *Cursus mathematici pars I. Elementa arithmeticae, geometriae et calculi geometrici*. Halle: Renger.

———. 1764. *Anfangsgründe der Arithmetick, Geometrie, und der geometrischen Berechnungen*. Halle: Renger.

Seyffer, Karl Felix. 1801. Review of Schwab's *Tentamen*. *Göttingische Anzeigen von gehlerten Sachen* 39: 377–389.

Staudt, Karl Georg Christian von. 1847. *Geometrie der Lage*. Nürnberg: Bauer.

Vermehren, Carl Christian Hermann. 1816. *Versuch die Lehre von der parallelen und convergenten Linien aus einfachen Begriffen vollständig herzuleiten*. Güstrow: Ebert.

Voelke, Jean-Daniel. 2005. *Renaissance de la géométrie non euclidienne entre 1860 et 1900*. Bern: Lang.

Volkert, Klaus. 2013. *Das Undenkbare denken. Die Rezeption der nichteuklidischen Geometrie im deutschsprachigen Raum (1860–1900)*. Berlin: Springer.

Webb, Judson. 1987. Immanuel Kant and the Greater Glory of Geometry. In *Naturalistic Epistemology*, ed. D. Nails and A. Shimony, 17–70. Dordrecht: Reidel.

Wolff, Christian. 1710. *Anfangsgründe aller mathematischen Wissenschaften*. Halle: Renger. Revised ed. by K.C. Langsdorf. Marburg: Akad. Buchhandlung 1797.

Wood, David W. 2012. *"Mathesis of the Mind". A Study of Fichte's Wissenschaftslehre and Geometry*. Amsterdam: Rodopi.

Chapter 21
The Geometer's Gaze: On H. G. Zeuthen's Holistic Epistemology of Mathematics

Nicolas Michel

I'm an eye. A mechanical eye. I, the machine, show you a world the way only I can see it. I free myself for today and forever from human immobility . . . Freed from the boundaries of time and space, I co-ordinate any and all points of the universe, wherever I want them to be. My way leads towards the creation of fresh perception of the world. Thus I explain in a new way the world unknown to you.

Dziga Vertov, Kinoks: A Revolution (1923), cited in Berger (1972, p. 17).

Abstract This chapter explores the epistemology of geometry expounded by Danish mathematician Hieronymus Georg Zeuthen towards the end of his career. Zeuthen's views, a defense and reinvention of the role of intuition in geometry, are placed against the backdrop of episodes in nineteenth-century history of algebraic enumerative geometry, the subject of the vast majority of Zeuthen's scientific output; but also that of the holistic psychologies and philosophies of Harald Høffding and Henri Bergson. In so doing, this chapter presents Zeuthen's late writings on mathematics as motivated by the construction of a novel way of perceiving geometrical figures, a gaze freed from the limitations of symbolic reasoning or mechanical computations.

N. Michel (✉)
Fakultät für Mathematik und Naturwissenschaften, Bergische Universität Wuppertal, Wuppertal, Germany
e-mail: michel@uni-wuppertal.de

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_21

21.1 Introduction: An Address at The Danish Royal Academy

On October 16th, 1914, the mathematician Hieronymous Georg Zeuthen (1839–1920) stepped foot into the main lecture hall of the Danish Royal Academy of Science to address a mixed audience composed equally of philologists and philosophers, of psychologists and physiologists (see Fig. 21.1).[1] Such an address by Zeuthen was a common occurrence in Copenhagen, where he was often seen expounding the importance of the history and epistemology of ancient mathematics—for modern practitioners and outsiders to the field alike. On this particular day, however, Zeuthen was instead scheduled to present his newly published book, a *Textbook on the Enumerative Methods of Geometry* ("Lehrbuch der abzählenden Methoden der Geometrie"). This was the branch of mathematics on which he had spent the vast majority of his professional career, starting with his 1865 dissertation. It was also a discipline in which he was widely regarded as a leading specialist, well beyond the borders of Denmark. His expertise had translated into a massive research output, disseminated in journals in France, Germany, Italy, and England;[2] he had also been called upon by Felix Klein to arbitrate matters of dispute in the pages of the *Mathematische Annalen* and to write an authoritative survey for the *Encyklopädie der mathematischen Wissenschaften*.[3] Lastly, Zeuthen had made these methods the subject of his more advanced teaching at the University of Copenhagen since the late 1870s.[4] This much-awaited textbook, the content of which Zeuthen had already been collecting and organising at the turn of the century, was to be his final word on geometrical theories of growing importance to mathematicians at large, and on the proper way to teach them.

So that his fellow academicians might appreciate the significance of this book, however, Zeuthen had to leave geometrical technicalities aside. Instead, his presentation would focus on the connections between the methodological principles underlying his own research and a question central to the "trends in the scientific world of the time," namely, that of the relation between "the scientific use of symbols and

[1] On the institutional organisation of Danish science at this time, and the central place this Academy occupied therein, see Kragh (2015).

[2] Beyond the *Tidsskrift for Mathematik*, of which he was chief editor between 1871 and 1889, Zeuthen mostly published in the *Mathematische Annalen* and in the *Comptes Rendus de l'Académie des Sciences*. A general biography can be found in Kleiman (1991).

[3] Zeuthen (1890, 1921). This survey was for the most part written and handed down to the editors of the *Encyklopädie* in 1899, then revised and finalised in March 1905. As such, contrary to what the dates of publication indicate, it is the preparation of this survey which informed much of the content of Zeuthen's 1914 textbook and not the other way around.

[4] Zeuthen retired from his position at the University of Copenhagen in 1910, having secured funding from the Carlsberg Foundation to finish writing his textbook. This foundation played a major role in the funding of Danish science at this time; see Kragh et al. (2008, pp. 320–324; 399–403).

Fig. 21.1 P. S. Krøyer, *Et møde i Videnskabernes Selskab* (1897). Zeuthen can be seen seated on the right, while Høffding is on the front left (both circled in red). Reproduced with generous permission from the Royal Danish Academy of Sciences and Letters

intuition."[5] The term 'symbol,' Zeuthen immediately added, was to be understood in the extended sense given to it by the French philosopher Henri Bergson (1859–1941): beyond the sole characters of algebraic calculations or specious arithmetic, it encompassed all technical images or nomenclatures which may be combined mechanically, in "law-bound" [*lovbundne*] fashion, in order to enable intelligence to represent and analyse objects of knowledge. The articles of a legal code [*Regler*], for instance, constitute such symbols; and their formal prescriptions in the context of judicial procedures can be distinguished from the manner in which actual judgments [*Skøn*] are delivered, a process which involves much more flexibility and attention to particular circumstances. The same goes for mathematics, according to Zeuthen, wherein one must distinguish between **computations** [*Regning*]—regulated by rigid symbolic laws—and **reasoning** [*Ræsonnement*]—a faculty which ultimately relies on intuition.

[5] Zeuthen (1914b, p. 273). All translations mine, unless otherwise noted. Zeuthen gives no indication as to what those 'scientific trends' are. The second half of this chapter, however, will suggest various plausible examples as we turn to Harald Høffding's lectures on the psychology and philosophy of science and explore their impact on Zeuthen's own epistemology of geometry. This distinction between symbolic and intuitive cognition (in the sciences) was by no means new. In his famous 1684 *Meditations on Knowledge, Truth, and Ideas*, G. W. Leibniz had already described adequate knowledge as being either "symbolic" and "intuitive"; see Leibniz (1684, p. 537). Furthermore, Kant's theory of human cognition as stemming from both sensibility (*Sinnlichkeit*) and understanding (*Verstand*) and ensuing questions regarding the connection between both stems fuelled a long tradition of commentaries all through the nineteenth century amongst scientists, artists, and philosophers alike; see Huehn and Vigus (2013).

Zeuthen viewed his own work as firmly rooted in the camp of the latter. To an image of computations as the paragon of epistemic safety and exactness in mathematics, he opposed an agent-centred epistemology which does not underplay the risk of human-induced mistakes and the array of strategies routinely deployed by computers to avoid them, from active training to the verification of results through alternative computational methods. One particularly important source of computational error, for Zeuthen, was the use of symbolic computations outside of their proper domain of applicability—something which arises especially often when "the computer [*Regneren*], busy operating with **computing symbols** [*Regnesymbolerne*], forgets to **reason** about their proper use."[6] As a historian of mathematics, Zeuthen viewed this as a routine occurrence in the early development of analysis, especially in the context of computations with possibly divergent series. It was also a tendency that needed combatting in his own geometrical provinces; hence his writing of a textbook revolving around not calculations but rather methods, their rational use, and the intuition that undergirds them.[7]

Prima facie, it might appear somewhat unsurprising that a mathematician writing in the 1910s would elect to reflect on the limits of symbolic means of representation and on their relation to computing and reasoning. Much has been written about the so-called 'conceptual' approach of Göttingen-based mathematicians who, after Peter Gustav Lejeune-Dirichlet's famous *dictum*, sought to "replace calculations by thought," and to construct theories reliant not on external forms of representation, but on concepts themselves.[8] The years following Zeuthen's address also saw the emergence within mathematics of broader, sharper critiques of language (symbolic or otherwise) and of its ability to express what matters most to mathematical life, from L. E. J. Brouwer to Ludwig Wittgenstein.[9] More generally, the 'modernist transformation' of mathematics as a discipline between the 1880s and the 1920s has largely been described as a reconfiguration of the relation between mathematics and its language(s), the roles of intuition and symbols being at the core of well-documented disputes related to this transformation.[10]

And yet, Zeuthen's reflections on intuition and computations escape the Procrustean bed formed by these classical narratives. For his intuition was neither Felix Klein's *Anschauung* nor Brouwer's mystical introspection, but rather what he somewhat mysteriously defined as a form of 'holistic perception' [*Helhedsopfattelse*], an ability to perceive at once a connected whole where symbolic cognition and

[6] Zeuthen (1914b, p. 272). Emphasis mine.

[7] On the separation of intelligence and calculation at the turn of the nineteenth century, see Daston (1994).

[8] Ferreirós (2007, p. 28). For an example of a mathematical practice developed in keeping with these methodological principles, see Haffner (2017).

[9] Brouwer (1975, pp. 6–8; 72–75), Janik and Toulmin (1973, pp. 177–184). One may also think of Hermann Weyl, who shared certain holistic commitments with Zeuthen, albeit in a very different mathematical, political, and philosophical context Schappacher (2010, pp. 3269–3276), Eckes (2018).

[10] See Mehrtens (1990), Gray (2008), as well as the chapters in part IV of the present volume.

rigid computational rules present one's mind with disconnected particulars[11]. This holistic intuition, as will be shown, constituted for Zeuthen the faculty that grounded not only proper understanding of mathematical theorems and their proofs, but also the critical and rigorous use of methods in geometry.[12] And Zeuthen's critique of symbolic expressions, as his borrowing of Bergson's definition already implies, runs askew of traditional discourses on the shortcomings of algebraic notations or computations. At its core, it is a wide-reaching critique of any and all attempts at a mechanical, impersonal use of language that pretends to do away with the critical subject perceiving the proof or the figures featuring in it. Lastly, while commentators have often sought to confine Zeuthen's philosophy of mathematics within the vague category of 'Platonism', such qualifications only obscure his more original claims regarding the existence of various perspectives on a given geometrical figure, as well as his sensitivity to the historical and cognitive processes involved in the formation of such figures.[13]

To fully understand these admittedly abstract pronouncements, one must first contextualise them within Zeuthen's scientific practice and milieu. Contextualising, here, necessitates the conjunction of two localisations. First, Zeuthen's epistemology of mathematics must be situated against the specific backdrop of the (algebraic) enumerative geometry he had practised throughout his whole career; a theory often ignored in the historiography of modern(ist) mathematics. Second, Zeuthen's conception of intuition must be traced back to the Bergsonian philosophy of the Danish philosopher and psychologist Harald Høffding (1843–1931), and more largely to the particular brand of rational holism that was increasingly gaining traction amongst members of the Danish Royal Academy of Science. These two localisations, carried out respectively in Sects. 21.2–21.3 and 21.4–21.5, will then reveal a much more precise and potent meaning behind Zeuthen's philosophical forays. In Zeuthen's reflective writings, I shall argue, these commitments amounted to the creation, justification, and teaching of a novel way of seeing and forming geometrical figures, a form of mathematical subjectivity informed by a productive dialogue between mathematics, philosophy, and psychology.[14]

[11] The Danish word 'Helhed' translates to 'whole' or 'totality' in English. In Zeuthen's texts, it is constantly paired with the term 'Enkeltheder', or 'particular'.

[12] Of course, the idea that there can be a "method of intuition" is itself a core Bergsonian commitment.

[13] Lützen and Purkert (1989), Sigurdsson (1992).

[14] On the relation between geometry and psychology at the turn of the twentieth century in a rather different context (both scientific and cultural), see Gray (2008, pp. 388–405).

21.2 The Theory of Relative Generality

Weaving the mathematical and the philosophical was not something Zeuthen reserved exclusively for his Danish audience, nor something he discovered late in his life. While overt references to holistic psychology and philosophy would only appear in his writings in the 1910s, as a reaction to scientific conversations only then developing in Copenhagen, this conceptual encounter had been made possible by past decades of reflexive and thoughtful engagement with algebraic geometry and the history of mathematics. In both endeavours, one theme of constant importance to Zeuthen was the definition and comparison of the "general" and the "particular" in geometry—a theme we need to first explore in order to make sense of his later defence of geometrical intuition.[15]

To Zeuthen, the distinction between the general and the particular was both a technical and an epistemological one, having to do with the very validity of certain theorems as well as with the sort of knowledge one could or should strive for in the course of mathematical research. The first sections of the *Lehrbuch* are in fact entirely devoted to the clarification of this distinction, as a necessary propaedeutic for the teaching, usage, and interpretation of all enumerative methods in geometry.[16] Of such crucial importance was the discussion of generality that one also finds traces of it all through Zeuthen's correspondence with the French mathematician Georges-Henri Halphen (1844–1889).[17]

The lesson Zeuthen drew from his decades-long engagement with algebraic geometry was that the distinction between the general and the particular was essentially a "relative" one: a figure, a formula, or a theorem could all be equally found to be general or particular depending on the selection of a specific point of view. This relativity, in return, made it crucial and even necessary to teach "methods rather than calculations," for an emphasis on the latter obfuscates the need for a specification of the sense in which their (symbolic) results may be viewed as generally valid.[18]

To explicate these claims, Zeuthen had a plethora of examples at his disposal. In 1909, while presenting an early draft of his *Lehrbuch* at the first Scandinavian Congress of Mathematicians, Zeuthen elected to discuss a result at the heart of the 'Duality controversy' between Joseph-Diez Gergonne (1771–1859), Jean-Victor Poncelet (1788–1867), and Julius Plücker (1801–1868); a result he viewed as a

[15] On generality as an epistemic value in the history of mathematics, see Chemla et al. (2016).

[16] Zeuthen (1914a, pp. 1–17).

[17] Halphen was not only a renowned expert in (among other things) enumerative geometry, but also a close friend of Zeuthen's, whose letters to Halphen can be found at the Bibliothèque de l'Institut, Paris (Cod Ms 5624). A portion of this correspondence is transcribed in Michel (2020, pp. 465–497).

[18] Letter from Zeuthen to Halphen dated Nov. 5th 1879, Bib. de l'Institut, Paris, Cod Ms 5624/231; Michel (2020, p. 480).

paradigmatic example of the sort of errors which may arise in geometry from an "inconsiderate usage of the words 'general' and 'particular'."[19]

This controversy, it is known, was initially triggered by a plagiarism charge repeatedly thrown by Poncelet towards a young Plücker in 1827 and 1828.[20] This accusation, it turned out, was largely the result of a creative rewriting and reorganization of one of Plücker's papers by Gergonne, acting as editor of the journal to which this paper was destined. Gergonne had disposed Plücker's geometrical propositions in parallel columns, so as to highlight the duality between geometrical objects and propositions through textual symmetry (e.g., interverting the words 'points' and 'lines,' 'intersections' and 'tangents' etc. so as to turn the statement of Pascal's theorem into that of Brianchon's). In so doing, Gergonne had also made Plücker an unwitting party and ally in a dispute of his own with Poncelet, a dispute that revolved around the nature and scope of the notion of duality in geometry as well as around matters of priority.

Part of Poncelet's rhetorical strategy, in this controversy, was to accuse Gergonne of having exaggerated the importance of duality as a general mechanism for the transformation and duplication of geometrical theorems through linguistic substitutions (as opposed to his own conception of duality, centered around pole/polar relations). In so doing, Poncelet claimed, Gergonne had been led to grave errors. In one memoir, starting with the classical definition of the order m of a curve (i.e., the number of points at which the curve intersects an arbitrary fixed straight line), and applying to it the aforementioned linguistic transformation, Gergonne had mistakenly concluded that, from an arbitrary fixed point, the same number m of tangents to the curve could be drawn.[21]

As Poncelet was eager to point out, however, this theorem is only true for conics. Instead, the number of tangents one can draw to a curve of order m from an arbitrary fixed point is, in general, $m(m-1)$.[22] Acknowledging his mistake, Gergonne thus introduced the notion of 'class' (denoted by later authors with the symbol m^*) to describe this latter number of tangents, a correction which Poncelet rejected on the grounds that it introduced two competing classifications for the same geometrical objects and thus had little scientific value.[23]

[19] Zeuthen (1910, pp. 33–34). On this Congress, of which Zeuthen was vice-president, and on the rise of a Scandinavian identity for mathematicians more broadly, see Turner and Sørensen (2013).

[20] For a thorough analysis of the social and institutional aspects of the Duality Controversy, from which the following paragraphs borrow, see Lorenat (2015b). In Zeuthen's view, Poncelet's projective geometry, and especially his "principle of continuity," was the main precursor behind the principles that would lie at the core of the enumerative methods in the second half of the nineteenth century (Zeuthen 1917, p. iii). This might explain the use of this specific example while discussing the epistemological foundations of a later, distinct theory. In the final version of the *Lehrbuch*, however, Zeuthen would eventually elide most of these historical considerations.

[21] Gergonne (1827, p. 216).

[22] A simple algebraic proof of this result can be found in Salmon (1852, p. 62). See also Gray (2007, pp. 53–61; 165–170).

[23] On the history of these classifications, see Lê (2023).

But the generality of Poncelet's correction, Zeuthen noted, can also be questioned further. While it is true that, "in general," a curve of order m is a curve of class $m(m-1)$, it is not the case that a "general" curve of order m is also a "general" curve of class $m(m-1)$. Indeed, viewed as a curve of class $m(m-1)$, the aforementioned general curve of order m will necessarily possess several singular tangents (e.g., bitangents or cuspidal tangents)—and therefore be a special member of this family of curves, one that presents singularities. Accordingly, the dual of a general curve of class n is only a special curve of order $n(n-1)$, i.e., one that has certain singular points (e.g., double points or cusps). That this is necessarily the case derives from an elementary argument: considering a generic smooth curve of order m, suppose its dual curve is a general element of the family of curves of class $n = m(m-1)$ (i.e., one without singular tangents). One can then form the dual of this latter curve, which, per hypothesis, is of order $n(n-1) = m(m-1)(m^2-m-1)$. Now, this immediately contradicts the involutive character of duality, as this dual curve is in fact the original curve of order m, and such explosions of the order do not happen under repeated dualisation.

This is why, Zeuthen concluded,

> the definition of generality is different depending on whether we speak of a curve of a certain order or of a curve of a certain class. A general curve of order m is a curve represented in a usual system of coordinates through an equation of degree m, and its general properties are those which belong to it regardless of any relations between its coefficients ... Then, tangents are the lines which intersect the curve at two coinciding points. In the event the curve has multiple points, the number $m(m-1)$ of tangents passing through an arbitrary point will then also include, a certain number of times, the straight lines joining this point to the multiple points; but these lines will be 'foreign or improper solutions' [to the problem of determining these tangents] if we look at the same curve as belonging to the curves of a certain class, and to obtain in reality this class, one must still subtract the number of these foreign solutions to the number $m(m-1)$.[24]

In other words, there exist two different ways of looking at a given curve: one may view it as a member of the collection of curves of a certain order or of curves a certain class; and, crucially, different theorems appear as general depending on which of these perspectives is chosen. A first major issue with blind reliance on computations thus appears: their proper, scientific use requires the clear specification of a viewpoint on the objects the symbols occurring throughout said computations represent—or, equivalently, the scientific use of symbols must be supplemented by a constant awareness of the genetic processes from which they originate, e.g., a awareness of whether they represent curves generated via point- or line-motions. To mobilise the equation $m^* = m(m-1)$ in the course of a proof is to already fix one perspective on curves and their generation, namely a point-centric one. And to forego critical reasoning, that is to say to rely exclusively on symbols without reflecting on their provenance and the assumptions they carry with them, may in turn lead to the counting of unintended objects, such as the artificial tangent lines described by Zeuthen above. Relying on symbols, equations,

[24] Zeuthen (1910, pp. 34–35).

and computations, thus requires a form of active oversight, a "personal control" over the origin of said symbols and the assumptions they carry—a task for which, as we will see below, Zeuthen found intuition (i.e., the faculty presiding over the rational use of methods) to be particularly well suited.[25]

Of course, readers already familiar with this historical episode will know that more refined equations to treat such questions would soon become available. Indeed, though he had been forced into these debates by Gergonne, Plücker was no passive observer thereof. Building on Poncelet's idea of regarding curves alternatively as a locus generated by the motion of a point or as an envelope generated by the motion of a (tangent) straight line, he analysed the impact of each type of singularity on the dual curve from each viewpoint. A double point, for instance, was found to cause a double reduction in the order of the dual; whereas a cusp counted for three. And conversely, for each bitangent a curve may have, its class would have to be reduced by two upon taking its dual.[26] These insights would quickly be put in symbolic form and become what are now called "Plücker's formulas," i.e.:

$$m^* = m(m-1) - 2\delta - 3\kappa$$
$$m = m^*(m^*-1) - 2\delta^* - 3\kappa^*$$

where m, δ, κ represent respectively the order, number of double points, and number of cusps of a curve described in point-coordinates, and m^*, δ^*, κ^* represent the analogous quantities for the dual curve, described in line-coordinates.[27]

These more refined formulas are, without a doubt, a net improvement over Poncelet's. But, to Zeuthen's eyes, even they remain imperfect for the teaching and practice of geometry. Whilst acknowledging "[their] great utility," Zeuthen argued that therein is still nested an "obfuscat[ion] of the distinction between the general and the particular," stemming from "the[ir] double starting point in the representation via point and line coordinates," thereby giving particular significance to certain singularities (and only them).[28]

Indeed, for symbolic computations to be exact, one must always employ them within the bounds of their range of applicability. In the theory of functions, this might mean that one has to handle with care the convergence of specific series; in algebraic geometry, it means ensuring that the sense in which a figure is taken to be general—in other words, the viewpoint from which they are looked at—suits the symbolic formulas one elects to use. And where the computer may lose sight of this imperative in the course of their symbolic activity, the geometer who relies chiefly

[25] Zeuthen (1910, p. 40).

[26] Plücker (1834).

[27] In 1834, Plücker did not write these results as equations but rather spelt out how to reduce the order of a given curve to obtain that of its dual. Five years later, however, while systematising and expanding on these findings in his treatise on algebraic curves, he eventually wrote very similar equations (Plücker 1839, pp. 207–212).

[28] Zeuthen (1910, p. 35).

on methodical reasoning (i.e., ultimately, on intuition) always keeps this perspective in mind and in check, for theirs is a deductive practice which cannot be developed mechanically and detached from the figures under study.[29] For this reason, formulas such as Poncelet's or Plücker's should not be viewed as ready-made equations to be plugged uncritically within geometrical reasoning, but as methods for the study of the correlation (or duality) of algebraic curves or surfaces, always to be employed with a clear view of the relative generality of the results they imply.[30] It is in this sense that Zeuthen's was a teaching of the methods and not of the computations of geometry, and that he viewed intuition, rather than computations, to be the surer path to rigour.[31]

21.3 What Formulas Cannot Express

Zeuthen's subordination of calculations and symbols to reasoning and intuition went beyond the pedagogical and methodological concerns outlined in the previous section. In fact, Zeuthen proved even skeptical of the very possibility for symbolic language (in the extended sense previously given to this adjective) to precisely

[29] Zeuthen expressed a related idea whilst discussing the history of early-modern mathematics: "The modern method will generally have far more extensive uses, and it will be stated in rules that can be applied purely mechanically; but precisely because the related older methods were not so designed, but rather had to be adapted to each case, they could lead to a deeper penetration and a more versatile investigation than the corresponding modern treatment. The mere fact that the less developed form required a greater effort of thought led to observations that would easily escape one who now reach the same main result with railroad speed" Zeuthen (1903, pp. 554–555), cited in Blåsjö (2021, pp. 11–12). Skúli Sigurdsson reads such passages as indicative of Zeuthen's "distrust of mass-produced knowledge" and "scientific elitism" (Sigurdsson 1992, p. 112). This interpretation, however, severely downplays the intrinsic epistemological and mathematical concerns that Zeuthen developed through his geometrical practice, which this chapter shows were much more substantial and coherent than whatever superficial disdain Zeuthen may have had for mass education.

[30] In his *Encyklopädie* entry, Zeuthen made a point of noting that Halphen's rewriting of Plücker's formulas into methods suitable for all kinds of singularities (as opposed to Plückerian singularities exclusively) is "more convenient for enumerations." It was well known at the time that all (higher) singularities of an algebraic curve could be reduced to certain Plückerian invariants, and so this rewriting did not result in a net gain in mathematical knowledge but rather provided (to Zeuthen's eyes) a more methodical approach to the geometry of these curves. A similar point is made regarding Halphen's rules for the intersections of two algebraic curves at singular points and the formula corresponding to Bézout's theorem (Zeuthen 1921, pp. 260–262).

[31] To be sure, Zeuthen was not alone in seeking to justify a form of generality that could accommodate the existence of possible counter-examples. In his discussion of Weierstrass' theory of elementary divisors, for instance, historian Thomas Hawkins proposed the term "generic reasoning" to capture a certain disregard for problematic cases which may occur in certain instances of specification of algebraic symbols (Hawkins 1980, p. 295).

express and delineate the domain of applicability of profound geometrical truths, such as Poncelet's principle of continuity:

> The exactness of the 'principle' will depend on the formal statement that is given to it, and it might be difficult to find one which leaves room for no objections and no abuses; but it will be enough to teach the *method* of which the 'principle' is but a formal condensation, in a manner that ensures its just application to all particular cases.[32]

Not only do symbols conceal specific assumptions regarding the distinction between the general and the particular (and, therefore, specific viewpoints on geometrical figures); this limitation turns out to be beyond salvation by any sort of *ad hoc* refining of the formal expressions constructed therewith. The solution to the risks that symbolic practices carry with them is not more, better symbols; but a (re)turn to intuition and methodical reasoning.

This skepticism towards symbolic practices, and the ensuing recommendation to focus on teaching the methods, which are only ever imperfectly encapsulated in formal expression(s), echoed decades of infighting amongst the protagonists of the early development of enumerative geometry. Further down the same paragraph, Zeuthen would predict that the fate of all attempts at reducing Poncelet's principle down to a concise formula would also befall the 'principle of conservation of number' of Hermann Schubert, a leading figure in the development of this branch of geometry. More crucial to the formation of Zeuthen's epistemology of geometry, however, was the series of disputes that had punctuated the emergence of the first enumerative theories of conic sections; an episode he had directly witnessed and participated in from his first forays into mathematical research onwards.[33]

In the fall of 1863, a young Zeuthen left Denmark for Paris on a stipend in order to study with Michel Chasles (1793–1880), the world-famous geometer whose books had constituted an important part of his scientific education. The timing of this visit was auspicious: Chasles was putting the finishing touches on his latest grand theory, one that would garner much praise from mathematicians across Europe and convincingly exhibit the strength of his geometrical methods by solving problems that had theretofore eluded the most skilled of algebraists. This new theory, which

[32] Zeuthen (1910, p. 38). As is well known, Poncelet's principle had come under heavy fire from Augustin-Louis Cauchy and others (Gray 2007, pp. 47–50).

[33] On these disputes more broadly, see Michel (2021). In fact, one could read most of Zeuthen's attacks on computational practices as a direct critique of Schubert's own approach to enumerative geometry and especially his 1879 *Kalkül der abzählenden Geometrie*. Where Schubert laid out the computational rules at the heart of this branch of geometry and wrote a book full of symbols and numbers (the *Kalkül*), Zeuthen elected to rely on words, diagrams, and reasonings to teach individual mastery and control over the figures being enumerated (the *Methoden*). Only this, Zeuthen effectively argued, would ensure that future geometers avoid the pitfalls Schubert had been misled into by undue reliance on his symbolic calculus. Indeed, while containing a great deal of novel and impressive results, Schubert's *Kalkül* had sustained substantial critiques by Halphen and several others for its lack of rigorous justifications and its reliance on principles against which many counter-examples could be levied. At the time of Zeuthen's writing, acceptable justifications for Schubert's most important results were still very much lacking, and in fact would not emerge for decades.

Chasles called 'the theory of characteristics,' served to enumerate all the conic sections of a plane that simultaneously satisfied five independent conditions, such as passing through a given point or touching a given curve.[34] Chasles began publishing his method for solving these difficult problems in February 1864, and continued doing so for over two years. Zeuthen, meanwhile, had had to return to Denmark in April of that same year due to the outbreak of the Second Schleswig-Holstein War. Soon thereafter, he began writing his doctoral dissertation: an amelioration of Chasles's aforementioned theory, and Zeuthen's first contribution to a topic he would continuously revisit until his retirement.

Chasles's theory not only offered plenty in the way of novel results and methods; it also definitively refuted a series of previous results. In 1848, while working on unrelated geometrical questions and observing regularities in the number of conics satisfying contact problems, Jakob Steiner (1796–1863) had conjectured that there were $6^5 = 7776$ conics tangent to five other given conics (all in the same plane). This conjecture was later justified and integrated into more general formulas by, among others, the naval officer Ernest de Fauque de Jonquières (1820–1901), a self-styled disciple of Chasles.

To that end, De Jonquières had proposed to define a 'series' [*série*] of conic sections as a collection of such curves satisfying simultaneously four independent conditions. In other words, a series is a 1-parameter family of conics, which De Jonquières thought could be represented by a quadratic equation $F(x, y) = 0$, with the coefficients of F all being rational functions of one common variable λ (i.e., the parameter of the series).[35] De Jonquières also defined the index [*indice*] N of such a series as the number of conics in it that pass through an arbitrary fixed point; the number being invariant due to the principle of continuity (so long as one accepts intersections which may be imaginary or at infinity, and as one counts with multiplicity). By fixing the values of x and y in the aforementioned quadratic equation, one can see that this number is none other than the maximal degree in λ of the coefficients of F.

One key result of De Jonquières's theory was that, in a series of index N, there are $2N$ conics touching a given arbitrary line L.[36] Suppose, indeed, that the line L is the horizontal line $y = 0$ (the general case presenting little added difficulty). Then, a curve tangent to L corresponds to a value of the parameter λ such that $F(x, 0)$ is a quadratic polynomial with a double root, i.e., whose discriminant is zero. From elementary algebra and the previously-given characterisation of N it follows that this discriminant is of order $2N$, hence the result. More generally, for any geometrical condition Z, De Jonquières thought he could find a number α such

[34] Chasles (1864), Michel (2020, pp. 127–184).

[35] de Jonquières (1861, pp. 1134). De Jonquières's definition in fact extends to curves of order m. His assertion concerning the representability of such series via rational equations would later be shown to be erroneous by Cayley, but this mistake is of little importance for the present discussion.

[36] More generally, De Jonquières claimed that, in a system of curves of order m, the number of elements touching a given straight line was $2(m - 1)N$.

that the number of curves satisfying Z in a series of index N would be αN; and it is through such a result that he justified and generalised Steiner's conjecture.[37]

In modern mathematical terms, we may understand this result as a direct application of Bézout's theorem in the projective variety $\mathbb{P}^5$. This latter variety is an obvious moduli space for plane conics, since each of these curves is defined by an (homogeneous) equation

$$aX^2 + bY^2 + cZ^2 + 2dXY + 2eXZ + fYZ = 0,$$

that is to say by six coefficients (a, b, c, d, e, f), of which at least one is non-zero. Since two proportional sets of coefficients represent the same curve, a conic is associated not to a sextuplet directly, but to its equivalence class under the relation $(a, b, c, d, e, f) \sim (\lambda a, \lambda b, \lambda c, \lambda d, \lambda e, \lambda f)$, i.e., an element of $\mathbb{P}^5$. Within this framework, a series of conics of index N is thus represented by a subvariety of dimension 1 and of order N, while a condition corresponds to the hypersurface formed by the conics satisfying it: it is of dimension 4 and of a certain order α. The intersections of these subvarieties, of which there are αN per Bézout's theorem, then represent the conics in the series that satisfy the condition, as stated by De Jonquières's theorem.[38]

Unfortunately, these results present several fatal flaws, which Chasles and others quickly pounced on. For one thing, De Jonquières's theory implies that the number of conics touching five given straight lines is $2^5 = 32$, when there should in fact only be one—this number following from the principle of duality and the fact that through five given points passes one and only one conic.[39] The direct reason for this excess, as Luigi Cremona was the first to point out, is that De Jonquières's methods count many double lines (i.e., 'flat' degenerate conics consisting of one line, counting twice) which do not properly satisfy the conditions under study, and which appear as mere computational artefacts.[40] In the language of the anachronistic framework sketched above, the issue stems from the fact that the hypersurface of conics tangent to a given curve or line always contains the subvariety $V \subset \mathbb{P}^5$

[37] de Jonquières (1861, pp. 115–116; 121).

[38] The modern mathematical explanations are adapted from Kleiman (1980).

[39] To see how this number 32 derives from De Jonquières's theory, consider the series of conics passing through four given points. Its index is 1, as it is none other than the number of conics passing through $4 + 1 = 5$ points. De Jonquières's theorem above implies that the number of conics in this system touching a given straight line L_1 is $2 \times 1 = 2$. It follows that the index of the series of conics passing through three given points and touching L_1 is 2, per definition of this quantity. Therefore, De Jonquières's theorem can once again be used to find that there are 2×2 conics touching L_1, another given straight line L_2, and passing through three given points. Repeating this reasoning three more times yields the stated result.

[40] Cremona (1863).

formed by the double lines.[41] As it turns out, this very argument also can be levied against Steiner's 7776 conics, thus rendering this number ultimately meaningless.

Chasles's theory of characteristics, in a nutshell, rectifies these methods by restoring duality in De Jonquières's concepts.[42] Discarding De Jonquières's terminology, Chasles proposed to consider systems of conics, similarly defined as collections of conics satisfying four independent conditions, and to attach to them two numbers μ and ν, which he called 'characteristics' of the system. μ is none other than de Jonquières's index, while ν denotes in dual fashion the number of conics in the system which touch a given straight line. Chasles's central observation, borne out of the application of a uniform counting method over hundreds of particular cases, was the following: for any condition Z, one can find two numbers α and β such that, in any system S of characteristics μ and ν, the number of conics satisfying Z be

$$\alpha\mu + \beta\nu.$$

This formula thus replaces and complicates De Jonquières's αN by adding a second, dual factor. In so doing, it manages to keep at bay the unwanted double lines that ruined De Jonquières's results (and it does so at a very low computational cost, since only one term is added). Indeed, such degenerate conics can now also be viewed as enveloped by two pencils of tangent lines, whose centres must be on the double line. To visualise this phenomenon, picture an ellipse on the verge of flattening whilst its endpoints are being fixed in place. As the ellipse turns into a double line, its tangents turn into two pencils centred around said endpoints (see Fig. 21.2). These centres thus play a specific role on the double line, as all tangent lines to the conic must pass through one of them. This added specification of what it means to be tangent to a (degenerate) conic makes it so that a double line does not appear tangent to every

[41] This subvariety can be identified with what is now called the Veronese surface, hence the choice of the letter V. This surface is the image of the map

$$\begin{aligned} v\colon \mathbb{P}^2 &\longrightarrow \mathbb{P}^5 \\ [x, y, z] &\longmapsto [x^2, y^2, z^2, xy, xz, yz] \end{aligned}$$

Now, the (homogeneous) equation of a double line is

$$(aX + bY + cZ)^2 = a^2X^2 + b^2Y^2 + c^2Z^2 + 2abXY + 2acXZ + 2bcYZ = 0$$

Hence, a double line corresponds in $\mathbb{P}^5$ to a point of coordinates $(a^2, b^2, c^2, ab, ac, bc)$ and thus belongs to the image of the map f (that is, to the Veronese surface). Conversely, if a point belongs to this surface, computations of the minors of its corresponding quadratic form show that it represents a double line in the space of conics.

[42] The exact relation between Chasles's and De Jonquières's theories is more complex than this chapter can really suggest, and was the object of a fierce priority dispute between 1866 and 1867 (Michel 2020, pp. 185–224).

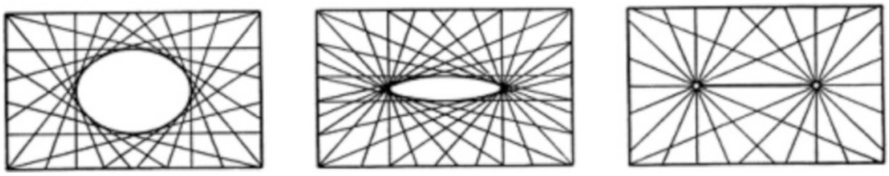


Fig. 21.2 A degenerating ellipse and its tangents; Klein (1928, p. 84)

single straight line in the plane, thereby excluding the artificial solutions which had plagued De Jonquières's methods.[43]

From a modern viewpoint, Chasles's theory amounts to replacing the moduli space $\mathbb{P}^5$ by a more involved parameter space which properly tracks point-conics and their duals, namely the variety of complete conics. One way of viewing it is to start with the subvariety $A \subset \mathbb{P}^5 \times \mathbb{P}^{5*}$ composed of pairs $(\mathcal{C}, \mathcal{C}^*)$ of smooth conics and their duals. This is only a quasi-projective subvariety, however, and one cannot directly apply Bézout's theorem in it to obtain Chasles's formula. Instead, in order to enumerate intersections of systems and conditions, one must still compactify A by taking the closure in $\mathbb{P}^5 \times \mathbb{P}^{5*}$ of the graph of map sending a smooth conic $\mathcal{C}$ to its dual $\mathcal{C}^*$. Modern intersection theory, in this new moduli space, yields Chasles's theorem.[44]

Zeuthen's own analysis of this historical episode, however, departs from unnuanced assessments of Chasles's theory as a mere correction and improvement upon De Jonquières's. Instead, Zeuthen frames it as a complication of the point of view adopted on geometrical figures, akin to what had happened to Poncelet's relation between the order and the class of a curve. In essence, Zeuthen views Steiner's and De Jonquières's results as "general theorems" fully worthy of this title, so long as one keeps in mind that they derive from a "representation of conic sections via point coordinates only."[45] From this perspective, a tangency is thus merely a double intersection, and these double intersections are precisely what was enumerated in the algebraic proof sketched above, when forming and cancelling discriminants. Double lines, however, always intersect a straight line at one such double point, thereby artificially appearing amongst the conics tangent to said straight line.[46] This is precisely what seemingly lead to an 'explosion' of the order of a curve when Poncelet's formula was applied twice to it, as singularities gave rise to

[43] Note that, from any arbitrary fixed point in the plane, two tangents to the conics can be drawn: namely, the two straight lines joining the point and the centres of the pencils. With this stipulation of the new meaning of tangency, the curve is still of class 2.

[44] See Semple (1982) for details. This compactification is equivalent to a blow-up of $\mathbb{P}^5$ along the Veronese surface mentioned previously.

[45] Zeuthen (1910, p. 35).

[46] Note that, had we taken a purely lineal viewpoint on conics, artificial solutions would have arisen in the form of line pairs, that is to say the other kind of degenerate conics in classical projective geometry.

similar artificial intersections or tangencies. Systems of conics that include such degenerate curves, therefore, are no longer 'general' elements of this family of geometrical objects, and De Jonquières's result, when wielded critically rather than mechanically, can be accommodated—with the distinction between the general and the particular from De Jonquières's viewpoint being particularly clear, as Zeuthen noted, since it suffices to determine whether or not double lines are present in a given system.[47]

Chasles's theory, in Zeuthen's retelling, thus appears as the move from De Jonquières's limited, point-centric perspective on conics to one that equally encompasses point- and line-coordinates; a move which the anachronistic framework outlined above makes explicit as we depart from the naive moduli space $\mathbb{P}^5$ to two copies thereof, coupled in a way that captures the duality so essential to the theory of conic sections. By so jointly taking into account both viewpoints on conics, and by consequently adding supplementary symbols to De Jonquières's formula, Chasles seemingly accomplished a progress akin to the move from Poncelet's theorem to Plücker's formulas. And here, too, an adjunction of symbolic terms would prove insufficient. For one, Zeuthen argued that this dual viewpoint, and the distinction between the general and the particular therein, were much harder to properly define than De Jonquières's. But more importantly, a series of new counter-examples to Chasles's theorem discovered by Halphen in 1876 would reveal looming deficiencies even in this refined formula.

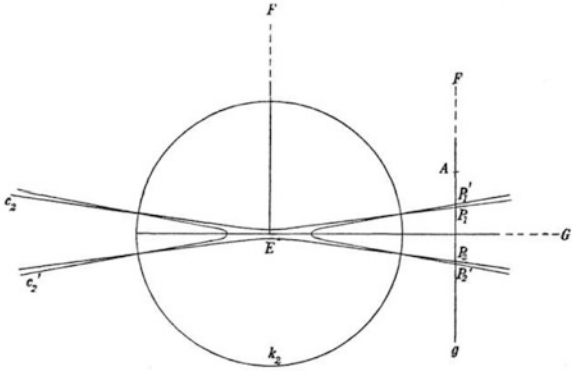
What Halphen had found was that systems of conics may in fact present singularities—that is to say, degenerate curves—of a kind not fully captured by Chasles's double representation as loci and envelopes.[48] These conics always appear as formed by one straight line and one point on it, though they may arise in a variety of ways (see Fig. 21.3). For instance, Zeuthen explains, we may view them as a further degeneration of a double line in which the centres of the pencils of tangents coincide; or as a line pair in which the two lines coincide. And a key feature of Halphen's discovery was that there was an infinite number of ways to view such degenerate conics as limits of general conics, thereby barring all hopes of representation and enumeration via finite symbolic expressions.

To explain this phenomenon, a new representation of conics and their degenerate forms was required. To that end, Halphen attached a self-polar triangle to each conic of a system; that is to say a triangle in which each side is the polar line of the opposite vertex with respect to the conic, and vice-versa. This can be done

[47] Zeuthen (1910, p. 36).

[48] The question of whether these degenerate conics constitute legitimate counter-examples to Chasles's theorem is rather delicate, and I will limit myself in the present essay to Zeuthen's understanding thereof. Suffice it to say that is very much possible to rigorously formalise and justify either Chasles's and Halphen's theory in the framework of modern algebraic geometry, and that choosing between them hinges upon meta-theoretical factors (computational tractability, geometrical significance, etc.). Note that in the case of Plücker's formulas, representation theorems show that higher singularities can be reduced to those enumerated by said formulas, thereby ensuring the absence of such difficulties.

Fig. 21.3 Halphen's degenerate conic (Zeuthen 1914a, p. 316)



in such a way that, in the neighbourhood of a singularity of the system (i.e., of a degenerate conic), these triangles remain proper.[49] In return, these triangles give rise to particularly useful systems of homogeneous coordinates in which to represent the conics of the system, since expressing the (punctual) equation of a conic in the homogeneous coordinates of a self-polar triangle amounts to diagonalising the matrix of the conic.[50] Thus, in the neighbourhood of a degenerate conic, each conic of the system can be represented via a punctual equation of the form:

$$g_1 Q^2 + g_2 R^2 + g_3 S^2 = 0$$

where $Q = 0, R = 0, S = 0$ are the equations defining the three sides of the self-polar triangle attached to the conic (whose vertices I denote correspondingly q, r, s) and the g_i's are coefficients determining the conic, but also functions of the parameter which describes the entire system of conics.

To investigate the singularities of a system of conic, Halphen explained, was to investigate the behaviour of the g_i's in the neighbourhood of a degenerate conic, that is to say when at least one of these coefficients vanishes. Note that, at all times, at least one of g_i's must be different from zero; and so Halphen assumed, without loss of generality, that this was the case for g_3. The remaining possibilities are thus:

- Only g_1 vanishes; in which case the degenerate conic is a line pair whose intersection coincides with q.[51]
- Both g_1 and g_2 vanish, but at the same rate (i.e., the ratio $g_1 : g_2$ converges); in which case the degenerate conic is a double line coinciding with S.[52]

[49] Halphen (1878, pp. 33–35).

[50] See Semple and Kneebone (1952, pp. 109–112) for justifications.

[51] Indeed, note that the (punctual) limit-equation is $g_2 R^2 + g_3 S^2 = 0$, which is only satisfied by the two straight lines of equations $\sqrt{g_2} R + \sqrt{g_3} S = 0$ and $\sqrt{g_2} R - \sqrt{g_3} S = 0$. These lines intersect at the point of coordinates $(1, 0, 0)$, that is to say q.

[52] To see that, one can form the dual equation $g_2 g_3 Q^2 + g_1 g_3 R^2 + g_1 g_2 S^2 = 0$, which defines the tangents to the conic. Dividing both sides by g_1, the limit equation becomes $\alpha \cdot g_3 Q^2 + g_3 R^2 = 0$ for some finite, non-zero coefficient α. It follows that the tangents to the degenerate conic are the two pencils of equation $\sqrt{\alpha g_3} Q + \sqrt{g_3} R = 0$ and $\sqrt{\alpha g_3} Q - \sqrt{g_3} R = 0$ (those are linear equations

- Both g_1 and g_2 vanish, but at different rates. Supposing g_1 vanishes at a higher rate, the degenerate conic will be formed of one line coinciding with S, with only one special point (namely, q).[53]

This third case yielded a new type of degenerate conic which Chasles's theory proved as incapable of differentiating as De Jonquières's had been in the presence of double lines—that is to say, these conics could appear as "improper solutions, which do not answer the question originally asked, but which are introduced by the instruments (analytical or geometrical) used to solve it."[54] Indeed, in one of his earliest communications on the matter, Halphen had shown how to construct at will specific systems of conics and conditions that would put the lie to Chasles's $\alpha\mu+\beta\nu$ formula due to the very presence of one of these degenerations—which, following Zeuthen, we will call Halphenian degenerations [*Halphenschen Ausartungen*].

To account for the presence of one such degeneration, Zeuthen explained, one could always add a third term '$+\gamma\rho$' to Chasles's formula. However, unlike double lines and line pairs, Halphen's third category of conics contains a multitude of degenerate figures: one for each possible ratio between the vanishing rates of g_1 and g_2.[55] To accommodate Chasles's formula to a given system of conics, one would therefore have to add one new factor for each such ratio present in the system. In sum, no (finite) amount of formal tweaking would save computations in the enumerative geometry of conics, though Halphen's memoir provided the elements of a method for solving each individual problem without resorting to a general expression.[56]

Halphen's conclusions, therefore, were not entirely negative—at least in Zeuthen's reading. Rather than establishing a formal expression constrained by a specific viewpoint on conics (and a related distinction between the general and the particular), Halphen's results constituted a new kind of result, which Zeuthen named an "absolute theorem." Such a theorem, Zeuthen analysed, is one in which "it is no longer required to say or specify 'in general,' words which mean that a theorem is also true in every particular case, so long as it is viewed as a particular case."[57] In

in dual coordinates, hence the equations of pencils of tangents rather than lines of points). These pencils intersect at the line of coordinate (0,0,1), that is to say S, which is in agreement with the fact that the limit point-equation is $S^2 = 0$.

[53] That the conic as locus of points turns into the double line $S^2 = 0$ is obvious from the general equation. Furthermore, dividing the dual equation by g_2, one obtains $g_3 Q^2 + \frac{g_1}{g_2} g_3 R^2 + g_1 S^2 = 0$. Since g_1 vanishes at a greater rate, the limit form of this equation is $g_3 Q^2 = 0$; hence the pencil of lines centred around q forms a double pencil of tangents to this degenerate conic.

[54] Zeuthen (1890, p. 462).

[55] Equivalently, one may associate each mode of degeneration to the ratio $\frac{m}{n}$, where m and n are the (different) rates at which the semi-axes a and b of a degenerating conic are vanishing (Zeuthen 1914a, pp. 318–322).

[56] Zeuthen (1910, p. 38).

[57] Letter from Zeuthen to Halphen dated Dec. 18th 1879, Bib. de l'Institut, Paris Cod Ms 5624/237; Michel (2020, p. 494).

lieu of a equation constrained by hidden assumptions and a domain of applicability, Halphen's theory had resulted in an enumerative method which encompassed all possible singularities of systems of conics, and therefore all possible determinations of the meaning of the word 'general.'

And this is exactly what one finds, if not directly in Halphen's own memoirs, in the chapter on systems of conics in Zeuthen's *Lehrbuch*. Given a certain system of conics and a geometrical condition, Zeuthen teaches therein how to methodically investigate which degenerate forms occur in said system and whether or not the condition can be satisfied by some Halphenian degenerations. More importantly still, he shows how to determine the genres [*Gattungen*] of Halphenian degenerations present in the system and satisfying the condition, that is to say the values corresponding to the ratios of degeneration coefficients described above. Only when this is done does he provide a simple computational rule to subtract from Chasles's $\alpha\mu + \beta\nu$ formula the number of foreign solutions (with multiplicity), thereby solving the enumerative problem at hand.[58]

Such a method does not proceed by the mechanical application of a single formula. Instead, it requires one's constant surveying of the figures of a system and acute awareness of all the particulars cases one can find amongst them—that is to say, all the ways curves can degenerate. In Zeuthen's view, however, this slow and methodical examination has its own rewards. Not proceeding from symbols, it does not embark implicit assumptions about what counts as general or as particular. Rather, it provides an overview from which such specifications become superfluous, a gaze that encompasses at once the infinite variety of particulars that can present themselves in the course of the application of a method to a certain figure. It is the conceptualisation of this gaze, borne out of practical and pedagogical necessity to overcome the failures of symbolic expression, that form the basis for Zeuthen's conceptualisation of holistic intuition—to which we now turn.

21.4 Bergson in Copenhagen

As a matter of fact, the notion of intuition scarcely features in Zeuthen's writings prior to 1914, whether it be his correspondence, his publications, or his public addresses. Perhaps wary of associations with unrigorous reliance on visual reasoning, Zeuthen instead contrasted his use of geometrical methods to the analysts' computations. In the pages of the *Lehrbuch* and in his address to the Academy in Copenhagen, however, Zeuthen began embracing a novel rhetoric of intuition; and

[58] These methods and their application to several examples are detailed in Zeuthen (1914a, pp. 315–328). Note that, except in trivial cases, the number of degenerate conics in a system of conics is always finite. In fact, in a system of characteristics (μ, ν), there are in general $2\mu - \nu$ double lines and $2\nu - \mu$ line pairs (a result which Chasles had already found in 1864, and which Zeuthen placed at the centre of his 1865 dissertation). Consequently, there can only be a finite number of Halphenian degenerations in a given system, though this number can be arbitrarily large.

he did so after revisiting this concept and placing it at the centre of a decidedly holistic epistemology of mathematics. This transformation, it turns out, reflects Zeuthen's engagement with his immediate scientific entourage in Copenhagen, and especially with the philosopher and psychologist Harald Høffding. In the years directly preceding the publication and presentation of his textbook, Zeuthen had attended Høffding's lectures on the history of philosophy and on experimental psychology. In particular, Høffding's positivist twist on the philosophy of Henri Bergson and his holistic proposal to think anew scientific cognition had left a strong impression on Zeuthen. No passive listener, Zeuthen found creative ways to connect these lectures with his own expertise in algebraic geometry, its history and its practice; and through this encounter he re-centred his epistemology of mathematics around the distinction between symbolic and intuitive cognition.[59]

Indeed, key to Zeuthen's late epistemological writings is a distinction first introduced by Bergson in his 1903 *Introduction à la Métaphysique* between "two profoundly different ways of knowing a thing." Respectively, those forms of knowledge were analysis, which relies on symbolic representations and a specific viewpoint on said thing, and intuition, which eschews symbols and operates from within the thing itself.[60] Or, to frame the distinction in the terms Bergson would employ in his 1907 *L'Évolution créatrice*, our epistemic life can take either one of two directions: that of "intelligence" or that of "instinct."[61] The first form of cognition can only yield "relative" knowledge, Bergson argued, due to its being obtained from a fixed location outside the object of study—knowledge of a thing being acquired through the study of what this thing *isn't*, namely symbols standing in its place. By contrast, the second form yields "absolute" knowledge, for it is deployed without accommodating for any biased perspective or any translation into foreign symbolism—it is direct knowledge of a thing acquired from within it. What does it mean to really know a town, for instance? Suppose one uses a camera to capture it through an arbitrarily large number of still photographs, possibly even with atomic precision. Each of these photographs will depict but one fixed, immobile perspective on its streets and its monuments. The knowledge thus obtained, Bergson argued, can only be qualitatively different from (and inferior to) the inner comprehension of the specific organisation and dynamics of that same town which one will grasp by wandering and living in it—by inhabiting it as one

[59] Høffding lectured on Bergson's philosophy during the year 1913, that is to say the very year preceding Zeuthen's address in which the concepts of (Bergsonian) intuition and symbols are made central. These lectures were later published and translated in several languages, including French (1917) and English (1920).

[60] This distinction, of course, was not entirely new, and Bergson framed it also as a reaction to certain neo-Kantian epistemologies and metaphysics; a topic well outside the scope of the present chapter.

[61] Bergson (1907, pp. 146–147). Note that this distinction is not a dichotomy: Bergson immediately followed its introduction by the acknowledgement that "neither intelligence nor instinct can be found in its pure state . . . there is no intelligence in which one cannot find traces of instinct; no instinct which isn't surrounded by a halo of intelligence" (Bergson 1907, pp. 147–148).

inhabits a cohesive and mobile whole. No matter how many photographs of this faraway town are obtained, and the reconstruction of its layout one can obtain through the combination of these motionless *clichés*, its inner life and distinct ambience remain a mystery to those who cannot experience it directly and fully.[62]

The similarity between Bergson's distinction and Zeuthen's epistemological assessment of the historical development of enumerative geometry is striking. The passage from De Jonquières's $\alpha\mu$ formula to Chasles's $\alpha\mu + \beta\nu$ formula to accommodate double lines, as well as all the supplementary terms '$+\gamma\rho$' potentially introduced to account for each individual Halphenian degeneration, may all be regarded as constitutive of an 'analysis' of the enumerative geometry of conics. That is, seeking to patch faulty formulas by adding these terms is akin to the endless conjunction of yet more viewpoints on one same object of knowledge, which always remains imperfectly encoded into symbols. Such finite symbolic expressions, as we have seen, can only yield a 'relative' knowledge of the enumerative geometry of conics, for they enforce specific distinctions between the particular and the general—they disregard the fact that certain degenerations give rise to artificial solutions to enumerative problems. Halphen's 'absolute theorem,' by contrast, is constitutive of a fuller understanding of the theory of these curves that encompasses all possible perspectives on them, precisely because its methodical principle is located within the degenerating conic itself, with its infinity of subtle variations. It is perhaps for this reason that Zeuthen, in a letter to Halphen directly following the publication of the latter's counter-examples, argued that while "[Halphen] attributed to computations [*au calcul*] the honour of putting him on the right path regarding $\alpha\mu + \beta\nu$ ... [Halphen] seemed [to Zeuthen] to be the most geometrical amongst those who tackled this question—as he was the only one who had managed to shed complete light on it."[63] Halphen's theory, as read by Zeuthen, thus overcomes the limitations of discursive representations and provides an intuitive knowledge of the theory of conics.

But this comparison of Zeuthen's epistemology of (enumerative) geometry and of Bergson's metaphysics is not without issues. When he described Halphen's theorem as "absolute," Zeuthen could not have possibly been aware of Bergson's association of this word with intuitive knowledge, for obvious chronological reasons. More substantial is the apparent tension nested within Zeuthen's repurposing of the conceptual categories of a philosopher so uninterested in the possibility of genuine intuitive knowledge in geometry—or mathematics more broadly.[64] The way out

[62] Bergson (1903, pp. 1–3). At the time of Høffding's lectures, Bergson's theory of intuition was the subject of much discussion, with scholars in various European countries but also in China as well as in the U.S. producing both interpretative and critical assessments. Various recent publications in *Bergsoniana*, the journal edited by the *Société des amis de Bergson*, explore this multi-faceted and global reception.

[63] Letter from Zeuthen to Halphen dated Dec. 15th 1877, Bib. de l'Institut, Paris, Cod Ms 5624/226, Michel (2020, p. 475).

[64] In various texts, Bergson characterises geometry as a special domain of knowledge in which our (analytical) intelligence can triumph alone, and where intuition is not only superfluous, but

of these difficulties is to recognise that, as reception history has amply shown, one should not seek to confront Zeuthen's use of the words 'symbol' and 'intuition' with an hypothetical, undiluted Bergsonism, but rather with the local interpretation and reinvention thereof to which he had been exposed.[65] And Bergson in Copenhagen was a figure whose views had been mostly mediated and reshaped by Høffding—a self-avowed empiricist in constant conversation with the local scientific community, from Zeuthen himself to the members of the Bohr family, and who sought to accommodate Bergsonism to what he understood to be the scientific practice of his contemporaries.[66] For this reason, we must first make a detour through Høffding's lectures before returning to Zeuthen's defense of intuition.

While finding plenty to praise and promote in Bergson's epistemology and metaphysics, especially in its polemical charge against mechanical conceptions of human psychology and cognition, Høffding nonetheless raised one central objection against the theses outlined above. Analysis and intelligence, he argued, should not be pitted *against* intuition and instinct, but rather these faculties should be viewed as working in tandem, on a historical as well as on individual scale.[67] More precisely, while he agreed with Bergson that intuition constituted the ideal form of knowledge one should strive for in most, if not all, epistemic endeavours, Høffding argued that instinct and intelligence had in fact complementary roles to play towards this very goal. "Pure instinct," he explained, "points immediately and exclusively in a single direction," namely that of basic needs and interests. Only the awakening of intelligence allows one to go past this stage, for it "delivers instinct from need, makes it entirely disinterested, and thereby plunges it into the ... whole of our existence": it opens the very possibility of a richer understanding and appreciation of the world, which intuition will eventually provide to those who attain it. In sum, Høffding argued, "there was an ascent from instinct, through intelligence, to

made impossible by the spatialisation of duration and motion. See for instance Bergson (1907, pp. 174–175).

[65] To discuss the nuances of reception theory and historiography is well beyond the scopes of the present chapter. To fix ideas, it is enough to cite one particularly influential presentation of this approach: "Our current interpretations of ancient texts, whether or not we are aware of it, are, in complex ways, constructed by the chain of receptions through which their continued readability has been effected. As a result we cannot get back to any originary meaning wholly free of subsequent accretions. Meaning is produced and exchanged socially and discursively, and this is true of reading, even in a society like ours, in which it has become, to a greater or lesser degree, a 'private' activity. In order to be read, a text has to be made readable, in a complex process which begins with the acculturation of children and continues through educational institutions to wider interpretative groups" (Martindale 1994, pp. 7–8).

[66] On Høffding's holistic epistemology and its influence on Niels Bohr's understanding of quantum mechanics, see Faye (1991, pp. 77–109). One would be remiss not to note that intellectual exchanges went both ways, and Høffding frequently engaged with the scientific practice of his colleagues to substantiate his philosophical positions, including Zeuthen's cognitive history of mathematics; see for instance Høffding (1915, p. 315).

[67] For related reasons, Høffding is equally skeptical of Bergson's sharp demarcation between science and art (Høffding 1915, p. 299).

intuition."[68] Through the description of such an ascent, Høffding effectively sought to adapt Comtean narratives for the historical progress of scientific knowledge to Bergsonian distinctions between forms of cognition. He did so through the postulation of another law of three stages: the law that presides over the transition from unrefined instinct to a higher form of intuition through the necessary mediation of intelligence—that is, of a symbolic grasp on things.[69]

Elaborating on this critique of Bergson, Høffding thus proposed a classification of the types and levels of intuition, so as to separate the "original and immediate intuition" from that "at which we arrive having passed through the work of analysis," and to identify various transitional phases.[70] Of this classification, "concrete intuition" forms the most elementary level of intuition: an example of it is "the perception of a sensible image upon opening one eye's, image which at this stage forms a certain totality."[71] Other faculties, such as memory or imagination, contribute to this concrete intuition; the common factor being that they summon to one's consciousness a totality in which no parts can yet be analysed, compared, or abstracted—tasks which will be accomplished through the use of intelligence. A second, intermediary level of intuition is then introduced, which in fact relies explicitly on intellectual activity: "analytical intuition," Høffding explains, provides "immediate knowledge of a relation" between two totalities grasped by the first, concrete intuition.[72] Through reflexive comparison of successive perceptions or representations, analytical intuition for instance enables one to grasp the similarity or identity between two objects of concrete intuition, and thus marks the "passage from perception to analysis."

Intelligence may then grasp these very objects through symbols and the method of analysis, but this is not enough to achieve the highest form of knowledge. As Bergson had argued, intelligence compartmentalises the comprehension of an object into the symbolic grasp of its particulars, thus not providing the complete

[68] Høffding (1915, p. 252). Note the presence, here as in several other parts of this text, of then-fashionable evolutionary tropes. At a similar period, Høffding had extensively lectured on Charles Darwin's theory of evolution: "Høffding revealed some reservations about Darwin's scientific theories ... In spite of these reservations, Høffding argued that Darwin's work had a profound impact on philosophy. In Høffding's eyes, Darwin brought to victory a developmental and evolutionary view of life that also characterised the work of Spencer" (Hjermitslev 2014, pp. 134–135).

[69] Høffding made no secret of his admiration for Comte, whom he described in his lectures as "the greatest figure of the nineteenth century" (Høffding 1915, p. 71).

[70] Høffding (1915, p. 253).

[71] Høffding (1915, p. 255).

[72] Høffding (1915, p. 257). Høffding in fact identifies some other intermediary forms of intuition, including a "metaphysical intuition" whose relation to Bergson's intuition is discussed at length. I shall not enter these nuances here, as Zeuthen mostly ignored them and they only marginally concern the present discussion.

understanding one should strive for. This highest form of knowledge, in Høffding's epistemology, is called "synthetic intuition" and defined as:

> The immediate perception of a connection or a totality which may be acquired by going through a series or a group of members or parts, if one has a certain comprehension of their mutual relations. Thus, the action of following a complicated demonstration, or of observing the connection between different points of view under which one and the same subject can be considered, often becomes the object of a comprehensive glance ... Even the work of thought, analysis and demonstration, is useful to the intuition of totality ... The view of the totality is conditioned by the regular connection discovered by the aid of thought ... [This intuition] differs at once from the observation of particular subjects and from the abstract knowledge of the laws which govern their appearance. It sees concrete existence in its individuality, and at the same time as a whole, as a folding together of the general laws by which it has inner connection with the rest of existence.[73]

Analysis, which decomposes the totality first given by concrete intuition into symbols, is thus resolved into a new synthesis. The whole thus obtained is a much more potent object of knowledge as the one provided by concrete intuition, however, as intelligence has brought to the fore the properties of its individuals and allowed for the emergence of the patterns and connections that now structure the whole and its proper understanding. In fact, for Høffding, the sharper the analytical decomposition the clearer the intuitive synthesis; scientific practice thus synergizing with holistic epistemology, *pace* Bergson.

And indeed, like Bergson himself, Høffding spent considerable amounts of time reading and engaging with the sciences of his time, whether it be psychology (to which he himself contributed), linguistics, or even mechanics. In this regard, his rational holism and his theory of the developmental stages of intuition were not merely to be viewed as correctives to Bergson's philosophy, but rather as empirically supported description of the development of each and every form of scientific activity.

In his 1882 *Outlines of Psychology* already, Høffding had borrowed the model for the formation of human languages proposed by the German philologist Friedrich Max Müller to argue for strikingly similar conceptions of the historical development of human cognition. This model, as Høffding recalled it, proposes that "human speech seems to have passed through three stages of development." At the first one of these stages, words consist merely of roots; that is, of primitive sounds which operate as inseparable totalities and in direct correlation with individual sensations. At the second stage, roots coalesce two-by-two to form words—Høffding gives the artificial example of the composed word 'meat-broth,' which serves to distinguish in consciousness between various types of broths, but also to symbolise the conjunction of two kinds of food. And at the third, final stage, roots become so intractably intertwined that words form complex totalities, which only trained philologists can analyse back to their original components—a situation supposedly exemplified by the Aryan and Semitic languages. These stages are perfectly analogous to the levels of intuition discussed above, with words being

[73] Høffding (1915, pp. 257–258).

first associated to unrefined, instinctual totalities; intelligence forming then the first analyses and associations of these first totalities; and a higher intuition finally producing linguistic systems consisting of complex wholes which only a select few can analyse and decompose into the particulars involved in their genesis. Concluding his summary of Müller's views, Høffding then stated in no uncertain terms the vast generality he ascribed to this sort of explanatory framework:

> The same holds good of every conception of a totality, which has been reached through the laborious working up of details; the totality stands out as the object of immediate intuition, of an "intuitive knowledge," from which all discursive elements and processes have vanished. Here custom co-operates; the oftener we have gone through the details, the more completely and easily can the totality come instantaneously before us. Successive apprehension precedes simultaneous.[74]

At this stage, it is worth pausing to consider how we moved from a discussion of *ways of knowing*, starting with Bergson's distinction between analysis and intuition, to *ways of forming* objects of knowledge—such as linguistic systems, their syntax, and their semantics. This is no accidental slip: in Høffding's dynamical conception of truth, forms and objects of cognition mutually determine each other. Colours, for instance, depend not only on the properties of things in the world but also on the beholder's visual sensibility: their visual organs, the range of colours they are used to experiencing, the degree of effort they put into distinguishing between them etc.[75] In similar language, objects of knowledge in general are not merely given but rather form the "result of a process"—a process whose paradigmatic history is provided by the theory of the levels of intuition outlined above.[76] Understanding and recalling the historical and cognitive processes through which totalities have emerged thus appears key in Høffding's philosophy of science. To know how a language functions is not only to know how its components (nouns, verbs, grammatical rules, etc.) relate to each other, but also from which analytical decompositions and operations they arose—that is, their etymology, their relation to ancient tongues, etc.

And this imperative was not only reserved to sciences dealing with human constructions such as language or colour perception: discussing James Clerk Maxwell's analysis of the idea of motion, Høffding stressed its importance for geometry itself:

> Geometry, in its relation to the doctrine of motion is, in fact, derivative, or is a part of the latter; for geometry is peculiarly concerned with the process by means of which figures are produced in space. A line is not originally a mark on the blackboard, which can equally be called BA as AB, but it is the locus of a motion from A to B. The idea of motion lies back of the idea of form . . . We are far too much accustomed to hold fast to ready-made symbols and figures, and to overlook the genetic process of sense-perception and of thought.[77]

[74] Høffding (1891, pp. 163–164). In Müller's theory, these stages are respectively called the *radical*, *terminational*, and *inflexional* stages.

[75] Høffding (1891, pp. 103–104).

[76] Faye (1991, pp. 83–85).

[77] Høffding (1915, p. 111).

One can very well picture how struck Zeuthen must have been upon hearing Høffding establish such connections between holistic philosophy and geometrical practice. Forgetting from which motions curves derive, manipulating symbols whose origin is lost—those were the very flaws he had detected in the work of predecessors and contemporaries, the very dangers against which he had warned students and colleagues. When using formulas such as De Jonquières's $\alpha\mu$ or Poncelet's $m^* = m(m-1)$, Zeuthen had shown, one must at all times remember that one is using symbols representing conics that derive from a specific "genetic process," namely the view that curves are generated through the continuous motion of a point. But where Høffding in his lectures devoted only a few side-remarks to the mathematical sciences, Zeuthen saw an invitation to develop a novel account of geometrical intuition, both as a way of knowing and a way of forming geometrical objects.

21.5 Perceiving Totalities

Having explored Zeuthen's engagement with enumerative geometry as well as Høffding's rational holism, we are finally able to return to Zeuthen's 1914 address (as well as to a few later, related texts) and to give a fuller account of the theses expounded therein. In essence, the main thrust of this address was to sketch the outlines of a 'method of intuition' (to borrow from another key term of Bergsonian philosophy) in modern geometry, and to appraise the epistemic benefits one could derive from embracing such a method in lieu of symbolic (that is, computational) ones.

A first component of this project involved distinguishing between two forms of understanding of a mathematical proof along the lines of demarcation outlined by Bergson and Høffding. By intuitive understanding of a proof, accordingly, Zeuthen understood a "holistic" understanding of a proof—i.e., a total overview of an organically connected chain of propositions, each of which supports and verifies the others. Zeuthen then contrasted such an overview with a "logical" understanding of that same proof, an understanding which is limited to the mechanical derivation of each isolated symbolic step thereof, be it a computation or the invocation of a certain axiom. In other words, while reading a long and complicated proof, one may verify that each individual lemma in it is correctly established or that each individual computation in it is valid, and therefore be forced to acknowledge that the final statement has been proven. These verifications, however, are isolated and independent from each other (although one must ensure that they follow on a local level). Possessing an intuitive understanding of that same proof, by contrast, means being able to "encompass [at once] a recollection of all the particular inferences" present in it; that is, being able to account for the mutual dependence of its lemmas, its definitions, and its computations. From a foundational perspective, logical understanding may be enough to ground mathematics as a body of theorems; but Zeuthen argued that it leaves the working mathematician exposed to failure and "personal errors" [*personlig Fejl*], and therefore cannot provide the level of certainty and comprehension that a total intuition does.[78] Already in 1877, Zeuthen had privately confessed to Halphen that "a mathematical truth [was] to him somewhat frightening, so long as he [did] not see its connection to other truths, or a path that [could] lead him to it."[79]

Zeuthen's worry makes sense when viewed in light of the above discussion of theorems of algebraic geometry such as Poncelet's $m^* = m(m-1)$ formula. Suppose such a formula is invoked in the course of a complex proof, possibly even as a lemma with a specific subproof, and suppose that its symbolic expression is correctly employed to derive another result. At the local level of mechanical verification, both steps might appear as perfectly valid, but one may still worry that, in the course of the larger proof, Poncelet's formula be implicitly used in cases where its generality is faulty—i.e., that the assumptions it carries regarding the particular/general distinction in the theory of algebraic curves be no longer acceptable in other parts of the proof. Or, to borrow another domain of mathematics which Zeuthen frequently cites in relation to such epistemic fears, one may locally use valid computations with infinite series only to mistakenly use results of said computations outside of their initial domain of applicability—for instance, in domains where these series no longer converge. More generally, Zeuthen explains, in the absence of intuitive oversight, he does not "feel altogether secure regarding the possibility that the result obtained with one inference had not been used in a subsequent one with a greater extension than had been proved."[80] Intuitive understanding prevents such issues precisely because the encompassing gaze it provides connects at once all particular elements of the proof, thereby making it immediately apparent whether or not such errors are being committed or not.[81] In this view, the certainty of mathematical knowledge has less to do with its logical foundations and mechanisms than with the trained form of human intuition that can perceive it as a whole, control its internal coherence, and understand its organisation.[82]

But Zeuthen's concept of holistic intuition was not limited to the comprehension of mathematical proofs. Borrowing from Høffding's general theory of intuition, Zeuthen relied on this latter concept to construct an account of the historical and psychological formation of mathematical objects, and especially geometrical figures. And in such an endeavour, Høffding's comparison of various levels of

[78] Zeuthen (1914b, p. 277).

[79] Letter from Zeuthen to Halphen dated Dec. 1st 1877, Bib. de l'Institut, Paris, Cod Ms 5624/225; Michel (2020, p. 473).

[80] Zeuthen (1914b, p. 277).

[81] In this sense, intuition is the faculty that undergirds the rational use of methods as described previously.

[82] In his final volume on the history of Ancient Greek geometry, Zeuthen would similarly argue that what makes mathematics a "reasoned science" is not merely "logical conclusions derived from arbitrary suppositions," but the constitution of a "whole, in which hypotheses and conclusions are equally accounted for" (Zeuthen 1917, p. 174).

Fig. 21.4 Edgar Rubin's vase, from his doctoral dissertation *Synsoplevede Figurer* (1915), a paradigmatic illustration of Gestalt psychology

intuition would prove key in identifying different ways to perceive and apprehend geometrical totalities.

The most primitive form of geometrical knowledge (and thus the most primitive geometrical figures), Zeuthen asserted, had emerged through the use of a primitive form of intuition: the "apprehension of a total image, through a combination of bodily and mental faculties." Vision, of course, but also tactile perception, memory, and the ability to bring together or collate recollections all partake in the mostly unconscious formation of the first intuitive figures, which appear as complex totalities, out of reach of conscious (e.g., linguistic) description or analysis. As Edgar Rubin's well-known studies of figure-ground perception (see Fig. 21.4) had shown, for Zeuthen, the identification of shapes and borders within the complete image perceived by one's eyes could only derive from the formation and use of an intuitive sense.[83] Moreover, Zeuthen noted, such a faculty was even subject to historical change, as painters of the Renaissance would increasingly appreciate perspective and thus identify new intuitive features within perceived images—just as perceptions of colours in Høffding's psychological theories partially depend on the biological and cultural environment of the beholder.[84] Through a historical and cognitive study of this process, Zeuthen thus depicted the emergence of geometry as the deployment and refinement of a "holistic sense" [*Helhedsfornemmelse*], that is to say a primitive form of intuition. Such a sense may be said to be holistic, he argued, because it "encompasses the particulars of the whole, none of which yet appear to consciousness": it gives rise to intuitive images in which one can perceive shapes or borders whilst keeping subconscious the process of aggregation of memories and sensations that led to this perception.[85]

[83] Rubin defended his dissertation at the University of Copenhagen in 1915. By this point, he had been studying psychology and philosophy under (among others) Harald Høffding for some seven years. The cousin of Niels Bohr, he was already well integrated into the Danish scientific elite. His work would be rapidly co-opted by major proponents of Gestalt psychology such as Max Wertheimer (to whom I return shortly below). On Rubin's scientific training and psychological contributions, see Pind (2014).

[84] Zeuthen (1914b, p. 53).

[85] Zeuthen (1914b, p. 274). This psychological account of the formation of geometrical objects is expressed in greater detail in Zeuthen (1917, pp. 46–54; 373).

This pre-scientific intuitive grasp on geometrical objects and knowledge, Zeuthen explained, was allegedly that of "Eskimos" who "should be able to draw a very reliable coastline," for they had—out of practical necessity, presumably—trained their senses so as to identify and represent the shape of the surrounding coastlines and to form an intuitive sense of the mathematical concept of "similarity" between shapes.[86] Such a sense could prove extremely useful, as the intuitive knowledge that these imaginary Eskimos possess of their surroundings is sufficient for "the drawing of a very reliable coastline" by simple homothety; that is, without a conscious (let alone symbolic) grasp on the particulars that form the relation of similarity between the coastline and their intuitive representation of it.[87]

However, Zeuthen continued, this sense is not a sufficient basis for the emergence of mathematics as a rational endeavour. This is due precisely to the fact that intuitive images thus obtained erase the particulars from which they were formed. In drawing intuitive maps of coastlines, for instance, Zeuthen's Eskimos are incapable of introducing symbols for the magnitudes and angles present in the similar representations (i.e., maps) of the borders of their land, thereby not allowing for the development of a general theory of these figures.[88] The introduction of such symbols, whether it be numbers for lengths, algebraic equations, or technical words like 'similarity' or 'plane', thus marks the passage from a primitive form of intuitive certainty to scientific knowledge—mirroring the first transition of Høffding's own law of three stages, i.e., the first ascent from naive intuition through intelligence.

[86] Zeuthen (1914b, p. 276). Note that the term Eskimo is no longer used to describe the ethnic groups Zeuthen was referring to and is in fact regarded as derogatory by many. One speaks instead of the Inuit and the Yupik. Zeuthen's attribution of various levels of intuition to various peoples echoes Klein's infamous 1893 lecture on spatial intuition during the Evanston Colloquium, in which a distinction between naive and refined intuition is drawn and distributed across racial lines (Klein 1894, pp. 41–42). Quite strikingly, Klein refers to Zeuthen's historical research to position Euclid as a pivotal figure in the passage between these two forms of intuition. In fact, in another of these lectures given on American soil, Klein also presented Zeuthen's work on plane quartics in the context of his defence of *anschauliche Mathematik*, though he ascribed to this latter term a rather different meaning. On Klein's racial theory of intuition, see Jemma Lorenat's chapter in this very volume.

[87] Elsewhere, Zeuthen described in similar fashion the emergence of an intuitive image of space in pre-scientific Greek geometry (Zeuthen 1917, p. 49). Zeuthen also noted that one can of course be taught to perceive such forms, something he viewed as routinely achieved in modern classrooms. However, for historical and epistemological purposes, he elected to focus on the spontaneous, organic emergence of such holistic intuitions (Zeuthen 1917, p. 51).

[88] It must be noted here that, as with his repurposing of Rubin's work, Zeuthen was effectively connecting his philosophy of geometry to salient cultural references adapted to his Copenhagen-based audience. Indeed, artefacts coming from Greenland (and thus from the hands of otherised 'Eskimos') were of great interest to the Danish cultural elite, with explorers' travel reports and public displays in ethnographic museums such as the National Museum of Denmark (*National-museet*) garnering much success. And Inuit maps, whether drawn on flat surfaces or even carved on wood in three dimensions, had prominently featured in such displays (see Figs. 21.5 and 21.6). On these latter carved maps, which were brought to Copenhagen by the Danish explorer Gustav Holm in 1885 (De Jonghe 2022, pp. 60–63).

Fig. 21.5 R. J. Flaherty & Wetallok, *Map of Belcher Islands* (1909). US Library of Congress, www.loc.gov/item/2021668436/

Fig. 21.6 Wooden Inuit Maps, Ammassalik Coast (Eastern Greenland), Greenland National Museum and Archives

For reasons explained above, still, such symbolic encoding of intuitive images cannot suffice to buttress the mathematical practice of actual human agents. "From a logical standpoint," Zeuthen stressed, "it is indeed enough to possess this mechanically accumulated sum of knowledge [*mekanisk opdyngede Sum af Viden*] expressed via symbols, but from a psychological standpoint, it is important ... to gather all this knowledge into an organic and connected whole [*organisk sammenhægende Hele*]." Having symbolically analysed original intuitive totalities, one must still return them as wholes to intuition—but to a higher form thereof, which Zeuthen calls "holistic cognition" [*Helhedserkendelse*].[89]

This holistic cognition, corresponding to the final stage of Høffding's developmental model, consists in a form of holistic intuition wherein particulars can freely be conjured up before our consciousness, through the controlled use of symbols if necessary. It thus keeps the holistic quality of intuitive sense, thereby guaranteeing the clarity and understanding that mechanical reasoning fails to provide, but it also possesses the scientific quality that only symbols could bring by analysing the intuitive image *qua* totality into individual parts. This level of

[89] Zeuthen (1914b, p. 276).

intuitive understanding is what one should strive for while mastering a formula, Zeuthen explains:

> Such a formula is entirely composed of symbols. They who know [these symbols] may apply this formula to such individual cases [*Enkelttilfælde*] as are obtained by giving determined numerical values [*bestemte Talværdier*] to symbols of magnitude [*Symbolerne for Størrelser*] ...The wholly new investigations in which [this formula] may be useful, however ...may only be carried out by one who possesses an **intuitive overview**, which regroups in one **all the particulars of the formula**, preferably with the addition of **a recollection of the role played by these particulars in its genesis**.[90]

Such a process is not limited to ancient or elementary geometry: in fact, Zeuthen notes, it can still be observed in the latest branches of algebraic geometry, such as the theory of curves of the fourth order.[91] This general description of what it means to master a formula must then be put in perspective with Zeuthen's diagnosis regarding the imperfections and uses of Chasles's $\alpha\mu + \beta\nu$ formula. This is, after all, a formula composed entirely of symbols and which may be applied (in fact, which had been applied) to a great many individual cases in order to determine numbers of conic sections satisfying certain conditions. As Halphen's memoirs had shown, however, to employ this formula safely and fruitfully in further research, one had to master more than the symbols featured in it or the elementary rules of arithmetic. Rather, one had to maintain constant overwatch over the infinite variety of particulars that can arise in the use of this formula, that is to say, all of the ways conics being enumerated could degenerate and cease being 'general' curves. To master the formula is therefore to remember which viewpoints on conics these symbols originated from (i.e., the double viewpoint on conics as point- and line-

[90] Zeuthen (1914b, p. 278). Emphasis mine. At a similar period, though with different mathematical and philosophical references in mind, Henri Poincaré also critiqued the insufficiency of a purely formal (or logical) conception of mathematical knowledge while stressing the importance of understanding the genesis of mathematical proofs and concepts. This is why Poincaré, like Zeuthen, emphasised the necessary role played by intuition in the formation *and* the justification of said knowledge. One reconstruction of Poincaré's views holds for instance that "pure intuition is necessary [to] *understand* [pure mathematics'] justifications and its proofs. The context of justification is pragmatically connected with the logical reconstruction of the genesis because 'to understand mathematics means to learn their development. So, if S has a true belief for p, justified by a formal proof, it does not follow that he is understanding p" (Heinzmann 1999, p. 45). Repurposing language quite close to Zeuthen's, Poincaré also placed mathematics on both sides of the intuition/intelligence divide by stressing the importance of "aesthetic sensibility" in mathematical practice as well as the inability of logic to provide cohesive understanding of the unity of a theory or proof: "When the logician has resolved each demonstration into a host of elementary operations, all of them correct, he will not yet be in possession of the whole reality; that indefinable something that constitutes the unity of the demonstration will still escape him completely. What good is it to admire the mason's work in the edifices erected by great architects, if we cannot understand the general plan of the master? Now pure logic cannot give us this view of the whole; it is to intuition we must look for it" (Poincaré 1952, p. 126). Unlike Zeuthen, however, Poincaré did not go as far as theorising a use and a form of intuition that could lead to the sort exactness and certainty that formal arguments provide.

[91] Zeuthen (1914b, p. 276).

figures) so as to know when it can be mechanically applied and when a certain number of foreign solutions had to be subtracted. Whether it be geometrical proofs or objects, proper knowledge requires both holistic overview and acute awareness of the genetic processes they derive from—the trained gaze which first constituted them.

21.6 Conclusion: Geometry as a Way of Seeing

Zeuthen's epistemology and ontology of mathematics, now situated at the intersection of two specific locales—the emerging community of (algebraic) enumerative geometers and Harald Høffding's lecture halls—thus appears indissociable from scientific and intellectual trends on a much wider scale, both in a geographical and a disciplinary sense. Alongside a new generation of scientists lamenting and reacting to the rise of atomising, mechanical explanations of the world, Zeuthen sought to construct in his own mathematical parcel an alternative form of objectivity and cognition: one that would rely on critical forms of reasoning rather than automatic calculations, on one's refined intuition rather than on supposedly trustworthy symbolic rules, and on an appreciation for the historical and psychological origins of the figures under study. Or, in other words, Zeuthen sought to theorise a conception of mathematics, its epistemology as well as its ontology, that does not elide the figure of the mathematician at their centre: that is, of the epistemic agent who perceives, connects, understands, and constructs objects of knowledge.

That such concerns and reactions were quite topical at this period is immediate if one compares the claims made above with the definition of "trained judgement" proposed by Lorraine Daston and Peter Galison in their history of scientific objectivity, an ideal which they argue emerged around the turn of the twentieth century and define as

> an act of cultivated perception and cognition ... that was both anti-algorithmic and anti-mechanistic. Trained judgement ... [stands] opposed to—or perhaps on top of—the fragmented building-up, the mechanically calculated, automated, protocol-driven set of procedures. Scientific image judgement had to be acquired through a sophisticated apprenticeship, but it was a labour of a very different sort from the rehearsed moves of the nineteenth-century mechanical objectivist. Interpreted images got their force not from the labour behind automation, self-registration, or absolute self-restraint, but from the expert training of the eye, which drew on a historically specific way of seeing.[92]

One paradigmatic example of this specific conception of and quest for scientific objectivity lies in the core tenets of Gestalt psychology as they had been formulated in the early 1910s, e.g. by Austro-Hungarian psychologist Max Wertheimer. Following Gabriele von Wartensleben, one of Wertheimer's first and most acute students,

[92] Daston and Galison (2006, p. 331).

one can encapsulate those tenets in the following theses:[93]

- "The contents [*Erlebnisinhalte*] of our consciousness are mostly not summative, but constitute a *particular characteristic togetherness* [*Zusammensein*]."
- "Most impressions are grasped as chaotic masses ... on the way to sharper formations [*Gestalten*]. What is finally grasped are 'impressions of structure' [*Gebildefassungen*] ... They are something specifically *different* from and more than the summative totality of the individual components. Often the 'whole' is grasped even before the individual parts enter consciousness."
- "The process of knowing is very often a process of 'centring,' [*Zentrieren*] i.e., of 'structuring' [*Gestalten*] that particular aspect which provides the key to an orderly whole, a unification of the particular individual parts that happen to be present."
- "The entity that results from the knowledge process depends in many respects not only on the object, but also on the observer. Thus there are several ways of grasping many phenomena, but generally only one can be correct: that which makes all states understandable and derivable from the central 'idea' and thus gives meaning to the entire given."

Drawing out the parallel to Zeuthen's reconstruction of the origins of geometrical cognition thus yields a striking match. In it, geometrical figures first appear to our consciousness not as collections of shapes and magnitudes, but as wholes which we progressively learn to comprehend by identifying borders and similarities. These identifications amount to grasping structures on the whole, through which individual parts (the interior and the exterior of a given line, for instance) are brought to consciousness—an activity we may also regard as a centring of the perspective on totalities. Through such centring, the whole is made orderly: its particulars are brought to consciousness, but one's intuitive grasp still oversees their mutual connection as well as the perspective from which they originate. Crucially, the mathematical object resulting from this process (e.g., the general conic section) depends on the perspective from which the data of unrefined intuition was centred: this is exactly what Zeuthen had explained when dissecting the history of algebraic curves and systems of conics.[94]

[93] von Wartensleben (1914, pp. 1–3), emphases in the original, cited in Ash (1995, pp. 123–124) (whose translation I reproduce with minor alterations). Ash describes this summary as "[containing] nearly all of the fundamental principles of Gestalt theory, with their implications for logic and the theory of knowledge." For a study of the cultural history of holism and Gestalt psychology in Germany at this period, see also Harrington (1996). It must be noted here that the literature on scientific holism suffers from a quasi-exclusive focus on the German-speaking territories, thereby not fully restoring the specificities of the brand of holism developed simultaneously by the Danish authors discussed here and in the previous pages. This gap is yet to be filled.

[94] This parallel should not be misconstrued as an indication that Høffding, Zeuthen's foremost psychological reference, was a Gestaltist. Rather, his own psychological theory is closer to that of the English associationists such as Mill and Spencer; see Pind (2014, pp. 32–37). Furthermore, the peak of Høffding's engagement with psychological research predates the rise in popularity of

Gestalt psychologists were no relativists, however. Rather, their quest was one "for objective order that lies not behind, but within the flux of experience."[95] This is why von Wartensleben, upon concluding her summary, stressed the existence of one superior perspective on the totalities present to our consciousness. It is not enough to say that one can identify various structures in them; one must still seek to join them into one all-encompassing gaze. In Zeuthen's holistic epistemology of geometry, such a perspective was to be found in the 'absolute' theorems of sort exemplified by Halphen's enumerative methods. Absolute knowledge, here, is one in which intuitive apprehension of the totality is maintained (thereby eschewing the ruinous reduction to a series of symbolic inferences that computers limit themselves to), but in which particulars can always be brought up to consciousness and critically assessed if necessary, the modalities of their genesis being equally transparent to the knowing agent. The art of the intuitive geometer, in sum, is the art of perceiving totalities; and while axioms and inferential rules may found geometry as a logical system, it is this art which, for Zeuthen, founds geometry as mathematical practice.

Understood in this way, Zeuthen's intuitive cognition appears constitutive of a 'way of seeing' geometrical figures, a technique for the perception of a totality that allows for specific modes of (symbolic) individuation, a gaze that requires individual training, self-control, but also slow and massive historical progress.[96] Beyond, or rather beneath epistemological discussions of the respective merits of intuition and computations, Zeuthen's was an attempt at reconfiguring the relationship between the mathematician and their objects, and at reinventing the practices of objectivity which modern enumerative geometry called for.

Acknowledgments My gratitude goes to Viktor Blåsjö for his linguistic help with the Danish sources; all remaining errors being of course solely mine. Previous versions of this paper have been significantly improved thanks to remarks from Karine Chemla, Chanchan Guo, Eric M. Gurevitch, Francois Lê, Jemma Lorenat, and David Waszek. I also wish to thank audiences at the Toeplitz Kolloquium in Bonn and the Geometry Center in Utrecht for their useful feedback on early presentations of this material.

Gestaltist theories. Many of Høffding's insights, however, would later be put to good use by Gestalt theorists, including but not limited to Rubin; see Ash (1995, pp. 84–85).

[95] Ash (1995, p. 2). See also, a few paragraphs below: "[The Gestalt theorists] challenged the empiricist assumption that 'sense data' are the 'atomic facts' of experience by arguing that there are no such unambiguous 'data.' Rather, they maintained, the objects we perceive are always located in what would now be called self-organising systems—constantly changing dynamic contexts or situations, of which our phenomenal selves, too, are parts."

[96] I borrow this expression from Berger (1972), though Berger himself shaped it on the back of his reading of Walter Benjamin's 1935 essay *The Work of Art in the Age of its Mechanical Reproduction*, and especially the suggestion expressed therein that modes of human perception (of works of art) are not ahistorical, but in fact did and do change in keeping with transformations in culture and civil society. For a comparable perspective on the historicisation of ways of perceiving geometrical figures, see Lorenat (2015a).

References

Ash, M. G. 1995. *Gestalt Psychology in German Culture, 1890–1967. Holism and the Quest for Objectivity*. Cambridge: Cambridge University Press.

Berger, J. 1972. *Ways of Seeing*. London: Penguin Books.

Bergson, H. 1903. Introduction à la métaphysique. *Revue de métaphysique et morale* 11: 1–36.

Bergson, H. 1907. *L'Évolution créatrice*. Paris: Félix Alcan.

Blåsjö, V. 2021. Historiography of Mathematics from the Mathematician's Point of View. In *Handbook of the History and Philosophy of Mathematical Practice*, ed. B. Sriraman. Berlin: Springer.

Brouwer, L. E. J. 1975. *Collected Works Of L. E. J. Brouwer. Vol. 1: Philosophy and Foundations of Mathematics, edited by A. Heyting*. Amsterdam, Oxford: North Holland Publishing Company.

Chasles, M. 1864. Construction des coniques qui satisfont à cinq conditions. Nombres des solutions dans chaque question. *Comptes rendus de l'Académie des sciences* 58: 297–308.

Chemla, K., R. Chorlay, and D. Rabouin, eds. 2016. *The Oxford Handbook of Generality in Mathematics and the Sciences*. Oxford: Oxford University Press.

Cremona, L. 1863. Sulla teoria delle coniche. *Giornale di Matematiche* 2: 17–20.

Daston, L. 1994. Enlightenment Calculations. *Critical Inquiry* 21: 182–202.

Daston, L., and P. Galison. 2006. *Objectivity*. Cambridge: Zone Books.

De Jonghe, B. 2022. *Inventing Greenland: Designing an Arctic Nation*. New York and Barcelona: Actar Publishers.

de Jonquières, E. 1861. Théorèmes généraux concernant les courbes géométriques planes d'un ordre quelconque. *Journal de mathématiques pures et appliquées* 6: 113–134.

Eckes, C. 2018. Weyl's Philosophy of Physics: From Apriorism to Holism (1918–1927). *Philosophia Scientiæ* 22: 163–1927.

Faye, J. 1991. *Niels Bohr: His Heritage and Legacy. An Anti-Realist View of Quantum Mechanics*. Dordrecht: Springer.

Ferreirós, J. 2007. *Labyrinth of Thought. A History of Set Theory and Its Role in Modern Mathematics. Second revised edition*. Basel: Birkhauser.

Gergonne, J. D. 1827. Géométrie de situation. Recherches sur quelques lois générales qui régissent les lignes et surfaces algébriques de tous les ordres. *Annales de mathématiques pures et appliquées* 17: 214–252.

Gray, J. 2007. *Worlds Out of Nothing. A Course in the History of Geometry in the 19th Century*. London: Springer.

Gray, J. 2008. *Plato's Ghost. The Modernist Transformation of Mathematics*. Princeton and Oxford: Princeton University Press.

Haffner, E. 2017. Strategical use(s) of arithmetic in Richard Dedekind and Heinrich Weber's *Theorie der algebraischen Funktionen einer Veränderlichen*. *Historia Mathematica* 44: 31–69.

Halphen, G.-H. 1878. Sur les caractéristiques des systèmes de coniques et de surfaces du second ordre. *Journal de l'École Polytechnique* 45: 27–89.

Harrington, A. 1996. *Reenchanted Science. Holism in German Culture from Wilhelm II to Hitler*. Princeton: Princeton University Press.

Hawkins, T. 1980. Non-Euclidean Geometry and Weierstrassian Mathematics: The Background to Killing's Work on Lie Algebras. *Historia Mathematica* 7: 289–342.

Heinzmann, G. 1999. Poincaré on Understanding Mathematics. *Philosophia Scientiæ* 3: 43–60.

Hjermitslev, H. H. 2014. The Danish Commemoration of Darwin in 1909. In *The Literary and Cultural Reception of Charles Darwin in Europe, vol. 3*, ed. T. F. Glick and E. Shaffer, 128–159. London and New York: Bloomsbury Academic.

Høffding, H. 1891. *Outlines of Psychology*. London: MacMillan. Translated by Mary E. Lowndes.

Høffding, H. 1915. *Modern Philosophers and Lectures on Bergson*. London: MacMillan. Translated by Alfred C. Mason.

Huehn, H., and J. Vigus, eds. 2013. *Symbol and Intuition. Comparative Studies in Kantian and Romantic-Period Aesthetics*. New York. Routledge.

Janik, A., and S. Toulmin. 1973. *Wittgenstein's Vienna*. New York: Simon and Schuster.

Kleiman, S. L. 1980. Chasles's Enumerative Theory of Conics: A Historical Introduction. In *Studies in Algebraic Geometry*, ed. A. Seidenberg, 117–138. Providence: American Mathematical Society.

Kleiman, S. L. 1991. Hieronymus Georg Zeuthen (1839–1920). In *Enumerative Algebraic Geometry. Proceedings of the 1989 Zeuthen Symposium*, ed. S. L. Kleiman and A. Thorup, vol. 123, pp. 1–13. Contemporary Mathematics. Providence: American Mathematical Society.

Klein, F. 1894. *The Evanston Colloquium. Lectures on Mathematics, delivered from Aug. 28 to Sep. 9, 1893*. New York and London: MacMillan and Co.

Klein, F. 1928. *Vorlesungen über Nicht-Euklidische Geometrie*. Berlin: Springer.

Kragh, H. 2015. From Ørsted to Bohr: The Sciences and the Danish University System, 1800–1920. In *Sciences in the Universities of Europe, Nineteenth and Twentieth Centuries. Academic Landscapes*, ed. A. Simões, M. P. Diogo, and K. Gavroglu, vol. 309, 31–47. Boston Studies in the Philosophy and History of Science. Dordrecht: Springer.

Kragh, H., P. C. Kjærgaard, H. Nielsen, and K. H. Nielsen. 2008. *Science in Denmark. A Thousand-Year History*. Aarhus: Aarhus University Press.

Lê, F. 2023. Des taxons et des nombres: Quelques remarques sur les ordres, classes, et genres des courbes algébriques. *Revue d'histoire des sciences* 76: 85–134.

Leibniz, G. W. 1684. Meditationes de Cognitione, Veritates, & Ideis. *Acta Eruditorum*: 537–542.

Lorenat, J. 2015a. Figures Real, Imagined, and Missing in Poncelet, Plücker, and Gergonne. *Historia Mathematica* 42: 155–192.

Lorenat, J. 2015b. Polemics in Public: Poncelet, Gergonne, Plücker, and the Duality Controversy. *Science in Context* 28: 545–585.

Lützen, J., and W. Purkert. 1989. Conflicting Tendencies in the Historiography of Mathematics: M. Cantor and H.G. Zeuthen. In *The History of Modern Mathematics, vol. III: Images, Ideas, and Communities*, ed. E. Knobloch and D. E. Rowe, 1–42. Boston: Academic Press.

Martindale, C. 1994. *Redeeming the Text. Latin Poetry and the Hermeneutics of Reception*. Cambridge: Cambridge University Press.

Mehrtens, H. 1990. *Moderne Sprache, Mathematik: Eine Geschichte des Streits um die Grundlagen der Disziplin und des Subjekts formaler Systeme*. Frankfurt am Main: Suhrkamp.

Michel, N. 2020. *Of Words and Numbers. The Writing of Generality in the Emergence of Enumerative Geometry (1852–1893)*. Ph. D. thesis, Université de Paris.

Michel, N. 2021. Mathematical Selves and the Shaping of Mathematical Modernism: Conflicting Epistemic Ideals in the Emergence of Enumerative Geometry (1864–1893). *Isis* 112: 68–92.

Pind, J. L. 2014. *Edgar Rubin and Psychology in Denmark: Figure and Ground*. Berlin: Springer.

Plücker, J. 1834. Solution d'une question fondamentale concernant la théorie générale des courbes. *Journal für die reine und angewandte Mathematik* 12: 105–108.

Plücker, J. 1839. *Theorie der algebraischen Curven, gegründet auf eine neue Behandlungsweise der analytischen Geometrie*. Bonn: Adolph Marcus.

Poincaré, H. 1952. *Science and Method; Translated by Francis Maitland; Preface by Bertrand Russell*. New York: Dover Publications.

Salmon, G. 1852. *A Treatise on the Higher Plane Curves: Intended as a Sequel to A Treatise on Conic Sections*. Dublin: Hodges and Smith.

Schappacher, N. 2010. Rewriting Points. In *Proceedings of the International Congress of Mathematicians 2010*, ed. R. Bhatia, vol. 4, 3258–3291. New Delhi: Hindustan Book Agency.

Semple, J. G. 1982. Complete Conics of S_2 and their Model Variety $\Omega_5^{102}[27]$. *Philosophical Transactions of the Royal Society of London A* 306: 399–442.

Semple, J. G., and G. T. Kneebone. 1952. *Algebraic Projective Geometry*. Oxford: Clarendon Press.

Sigurdsson, S. 1992. Equivalence, Pragmatic Platonism, and Discovery of the Calculus. In *The Invention of Physical Science. Intersections of Mathematics, Theology and Natural Philosophy Since the Seventeenth Century*, ed. M. J. Nye, J. L. Richards, and R. H. Stuewer, 97–116. Dordrecht: Springer.

Turner, L. E., and H. K. Sørensen. 2013. Cultivating the Herb Garden of Scandinavian Mathematics: The Congresses of Scandinavian Mathematicians, 1909–1925. *Centaurus* 55: 385–411.

von Wartensleben, G. 1914. *Die christliche Persönlichkeit im Idealbild: eine Beschreibung sub specie psychologica*. Kemptem and München: Joseph Kösel.

Zeuthen, H. 1890. Sur la révision de la théorie des caractéristiques de M. Study. *Mathematische Annalen* 37: 461–464.

Zeuthen, H. 1903. Ved forelæggelsen af "Mathematikkens Historie i 16. og 17. Aarhundrede". *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger*: 553–572.

Zeuthen, H. 1910. Introduction à un traité didactique des méthodes énumératives de la géométrie. In *Compte rendu du congrès des mathématiciens tenu à Stockholm, 22–25 septembre 1909*, ed. G. Mittag-Leffler and I. Fredholm, 32–42. Leipzig and Berlin: Teubner.

Zeuthen, H. 1914a. *Lehrbuch der abzählenden Methoden in Geometrie*. Leipzig: Teubner.

Zeuthen, H. 1914b. Om Anvendelse af Regning og af Ræsonnement i Mathematiken. *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandliger*: 271–286.

Zeuthen, H. 1917. *Hvorledes Mathematiken i Tiden Era Platon Til Euklid Blev Rationel Videnskab*. København: Andr. Fred. Høst & Søn.

Zeuthen, H. 1921. Abzählenden Methoden. In *Encyklopädie der mathematischen Wissenschaften*, ed. F. Meyer and H. Mohrmann, vol. 3-1-2, 257–312. Leipzig: Teubner.

Chapter 22
Variations on Enriques' 'Scientific Philosophy'

Umberto Bottazzini

Abstract In the 1930s Federigo Enriques presented the essential elements of his 'historical epistemology'. This was the result of a long philosophical journey that began in the last decade of the nineteenth century, and from his early interest in the philosophical problem of space and in the philosophy of geometry grew up to his critical approach to the nature of scientific knowledge as discussed in his *Problemi della scienza* (Enriques, 1906). Eventually, through his criticism of both kantianism and logical empiricism he elaborated his new epistemology founded on historical criticism of scientific theories.

22.1 Introduction

Paris, September 15, 1935. Congrès international de philosophie scientifique.

> J'appartiens – moi – à la génération de ceux qui, élevéés dans le milieu de la philosophie positive, ont vu, dans leur jeunesse même, se relever l'étendard de l'idéalisme métaphysique et engager une lutte violente contre l'esprit positif. Après trente années dominées par ces courants de la pensée, j'assiste aujourd'hui au renouveau de la philosophie scientifique, qui – à la vérité – n'a jamais cessé d'exister et d'être affirmée, pendant cette période, par des penseurs sortis du domaine des sciences particulières, mais qui, – depuis quelque temps – semble reprendre force, tendant à une domination nouvelle sur la culture. C'est là un événement que je salue de tout mon coeur. (Enriques, 1936b, p. 23)

There was an evident autobiographical flavor in the words with which Federigo Enriques, in the opening talk of the first section of the Congrès, devoted to *Rationalisme empirique et empirisme logique*, hailed "the revival of scientific philosophy". Not only did he have to fight 'metaphysical idealism' in his youth, but he was one of the "thinkers from the field of particular sciences" (namely, mathematics) who in the last decades has steadily risen as a defender of scientific

U. Bottazzini (✉)
Mathematics, Università degli Studi di Milano, Milan, Italy
e-mail: umberto.bottazzini@unimi.it

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_22

philosophy. What was meant by "scientific philosophy"? How could it be defined? In Enriques' words,

> La philosophie scientifique, en tant qu'elle aspire à établir une discipline supérieure de la pensée rationnelle, ne saurait se réduire à un système philosophique particulier, résolvant en un sens déterminé les oppositions traditionnelles des écoles. (Enriques, 1936b, p. 24)

But he was not the only one to offer a definition of it. For instance, in the opening address of the Congress Louis Rougier stated:

> Nous croyons que la philosophie peut devenir scientifique, en prenant pour objet la science elle-même, et pour méthode l'analyse logiques de ses notions, de ses propositions, de ses théories, de ses démonstrations. Ainsi comprise, la philosophie constitue ce que on a proposé d'appeler *la syntaxe et la sémantique du language scientifique*. (Rougier, 1936, p. 8)

This conception, developed mainly by the followers of the Vienna circle, reduces the philosophers to play the role of "grammairiens de la science", Rougier continued. By establishing the tautological character of thought thanks to logic, they were able to "achever la désagrégation de l'apriorisme au profit de l'*empirisme logique*", he claimed before ending his address with the invitation to pay homage on one side to the ancestors of the Vienna school Frege, Peano, Hilbert, and above all to Russell, "notre maître à tous", as well as on the other side to Poincaré, Mach and Duhem.

Following Rougier's talk Russell, who attended the Congress, took the floor to say he was "glad that Frege and Peano received due honors". Then he mentioned Leibniz, Wittgenstein and both the Vienna and the Polish schools, and concluded his address stating that "modern science arose from the marriage of mathematics and empiricism; three centuries later, the same union is giving birth to a second child, scientific philosophy" prophesying that the latter "is perhaps destined to as great a career" (Russell, 1936, p. 11).

In turn, Russell's speech was followed by Enriques's. In his short welcome address the Italian mathematician/philosopher argued that philosophical freedom must be preserved within scientific thought adding that the danger threatening the revival of scientific thought was logicism "d'où pourrait bien sortir une nouvelle scolastique" (Enriques, 1936a, p. 12).

Enriques' talk *Philosophie scientifique* opened the first section of the Congress devoted to "Rationalisme empirique et empirisme logique". After his confession of juvenile philosophical faith quoted above, Enriques accused positive philosophy of having lost "le sense de l'unité du savoir", and warned against the risks of dogmatism of the positivist philosophers who cite scientific manuals "avec l'assurance des théologiens qui se réclament de la Bible" (Enriques, 1936b, p. 25) while the *savant chercheur et critique* (maybe, Enriques himself?) could not find in their considerations the atmosphere of doubt needed for the life of reason.

According to Enriques, his criticism of Kantian apriorism could also be addressed to the renewed currents of thought that the program of the Congress "semble confondre un peu avec la 'philosophie scientifique' tout court" (Enriques, 1936b, p. 26), namely the empirical logicism. In spite of his great interest in the ideas and critiques of the followers of this philosophical school he stated openly he would be less willing to admit that "leur système constitue la seule philosophie vraiment scientifique". Enriques continued:

> Je me défie davantage du logicisme. La raison qui construit la science, et qui se révèle par l'évolution historique de la pensée, ne saurait s'expliquer par une analyse purement logique. (Enriques, 1936b, p. 26)

In other words, "un système de signes vide et tautologique" could not "satisfaire notre raison scientifique" Enriques added. In his opinion, logic itself posed a problem. Actually, what is meant by logic? Enriques asked. If it is viewed as an analysis of the operations of the mind there is the risk of falling into psychologism. If, on the contrary, we consider logic as aiming at relationships that are somehow outside our mind we are getting very close to the medieval position oscillating between metaphysical realism and nominalism. Summing up, he concluded, "des deux côtés je vois surgir devant nous le spectre d'une nouvelle scolastique". (Enriques, 1936b, p. 27)

The same view was shared by Lautman, who attended the Congress. In his talk *Mathématiques et réalité* he stated:

> Les logiciens de l'Ecole de Vienne prétendent que l'étude formelle du langage scientifique doit être le seul objet de la philosophie des sciences. C'est là une thèse difficile à admettre pour ceux des philosophes qui considèrent comme leur tâche essentielle d'établir une théorie cohérente des rapports de la logique et du réel. (Lautman, 1936, p. 24)

There is a physical reality—Lautman continued—and 'a miracle' should be explained, namely the fact that the most developed mathematical theories are needed to interpret such a reality. (To the modern reader Lautman's words recall Wigner's claim about "the miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics" as he stated in his celebrated paper (Wigner, 1960)). A philosophy that would not be entirely concerned with this solidarity between domains of reality and methods of investigation would be singularly devoid of interest, was Lautman's conclusion.

As for Enriques, this kind of conclusion was not new. In recent years he repeated it in his reviews of texts by logical positivists published in *Scientia*, the journal he founded in 1907. Reviewing Carnap's essay *L'ancienne et la nouvelle logique* (1933) Enriques warned that

> Nous devons faire les plus grandes réserves sur la valeur que l'éminent auteur attribue à la tendance de la critique logique. Car une philosophie qui tirerait sa norme exclusivement de cette logique risquerait de revenir au point où se trouvaient les écoles médiévales, dont l'idéal des nouveaux logiciens se rapproche si singulièrement. (Enriques, 1935, p. 69)

He also disagreed with Carnap's claim that the precise logical method was going to give the *coup de grâce* to all forms of metaphysics. Instead, Enriques objected that precisely starting from his own research in logic Russell arrived at "une nouvelle métaphysique *réaliste*", something that according to Carnap's critique of logic should turn out to be meaningless. And finally he challenged Russell's position, adopted by Carnap, which reduced mathematics to be a branch of logic as “un stérile jeu de combinaisons tautologiques”. (Enriques, 1935, pp. 69–70)

In the same issue of *Scientia* Enriques published his review of Ph. Frank’s essay *Théorie de la connaissance et physique moderne* (1934). As “ce nouveau positivisme a beaucoup points de contact avec notre pensée” Enriques admitted (Enriques, 1935, p. 227), it was interesting to highlight the differences. To this aim it was necessary to go back to Mach, whose disciple “dans un sens large” he was pleased to proclaim himself. However, according to Enriques, it was too limiting for the scientist’s activity to reduce science to the economic description of facts by excluding “la répresentation imaginative d’une réalité plus vaste, construite pour satisfaire les exigences de la raison”. His beliefs—Enriques added—matured as a reaction against certain tendencies in which the conception of truth dissolves either into forms of pragmatism as W. James’s that will eventually lead to a new “idéalisme métaphysique”, or into “positions scolastiques” as Duhem’s (Enriques, 1935, p. 228).

Enriques repeated essentially the same criticism in his review of Carnap’s essay *La Science et la Métaphysique devant l’analyse logique du language* (1934). Although in his essay Carnap offered “un spécimen savoureux de la philosophie de Heidegger (‘le Néant néante’)” (Enriques, 1936c, pp. 109–110) Enriques claimed to have no illusions that it will contribute to “refroidir l’enthousiasme de tant de fanatiques de ce genre de spéculations”. A more serious objection to Carnap’s logicism was that “tendances naturelles à l’ontologie” could be expressed through mathematical logic as Russell’s renewed realism showed. In the end, he invited to look with a certain mistrust at this philosophical conception that “nous ramène assez près de la mentalité médievale”. In Enriques’ opinion, one of the clearest and most in-depth expositions of the ideas of the Vienna circle (Mach-Verein) was provided by H. Hahn’s essay *Logique, mathématique et connaissance de la réalité* (1935). But, against Hahn’s conception, Enriques stated once more: “nous devons protester contre le jugement d’après lequel la pensée logique ne serait que tautologique”. (Enriques, 1936c, p. 176)

As Lolli (Lolli, 2018, p. 121) has remarked, this notwithstanding, Enriques was appointed to the International Committee of the Congress for the Unity of Science, planned to take place in Paris, along with Carnap, Frank, Neurath, Reichenbach, Schlick, Morris, Niels Bohr, Rougier, and Russell among others. He had already been listed among the ancestors of the neo-positivist philosophers in the manifesto of the Vienna Circle, written by Hahn, Neurath, and Carnap (Hahn-Neurath-Carnap, 1929). In addition to Enriques, under the heading “Foundations, goals and methods of empirical science (hypotheses in physics, geometry, and so on)” their list included Helmholtz, Riemann, Mach, Poincaré, Duhem, Boltzmann and Einstein. In fact, the reflections on the nature of space and geometry had been at the basis of Enriques’ interest in philosophy.

22.2 Early Philosophical Hints

"A train of thought which gradually came to maturity during the ten years from 1890 to 1900 has resulted in a critical study of certain problems relating to the logical and psychological development of scientific knowledge" (Enriques, 1914, p. XIV). This was the philosophical path followed by Enriques as he confessed in the opening lines of his *Problemi della scienza* (Enriques, 1906), the book that marked his entry into the philosophical scene.

At the beginning of the twentieth century, when Enriques entered the field of philosophy, he was one of the leading mathematicians of his days, a great geometer internationally recognized who achieved fundamental results in algebraic geometry, including the classification of algebraic surfaces obtained through joint work with Guido Castelnuovo. This notwithstanding, as his correspondence with the latter shows, Enriques has been involved in philosophical issues ever since he began teaching projective geometry at the university of Bologna. Or even many years before, according to the testimony of Gaetano Scorza-Dragoni, who once told of a conversation with Enriques during a walk in Rome: "We had both found ourselves led to the study of sciences, let's say exact, by a high school philosophical infection and by the conviction that only in natural philosophy could we find an answer (admittedly, partial) to the problems that had fascinated us during the years of high school (*liceo*)" (In: (Enriques, 1958, p. 7))).

Apparently, Enriques was provided by his students with the opportunity to resume his philosophical interests openly. On November 23, 1894 he wrote to Castelnuovo:

> Some young pupils ask me to give a course in Higher Geometry. I am not alien to the idea of partially satisfying them with a series of weekly conferences. [...] In that case, I will tell you the plan: they would be inspired by a general principle that completes that of Klein (Programm) in order to include various other types of research. (Bottazzini-Conte-Gario, 1996, p. 151)

A few weeks later he wrote that he had begun preparing the conferences, and explained: "I wanted to start with philosophical-mathematical reflections, and with the development of the concept of abstract geometry" (Bottazzini-Conte-Gario, 1996, p. 164). Eventually, his conferences were collected in the lithographed volume *Conferenze di geometria* (Enriques, 1894–95). Under the heading "Foundations of a geometry of hyperspace" there Enriques began by stating that "Geometry has as its object the study of the relations inherent in the concept of space as it arises in our mind from the order of external sensitivity, that is, as it is presented to us by *intuition*" (Enriques, 1894–95, p. 2).

As Peano argued, even according to Enriques "there is something arbitrary in the choice of the fundamental entities of space", among whom there are relations "given *a priori* as postulates". These "are deduced from intuition and their complex takes the place of definition for the fundamental entities", that is, it implicitly defines the fundamental entities "by establishing those mutual relations among them, which serve to fix them as much as necessary for geometric developments" (Enriques, 1894–95, p. 3). Therefore, in Enriques' eyes geometry appeared "in its principle and in its development" to be a "subjective" science, since the postulates reflect the concept of "intuitive space" present in our mind, the definitions and demonstrations being "only logical operations". Whether "real" space corresponds to this "intuitive space" is a philosophical question "closely related to the problem of knowledge", which Enriques only hinted at there, but which within a few years will become dominant in his interests.

However, Enriques went on, the development of geometry is independent of its relationship with external reality. The distinction between "physical" and "subjective" geometry allows us to found various "more general" geometries, in which some postulate is disregarded through an analysis of the postulates conducted following (a) the "physical" criterion; (b) the "physical-psychological" criterion; and finally (c) the "logical" criterion. Enriques favored the last two criteria that are based on intuition and logic, since relying on logic alone "as some believe" would end up reducing "mathematics to a mere syllogistic exercise"(Enriques, 1894–95, p. 6).

His aim was "to reconcile the needs of the logical spirit with the advantages and attractions that intuition confers on geometric studies", as he would write in the introduction to *Lezioni di geometria proiettiva* (Enriques, 1898, p. V). (At Klein's suggestion Enriques' *Lezioni* were translated into German in 1903.)

The fact that the postulates of geometry are "derived from intuition" allowed Enriques to hold *a priori* that they are compatible with each other on the basis of the "principle of reason", according to which "several truths conceived together as elements of the same concept are compatible" (Enriques, 1894–95, pp. 7–8). As for their independence, Enriques recognized its meaning both in the order in which the postulates are stated (as Peano claimed), and in their composition (since a postulate can "split into others, some of which can be deduced from the preceding ones").

Enriques' interest was aimed at "abstract" geometry that "can be interpreted in infinite ways as a concrete (intuitive) geometry by fixing the nature of its elements" (Enriques, 1894–95, p. 9–10). Thus, for example, abstract plane geometry can be "interpreted indifferently as intuitive geometry on the plane or as that on developable surfaces", and abstract projective geometry of space can be interpreted as a geometry of linear systems ∞^3 of algebraic plane curves of given order, or as a geometry of the involutions of order $n(> 2)$ and of 3rd kind on the straight line (Enriques, 1894–95, p. 15). Similarly, the "abstract geometry of hyperspaces will thus receive an infinity of various interpretations", some of whom Enriques exemplified in his *Conferenze*.

In this connection, commenting on the foundation of the geometry of hyperspaces he stated in (Enriques, 1894) that his aim had been "to establish the postulates derived from the experimental intuition of space that appear simpler to define the object of projective geometry" (Enriques, 1894, p. 142) adding in a footnote: "As for those intuitive concepts, we do not intend to introduce anything more than their logical relations so that geometry thus founded can still receive an infinite number of interpretations [...]. It only seems to us that the experimental origin of geometry should not be forgotten in the search for the hypotheses on which it is based".

On the other hand, in the *Appendix* to his *Lezioni di geometria proiettiva* he made it clear once more: "*Projective geometry can be considered as an abstract science, and it can therefore be given interpretations different from the intuitive one, by stating that its elements (points, lines, planes) are concepts determined in whatever way satisfy the logical relations expressed by the postulates* (italics in original).

His latent interest in philosophical questions had clearly emerged a couple of years before. On May 4, 1896 Enriques confessed to Castelnuovo that

> while mathematical questions doze until the best time, for several days I have been dealing with another question that takes only the pretext from mathematics: hearing its name you will be more horrified than amazed. It is the "philosophical problem of space" (Bottazzini-Conte-Gario, 1996, p. 260)

The whole letter is worth reading for it sheds light on Enriques' early steps in the philosophical field.

> Books on psychology and logic, physiology and comparative psychology, criticism of knowledge etc. pass on my desk where I savor them with pleasure trying to extract the juice as far as my problem is concerned. At another time I will show you the method and the plan of my work, for which I certainly do not expect your full adherence. However I hope you will be convinced that the issue I am dealing with is not metaphysical but positive-critical, and some aspects may have at least a benevolent neutrality on your part.
>
> For in my program there is the question of the genesis of the concepts of space from the data of physiological psychology (especially of the eye and of touch) of Helmholtz, Wundt etc. And I would like to draw proof of the concept already stated in my *Conferenze* that the two main lines of development (metric-analytic and projective-metric) of the treatment of postulates have their foundation in the differentiation of senses.
>
> But this is not the only aspect in which I consider the problem. There is first the critical side with respect to the theory of knowledge, and the question of the genesis of the concept of space in evolution: then, the question of innate ideas and reattachment to the controversial biological problem of inheritance.
>
> As you see there is fun. For my part, I bring to research an enthusiasm, which you will deem worthy of a better cause, but which is certainly greater than I have ever felt for any other matter. (Bottazzini-Conte-Gario, 1996, pp. 260–261)

Enriques concluded his letter with a enthusiastic appreciation of Wilhelm Wundt, whom he did not hesitate to hail as the "the most marvellous philosophical, physiological, psychological, mathematical, etc. intelligence". At the same time he invited Castelnuovo "to read the 'Logik' of Wundt, at least that part about the methods of mathematics".

A few years later the ideas sketched in this letter will be taken up by Enriques in an elaborate way in the opening article of his own (Enriques (1900a)). "The primitive data of geometry and physics are fundamentally acquired in the same way on the basis of certain *immediate sensations* or certain very simple *elementary experiences*, interpreted in accordance with the *logical structure* of our mind" (italics in original) (Enriques, 1900a, pp. 4–5). In the end, Enriques stated, the common foundation of both physics and geometry lies in their empirical basis that gives geometry the character of an experimental science. The strictly Kantian point of view that the postulates express *a priori* conditions of subjective sensitivity, almost structural laws of the psyche, seems to be outdated, Enriques commented.

Starting to discuss the fundamental concepts and propositions of geometry, he stated that "it is desirable" that primitive concepts and postulates be "*absolutely independent*". This is an "essential condition" for geometry to be treated in a "*rigorously*" logical way as an "*abstract logical theory*" (italics in original) worthy of receiving various interpretations, as he had shown in his *Conferenze di geometria* (Enriques, 1894–95). At the same time he warned against an excess of formalism: going beyond the limits of reality to follow only the laws of symbols there is a risk of falling into the void, "the thought fades and disappears into nothingness, like a vague and incoherent fog" (Enriques, 1900a, p. 12). Instead, when it comes to geometry, *intuitive evidence must shine brightly in principles* (italics in original), Enriques emphasized.

It is interesting to remark that when he wrote these pages on the principles of geometry, and the independence of the postulates, he had already got to know (and studied) Hilbert's *Grundlagen der Geometrie* just appeared a few months before. In fact, writing to Castelnuovo on October 2, 1899 he raised an objection against the way in which Hilbert "claims to establish the independence of his postulates of congruence from the previous ones" (Bottazzini-Conte-Gario, 1996, p. 428). Enriques resumed his criticism in his review (Enriques, 1900b) of Hilbert's booklet: "The independence of group IV [axioms of congruence] does not seem us to be established in a satisfactory way", he claimed before explaining why.

Indeed, Enriques did not share Hilbert's formalistic approach. However, in the concluding remarks he stated that "by studying abstract questions of a purely logical nature with abstract methods" Hilbert showed to "understand all the value of geometrical intuition" (Enriques, 1900b, p. 7). Instead, Enriques was openly inspired by the physiological, psychological approach of Helmholtz and Wundt as exemplified by Klein for geometry, for instance in (Klein, 1890). This was also the approach to the foundations of geometry that Enriques followed in the article on the principles of geometry that he wrote for the *Encyclopädie der mathematischen Wissenschaften*. (In 1897 he was asked by Burkhardt to write it, but the article appeared in print only ten years later, in 1907.)

Enriques discussed the matter in detail in *Sulla spiegazione psicologica dei postulati della geometria*, the first article of properly philosophical nature that he published in 1901 in a philosophical journal (Enriques, 1901). While waiting to expose his research in a larger work, in this article he presented the results of his study of the geometric postulates as seen in the genetic aspect of the psyche.

In his terms, the general problem he faced was that of deducing the spatial concepts that fall under the exact intuition of the mathematician from sensitive representations whose genesis is clarified by physiology. More precisely, his aim was "to explain the postulates of geometry that is thus subjectively constructed by re-attaching their necessity to the logical structure of thought". (Enriques, 1901, p. 146)

Referring to the works of Helmholtz and Wundt, Enriques underlined the importance of "a previous analysis of spatial concepts that only the mathematician is able of carrying out", adding that "the mathematician himself would be lost" trying to figure out the constitutive elements in the aforementioned concepts "if they did not appear already separated in the evolution of geometric science, in the various branches of projective, metric and differential geometry, whose common foundation is the general theory of the continuum or Analysis situs". (Enriques, 1901, pp. 148–149)

The latter is a conception that Enriques made his own from Klein (1890). "In what sense—Klein asked—is it psychologically correct to treat projective geometry before metric geometry, and even to consider the first as the foundation of the second?" (Klein, 1890, p. 570). The former has to do with the "optical" properties of space, the latter with the "mechanical" ones and, admittedly, there is an obvious distinction between them. The question Klein asked was: in the "methodical building" of the science of space which one has to be considered first? Helmholtz argued that we must begin with the mechanical properties that "find their mathematical expression in the free mobility of bodies". But "my works—Klein objected—show that we can start just as well with the optical properties".

Turning to the nature of the geometrical axioms, Klein claimed to consider the spatial intuition (*Anschauung*) as "*something essentially imprecise*" (italics in original). Thus, he went on, stating the axioms of geometry was a way of putting precise statements in the imprecise intuition. A few years later, Klein discussed the origin and nature of the axioms of geometry in the preliminary remarks to his report for the Lobačevskij Prize awarded to Sophus Lie. "Where do the axioms come from?", Klein asked . "A mathematician who knows the non-Euclidean theories will hardly want to hold on to the opinion of earlier times that the axioms, according to their concrete content, were necessities of inner intuition" (Klein, 1898, p. 584). Thus, "do the axioms come from experience?" was the further question he asked. No doubt that experience plays a great role, Klein admitted. However, while "*the results of observation are only valid within certain limits of accuracy and under particular conditions, by setting up the axioms we are making statements of absolute precision and generality*" (italics in original) (Klein, 1898, p. 585). Summing up, according to Klein "the real essence of the axioms lies in the idealization of the empirical data".

From the philosophical point of view, in the background there is the problem of the correspondence between logical and/or mathematical theories and reality, which Enriques discussed in his letters to Giovanni Vailati, a philosopher and logician trained at Peano's school. In a letter on April 16, 1901 after mentioning in general terms their respective criticism of Kantian philosophy, Enriques claimed not to care much for philosophers who are not scientifically educated, adding the programmatic statement: "I think philosophy should be done by scientific minds, and *in the service of science* (italics in original)" (in: Vailati, 1971, pp. 564–565).

And one month later (May 16, 1901), as a reply to Vailati's comments on his paper (Enriques, 1901) Enriques refused to continue the discussion on Kantianism, and turned instead to his own studies on logic. In his opinion there are two aspects to logic: (1) "The subjective aspect in which logic appears as the study of certain operations of thought" and (2) an objective aspect, in which "we can ask if anything real corresponds to the principles and logical axioms".

Then, according to Enriques, a third research was needed.

> "Is subjective logic experimentally acquired (as a set of data) - he asked - or has it to be considered as a reflection of the structure of thought? As you know, I'm for the second thesis.
>
> How then to explain the correspondence between structural logical laws and external reality? Here the research turns to physiology, and asks for a hypothesis on the functioning of the brain that satisfies the required explanation. You will not imagine that I am so bold as to ask what the physiological conditions or the physiological aspect of thinking are. But it can be admitted that the phenomenon of thought responds to a physiological phenomenon localized in a certain group of cells and nerve fibers of association" (in: Vailati, 1971, pp. 568–569).

Vailati commented to Giuseppe Vacca, another member of Peano's school, to have read Enriques' paper, and to have not hidden his own appreciation of the kind of 'philosophy' on which the paper was based but also to have criticized Enriques for his "taking psycho-physiology too seriously". Now—Vailati continued—Enriques "is dealing with logic, but I still haven't quite understood with what kind of 'logic': certainly not mathematical logic, as far as one can see (in: Vailati, 1971, pp. 188–189).

Indeed, Enriques was talking about how logical reasoning relates to reality, a subject that appears repeatedly not only in his *Problemi della scienza*, but also later on in the volume *Per la storia della logica* (1922) in which Enriques retraced, explained and exemplified this conception of logic, far from the contemporary developments of mathematical logic.

22.3 *Problemi della scienza*

The plan of this book—Enriques confessed in the Preface—has been basically settled since the year 1901, when he published the paper on the psychological explanation of the geometric postulates. Then, he explained that

> my faith in this *philosophy of science* has led me from the fields of geometry, where thought rests quietly in the security of acquired facts, to discuss the building up of a science of knowledge which may become the common possession of the studious and may tend to unify the various domains of knowledge in one synthetic view of the cognitive methods. (italics in original) (Enriques, 1914, p. XV)

Problemi della scienza begins with a long introduction in which Enriques' criticism is addressed to Kantianism and positivism. Then, in the search for "*a positive definition of reality*" he observed that "*our belief in the reality of a thing rests upon a totality of sensations which invariably follow under certain conditions arranged at will*" (italics in original) (Enriques, 1914, pp. 55–56). In the last analysis, Enriques commented, "the belief in reality rests upon *an associative relation* between our sensations" (italics in original). In his opinion, the true character of reality is constituted by "the correspondence between sensation and expectation" that "gives us the positive definition of reality" (Enriques, 1914, p. 56). At first sight, Enriques continued, this way of presenting things is shocking because it does not seem to fit with the ordinary view according to which reality "would not cease *to exist in itself*, even if all communication between our minds and the external world were broken off" (italics in original). The latter was *not* Enriques' view.

In a letter to Vailati on April 23, 1901 he had written that "the claimed 'things in themselves', the 'absolute real or independent of us', etc. are meaningless sentences, so that saying that the aforesaid things are unknowable is also nonsense" (in: (Vailati, 1971, p. 566)). In *Problemi della scienza* he repeated that the expression existence *in itself* "is devoid of sense, unless indeed it signifies the *impotence of the will to modify the sensations that refer to reality*, without changing the conditions with which these sensations are bound up" (italics in original) (Enriques, 1914, p. 56). Rejecting as void of sense what he called the transcendental idealism that attributed to reality an absolute significance in and for itself—Enriques went on—would lead nearer to Mach's phenomenalism or Vailati's interpretation of idealism.

It took Enriques some ten pages to clarify his position on realism, and to reach the conclusion (the postulate of knowledge) that "the real gets defined [...] as *an invariant in the correspondence between volition and sensation*" (italics in original) (Enriques, 1914, p. 65). The term 'invariant' was not chosen at random. Referring to its group-theoretical meaning Enriques explained that his view that the real was an invariant required to "carefully limiting" the *group of elements* (namely volitions and sensations) and the *group of transformations* (italics in original) to which the relevant invariant referred. In Enriques' view the postulate of knowledge thus formulated applied equally to common sense knowledge and to scientific knowledge as well.

As for the process of the acquisition of knowledge, he held that the latter was subject to a continual process of revision. In his opinion

> The progress of science is a process of *successive approximations*, in which new and *more precise, more probable* and *more extended inductions* result from partially verified *deductions*, and from those contradictions that correct the implicit hypotheses. In this process certain primary and general concepts, such as those of geometry and mechanics, give us some guiding principles that are but slightly variable if not *absolutely* fixed. Therefore we should turn our attention to these concepts in order to explain their actual value and their psychological origin. (italics in original) (Enriques, 1914, p. 166)

This was exemplified in the chapters devoted to Geometry and Mechanics, for in his view "we must give geometry a place of honor among philosophical studies!". All the more because "especially in the past century the progress of geometry has had a direct effect upon the development of a form of rationalism" (Enriques, 1914, p. 173).

Against Kant's thesis ("which denies the existence of a real object corresponding to the word 'space' ") Enriques (with Herbart) held the reality of "spatial relationships". To Poincaré's "new nominalism" (those relationships have no real meaning, absolutely independent of bodies) Enriques opposed a "more precise evaluation of Geometry as part of physics".

Moving on to discuss the psychological acquisition of geometric concepts ("the psycho-genetic development of the postulates of geometry"), he resumed in detail what he held in his paper (Enriques, 1901) on the relationships between the spacial data of sight and projective geometry, the spatial data of tactile and muscular sensations and metric geometry, and eventually the postulates of continuity.

Then Enriques turned to mechanics treated “as an extension of geometry”, and finally to physics treated “as an extension of mechanics”.

Problemi della scienza got contrasted receptions. In Italy it received fair reviews by Vailati, who pointed his criticism to Enriques’ approach to logic *et pour cause* (Vailati, 1906), and by Enriques’ colleagues and friends Severi and Levi-Civita (who limited himself to reviewing the chapter on mechanics).

Abroad, in an enthusiastic review appeared in *Mind* (**17** (1908), p. 132) Enriques’ book was hailed as “probably the most comprehensive study that has appeared in recent years on the concepts on which modern science is built”. According to the reviewer, its author shows “profound knowledge of the history and modern developments of science” as well as “critical grasp of the trends of philosophical thought and the underlying psychological questions rarely found united in one mind”.

On the contrary, not quite as enthusiastic was Broad’s review of *Problems of science*, the English translation of Enriques’ 1906 book, which also appeared in *Mind* (**24** (1) (1915), pp. 94–98). Broad began by remarking that the book “covers very much the same ground as Poincaré’s three books on the philosophy of science”. (For a comparison of Enriques’ and Poincaré’s works see Gray (2006)). Admittedly, Broad continued, Enriques’ book “gives an impression of very deep and wide learning [. . .]. Unhappily the style is very heavy, and one can never forget for a moment that one is reading a translation from a foreign tongue”. In addition, “the argument is obscure trough its condensation even to persons familiar with the problems under discussion; to others it must often be quite unintelligible”.

However, the fiercest criticisms came from the Italian idealist philosophers Croce and Gentile who viewed science as a practical activity without any philosophical value. In his long review of *Problemi della scienza* appeared in “La Critica”, (6 (1908), 430–446) the latter stated that Enriques, as “all the dreamers of a scientific philosophy [. . .] doing scientific philosophy never clashes with philosophy”. Gentile’s criticism was also addressed against “Scientia. Rivista di scienza”, the new journal for “a scientific synthesis” founded by Enriques and a group of scientists in 1907. According to Gentile, “it can only encourage scientific amateurism”. (It’s worth mentioning that in a short period of time, in addition to papers by Italian authors like Vailati, Castelnuovo, Volterra and Peano, “Scientia” published papers by Russell, Rutherford, Freud, Poincaré, Borel, Boutroux, Picard, Einstein etc.)

At the International Congress of Philosophy at Heidelberg (1908), attended by Enriques (and by Croce, who gave a talk on ‘The lyrical character of art and pure intuition’), the former—in his capacity of President of the Italian Philosophical Society—was charged to organize the next Congress to be held at Bologna in 1911. At Bologna’s Congress Enriques gave the opening talk on ‘The problem of reality’. With an image lent by the practice of honor duels still widespread at the time, Enriques recorded that during the Congress, attended by Croce and Peano, “he had the honor of crossing the foil of the word” about the nature of logic with Alessandro Padoa, a distinguished pupil of Peano. The Congress offered the occasion for a public controversy between Croce and Enriques. After the end of the Congress, on his way to Naples in an interview to a newspaper Croce called Enriques as "a willing professor, who with zeal but little preparation delights in philosophy". Furthermore Croce added ironically that Enriques "bears the burden of congresses of philosophers, which is as meritorious as mine would be meritorious and disinterested, if I were to organize congresses of mathematicians" (Croce, 1919). This was the beginning of a bitter controversy that has been told by historians many times (see e.g. Simili's preface to (Enriques, 2000)). After a series of interviews and polemical papers, eventually it ended with the victory of Croce's and Gentile's idealism.

Years later, in 1919 when he was Minister of Public Education, Croce reminded all that as "a small argument" with a mathematician who "had been taken by zeal for the abstractly rationalistic philosophy that arises easily in the brains of mathematicians, and seeks and finds fortune in democratic and Masonic circles. With their help, [Enriques] was able to put together the International Congress of Philosophy at Bologna in 1911" (Croce, 1919).

In this allusive way apparently Croce referred to the French "Revue de métaphysique et de morale", and to philosophers like Léon Brunschvicg, Émile Meyerson, or Xavier Léon who held positions close to Enriques's. As a matter of fact Enriques was highly appreciated in French philosophical circles. In the 1930s he was called by the publisher Hermann to take the scientific direction of the book series 'Philosophie et histoire de la pensée scientifique', and in 1936 after Meyerson's death he was elected as 'Correspondant de l'Académie des sciences morales et politiques' (Section de Philosophie) de l'Institut de France.

There is hardly any doubt that Enriques was referring to this thwarted period of his life when, in his 1935 talk quoted at the beginning of the present paper, he reminded the "lutte violente contre l'esprit positif" undertaken in his own youth by the supporters the "idéalisme métaphysique".

22.4 Towards a 'New, Historical Epistemology'

In *Problemi della scienza* Enriques had clarified how reality is not a pure fact, but the result of the rational construction in cognitive activity, which coordinates suitably associated data of the senses. His "critique of science" had shown that this is the case both with crude reality and with scientific reality as well. "Like the reality that belongs to common life, scientific reality is also a rational construction that coordinates the data of the senses", he held in *Scienza e razionalismo* (Enriques, 1912, p. 20). There he took up and clarified the model of science and, from this, that of scientific knowledge deriving from his epistemological analysis of 1906:

> The concept constructed by science represents the facts in an approximate way; therefore in its determination there is - it is true - an arbitrary element and an economic choice; but the arbitrariness is contained within the limits of the approximation marked by the experience, and with regard to the progress of the scientific construction it must be considered not a

> *convention* but a *hypothesis*, that is, a preordained disposition of future experiences. Thus, in the scientific relationship between hypothesis and experience we find in a higher form the invariant relationship between voluntary act and sensation, which constitutes the common meaning of reality. Science is not only approximate but also relative. This implies that the meaning of a scientific fact must be subordinated at all times to the set of all acquired knowledge. (italics in original) (Enriques, 1912, p. 20)

In this connection, one most characteristic elements of Enriques' scientific philosophy, including his philosophical continuism, namely the confidence in the continuity of scientific development, emerged clearly in the talk (Enriques, 1921) he gave in 1921 to introduce Einstein to the crowd of students and scholars who flocked in Bologna to attend the famous scientist's lectures.

"Einstein is presented to the public as a revolutionary", Enriques began. The theory of relativity "has brought a new opportunity to cry out for the bankruptcy of science". Someone has rejoiced or grieved that "even the firmest truth that for two centuries we have learned to revere as the triumph of human reason", i.e. the Newtonian law of universal gravitation, "must now be recognized as not exact" because "reason cannot admit intermediate term to the alternative of true or false" (Enriques, 1921, p. 271). Nothing is further away not only from Einstein's thought, Enriques protested, but also from the "historical concept of science" that is now accepted by scientists and especially by "mathematical thinkers". For, "no theory today claims to absolute exactness, but each one is given as a perfectible degree of truth, which unfolds and grows with the progress of reason". Thus, Einstein's theory does not mark the death of Newton's theory, but it represents "the conquest of a truer truth, in front of which the previous one will always appear as a degree of approximation".

It is interesting to remark that in Enriques' opinion Einstein had accomplished a "philosophical" revolution by bringing to an end a process that began in classical Greece and lasted many centuries. In his words

> The philosophical revolution that Einstein brought to completion is shown as the result of an evolution of thought over several centuries. It began 500 years before the Common Era with Parmenides of Elea, first advocate of the relativity of movement. Einstein is not belittled by saying that he concludes, in a broader cosmological synthesis, the work of a long series of philosophers, mathematicians and physicists, from whom he has collected disparate elements to merge them into his own construction. (Enriques, 1921, p. 274)

As for his reference to Parmenides, it is worth mentioning that Enriques wrote to Xavier Léon on October 27, 1918 that he had been working "for a year on the history of Greek philosophy and above all on the Eleatic school and on Democritus" in the hope "to draw new conclusions that are not without interest and will also provide me with clearer indications on modern philosophy".

Indeed, in his view, the historical understanding of knowledge was the only way to gain a deep understanding of a theory, not only in philosophy but also in science. "A dynamic view of science naturally leads to the terrain of history", he stated in the opening lines of his 1915 treatise on the geometric theory of algebraic equations and functions (Enriques, 1915, p. XI).

Thus, history becomes an integral part of science and finds its place in the exposition of theories. In a programmatic passage worth quoting *in extenso* from his Preface to (Rufini, 1926) Enriques wrote:

> The path of the history of knowledge is so hard that one must be well prepared to walk it. Minute erudition, the tool of tongues, the diligence in collecting and arranging study materials, are required as a preliminary ruling; but above these qualities the historian is asked for that intrinsic interest in the subject, which is true scientific and philosophical intelligence, addressed not so much to the development of the results as to the position of the problems and the inspiring ideas of the doctrines, without which the erudite remains only a erudite, a translator, a collector, an organizer and never becomes a historian; for, unable to understand science in its being, much less can he grasp its becoming, that is, reconstruct and evaluate its progress. (Rufini, 1926, p. 15)

In other words, history "ne peut évidemment pas se réduire à une collection et à une collation de textes et de notices savantes. I faut qu'elle soit construite par la pensée de l'historien" (Enriques, 1934, p. 48) Enriques will hold almost 10 years later in his 1934 pamphlet *Signification de l'histoire de la pensée scientifique*. "La science n'est pas simplement le reflet d'un ordre de choses en dehors de nous, c'est au contraire, la construction de la réalité par l'intelligence" (Enriques, 1934, p. 22), and the construction is built according to experience data, and even the principles evolve to adapt a larger reality. To the effect that, in conclusion "la vérité n'est qu'un acheminement vers le vrai" (Enriques, 1934, p. 7).

In his 1934 pamphlet his criticism was addressed against both Kantianism and pragmatism and, above all, idealism.

> La philosophie de la nature est chue dans le néant et les idéalistes de formation nouvelle croient s'être débarassés de ce poids mort en declarant que, sous n'importe quelle forme, l'étude de la nature est una activité d'ordre pratique, indifférente à la pensée. Ce faisant, ils n'ont pas seulement anémié l'idéalisme en méconaissant les raisons profondes de l'expression romantique [...] mais – fait bien plus grave pour des penseurs aux yeux desquels tout est dans l'histoire – ils ont péché contre la vérité historique. Car depuis ses plus lointaines origines la philosophie, ou tout au moins la philosophie occidentale, s'est inspirée et s'est conformée à la pensée naturaliste. (Enriques, 1934, pp. 32–33)

Nonetheless, at that time Enriques was considered by the heirs of the Vienna Circle as one of the thinkers who "prepared the ground for a modern theory of scientific empiricism". This is what Neurath would say in a letter to Enriques on June 18, 1937 reminding the "truly touching way" in which Enriques remembered his battles in favor of empiricism against metaphysical idealism in his talk at the 1935 Congress in Paris (quoted above). In a letter of February 4, 1937, Neurath asked Enriques to contribute the *International Encyclopaedia of Unified Science* with a seventy pages paper on the history of science (which in the end Enriques never wrote). In addition, Enriques was also invited by Neurath to write a short introductory text for the first issue of that *Encyclopaedia*.

In the same year 1937 Enriques took part in the Congrès Descartes with a talk on *Le problème de la raison* in which he summarized the essential features of his philosophical conception. He summed up his thoughts by saying that "la raison ne peut être conçue à la manière kantienne; elle n'est pas intuition capable de jugements sintentique a priori ni simple intellect discursif" (Enriques, 1937, p. 3). In his view, this 'new epistemology' will have both its foundation and its method in the historical critique of scientific concepts.

Nowadays, Enriques continued,

> on a compris que toute théorie scientifique ne renferme que des vérités partielles et nécessairement approchées. La Vérité adorée par les hommes ne descend pas de son autel, mais devient un terme idéal du progrès que la raison humaine ne saurait atteindre et qu'elle tend à réaliser par la construction historique d'une science toujours plus parfaite. (Enriques, 1937, p. 6)

Enriques would not have had much time to elaborate his ideas further. The racial laws in Italy (1938) that excluded him from universities and academies, and prevented him from publishing his writings, were followed by WWII. His research on Democritus carried on during the German occupation of Rome with the young Massimo Mazziotti was only published posthumously, after his death in 1946.

22.5 Conclusion

There is scarcely any doubt that Enriques was a somehow emblematic figure of his days, as was for example Poincaré, to whom often he has been compared. And, like Poincaré it is also true that Enriques sowed seeds that were to bear fruit even several years after his death.

The "meaning of the history of scientific thought", the relationship between science and history, between reason and history, between truth and error, the experimental rationalism "enlarged by the coordination of historical experience", the rational needs in the "scientific construction", the "construction of history", all these are elements that Enriques worked out in the mid-1930s and that in 1937 allowed him to present his "new" epistemology at the Congrès Descartes. His "historical epistemology" was the end point of a long philosophical journey begun with the *Problemi della scienza*. Thus, the image of Enriques as anchored to the philosophical themes and options of the early twentieth century, and therefore irremediably outdated, seems to me to be no longer sustainable.

References

Bottazzini, U., A. Conte, P. Gario. 1996. *Riposte armonie. Lettere di Federigo Enriques a Guido Castelnuovo*. Torino: Bollati Boringhieri.

Croce, B. 1919. *Pagine sparse*, ed. G. Castellano. Ricciardi. Napoli.

Enriques, F. 1894. Sui fondamenti della geometria proiettiva. *Rendiconti del Regio Istituto Lombardo di Scienze e Lettere, 27*: 550–567.

Enriques, F. 1894–1895. *Conferenze di geometria*, lith. Bologna.

Enriques, F. 1898. *Lezioni di geometria proiettiva*. Bologna: Zanichelli.

Enriques, F. 1900a. Sull'importanza scientifica e didattica delle questioni che si riferiscono ai Principii della Geometria. In *Questioni riguardanti la geometria elementare*. Bologna: Zanichelli, pp. 1–31.
Enriques, F. 1900b. *Grundlagen der Geometrie* di David Hilbert. *Bollettino di Bibliografia e Storia delle Scienze Matematiche, 3*: 3–7 (Repr. in Enriques 2000, pp. 317–321).
Enriques, F. 1901. Sulla spiegazione psicologica dei postulati della geometria. *Rivista di filosofia, 4*: 171–195 (Repr. in: F. Enriques, *Memorie scelte di geometria*, Vol. II: 1898–1910, Zanichelli, Bologna, 1959, pp. 145–161).
Enriques, F. 1906. *Problemi della scienza*. Bologna: Zanichelli.
Enriques, F. 1912. *Scienza e razionalismo*. Bologna: Zanichelli.
Enriques, F. 1914. *Problems of Science*. Chicago: The Open Court. Engl. Transl. of (Enriques, 1906)
Enriques, F. 1915. *Lezioni sulla teoria geometrica delle equazioni e delle funzioni algebriche*, ed. Chisini, O., vol. I. Bologna: Zanichelli.
Enriques, F. 1921. Le conferenze di Alberto Einstein a Bologna. Parole di presentazione. *Rivista di filosofia, 13*: pp. 271–274 (Repr. in Enriques 2000, pp. 329-332).
Enriques, F. 1934. *Signification de l'histoire de la pensée scientifique*. Paris: Hermann.
Enriques, F. 1935. Recensioni – Comptes rendus. *Scientia, 57*: pp. 69–71, pp. 227–229.
Enriques, F. 1936a. Allocution de M. F. Enriques. In *Actes du Congrès international de philosophie scientifique*, vol. I: Philosophie scientifique et empirisme logique, p. 12. Paris: Hermann.
Enriques, F. 1936b. Philosophie scientifique. In *Actes du Congrès international de philosophie scientifique*, Vol. I: Philosophie scientifique et empirisme logique, pp. 23–27. Paris: Hermann (Repr. in Enriques 2000, pp. 219–222)
Enriques, F. 1936c. Recensioni – Comptes rendus. *Scientia, 60*: pp. 109–110, pp. 175–176.
Enriques, F. 1937. Le Problème de la raison. In *Congrès Descartes. Travaux du IXe Congrès International de Philosophie*, pp. 3–6. Paris: Hermann.
Enriques, F. 1958. *Natura, ragione e storia*, ed. Lombardo Radice L. Torino: Boringhieri.
Enriques, F. 2000. *Per la scienza. Scritti editi e inediti*, ed. Simili, R. Napoli: Bibliopolis.
Gray, J. 2006. Enriques: popularising science and the problems of geometry. In *Interactions. Mathematics, Physics and Philosophy, 1860–1930*, ed. Hendricks, V.F., Jørgensen, K., Lützen, J., Pedersen, S. Boston Studies in the Philosophy of Science, vol. 251, pp. 135–153. Dordrecht: Springer.
Hahn, H., O. Neurath, and R. Carnap. 1929. *Wissenschaftliche Weltauffassung. Der Wiener Kreis, Vienna*. Berlin: A. Wolf Verlag. Reprint in [Stadler & Uebel 2012].
Klein, F. 1890. Zur Nicht-Euklidischen Geometrie. *Mathematische Annalen, 37*: 544–572.
Klein, F. 1898. Gutachten, betreffend den dritten Band der Theorie der Transformationsgruppen von S. Lie anlässlich der ersten Vertheilung des Lobatschewsky-Preises. *Mathematische Annalen, 50*: pp. 583–600.
Lautman, A. 1936. Mathématiques et réalité. In *Actes du Congrès international de philosophie scientifique, Sorbonne, Paris 1935. VI. Philosophie des mathématiques*, pp. 24–27. Paris: Hermann.
Lolli, G. 2018. Federigo Enriques at the 1935 international congress for scientific philosophy in Paris. *Philosophia Scientiae, 22*(3): pp. 119–134. Online since 25 October 2020, connection on 31 March 2021. http://journals.openedition.org/philosophiascientiae/1583. https://doi.org/10.4000/philosophiascientiae.1583
Rougier, L. 1936. Allocution d'ouverture du Congrès. In *Actes du Congrès international de philosophie scientifique, Sorbonne, Paris 1935. I. Philosophie scientifique et empirisme logique*, pp. 7–9. Hermann: Paris.
Rufini, E. 1926. *Il 'metodo' di Archimede e le origini del calcolo infinitesimale nell'antichità*. Roma: Stock.
Russell, B. 1936. The congres of scientific philosophy. In *Actes du Congrès international de philosophie scientifique, Sorbonne, Paris 1935. I. Philosophie scientifique et empirisme logique*, pp. 10–11. Paris: Hermann.

Stadler, F.. and T. Uebel, eds. 2012. *Wissenschaftliche Weltauffassung. Der Wiener Kreis, Vienna*. New York: Springer.

Vailati, G. 1906. I Problemi della scienza. *Leonardo, IV*: pp. 375–379.

Vailati, G. 1971. *Epistolario, 1891–1909*, ed. G. Lanaro. Torino: Einaudi.

Wigner, E.P. 1960. The Unreasonable effectiveness of mathematics in the natural sciences. *Communications on Pure and Applied Mathematics, 13*: pp. 1–14.

Part VI
Philosophical Issues

Chapter 23
Who's Afraid of Mathematical Platonism?—An Historical Perspective

Dirk Schlimm

Abstract In *Plato's Ghost* Jeremy Gray presented many connections between mathematical practices in the nineteenth century and the rise of mathematical platonism in the context of more general developments, which he refers to as *modernism*. In this paper, I take up this theme and present a condensed discussion of some arguments put forward in favor of and against the view of mathematical platonism. In particular, I highlight some pressures that arose in the work of Frege, Cantor, and Gödel, which support adopting a platonist position. The aim of this discussion is to provide an historically informed introduction to the philosophical position of mathematical platonism and to point at some of its mathematical and philosophical roots in the nineteenth century.

23.1 Introduction

It is a general theme in the work of some historians and philosophers of mathematics that mathematical practice and philosophical reflections about mathematics should be seen as intertwined and that a better understanding of both the history and the philosophy of mathematics can be gained by studying their interactions. In particular, the connection between mathematical developments in the nineteenth century and the rise of mathematical platonism was highlighted in Jeremy Gray's *Plato's Ghost* (2008) in the context of more general developments that he refers to with the term *modernism*. In this paper, I take up this theme and present a condensed discussion of some arguments put forward in favor of and against the view of mathematical platonism (Sect. 23.2). In particular the claim that abstract mathematical objects exists *independently of human agents* lacks prima facie plausibility. To give some support to this aspect of mathematical platonism I present some motivations for it that arose in the work of Frege, Cantor, and Gödel (Sect. 23.3) as well as in

D. Schlimm (✉)
McGill University, Montreal, QC, Canada
e-mail: dirk.schlimm@mcgill.ca

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_23

the experience of engaging with mathematics (Sect. 23.4). Finally, how some more recent discussions are related to those earlier motivations is indicated (Sect. 23.5). The aim of these discussions is to provide an historically informed introduction to the philosophical position of mathematical platonism and to present some of the early motivations behind it, thereby pointing at some of its mathematical and philosophical roots.

Due to the wide range of historical developments and philosophical arguments that are being touched upon, as well as the aim of keeping the discussion generally accessible, many fine-grained distinctions that have been introduced in the vast amount of literature on mathematical platonism of the past decades and many nuances of the mathematical and philosophical details of the work that contributed to its rise are beyond the scope of this paper. Thus, it can only be a modest contribution towards a prehistory of mathematical platonism, i. e., towards a better understanding of the developments that led mathematical platonism to become the main position in philosophy of mathematics of the twentieth century.

23.2 Aspects of Mathematical Platonism

The term 'platonism' appeared prominently in philosophy of mathematics for the first time in Bernays's paper 'Sur le platonisme dans les mathèmatiques' (1935).[1] However, in the same volume of the journal in which Bernays's paper appeared, also Fraenkel referred to Plato's view of mathematical objects and, even earlier, Poincaré had connected Cantor's realism to 'Plato's theory of ideas' (Bouveresse 2005, 55 and 62). Nevertheless, mathematical platonism should not be confused with Plato's views.[2] To avoid this confusion, some philosophers prefer to use the term *realism*, while others use 'Platonism' with an upper-case 'P' to refer to Plato's views and (lower-case) 'platonism' to refer to the contemporary position. We shall adopt the latter convention here and, without delving deeper into these issues, consider mathematical platonism as a philosophical position regarding the nature of mathematical entities, characterized by the following three claims:[3]

1. Mathematical entities are *abstract*.
2. Mathematical entities *exist*.
3. Mathematical entities are *independent from human agents*.

[1] Bernays advances his position in reaction to constructivist and intuitionist critiques of analysis and set theory, and distinguishes between different versions of platonism according to the strength of their assumptions. For a discussion of Bernays's philosophy, see Parsons (2008). We leave aside Bernays's reflections, as they deserve a more thorough discussion than can be given here.

[2] See Landry (2023).

[3] For an introduction to platonism from a contemporary perspective that is nevertheless sensitive to historical developments, see Panza and Sereni (2013); cf. also Linnebo (2018a).

Paradigmatic mathematical entities, which have been in use since antiquity, are *numbers* and geometrical *points* and *lines*. Contemporary mathematics is replete with other quite different entities, but for the present paper focusing on numbers and geometrical objects will do, as they exhibit many of the features that are characteristic also of many other entities. To get started, let us briefly discuss in turn the claims of mathematical platonism and assess them with regard to their prima facie plausibility.

23.2.1 Abstractness

When being confronted with the symbol '3', it is not uncommon, in particular for children, to say that 'this *is* the number three'. However, a different symbol of the same shape would then also be the number three and so we would have two different entities being the number three, so that the use of the definite article 'the' would no longer be appropriate.[4] One way out of this conundrum would be to invoke the distinction between *types* and *tokens*: in this case, the concrete instances that we can see are tokens of the same type. And suddenly we are in the realm of the *abstract*, since types are generally considered to be abstract.

Another motivation for considering numbers to be abstract is the simple fact that we can use different words or notations to refer to the number three, such as 'trois', 'drei', or 'III'. Consequently, we should not identify the number three with particular words or symbols (or, more accurately, with their abstract types), but rather with what these words or symbols stand for or represent.

Finally, from a mathematical perspective there are infinitely many natural numbers, and geometric lines do not have any width (according to Euclid), which does not easily square well with the view that these objects are concrete entities in our physical universe. There are, of course, many ways to give substance to the notions of types and mathematical entities, but that they are abstract seems fairly uncontroversial. In this respect, mathematical entities are no different than everyday concepts, such as redness and justice.

The abstract nature of mathematical entities accounts for the fact that questions such as 'How many letters has the number three?' or 'What color is the number three?' are nonsensical, although one can meaningfully ask them about particular linguistic or mental representations of the numbers, say in a particular language or when thinking about them.[5]

The abstractness of numbers and geometric objects also raises some perplexing questions, most prominently: How do we interact with such abstract entities? In the literature this is often referred to as the *problem of access* to abstract entities. Note,

[4] This, and many of the other arguments against alternatives to mathematical platonism can be found in Frege (1884).

[5] In fact, some people with synesthesia associate specific colors with particular numbers or numerals, see Cytowic (2018).

however, that this question is not specific to mathematics, but can be asked about any abstract entity, whether it is the letter ‘A’, Beethoven’s fifth symphony, or the concept of justice.

23.2.2 Existence

Spelling out the notion of existence is a thorny issue when it comes to abstract entities. While one might hesitate to concede that ‘The natural numbers exist’, I would also expect that most readers would answer the question whether there is a natural number greater than 3, but smaller than 5, with an emphatic ‘Of course, there is: it’s 4!’. This observation brings out a tension with our ordinary uses of ‘existence’ and ‘there is’. More importantly for the topic of this paper, it shows that we might not be as reluctant to talk about the existence of abstract entities as one might be inclined to assume initially.

While it is commonly accepted to say that the objects one directly perceives with the senses, such as trees and trains, exist in the world,[6] the situation is different in the case of abstract entities. Clearly, the number three cannot be seen, touched, or smelled, so that it does not exist in the same sense (if it does at all) as the tree I see outside. Again, however, this is not a problem specific to mathematical entities: we might also be inclined to say that Beethoven’s fifth symphony exists, even though there might not be anybody in the world playing it at this time. A possible solution of this problem is to distinguish between *actual* existence in the physical world and a different kind of existence that applies to abstract entities and that needs to be fleshed out separately.

An alternative approach to the existence of abstract entities is to consider them as human creations, whose existence is maintained through practices (e. g., concert performances) and physical or mental representations (e. g., the score of Beethoven’s symphony or the memory of how to play it). Without any such practices and representations, we might well be inclined to declare the object in question as having ceased to exist—however, people disagree whether this also applies to mathematical objects.

23.2.3 Positions in Nineteenth-Century Philosophy of Mathematics

The above considerations show that the first two characteristic features of mathematical platonism discussed above are not immediately obvious. While only few reject the abstract nature of mathematics, the notion of mathematical existence is clearly in

[6] This does not rule out the possibility of optical illusions. However, the fact that we can recognize them as such, also presupposes that other perceptions are not illusory.

need of more discussion. Indeed, we can interpret the main positions in nineteenth-century philosophy of mathematics as attempts to address the reservations about the abstractness and existence of mathematical entities. In all of these accounts, however, their existence remains tied, in one way or another, to empirical objects or human thought. That it should be conceived of as being completely independent from human agents requires additional motivations, some of which will be discussed in Sects. 23.2.4 and 23.3.

Worries about the existence of mathematical objects, conceived as abstract, has stood behind the move made by some mathematicians to identify them with concrete inscriptions. In the nineteenth century this view was sometimes called *formalism* (not to be confused with the position attributed to Hilbert in the first decades of the twentieth century, presented in Sect. 23.3.3). However, while one can find individual expressions of such an attitude towards mathematics, these views rarely caught on. The following passage by Heine is one that is frequently referred to in this connection: 'I adopt a purely formal standpoint in the definition [of numbers] by calling certain tangible signs *numbers*, such that the existence of these numbers is beyond all question' (1872, 173; translation and emphasis by DS). One reason why this view has not been taken up by many is, presumably, because it would be difficult to maintain in light of questions such as: Given that only a finite amount of numerals have been written so far, does this mean that there are only finitely many natural numbers? A possible reply would be to not identify the numbers with actual concrete inscriptions, but with *possible* ones. However, such talk about possible entities takes us away again from the view that mathematical entities are concrete and tangible, thus defeating the original motivation for the position.

Worries about the existence of mathematical entities are related to two other philosophical positions that link them either to the physical or the mental world: empiricism and apriorism.[7]

If I see three oranges on the table in front of me and my peers confirm this, I am fairly confident that there exist three oranges. By taking empirical evidence like this as a starting point, one can try to give an account of abstract entities through the process of *abstraction*. For example, first, by abstracting away from all properties of the oranges and retaining only the fact that there are three different objects one can arrive at an understanding of what it means to say that there are *three* things, rather than, say, four. Second, by abstracting away from all imperfections and sensible qualities of one of the oranges and idealizing its shape to a set of points that all have the same distance from a center, one can arrive at an understanding of the geometrical entity of a *sphere*. Indeed, Aristotle gave an account of Plato's abstract Forms along similar lines. Now, how exactly this process of abstraction is supposed to work and what it yields has been a matter of debate since then. Two popular alternatives are the following. First, abstraction brings into being or generates abstract entities that then continue their existence in some sort of abstract

[7] This terminology is taken from Pasch, who considered 'empiricism' and 'apriorism' as the main philosophical positions of his time (1926, 138).

realm. Second, through abstraction we are able to conceive of the concrete objects as imperfect instantiations of abstract entities; in other words, we actually see an orange, but we mentally see it as a sphere. Without delving further into the details of these accounts, the point is that empiricism connects the existence of mathematical entities to the existence of concrete objects through an (admittedly somewhat mysterious) process of abstraction.

What makes empiricism a prima facie plausible account is that it goes together well with how children seem to actually learn mathematical terminology. We learn the word sequence 'one, two, three, ... ' and how to map them to individual objects that we want to count. Then, we realize that this process can be applied to any collection of objects and can, in principle, be carried out indefinitely. Geometric shapes are initially studied with the use of diagrams, which can be made more and more accurate with the use of a ruler and a thinner pencil. In both cases concrete physical objects are the starting point and there is a gradual development towards more abstract mathematical notions. Indeed, the learning process was put forward in favor of an empiricist account by nineteenth-century empiricists, such as Klein and Pasch.[8]

Some difficulties of empiricism are: (1) to find appropriate physical correlates for all mathematical objects, (2) how to know in advance (i. e., without already possessing the concept of a sphere) what properties of an orange I need to abstract in order to arrive at a sphere, and (3) to account for how it is possible to ground (seemingly) necessary truths on contingent foundations. These problems can be resolved more easily if the origins of mathematical entities are located directly in the mind instead of in the physical world. Particularly influential for the nineteenth-century was Kant's account that mathematics studies the *forms of intuition*, i. e., space and time, and that the objects of mathematics exist as long as they can be presented in *pure intuition*. Thus, without the need to establish a direct link between the physical and the mathematical worlds, the abstractness and existence of mathematical entities can be accounted for. The charge of making mathematics depend on our psychology is countered by Kant by conceiving of the forms of intuition as *transcendental*, i. e., as necessary conditions for both pure and empirical thought. However, developments in nineteenth-century mathematics, such as the emergence of non-Euclidean geometries and of continuous, nowhere differentiable functions, which both defied imagination and expectation, were interpreted as severe, if not devastating, challenges to Kant's account.

23.2.4 Independence

All of the philosophical positions discussed above make mathematical truths dependent on our choice of concrete means of representation (formalism), on facts

[8] See Schlimm (2010, 102).

about the world (empiricism), or on facts about our minds (apriorism). However, that $5 + 7 = 12$ seems to be true independently of our notations, the nature of the universe, and our minds. Moreover, we already noticed above that the actual presence of particular representations for mathematical entities might be too strong a requirement for their existence. The fact that the real numbers are uncountable, but that there are only countably many sentences in English, for example, only adds to this difficulty, as it makes it impossible to represent all real numbers by different English expressions. These concerns are addressed by the third characteristic claim of mathematical platonism, namely that mathematical entities exist independently from human agents.

There are different ways of interpreting what it means for an abstract entity to depend on human agents: it might be said to exist only when somebody actually thinks of it, or it might begin to exist when somebody thinks of it and then simply continue its existence, or it might be said to exist as long as it features in some practices, or in physical or mental representations. For mathematical platonism, in its strongest from, the independent existence of mathematical entities is often expressed counterfactually: even if there had never been any human beings, mathematical entities would still exist.

This aspect of platonism is often considered to be the most difficult to accept. Indeed, it runs counter to our knowledge of other abstract entities, such as cultural artifacts. Clearly, Beethoven was the originator of his symphonies. Likewise, it is common lore of the history of mathematics that numbers only emerged over time, first with the positive integers, then zero, etc. That their existence (however conceptualized) should be independent of any human activities is a hard pill to swallow, so that we need compelling reasons for accepting it. In particular, if the existence of mathematical entities does not depend on any humans or on physical circumstances, it should also be timeless, or, perhaps, eternal — properties that are usually attributed only to religious entities.

23.3 Mathematical and Philosophical Pressures

In order to find some motivations for the independent existence of mathematical entities, let us now trace some of the roots of mathematical platonism by examining briefly some historical developments in mathematics. The nineteenth century was a time of tremendous advances, laying the foundation for modern mathematics as we know it today. In particular, the developments which transformed geometry have, deservedly, received a lot of attention: The possibility of consistent theories of non-Euclidean geometry questioned Kant's aprioristic view that the subject matter of geometry is space, taken as one of the forms of our intuition, and nudged mathematicians such as Gauss and Riemann towards an empiricist stance towards mathematics; the duality of projective geometry questioned the tight link between the language of mathematics and its meaning, paving the way towards formal approaches. In addition, advances in analysis, such as Weierstrass' definition of continuous, nowhere differentiable functions, put pressure on the roles of intuition and visualizability, and developments in algebraic number theory introduced hitherto unknown mathematical entities.[9] These developments are well known and their connections to platonist positions in the philosophy of mathematics are mainly indirect, by questioning the plausibility of alternative accounts. In the following, I want to present three mathematicians, whose work led them directly to formulate a view of mathematical entities as existing independently of human thought. In other words, their reflections on mathematics provided the motivation for accepting the third characteristic claim of mathematical platonism.

23.3.1 Cantor on Set Theory and Criteria for Existence

The work of Dedekind and Cantor on set theory has certainly been one of the turning points in modern mathematics. It introduced a coherent way of reasoning about infinity and led to a theory that was based on a few primitive notions, but nevertheless powerful enough to be considered as the foundation of all of mathematics. In this section, I will focus on Cantor's account of the introduction of transfinite numbers and his philosophical principles, as identified by Hallett (1984), because these considerations give some motivation to Cantor's views on the existence of mathematical entities.[10]

Cantor distinguishes between the *power* [Mächtigkeit] of a set and its *ordinal number* [Anzahl], which is defined only if the set is well-ordered. Every well-defined set A has a *power*, denoted by $|A|$, and two sets A and B have the same power if there is a bijection between them. On this basis, Cantor was able to show that $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|$, but $|\mathbb{N}| \neq |\mathbb{R}|$, i. e., that the power of the set of natural numbers, which Cantor called $\aleph_0$, is different from that of the real numbers. A set A, together with a successor relation, is called *well-ordered* by Cantor if it satisfies the following conditions:

> (i) there is a *first* element of the set; (ii) every single element (provided it is not the last in the succession) is followed by another determinate element; and (iii) for any desired finite or infinite set of elements there exists a determinate element which is *their immediate successor* in the succession (unless there is absolutely nothing in the succession following all of them). (Cantor 1883, 884)[11]

In contemporary set theory, a set is well-ordered by a total order relation if every non-empty subset has a smallest element. According to both definitions, $(\mathbb{N}, <)$ is well-ordered, but the integers $(\mathbb{Z}, <)$ are not. Cantor attributes *ordinal numbers* to well-ordered sets in such a way that two sets have the same ordinal number if there

[9] For a discussion of some of these developments, see Gray (1992, 2008).

[10] More details about Cantor's work can be found in Ferreirós (2007).

[11] The page numbers refer to the translation in Ewald (1996), who leaves the original 'Anzahl', which is rendered as 'ordinal number' in the present text.

is a bijection between them that also preserves the order of the elements (i. e., is an isomorphism). Thus, while the power of a set gives us some information about *how many* elements it has (cardinal numbers), the ordinal number also gives us some information about the particular structure of the set that is determined by the order relation. In particular, two sets of the same power can have a different ordinal number.

A question that arises from these considerations is: What ordinal numbers are there? Cantor provides three principles for their 'generation', here the first:

> The sequence (I) of *positive integers* $1, 2, 3, \ldots, v, \ldots$ has its ground of origin in the repeated positing and uniting of underlying unities, which are regarded as alike; the number v is the expression for a definite finite ordinal number of such positings following one another in a sequence; it is also the expression for the unification of the posited unities into a whole. The formation of the finite real integers thus rests upon the principle of *adding a unity to an already formed and existing number*; I call this principle (which, as we shall soon see, also plays an essential role in the generation of the higher integers) the *first principle of generation*. The ordinal number of the numbers v of class (I) formed in this way is infinite and there is no greatest among them. (Cantor 1883, 907)

Leaving aside some problematic issues with Cantor's conception of unities, which have been pointed out and criticized by Frege,[12] the account is similar to that put forward by others and has some initial plausibility. We simply keep on adding units (represented in the following by '$\circ$') to form larger numbers:

$$\underbrace{\circ}_{1} \qquad \underbrace{\circ\,\circ}_{2} \qquad \underbrace{\circ\,\circ\,\circ}_{3} \qquad \text{etc.}$$

Since this process can be continued indefinitely, there is a (potential) infinity of numbers characterized in this way. It is at this point that Cantor's philosophical principle of actual infinity, or the *domain principle*, applies: 'Any potential infinity presupposes a corresponding actual infinity' (Hallett 1984, 7). After all, we cannot *actually* generate all positive integers in this way, but by accepting this process as never-ending, we must assume the existence of *all* such numbers in the first place. Otherwise, we might run out of units to add and the process could not be continued indefinitely. Cantor named the ordinal number of the infinite set of positive integers 'ω',

$$\underbrace{\circ\circ\circ\circ\circ\circ\circ\circ\circ\circ\circ\circ\circ\circ\cdots}_{\omega}$$

and described its introduction as follows:

> I call it the *second principle of generation* of integers, and define it more exactly thus: *if any definite succession of defined integers is put forward of which no greatest exists, a new*

[12] Cantor's conception of pure units has also found it's defenders, e. g., Hallett (1984) and Fine (1998).

> *number is created* by means of this second principle of generation, which is thought of as the *limit* of those numbers; that is, it is defined as the next number greater than all of them. (Cantor 1883, 907–908)

A crucial aspect of Cantor's account is that we can apply his two principles of generation of integers again and again:

$$\underbrace{\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\circ\circ\circ\circ\cdots\bigcirc}_{\omega+1} \qquad \underbrace{\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\circ\circ\circ\circ\cdots\bigcirc\bigcirc}_{\omega+2} \qquad \text{etc.}$$

Indeed, we can also add ω to ω (resulting, in modern notation, in $\omega \cdot 2$) and again repeat this process indefinitely, which yields the next ordinal number, ω^2:

$$\underbrace{\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\circ\circ\circ\circ\cdots\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\circ\circ\circ\circ\cdots\;\cdots\cdots\;\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\circ\circ\circ\circ\cdots\cdots\cdots}_{\omega\cdot\omega=\omega^2}$$

We notice how Cantor assimilates the transfinite ordinal numbers to the positive integers, an assimilation which is encapsulated in the *principle of finitism*: 'The transfinite is on a par with the finite and mathematically is to be treated as far as possible like the finite' (Hallett 1984, 7). This is clearly a metaphysical principle for Cantor, but (once we accept the existence of infinite sets) it has a mathemtical underpinning, namely though the notion of well-ordering. In this way, we obtain the sequence of ordinal numbers: $1, 2, 3, \ldots, \omega, \omega + 1, \omega + 2, \ldots, \omega \cdot 2, \omega \cdot 2 + 1, \ldots, \omega^2, \ldots, \omega^\omega, \ldots,$ which so far are all *countable*, i. e., belong to sets of cardinality $\aleph_0$.

To generate ordinal numbers of uncountable sets, such as ω_1, Cantor invokes a third principle, but we don't need to consider that here in order to get the main point.

As I have indicated, there is no limit for the generation of ordinal numbers, but by the domain principle, each potential infinity presupposes an actual infinity, which Cantor called *transfinite*. In addition, to have one place for *all* of these infinities to live without the need of having to go beyond it again, Cantor invoked a third philosophical principle, the *principle of Absolute infinity*: 'The Absolute infinite cannot be mathematically determined' (Hallett 1984, 7).

Cantor saw his considerations about ordinal numbers as a special case of a more general methodology in mathematics. The limitlessness of the growth of mathematics is perhaps one of the most well-known aspects of Cantor's view, — expressed in the famous quote 'the *essence of mathematics* lies precisely in its *freedom*' (Cantor 1883, 896)—but it should be kept in mind that he did formulate some requirements on the introduction of new mathematical concepts. He writes:

> Mathematics is in its development entirely free and is only bound in the self-evident respect that its concepts must both be consistent with each other and also stand in exact relationships, ordered by definitions, to those concepts which have previously been introduced and are already at hand and established.
>
> In particular, in the introduction of new numbers it is only obligated to give definitions of them which will bestow such a determinacy and, in certain circumstances, such a relationship to the older numbers that they can in any given instance be precisely

> distinguished. As soon as a number satisfies all these conditions it can and must be regarded in mathematics as existent and real. (Cantor 1883, 896)

Thus, for Cantor new mathematical concepts must be *consistent* and *connected to other concepts by exact relationships*. As soon as these conditions are met, the mathematical entities in questions are 'existent and real'. This view foreshadows that of Hilbert, formulated at the turn of the twentieth century in terms of the consistency of axioms instead of concepts, to which we will return in Sect. 23.3.3.

We've seen that Cantor formulated a view that assigns a reality, or existence, to mathematical entities that are independent of any direct link to the physical world, visualization, or representation in intuition, but that are defined purely conceptually and with no other restrictions than their consistency and rigorously defined relations to other entities. Thus, despite the fact that Cantor speaks of 'awakening the concept' which 'slumbered inside us' in his mathematical paper (Cantor 1883, 918) and relegates any further speculations to philosophy, his view satisfies the three characteristic claims of mathematical platonism. Most importantly, that mathematical entities exist independently of human agents is motivated by the need to presuppose an actual infinite domain for any potentially infinite domain one defines and an absolute infinite domain to make room for any possible extension. As Hallett put it, Cantor proposes 'a Platonic principle: the "creation" of a consistent coherent concept in the human mind is actually the uncovering or discovering of a permanently and independently existing real abstract idea' (Hallett 1984, 18). Indeed, it becomes particularly clear in his correspondence that Cantor grounds the existence of mathematical entities in the existence of an eternal God, which is independent from human beings.[13]

23.3.2 Frege on Logic and Independence

Frege's main philosophical project was to show that arithmetic can be developed out of logic alone and that, in consequence, no appeal to intuition is required, as was held by Kant. As part of the project, Frege studied the views on the nature of numbers that were held by his contemporaries and found them all lacking. He fiercefully criticized these attempts and concluded that numbers are 'not abstracted from things', nor 'a property of things', 'not anything physical', nor 'anything subjective (an idea)'; finally, in contrast to Cantor's account, 'Number does not result from the annexing of thing to thing' (Frege 1884). For Frege, numbers are not attributed to collections of things, but to concepts, such as 'moons of Venus'

[13] Such a view did not originate with Cantor. For example, Kepler wrote in 1611: 'Without doubt, the authentic type of these [geometric; added by DS] figures exists in the mind of God the Creator and shares His eternity' (quoted from Davis et al. (2012, 145)). For more on Cantor's linking of mathematics and theology, see Hallett (1984) and Tapp (2014).

or ‘being a round square’. The latter can be used to illustrate an application of the number zero, as in ‘The number of round squares is zero.’

Concepts also underlie logical reasoning for Frege. It is thus no surprise that he called his logical calculus a ‘Begriffsschrift’, i. e., *concept script* (1879), and devoted considerable attention to the study of concepts (Frege 1892). In order to be used in logical inferences, Frege demands of concepts to be *definite* and *fixed* (Frege 1884, xvii). The former means that it must be unambiguous for any object whether it falls under the concept in question or not, which avoids problems of vagueness which we find in concepts such as ‘being bald’; the latter means that the extension of a concept must remain fixed once and for all.[14] The reason for this requirement is the following. Since the time of Aristotle it was held that logical inferences do not depend on the particular content of the terms involved, but on the *form* of the inferences, such as:

$$\text{If} \quad (1)\ \text{all}\ A\ \text{are}\ B \quad \text{and} \quad (2)\ x\ \text{is}\ A, \quad \text{then} \quad (3)\ x\ \text{is}\ B.$$

This inference form is considered to be valid, because whenever the premises (1) and (2) are true, the conclusion (3) must also be true. However, if the extension of *B* could change over time, it would in principle be possible for the premises to be true at some time, but the conclusion false at some other time, thereby rendering the inference invalid. In fact, without the concepts involved remaining fixed over the course of the argument, we would not be able to establish the validity of any logical inference, and Frege’s project of grounding arithmetic in logic could not get off the ground.[15] Thus his insistence on the view of concepts as definite and unchanging. This also makes them *objective*, i. e., not subject to the accidental state of the world or the ‘fancies of the mental’, which, in turn, guarantees the objective character of arithmetic as built purely on logical concepts.[16]

The objective reality of arithmetical concepts is illustrated by Frege with analogies to the physical world. He writes:

> [...] number is no whit more an object of psychology or a product of mental processes than, let us say, the North Sea is. The objectivity of the North Sea is not affected by the fact that it is a matter of our arbitrary choice which part of all the water of the earth’s surface we mark off and elect to call the “North Sea”. There is no reason for deciding to investigate the North Sea by psychological methods. In the same way number, too, is something objective. (Frege 1884, §26)

Accordingly, Frege takes a clear side on the debate whether mathematics is invented or discovered:

> [E]ven the mathematician cannot create things at will, any more than the geographer can; he too can only discover what is there and give it a name. (Frege 1884, §96, 107–108)

14 See Schlimm (2012) for a discussion of the contrast between a Fregean and a Lakatosian conception of concepts.

15 For a similar argument, see Poincaré (1909, 461).

16 Frege considered his view to pertain only to arithmetic and not, for example, to geometry.

Thus, while mathematicians cannot bring mathematical entities into existence, they can nevertheless interact with them. Frege gives only a hint on how this can be done, namely that the access problem can be solved by recourse to *reason* as a link between the realm of concepts and human thought:

> It is in this way that I understand objective to mean what is independent of our sensation, intuition and imagination, and of all construction of mental pictures out of memories of earlier sensations, but not what is independent of reason,—for what are things independent of reason? To answer that would be as much as to judge without judging, or to wash the fur without wetting it. (Frege 1884, §26, 36)

Later in his life, when writing about the *senses* of sentences, which he called *thoughts*, Frege explicitly talked about a 'third realm', different from both the physical world of things and the mental world of ideas:

> A third realm must be recognized. Anything belonging to this realm has it in common with ideas that it cannot be perceived by the senses, but has it in common with things that it does not need an owner so as to belong to the contents of his consciousness. Thus for example the thought we have expressed in the Pythagorean theorem is timelessly true, true independently of whether anyone takes it to be true. It needs no owner. It is not true only from the time when it is discovered; just as a planet, even before anyone saw it, was in interaction with other planets. ('Thought', quoted from Beaney (1997, 337).)

Here, again, Frege argues for the existence of an abstract realm, which is timeless and independent from human agents.

While both Cantor and Frege based their positions on the nature of mathematics on concepts, they focused on different aspects of mathematics to motivate the independence of mathematics from human agents. For Cantor, it was the indefinite extendability of mathematics that required a large enough ontology to fit it all in and he concluded that such a realm must be independent from human agents. For Frege, it was the objectivity of mathematics that was at stake. Since he found no other way of accounting for it on the basis of the available philosophical alternatives (i. e., based on physical or mental worlds), he argued for a view of numbers as being based on logical concepts and a view of logic as being based on definite and fixed concepts. For Frege, then, these motivated the third claim of mathematical platonism, i. e., its independence from human agents.[17]

23.3.3 Gödel on the Limits of Formalization

With the development of formal systems with explicit, truth-preserving inference rules, an alternative route to account for mathematical objectivity seemed to open up without requiring the positing of a platonic realm: mathematics could be reconstructed as the formal deduction of theorems from accepted true axioms.

[17] The sense in which Frege was indeed a platonist has been the subject of much debate. For a nuanced discussion of Frege's views, see Reck (2005).

Granted, accounting for the truth of the axioms remained a problem, but perhaps one that could be solved. Echoing Cantor's requirement of the consistency of mathematical concepts, now the consistency of the axioms became the *conditio sine qua non* for accepting a formal system. This could be shown by providing an interpretation, typically by exhibiting a set of objects and relations, or, as was suggested by Hilbert, by providing a proof of the syntactic consistency. Such a proof would still require the acceptance of some, presumably innocent, base system, but promised to make Cantor's and Frege's expansive views on existence largely unnecessary. The famous correspondence between Frege and Hilbert illustrates their diametrically opposed views regarding consistency and truth. As Hilbert writes in a reply to Frege:

> I was very much interested in your sentence: 'From the truth of the axioms follows that they do not contradict one another', because for as long as I have been thinking, writing, lecturing, about these things, I have been saying exactly the reverse: If the arbitrarily given axioms do not contradict one another, then they are true, and the things defined by the axioms exist. This for me is the criterion of truth and existence. (Letter from Hilbert to Frege, December 29, 1899; Frege 1980, 42)

What came to be known as *Hilbert's Program* consisted in (1) the formalization of mathematical theories and (2) giving proofs of the consistency of the resulting formal theories on the basis of (3) a very small and unproblematic part of mathematics, which Hilbert called *finitary*. In this way, Hilbert hoped, mathematics could be given a rigorous foundation that guaranteed that it was free from hidden contradictions. However, Gödel's groundbreaking incompleteness results (1931) showed these aims to be unachievable. His First Incompleteness Theorem asserts that the theory of natural numbers cannot be formalized in an appropriate axiom system of first-order logic, such that *all* truths that are expressible in the formal language can be derived from the axioms. As a consequence, the formal reconstruction is not complete. Gödel's Second Incompleteness Theorem asserts that a sufficiently strong formal system cannot prove its own consistency, so that its consistency cannot be proved by an even weaker system. Thus, the aim of proving the consistency of formalized mathematical theories with a weaker, finitary system, is impossible. While Gödel's technical results can be expressed more rigorously, the general lesson that has been drawn from them is that the notions of mathematical truth and provability in a formal system do not coincide.[18]

Without the recourse to formal systems Gödel himself returned to Frege's analogy between mathematics and natural science, comparing the axioms of mathematics with laws in science. Neither need to be self-evident, but they can be investigated on the basis of their consequences, which can be independently judged through sense perception in the case of science and intuition in the case of mathematics. Gödel writes:

> But, despite their remoteness from sense experience, we do have something like a perception also of the objects of set theory, as is seen from the fact that the axioms force

[18] For further discussion of the context and results of Gödel's work, see Giaquinto (2002).

> themselves upon us as being true. I don't see any reason why we should have less confidence in this kind of perception, i. e., in mathematical intuition, than in sense perception, which induces us to build up physical theories and to expect that future sense perceptions will agree with them [. . .] (Gödel 1947, 483–484)

Gödel saw further similarities between mathematics and science. Scientific and mathematical theories both help us understand our sensory observations and mathematical intuitions, but they also introduce elements that go beyond them: unobservables in science and higher set theory in mathematics. Our confidence in such assumptions must rest in both cases on their fruitfulness, in deriving empirical consequences in science, and leading to simpler proofs and new theorems in mathematics.

From a philosophical perspective, Gödel was more interested in the epistemology of mathematics and the methodological question of how to choose the right axioms of set theory given that particular claims, such as Cantor's Continuum Hypothesis, could not be settled by the hitherto accepted axioms.

While Gödel himself reported that he 'was a conceptual and mathematical realist since about 1925' (Wang 1987, 20), that is, even before he proved his Incompleteness Theorems, the gap between provability and truth that they opened up might well have pushed him even further in this direction. The analogy Gödel drew between mathematics and science certainly supports attributing to him a realist conception of mathematical entities that satisfies our characterization of mathematical platonism.[19]

23.4 The Mathematical Experience

In addition to the previous arguments, a motivation for mathematical platonism can be found in the experience of actually engaging in mathematics. In particular, a recurring question about the nature of mathematics, which appeared already in our discussion of Frege, is whether mathematics is invented or discovered. The historical development of mathematics itself, including the fact that we tend to identify important mathematical theorems with individual mathematicians, such as Pythagoras' Theorem and Cantor's Theorem, is often presented in a way that emphasizes that mathematics is made by humans and thus invented. However, everybody who has struggled at some point through a mathematical proof or realized a surprising connection between seemingly disparate mathematical concepts (e. g., between the theory of groups and the solvability of polynomial equations) can easily get the impression that mathematical results are not up to them, but impose themselves on us. The analogies between mathematics and natural science put

[19] It seems that Gödel would not have used the label 'platonism' for his view, see Urquhart (2014, 505). For more nuanced discussions of Gödel's philosophical views, see Gödel et al. (2010) and Elsby and Buldt (2019).

forward by Frege and Gödel quoted above clearly speak to such an impression, and Maddy interprets many mathematicians as considering realism to correspond to the phenomenological experience of 'actual mathematical activity' (Maddy 1996, 492).

The power of the mathematical experience of dealing with a reality that is outside of us is also expressed by many mathematicians. For example, according to Monk's 'subjective evaluation', 'the mathematical world is populated with 65% platonists, 30% formalists, and 5% intuitionists' (Monk 1976, 3); Davis et al. (2012, 377) write that 'Platonism was and is believed by (nearly) all mathematicians'; and Maher (1996, 146) suspects 'that nearly all mathematicians are, insofar as they give the matter conscious thought, unreconstructed Platonists'. Despite the fact that he considers mathematical platonism a myth and ultimately argues against a view of mathematics as certain and based on truth, Hersh acknowledges how one can easily be tempted to accept platonism:

> The basis for Platonism is the awareness we all have that the problems and concepts of mathematics exist independently of us as individuals. The zeroes of the zeta function are where they are, regardless of what I may think or know on the subject. It is then easy for me to imagine that this objectivity is given outside of human consciousness as a whole, outside of history and culture. This is the myth of Platonism. It remains alive because it corresponds to something real in the daily experience of the mathematician. (Hersh 1979, 18)

Based on related considerations, Tait speaks of realism as the 'default position' held by many nineteenth-century mathematicians, such as Cantor, Dedekind, Frege, and Hilbert (Tait 2005, 91).

23.5 Mathematical Platonism in More Recent Years

Many philosophical discussions about platonism have taken place since the work of Cantor, Frege, and Gödel, and contemporary debates involve much more fine-grained conceptual distinctions and elaborated arguments, but these are well beyond the scope of this paper. Nevertheless, in order to present a glimpse of some later discussions, let us briefly take a look at the direction of some contemporary philosophical discussions. These are usually not presented in relation to their nineteenth-century and early twentieth-century predecessors,[20] so I'd like to take the opportunity here to point out some connections.

As we've seen, some arguments for mathematical platonism were initially based on specific mathematical domains, such as transfinite set theory in the case of Cantor, arithmetic and logic in the case of Frege, and meta-mathematical results in the case of Gödel. However, the view of mathematics that emerged from these considerations was frequently extended to all of mathematics. Recall Cantor's emphasis on the freedom of mathematics in general, according to which

[20] But, see Maddy (1989) and Panza and Sereni (2013).

no restrictions should be imposed on mathematics, other than it should not lead to contradictions. Cantor formulated this requirement in terms of concepts, Hilbert in terms of the syntax of the formal reconstructions of mathematics. In recent years, a related position has been developed by Balaguer under the name of 'plenitudinous platonism' or 'full-blooded platonism', which states that '*all possible mathematical objects exist*', where 'possible' is understood in the broadest sense as logical consistency (Balaguer 1998, 5).

The relation between the objects of mathematics and those of the natural sciences has also received considerable scrutiny. On the one hand, the view that science is the ultimate arbiter of what exists (i. e., the philosophical position of *naturalism*) has led Quine and Putnam to put forward their famous *indispensability argument*. According to it, we must accept the existence of those mathematical entities that are indispensable for the formulation of our current scientific theories—and perhaps a bit more, to round it off. A major drawback of this position is that it classifies those entities that are accepted by mathematicians into those that exist and those that do not on the basis of their use in science. Consequently, as accepted scientific theories change, so would the ontological status of the mathematical entities involved. On the other hand, some philosophers consider the analogy between mathematics and science to be too strong, in particular with regard of the notions of objects and existence, and developed so-called 'lightweight' versions of platonism that rely only on 'thin' objects (Linnebo 2018b).

An influential development in philosophy of mathematics has been the shift from considering individual mathematical objects, such as the number four, towards considering mathematical *structures*, such as the natural number structure, as the fundamental entities of mathematics—this move also has its roots in the nineteenth century (Reck and Schiemer 2020). While this shift avoids some of the difficulties that alternative views have, the basic ontological and epistemological questions and difficulties remain, but a realist (or platonist) view of structures, such as *in re structuralism* (Shapiro 1997), is again very popular.

23.6 Conclusion

The aim of this paper was to present some of the motivations behind mathematical platonism and to illustrate how these emerged from philosophical reflections about mathematical work in the late nineteenth century. While the claim of an independent existence of mathematical entities might appear to be initially the most difficult to accept, we have seen that its origins do not lie with dogmatic philosophers who speculate about mathematics from their armchair, but in careful considerations about the nature of mathematics by mathematicians themselves, such as Cantor, Frege, and Gödel. The considerations that led them towards a platonist conception of mathematics are: (a) The unboundedness of mathematical entities, which makes it impossible to locate them in the physical world and in actual mental constructions; (b) a view of mathematics as being true objectively and independently of human activities, which leads to the timeless character of mathematical entities. Finally, (c) the failure of a promising alternative, namely the attempt to capture all arithmetical truths in terms of provability in a formal system. A point to notice is that all three mathematicians discussed focused their attention on mathematical *concepts*, which might also have contributed to their leaning towards mathematical platonism.[21]

What also became apparent in our discussions is that one reason why mathematical platonism remains a popular but also contested position in philosophy of mathematics is that there are some strong and initially plausible intuitions about the nature of mathematics that stand in conflict with each other. On the one hand, there is the dynamic view of mathematics as a human enterprise that has developed over time and is continuing to change. According to this view, mathematics is invented. On the other hand, there is the static view of mathematics as a body of objective and unchanging truths that are there to be discovered. As we have seen in Sect. 23.4, the latter is the view that fits more naturally to the experience of engaging with mathematics, in particular by research mathematicians. A possible way to reconcile a static conception of mathematics with the dynamic process of mathematical practice and education is to distinguish between the realm of mathematical entities as the grand framework in which all of mathematics takes place from our engagement with this realm. In fact, the accounts of Cantor, Frege, and Gödel can also be recast from this perspective: For Cantor, the Absolute infinite provides the background for all definitions of transfinite numbers; for Frege, the immutable realm of concepts is there to be grasped and explored by reason; for Gödel, the realm of mathematical concepts is there to be perceived by intuition and new definitions are to be evaluated in terms of their fruitfulness to obtain new results. Such a division of labor between what constitutes the mathematical landscape and what is involved in exploring it can also be used to reconcile philosophy and pedagogy in such a way that mathematical platonism is not opposed to teaching mathematics in a way that encourages the exploration of creative ideas.

Because the above considerations pull in opposite directions with regard to accepting the independent existence of mathematical entities, mathematical platonism is still a polarizing position, finding both committed supporters and ardent opponents. Opponents to mathematical platonism are faced with two main options: Either (i) they have to give a different account of the objectivity and truth of mathematical statements, or (ii) they have to give up these strong claims about mathematics. With regard to the first option we have seen some major difficulties that alternative accounts are faced with in the discussions of the main philosophical positions in nineteenth-century philosophy of mathematics (Sect. 23.2.3) and of Gödel's incompleteness results (Sect. 23.3.3). With regard to (ii), some philosophers have argued for mathematics as being a social construction (Ernest 1998) or for mathematical truth as being time-dependent (Grabiner 1974), but these views have remained suggestive and their implications have not been found very convincing.

[21] I am indebted to Erich Reck for this observation.

Thus, mathematical platonism still seems to enjoy the popularity that was expressed in the quotations in Sect. 23.4.

The focus of this paper was mathematical platonism and some motivations for it that emerged in the mathematical work and philosophical reflections of the late-nineteenth and early-twentieth centuries. As such, it supports Ferreirós and Gray's claim that 'many pivotal mathematical contributions' in this time period 'were not philosophically neutral' and it can be seen as a contribution to the general project towards 'an adequate historical understanding of modern mathematics' (Ferreirós and Gray 2006, 1–2).

Acknowledgments I would like to thank Michael Hallett, Erich Reck, Tabea Rohr, David Waszek, Paul Xu, and an anonymous reviewer for helpful comments on a previous version of this paper. Work on this paper was funded by Social Sciences and Humanities Research Council of Canada (SSHRC).

References

Balaguer, M. 1998. *Platonism and Anti-Platonism in Mathematics*. New York: Oxford University Press, Oxford.

Beaney, M., ed. 1997. *The Frege Reader*. Oxford: Blackwell Publishers.

Benacerraf, P., and H. Putnam, eds. 1983. *Philosophy of Mathematics – Selected readings*, 2nd ed. Englewood Cliffs: Prentice Hall.

Bernays, P. 1935. Sur le platonisme dans les mathématiques. *L'Enseignement Mathématique* 34: 52–69. English translation: On platonism in mathematics, by Charles D. Parsons in Benacerraf and Putnam (1983), pp. 258–271.

Bouveresse, J. 2005. On the Meaning of the Word 'Platonism' in the Expression 'Mathematical Platonism'. *Proceedings of the Aristotelian Society* 105: 55–79.

Cantor, G. 1883. *Grundlagen einer allgemeinen Mannigfaltigskeitslehre*. Leipzig. Translation by William Ewald in Ewald (1996), vol. 2, pp. 878–920.

Cytowic, R. E. 2018. *Synesthesia*. Cambridge: MIT Press.

Davis, M., ed. 1965. *The Undecidable. Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions*. Hewlett: Raven Press.

Davis, P. J., R. Hersh, and E. A. Marchisotto. 2012. *The Mathematical Experience, Study Edition*. Basel: Birkhäuser.

Elsby, C., and B. Buldt. 2019. Gödel – Husserl – Platonism. *Meta: Research in Hermeneutics, Phenomenology, and Practical Philosophy* 11 (2): 358–401.

Ernest, P. 1998. *Social Constructivism as a Philosophy of Mathematics*. Albany: State University of New York Press.

Ewald, W. 1996. *From Kant to Hilbert: A Source Book in Mathematics*. Oxford: Clarendon Press. Two volumes.

Ferreirós, J. 2007. *Labyrinth of Thought: A History of Set Theory and Its Role in Modern Mathematics*, 2nd rev. ed. Basel: Birkhäuser.

Ferreirós, J., and J. Gray, ed. 2006. *The Architecture of Modern Mathematics*. Oxford: Oxford University Press.

Fine, K. 1998. Cantorian Abstraction: A Reconstruction and Defense. *The Journal of Philosophy* 95 (12): 599–634.

Frege, G. 1879. *Begriffsschrift. Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle a/S.: Verlag Louis Nebert. English translation: *Begriffsschrift, A formula Language, Modeled upon that for Arithmetic*. In van Heijenoort (1967), 1–82.

Frege, G. 1884. *Die Grundlagen der Arithmetik*. Breslau: Verlag von Wilhelm Koebner. English translation by J. L. Austin: *The Foundations of Arithmetic*, Oxford: Basil Blackwell, 1953.

Frege, G. 1892. Über Begriff und Gegenstand. *Vierteljahrsschrift für wissenschaftliche Philosophie* 16: 192–205. English translation by Peter Geach in Beaney (1997), pp. 181–193.

Frege, G. 1980. *Philosophical and Mathematical Correspondence*. Chicago: University of Chicago Press. Edited by Gottfried Gabriel, Hans Hermes, Friedrich Kambartel, Christian Thiel, and Albert Veraart.

Giaquinto, M. 2002. *The Search for Certainty*. New York: Oxford University Press.

Gödel, K. 1931. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik* 38: 173–198. English translation: On formally undecidable propositions of Principia Mathematica and related systems I. In Davis (1965), 4–38.

Gödel, K. 1947. What Is Cantor's Continuum Problem? *American Mathematical Monthly* 54: 515–525. Revised and expanded in Benacerraf and Putnam (1983), 470–485.

Gödel, K., S. Feferman, C. Parsons, and S. G. Simpson. 2010. *Kurt Gödel: Essays for his Centennial*, vol. 33. Lecture Notes in Logic. New York: Cambridge University Press.

Grabiner, J. V. 1974. Is Mathematical Truth Time-Dependent? *American Mathematical Monthly* 81 (4): 354–365.

Gray, J. 1992. The Nineteenth-Century Revolution in Mathematical Ontology. In *Revolutions in Mathematics*, ed. D. Gillies, 226–248. Oxford: Claredon Press.

Gray, J. 2008. *Plato's Ghost: The Modernist Transformation of Mathematics*. Princeton: Princeton University Press.

Hallett, M. 1984. *Cantorian Set Theory and Limitations of Size*. Oxford: Claredon Press.

Heine, E. 1872. Die Elemente der Functionenlehre. *Journal für die reine und angewandte Mathematik* 74: 172–188.

Hersh, R. 1979. Some Proposals for Reviving the Philosophy of Mathematics. *Advances in Mathematics* 31: 31–50.

Landry, E. 2023. *Plato Was Not a Mathematical Platonist*. Cambridge: Cambridge University Press.

Linnebo, Ø. 2018a. Platonism in the Philosophy of Mathematics. In *The Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta, Spring 2018 ed. Metaphysics Research Lab. Stanford: Stanford University.

Linnebo, Ø. 2018b. *Thin Objects: An Abstractionist Account*. Oxford: Oxford University Press.

Maddy, P. 1989. The Roots of Contemporary Platonism. *Journal of Symbolic Logic* 54 (4): 1121–1144.

Maddy, P. 1996. Set Theoretic Naturalism. *Journal of Symbolic Logic* 61 (2): 490–514.

Maher, P. 1996. Potential Space and Mathematical Reality. In *Constructing Mathematical Knowledge: Epistemology and Mathematical Education*, ed. P. Ernest, 145–151. London: Falmer Press.

Monk, J. D. 1976. *Mathematical Logic*. New York: Springer.

Panza, M., and A. Sereni. 2013. *Plato's Problem: An Introduction to Mathematical Platonism*. New York: Palgrave Macmillan.

Parsons, C. 2008. Paul Bernays' Later Philosophy of Mathematics. In *Logic Colloquium 2005*, ed. C. Dimitracopoulos, L. Newelski, D. Normann, and J. R. Steel, vol. 28, 129–150. *Lecture Notes in Logic*. Association for Symbolic Logic, and Cambridge University Press.

Pasch, M. 1926. Die axiomatische Methode in der neueren Mathematik. *Annalen der Philosophie* 5: 241–274.

Poincaré, H. 1909. La logique de l'infini. *Revue de Métaphysique et de Morale* 17 (4): 461–482. English translation in *Mathematics and Science: Last Essays*, 45–64. New York: Dover Publications, 1963.

Reck, E. H. 2005. Frege on Numbers: Beyond the Platonist Picture. *The Harvard Review of Psychology* 13 (2): 25–40.

Reck, E. H., and G. Schiemer, ed. 2020. *The Pre-history of Mathematical Structuralism*. Oxford: Oxford University Press.

Schlimm, D. 2010. Pasch's Philosophy of Mathematics. *Review of Symbolic Logic* 3 (1): 93–118.

Schlimm, D. 2012. Mathematical Concepts and Investigative Practice. In *Scientific Concepts and Investigative Practice*, ed. F. Steinle and U. Feest, vol. 3, 127–147. Berlin Studies in Knowledge Research. Berlin: De Gruyter.

Shapiro, S. 1997. *Philosophy of Mathematics. Structure and Ontology*. Oxford: Oxford University Press.

Tait, W. 2005. *The Provenance of Pure Reason. Essays in the Philosophy of Mathematics and Its History*. New York: Oxford University Press.

Tapp, C. 2014. Absolute Infinity. A Bridge Between Mathematics and Theology? In *Foundational Adventures. Essays in Honour of Harvey M. Friedman*, ed. N. Tennant, 77–90. London: College Publications.

Urquhart, A. 2014. Russell and Gödel. *Bulletin of Symbolic Logic* 22 (4): 504–520.

van Heijenoort, J. 1967. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*. Cambridge: Harvard University Press.

Wang, H. 1987. *Reflections on Kurt Gödel*. Cambridge: MIT Press.

Chapter 24
History of Mathematics Illuminates Philosophy of Mathematics: Riemann, Weierstrass and Mathematical Understanding

Jamie Tappenden

Abstract This paper explores two respects in which a study of the history of mathematics can enrich the philosophy of mathematics. First, central concepts in the informal methodology of mathematical research—understanding, explanation, the "proper context", the correct or natural definition of a concept, etc.—can in many cases only be identified, refined and adjudicated as the practice evolves over time. Second, the distinction between mathematics and philosophy in many cases is not sharply delineated. Many paradigmatically "philosophical" goals—identifying central concepts and providing rationales for such choices, analysing concepts, establishing criteria of "rigour" etc.—arise organically within mathematical practice itself. I illustrate these observations by exploring the historical sequence beginning with the rationales of Gauss and Riemann for studying real functions in the complex numbers, with special attention to the double periodicity of elliptic functions. I illustrate the profundity of some of the methodological issues that can arise via a contrast between the Riemann and Weierstrass approaches to elliptic functions and their generalisations.

24.1 Nineteenth Century Analysis as Philosophy of Mathematics

This paper aims to convey a sense of what the philosophy of mathematics can learn from the history of mathematics, and that Jeremy Gray's work (particularly, but not exclusively, in nineteenth century complex analysis) is an especially rich source. I draw inspiration from the title theme of Gray (2009): "Nineteenth Century Analysis

J. Tappenden (✉)
Department of Philosophy, University of Michigan, Ann Arbor, MI, USA
e-mail: tappen@umich.edu

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*, Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_24

as Philosophy of Mathematics".[1] Philosophical analysis and foundations are not just conceptual flying buttresses welded to an autonomous practice of mathematics. As nineteenth century analysis examplifies, demands for conceptual analysis, debates over the proper choice of core concepts, over the nature of rigorous argument, over the relation between mathematical objects and the language used to describe them, etc. arise within mathematical practice itself, shaping ongoing mathematical investigation.[2]

The paper will unfold as follows. In Sect. 24.2, I'll lay out the objective: explore the roles of a family of vague concepts—"understanding", "explanation", "correct definition", "proper context", "appropriate technique" . . . —that collectively inform what I'll call the "informal methodology" of mathematical practice. (The modifier "informal" signifies not merely "not expressed via some formal system" but also something like "not systematically articulated" or "sometimes tacit/not fully explicit".) In Sect. 24.3 I'll briefly consider one case for orientation: Gauss' and Riemann's rationales for viewing the complex numbers as the proper environment for investigating functions of real numbers. I'll then explore the way the basic insights were developed in profoundly different directions in the study of elliptic functions by Riemann (Sect. 24.4) and Weierstrass (Sect. 24.5). These two approaches were viewed as profoundly in opposition in the second half of the nineteenth century. In Sect. 24.6 I'll consider some of the reasons for this opposition as it stood at the time, and how those reasons were refined, addressed and adjudicated in the subsequent decades.

24.2 "Explanation", "Understanding", "Proper Setting", "Fruitfulness" and Other Concepts of Informal Methodology

24.2.1 Informal Methodology

Say we have a theorem that admits of two different proofs, each equally rigorous (or at least potentially so) and drawing from logically equivalent premises. Or we have two equivalent definitions of some concept or class. It can happen that one of the proofs or definitions will be preferred over the other. What reasons might be

[1] Indeed, you could see much of the present essay as an extended reflection on the discussion of Fourier series in Gray (2009, §3).

[2] Lawrence Sklar's Locke Lectures (Sklar, 2000) make a similar observation for physics: in many cases foundational philosophy arises organically *within* physics. Of course, there are also important cases where philosophy and mathematics interact at more of a distance, so to speak, such as Riemann's immersion in *Naturphilosophie* (cf. Bottazzini and Tazzioli, 1995) and Herbart's writings in particular (Scholz, 1982), (Ferreirós, 2006), and (Laugwitz, 1999, §3.3). Gray (2008) explores a wealth of other interactions between mathematics and philosophy.

cited for that preference? Well, one might be taken to provide an "explanation", or "understanding" but not the other. One might admit of a "natural" generalization supporting proofs of other "important" theorems. One proof might be judged to be "deeper", or "more fruitful". One environment may be judged to be "more natural" or "appropriately X [where X might be "algebraic", "geometric", "analytic", ...]".[3] Preferences may be justified in broadly aesthetic terms—"beautiful", "elegant", "rich", "deep"—that may in themselves seem opaque, but in a context where more substance can be discerned. (cf. Sect. 24.3.2)

Recent decades have featured attempts to address these aspects of mathematical research.[4] To be sure, this practice-oriented research stream faces distinctive challenges, not least of which is the vagueness of the expressions it seeks to illuminate: it is unlikely that any will admit sharp necessary and sufficient conditions. Also, efforts to explain one typically end up appealing to others; it's difficult to treat just one in isolation.

To make sense of these patterns, we need to clarify how they are bound up with the practice of formulating and proving theorems. Mathematicians are not just black boxes that input premises and coffee, then output rigorous proofs of clearly stated results. They engage in what I'm calling informal methodology: making and giving reasons for conjectures, proposing and debating proof strategies, identifying potentially representative special cases, etc. en route to cogent proofs of interesting results. Talk of "explanation", "understanding", "proper context", etc. is a part of this orienting discourse. However formless this talk of "understanding" etc. may sometimes seem at first encounter, we can get some purchase on the content of what practicing mathematicians *say* by looking to what they subsequently *do*. This will not, of course, give the sharpness, clarity and unequivocal character of the definition of ring or group, nor can we expect many, if any, exceptionless generalisations. But one plays the hand that's dealt: you can only get as much exactness as the phenomena support.

24.2.2 Objectivity? Fruitfulness and Prediction

Much of the support for any investigation of this type will consist in simple judgements of practitioners: Accomplished mathematician X states a preference for Y and cites reasons Z_1, Z_2 and Z_3. Of course, any time human beings get involved, there will be differences of opinion, and indeed this paper will explore some variations that revealed themselves over the years to be quite profound. It could

[3] "...the ring and the corresponding affine scheme are equivalent objects. The scheme is, however, a more natural setting for many geometric arguments" (Eisenbud and Harris, 1992, p. 5).

[4] Early stirrings of this orientation can be found in Lakatos (1976) and the work of Philip Kitcher (for example Kitcher, 1981 and Kitcher, 1984). More recent representatives are the essays in Mancosu (2008) and Ferreirós and Gray (2006).

be, and occasionally has been suggested that such disagreements undercut the value of studying informal methodology, in some cases even suggesting dismissively that it will amount to nothing more than just each individual mathematician capriciously judging "what I like"?[5] Of course, this is not in itself an objection to studying informal methodology if the preferences themselves are systematically articulated and can be to some extent objectively characterised. Furthermore, those preferences that do exist are often themselves grounded in, or at least can be defended with, reasons that can themselves be critically evaluated. Also, there are cases where a judgement of informal methodology is unanimous or at least close to it. A core example in this paper is one such case: the consensus (or near-consensus) judgement that a family of problems in real analysis are only fully understood when framed within complex analysis. In addition, there is a further consideration, that will be a submotif of this paper: mathematics is forward-looking, and whether or not a judgement of informal methdology is ultimately tenable may depend upon matters as yet unknown.

I'll use the label "fruitfulness" for this mathematical virtue: One way for a proof, or definition, or framework, (etc.) to be superior to a competitor is for it to better facilitate further discovery.[6] Judgements of "fruitfulness", under various descriptions and labels, reflect a bedrock attitude of the culture of mathematics: hard-nosed orientation toward getting results, by finding proofs (of a sought-after type) for theorems judged to be worth proving. Of course, this is a virtue that draws upon others in the circle of informal evaluations: proofs are not sought for just any old theorems, one wants ideally to discover proofs to "important" theorems (or at least "theorems that are not a complete waste of time"), that "explain", or convey "understanding", etc. But the question of whether or not there is a proof there to be found, and whether it *can* be found, is not a matter of mathematicians' preferences.

To view a framework, concept, etc. as fruitful is to make a tacit or explicit prediction that worthwhile discoveries will be supported by that framework/concept/etc. These predictions can be correct or incorrect: perhaps the results won't come after all and the research will stagnate. This could prompt a re-evaluation—"they thought X was the right formulation 100 years ago, but now we know it doesn't work outside restricted domain Y". The concept of "scheme" was much disputed upon its introduction in the mid-twentieth century. A few decades later it became "the language of modern algebraic geometry" (following the title of Eisenbud and Harris, 1992). What changed? Many things, of course, and no doubt psychological and social factors were involved. But a decisively important factor was just that the scheme concept supported proofs of hard unsolved problems like the Weil conjectures in a way that the available alternatives didn't.

[5] "[Many mathematicians may use expressions like "explanatory"] to mean little more than "of the kind I like". And different kinds of mathematicians like different kinds of proofs" (Burgess, 2015, p. 96 fn. 22).

[6] I have more to say on this topic in Tappenden (1995, 2008, 2012).

"Fruitfulness" is a useful reference point for studying informal methodology because it offers a foothold in a kind of objectivity. I use the word "objectivity" with some trepidation, as it is itself notoriously slippery, but I have something minimal in mind: to be objective in this minimal sense is for it to be possible to make a distinction between seeming to have a property and actually having it. It can't be that to be an explanation, or a proof that conveys understanding (etc.) is *merely* for some group of qualified people to feel that it is (though such judgements provide defeasible evidence that it is). It should make sense to say things like "this sure seemed like an explanation, and still seems that way, but for reasons p,q, and r, it now appears that appearance is mistaken."

The coming sections will develop an extended example of mathematical innovation explicitly motivated by fruitfulness considerations, with attention to the way that other bits of informal methodology were subsequently deployed, drawing extensively from Jeremy Gray's work, especially Bottazzini and Gray (2013), Gray (2015) and Gray (1989). The claim will not be anything as simple as (say) "such and such is an explanation if the concepts involved are fruitful", since the discourse we're looking to clarify is not as simple as that. As we'll see, much of the value of work like Bottazzini and Gray (2013) for students of informal methodology is that the phenomena under study, and the predictions made, can take a very long time to play themselves out, in intricate and subtle ways. Sometimes it isn't until after several generations of mathematicians have finished their work that we can evaluate whether not the original evaluations were correct and spell out why.

24.3 "Hidden Harmony": The Introduction of Complex Numbers

24.3.1 Examples for Orientation: A Method or a Trick?

Sometimes we encounter an account of a phenomenon that strikes us as making everything clear and intelligible. Our immediate reaction may be to take the phenomenon to have been explained. These responses can be critically evaluated: is this *really* an explanation, or does it just appear to be? It will be useful for orientation to glance at two small-scale, limited cases.

A familar example appears in Spivak's *Calculus*, a widely used introductory textbook.[7] In the chapter on sequences and series of real numbers, it's noted that

[7] This case is also discussed in connection with philosophical studies of explanation by Steiner (1978, pp. 18–9), Lange (2010, pp. 329–32), and Skow (2015, pp. 80–2). See also Needham (1997, pp. 64–70).

the Taylor series for $f(x) = \frac{1}{1+x^2}$ has a circle of convergence $-1 < x < 1$. The discussion continues:

> If $|x| > 1$ the Taylor series does not converge at all. Why? ... What unseen obstacle prevents the Taylor series from extending past 1 and -1? Asking this sort of question is always dangerous, since we may have to settle for an unsympathetic answer: it happens because it happens – that's the way things are! In this case there does happen to be an explanation, but this explanation is impossible to give at the present time; although the question is about real numbers, it can be answered intelligently only when placed in a broader context. (Spivak, 1967, p. 428)

Two chapters later the "broader context" has been developed—*complex* power series—and we read: "it is no accident that the circle of convergence contains the two points i and $-i$ at which the function f is undefined..."(p. 470) This is a psychologically satisfying explanation that dispels any sense that "it happens [just] because it happens". And the view that this *is indeed* an explanation does seem to be universal, or nearly so, among those with the necessary background.

But is it *just* psychologically satisfying? It's instructive to contrast another solution that has given countless students a delicious "aha" moment: Poisson's evaluation of $\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$ by squaring the integral and evaluating the resulting double integral in polar coordinates. "A miracle occurs" and everything quickly falls into place. But after the sugar rush wears off the student might wonder: there are other ways to evaluate this integral, does this quick, magic way tell us anything special?[8] This use of polar coordinates doesn't appear useful for many other integrals, which could prompt the title question of Bell (2010): is this a method or a trick? (Bell's answer: so much needs to be true of the equation for the Poisson calculation to work that it has essentially just this one application. So: a trick.)

In the Spivak explanation, there is similarly a passing of the burden to one's attitude toward $\mathbb{C}$: why should what happens in the $\mathbb{C}$-world explain what happens in the $\mathbb{R}$-world? We could certainly slap together any number of bizarre structures containing an object o, stipulating that $\frac{1}{1+o^2}$ is undefined and at a distance 1 from the origin. We wouldn't count such structures as explaining anything. So the question "Is this an explanation?" is partly dependent on the question "What makes $\mathbb{C}$ special? What reasons can be given for viewing the specific broader context of $\mathbb{C}$ as the "right" one within which to address the convergence of $f(x) = \frac{1}{1+x^2}$? And why should mathematicians write and choose textbooks that try to inculcate students with the instincts, intuitions and habits of thinking that prompt that reaction?

The consensus attitude of the mathematical community appears to be that studying the behaviour of a function in $\mathbb{C}$ is indeed indispensible for understanding its behaviour in $\mathbb{R}$. It's a method, not a trick. There are many reasons for this. Some of them might be seen as simple matters of convenience—for example,

[8] For an extensive collection of other ways, see Boros and Moll (2004, ch. 8). See also Iwasawa (2009).

(Fundamental Theorem of Algebra) it is quite handy that polynomials with complex coefficients always split into linear factors over $\mathbb{C}$. But other reasons appear deeper. It will be illuminating to reach back in history to the initial motivations for giving $\mathbb{C}$ such a distinguished status, and follow out the consequences.

24.3.2 Gauss and Riemann: "Beauty", "Simplicity", "Roundness", "Hidden Harmony"…

An early declaration for complex numbers collectively as the correct context for the study of functions of analysis comes from Gauss.[9] He first asks whether a new function is to be applied only to real numbers "[with] the imaginary values of the argument only an appendage", or …

> …[one] accedes to my principle that in the realm of magnitude one must regard the imaginary $a + \sqrt{-b} = a + bi$ as enjoying equal rights with the real. We are not talking about practical use here; rather to me analysis is an independent science, which by slighting imaginary magnitudes loses exceptionally in beauty and roundness [*Schönheit und Rundung*], and suddenly every truth that would otherwise be generally valid would have to be accompanied by highly tiresome restrictions.[10]

In 1832 Gauss made similar remarks about the "highest simplicity and natural beauty" emerging when arithmetic is extended to include $\sqrt{-1}$.[11]

What is Gauss gesturing at? Of course, it's far from clear from the words alone what the significance of "Beauty and roundness"/"highest simplicity and natural beauty" is for proof construction. I would guess that by "roundness" Gauss means something that could be described as "intellectual/thematic cohesion", since without this roundness, "every truth that would otherwise be generally valid would have to be accompanied by highly tiresome restrictions", but this interpretation is hardly forced. A lot remains to be filled in.

Gauss acknowledges that he is groping and probing, striving for a rigorous, explicit articulation of what he grasps only dimly. Here is a not atypical remark, from a letter of 1825, indicating both his evaluation of the importance of his speculations for research and the state of those speculations as work in progress:

> These investigations [into curved surfaces] penetrate deeply into many others, I may even say into the metaphysics of space, …for example the true metaphysics of negative and

[9] The point is not just that complex numbers should be used, but that they are collectively treated as an object of study in their own right, and as supporting rigorous proof, in contrast to (say) Poisson, who saw them as a tool for discovery but not rigorous proof (Bottazzini and Gray, 2013, p. 127.).

[10] Letter of 1811 to Bessel (Gauss, 1880, p. 156). Translation (slightly modified) from Bottazzini and Gray (2013, p. 72).

[11] Gauss (1832, p. 102). These and similar passages from Gauss are discussed in Bottazzini and Gray (2013) section 1.5.3 and *passim*.

> imaginary quantities. The true meaning of $\sqrt{-1}$ stands very vividly before my mind, but it will be very difficult to put it into words, which can only give but a vague fleeing image.[12]

We can be confident that Gauss isn't *just* making a disinterested aesthetic evaluation. He is also making a methodological recommendation: This is the context in which certain problems should be addressed. Gauss sees something, he urges its importance, and it's safe to say, given that this is Gauss we are talking about, that he expects at least some of the eventual payoff will be proofs of hard problems. But to hazard informed conjectures about what he's trying to express, we'll need to dig into the details of the mathematics itself.

Here is a passage a few decades later in the tradition Gauss initiated, opening article 20 of (Gauss student) Riemann's epochal PhD thesis (Riemann, 1851) on complex functions (containing "the germ" of "a major part of the modern theory of analytic functions" (Ahlfors)).[13] Riemann notes that for functions of real analysis, after extending the domain to $\mathbb{C}$ "an otherwise hidden enduring harmony and regularity [*Harmonie und Regelmässigkeit*] emerges", providing coherence to already known theorems, and "open[ing] the road" to further discoveries.

> If we give these [functions of analysis] an extended range by assigning complex values ... an otherwise hidden harmony and regularity emerges... [A]lmost every step taken here has not only given a simpler, more compact *Gestalt* to results established without the help of complex magnitudes, but opened a road to new discoveries, as attested by the history of the investigations into algebraic functions, circular and exponential functions, elliptic and Abelian functions. (Riemann, 1851, p. 69–70; 34–5)[14]

Riemann explicitly indicates some research directions fertilized by this "harmony and regularity". (I'll follow out the elliptic functions reference in this paper.) He points to both "new discoveries" and new proofs of known results, evaluating these new proofs as preferable in qualitative respects ("a simpler, more compact *Gestalt*"). In fact, this promise of fruitfulness was realised in many different ways, as I'll discuss in Sect. 24.3.3 and *passim* through the rest of the paper.

Gauss and Riemann are not just saying that their constructions have certain aesthetic properties, and that by lucky happenstance they are also fruitful, but rather that whatever it is that grounds the aesthetic reaction ("beautiful", "harmonious", ...) is also at least partly responsible for the fruitfulness, suggesting a picture of this particular flavour of mathematical beauty as a kind of "productive richness".[15]

---

[12] Gauss (1927, p. 8). Johnson (1979) drew my attention to this letter; the paper (pp. 112–14) has additional quotes and more extended discussion of Gauss's efforts to articulate the foundations of $\sqrt{-1}$ including the use of geometric imagery *faute de mieux*.

[13] Ahlfors (1953, p. 3). This passage is quoted more extensively in footnote 36. For a discussion of Riemann's thesis in its historical context, see Bottazzini and Gray (2013, pp. 263–76) and Laugwitz (1999, pp. 96–123).

[14] I've modified the translation in Riemann (2004), drawing partly on Abe Schenitzer's translation in Laugwitz (1999, p. 97).

[15] There is precedent for such an analysis of aesthetic judgement in Hutcheson (1725). I have some tentative remarks about fruitfulness and aesthetic judgements of productive richness in Tappenden (2012)

With over $1\frac{1}{2}$ centuries of mathematics to survey, there is virtually no dispute among mathematicians that Gauss and Riemann were right to view $\mathbb{C}$ as they did. No doubt they *liked* working in $\mathbb{C}$, but the preference was grounded in reasons that we can retrospectively evaluate as deep and well-founded. This in turn helps explicate Spivak's assessment of an "explanation" of the convergence of the Taylor series for $f(x) = \frac{1}{1+x^2}$: Behaviour over $\mathbb{C}$ is consistently, systematically, and profoundly relevant to our understanding of functions over $\mathbb{R}$. It's not a trick.

24.3.3 More Riemann: Concepts Versus Calculation, Objects Versus "Manner of Representation"

Continuing with the Gray (2009) theme of nineteenth century analysis as philosophy of mathematics, I'll note that Riemann (1851) not only introduced new definitions and proof techniques but also inaugurated a broader debate about the *nature* of mathematical investigation, between a "computational" approach associated with the Berlin of Kummer, Kronecker and Weierstrass, and another, described by its adherents as "conceptual", striving to have proofs "compelled not by calculations but by thought alone."[16] The latter was particularly associated with Göttingen and Riemann, tracing back to Gauss and Dirichlet, and forward through Dedekind, Hilbert and others. I and others have discussed the "conceptual"/"computational" debate elsewhere, so I'll refer the reader there for more details.[17] In the coming pages I'll take up just what is needed as background to the discussion of elliptic functions.

Dedekind, in his *Lebenslauf* of Riemann (1876, p. 576), relates that in the late 1840s the young Riemann attended Eisenstein's lectures on elliptic functions, and they engaged in many discussions. Riemann conveyed that the two fundamentally disagreed on principles, since Eisenstein "stopped with formal calculation" while Riemann held that partial differential equations gave the "essential (*wesentliche*) definition" of the targeted functions.[18] Dedekind suggests this as the moment when

[16] For example: "I have tried to avoid Kummer's elaborate computational machinery so that here too Riemann's principle may be realized and the proofs compelled not by calculations but by thought alone." Hilbert (1998, p. X).

[17] This "conceptual" dimension of Riemann's approach is a theme running throughout Laugwitz (1999). I discuss Riemann's reorientation in a philosophical context in Tappenden (2006) and more briefly in Tappenden (2013, §9.3). A useful brief historical overview of the rivalry is Rowe (2000). For details and philosophical discussion of Dedekind's Riemann-inspired methodology see Avigad (2006) and Tappenden (2008). Some of the antecedents of the "conceptual" interpretation in Gauss's work are explored in Ferreirós (2007).

[18] This does not mean that Riemann didn't carry out virtuoso calculations when needed; cf. Siegel (1932, esp. p. 771).

Riemann began developing ideas that were "decisive for his entire later career".[19] Some years later, Riemann refers back to §20 of his thesis as conveying "the principle" of his "method" which "yields all the earlier results nearly without calculation" (Riemann, 1857b, p. 85). In Riemann (857c, p. 99) he asserts a variation: his method produced improved proofs "almost directly from the definition" of results which were previously established with "tiresome calculations".

Riemann also approached mathematical objects as language-independent. Just before §20 of his thesis, Riemann wrote that his principles "open the way for the investigation of specific functions of a complex variable, independently of expressions for the functions" (Riemann, 1851, §19). At one point in Riemann (1857a, p. 90) he makes explicit that he means to reject definitions based on power series. He notes that such a definition would be possible, but adds: "However, it seems inappropriate to express properties independent of the mode of representation, by criteria based on a particular expression for the function."[20] Instead, as he did in his discussions with Eisenstein, Riemann fixes the object of investigation—the complex-differentiable functions—via what are called the Cauchy-Riemann equations, conditions on the functions themselves.[21]

This emphasis on objects themselves rather than modes of representation underwrites one of the strategies making possible Riemann's "simpler, more compact *Gestalt*": determining *specific* functions in a way that reduces the information required "to the necessary minimum", "avoiding superfluous components":[22]

> Previous methods of treating these functions always set down as a definition an expression for the function, giving its value for every argument. Our investigation shows that, due to the general character of a function of a variable complex quantity, in such a determination parts of the determination are consequences of the others [so the information needed can be reduced to what is necessary]. This significantly simplifies the treatment. For example, in order to show the equality of two expressions for the same function, one usually needed to transform one into the other, i.e. one had to show they agreed for every value ... now it is sufficient to establish their agreement for a much smaller [set of arguments] (Riemann, 1851, §20)

That is: problems are crucially simplified because often limited information suffices to determine a specific function. For example: Where is the function zero? Where is it infinite, and in what way? As just one illustration of this broad phenomenon, I'll use the elliptic functions we'll consider in Sect. 24.4. As we'll

[19] The Eisenstein calculation-based treatment resembled in important respects the work of Weierstrass that I'll consider below. Weil (1976) explores the power of the Eisenstein account. Bottazzini and Gray (2013, §4.2.3.2) puts it in historical context.

[20] I discuss this feature of Riemann's view, with particular emphasis on Frege's conception of mathematics as dealing with objects rather than symbols, in Tappenden (2005a, 2006), and in §10.6 of Tappenden (2019).

[21] Riemann's definition and the Weierstrass series-based definition I'll discuss in Sect. 24.5 are essentially equivalent, though it would be some decades before that was proven. The qualifier "essentially" is necessary here; note for example the Weierstrass objection in footnote 50.

[22] Riemann (1851, p. 41). The quoted words are from the description of §20 in the table of contents.

see in 24.5 such a function can fruitfully be defined with a series, in a way that "gives its value for every argument". On the other hand, in line with the Riemann approach, it can be shown that, after fixing a background structure (the "period lattice"), if two elliptic functions Φ_1 and Φ_2 have zeros and poles of the same order at the same points, then there is a non-zero constant $c \in \mathbb{C}$ such that for every $z \in \mathbb{C}$, $c \cdot \Phi_1(z) = \Phi_2(z)$.[23] The functions differing by a constant multiple are, from a mathematical point of view, essentially identical, so simply knowing the information about the zeros and poles suffices to determine the whole elliptic function.

Dedekind and other followers of Riemann spoke of this approach as "defining functions through essential characteristic properties", in one case tracing the idea back to Gauss, and reinforcing the language-independence of the approach:

> [Gauss remarks in the *Disquisitiones Arithmeticae*]: 'But neither [Waring nor Wilson] was able to prove the theorem, and Waring confessed that the demonstration was made more difficult by the fact that no notation can be devised to express a prime number. But in our opinion truths of this kind ought to be drawn out of notions not out of notations.' In these last words lies, if they are taken in the most general sense, the statement of a great scientific thought: the decision for the internal in contrast to the external. This contrast also recurs in mathematics in almost all areas; [For example] ... Riemann's definition of functions through internal characteristic properties, from which the external forms of representation flow with necessity (Dedekind, 1895, pp. 54–55).[24]

To get a more robust sense of the consequences of Riemann's informal methodology we'll need to look to the mathematics that was guided by it. In the coming Sect. 24.4 I'll continue following out Riemann's reference to elliptic functions.

24.4 Foothold: The Double Periodicity of Elliptic Functions, Downstream from Riemann (1851)[25]

With the hindsight of more than one and a half centuries, Riemann's tacit prediction that his approach would prove fruitful has been borne out. And though there have been refinements of definitions, more rigour, and various bits of adjusting, clarifying and error-correcting, Riemann's core insights retain their force. Here I'll consider one case that exemplifies this power to not only support proofs but convey

[23] Jones and Singerman (1987, p. 76).

[24] I discuss the "inner characteristic properties" versus "external forms of representation" theme, with other quotes from Dedekind and others, in Tappenden (2006, esp §II). A more systematic examination of this and other aspects of Dedekind's methodology is Avigad (2006). A more recent paper, which connects Dedekind's interpretation of Riemann's methods with contemporary "structuralism" in mathematics is Ferreirós and Reck (2020).

[25] In this section I'm greatly indebted to the discussion in Stillwell (1989, esp. §14.4, §15).

understanding of a question confronted at roughly the time of Riemann's writing: why do elliptic functions have two periods?[26]

24.4.1 Riemann Surfaces Invoked as Explanations

Here are some expressions of a widespread opinion, citing a construction originating in Riemann (1851) as explanatory:

> People who know only the happy ending of the story can hardly imagine the state of affairs in complex analysis around 1850. The field of elliptic functions had grown rapidly for a quarter of a century, although *their most fundamental property*, double periodicity, *had not been properly understood*; it had been discovered by Abel and Jacobi as an algebraic curiosity rather than a topological necessity. The more the field expanded, the more was algorithmic skill required to compensate *for the lack of fundamental understanding*. [Cauchy even came] to understand the periods of elliptic and hyperelliptic integrals, although not *the reason for their existence*. There was one thing he lacked: Riemann surfaces. (Freudenthal, 1975, p. 449) (my italics)
>
> *The reason for the double periodicity* becomes completely clear with the introduction of Riemann surfaces by Riemann [in Riemann (1851)]...
> One could not ask for *a more satisfying explanation* of the mysteries of elliptic integrals. It is as if an analyst innocent of geometry had begun with
>
> $$\int \frac{dx}{\sqrt{1-x^2}}$$
>
> discovered sin and cos, and finally discovered the circle.
> (Stillwell "Introduction" to Poincaré, 1985, p. 6) (my italics)
>
> This intuitive *explanation* of double periodicity is due to Riemann [in his 1851 thesis], who later [1857] developed the theory of elliptic functions from this standpoint. (Stillwell, 1989, p. 227) (my italics)
>
> If we take the integral [$y = \int_0^z \frac{dt}{\sqrt{(t-\alpha)(t-\beta)(t-\gamma)(t-\delta)}}$] as the starting point, then Riemann's construction of the Riemann surface corresponding to (6.85) *explains* the double periodicity of the inverse [= elliptic JT] function. This in turn provides a complete answer to Eisenstein's question about the inverse of an integral and its periodicity. (Roy, 2017, p. 182) (my italics)

In each case we find Riemann's approach credited with "explaining" the double periodicity; indeed we "could not ask for a more satisfying explanation". He discovered "the reason" for the double periodicity of elliptic functions, a "topological necessity" not an "algebraic curiosity". Only with Riemann surfaces did we attain "fundamental understanding" of this "fundamental property" of elliptic functions.[27]

OK, so what are elliptic functions, and what is double periodicity?

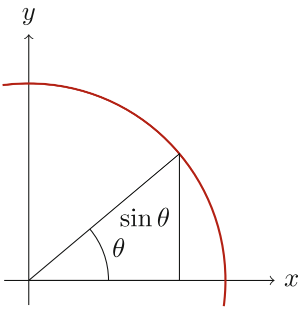
---

[26] Note Bottazzini and Gray (2013, p. 70).

[27] In connection with Sect. 24.5 below, note "Eisenstein was also pleased that his approach showed clearly why the elliptic functions are doubly periodic." (Bottazzini and Gray, 2013, p. 227).

24.4.2 An Aside for Orientation: $sin(x)$ Is Periodic

In Sect. 24.4.1 Stillwell makes an analogy with a familiar fact about trigonometric functions that will be useful for clarification. $Sin(x)$ is (singly) *periodic* (with period 2π), in that for arbitrary real x and integer n: $sin(x) = sin(x + 2n\pi)$. Geometric explanation: form a right triangle with x as one of its non-right angles. Setting the hypotenuse = 1 (and temporarily changing the argument variable to θ to avoid ambiguity), $sin(\theta)$ is the length of the line opposite angle θ. With hypotenuse as the radius of a circle, there's a familiar picture:



So why is $sin(x)$ periodic? Because as x increases you just go around a circle! Stillwell also points to this equation for the *inverse*:

$$\sin^{-1}(x) = \int_0^x \frac{dx}{\sqrt{1-x^2}}$$

(Taking the inverse of $\sin^{-1}(x)$ gives $\sin(x)$ back.)

24.4.3 Elliptic Integrals on a Torus

I'll approach elliptic functions in a nineteenth-century way, via the inverse of what can be seen as a generalization of the sin^{-1} integral: *elliptic integrals*.[28] It will suffice for our purposes to take the integral from the Roy quote above as a representative. Notationally it's convenient to ask how $\Phi(z)$ behaves, with:

$$\Phi^{-1}(z) = \int_0^z \frac{dt}{\sqrt{(t-\alpha)(t-\beta)(t-\gamma)(t-\delta)}}$$

[28] In contemporary textbooks, elliptic functions are often just *defined* as doubly periodic meromorphic functions of a complex variable. See for example Jones and Singerman (1987, p. 72) or Markushevich (1987, p. 138).

As with trigonometric functions, the behaviour of the inverse of Φ^{-1} is in key ways more tractable than the integral. When $\Phi(z)$ is obtained by inverting an elliptic integral, it is called an *elliptic function*. As advertised, elliptic functions are doubly periodic: given an elliptic $\Phi(z)$, there will be two independent $\omega_1, \omega_2 \in \mathbb{C}$ such that $\forall z \in \mathbb{C}, \forall m, n \in \mathbb{N}$:

$\Phi(z) = \Phi(z + m\omega_1 + n\omega_2)$.[29]

Why? With limited space I'll just sketch enough of the argument to convey its conceptual significance. (Much of the action in §§1–4 of Bottazzini and Gray, 2013 is the clicking into place of components of this explanation, with the denoument in §5.)

For the present purposes, we'll just think vaguely of $y = \int_{z_1}^{z} \Phi(t)dt$ as (with qualification) the length of a path extending from z_1 to z, and $y = \int_{\mathcal{C}} \Phi(t)dt$ as the length along the curve $\mathcal{C}$. I say "with qualification" because the "length" of a line in complex analysis diverges from geometric intuition. So long as nothing funny occurs, the "length" of the curve will depend only on its beginning and end points, not the specific path taken.[30] If a curve returns upon itself, starting and ending at one specific point z_a, the value of the integral—the "length of the curve"—will be 0. One way for "something funny" to occur is if the curve starts and ends at z_0, but doesn't divide the surface into distinct areas: inside, outside, and the boundary consisting of the curve itself. In that case, the integral can be non-zero.

The Riemann approach replants the original integral on a new surface, restructuring the problem in different, ultimately topological terms. This is valuable for many reasons; I'll just note a simple one. If $z^2 = y$ is understood as presenting y as a function of z it is unambiguous: each z corresponds to just one y. But the inverse, with z a "function" of y, assigns every y but 0 *two* values: $+\sqrt{y}$ and $-\sqrt{y}$. Riemann found it fruitful to turn such multifunctions into functions (strictly speaking), by dividing the plane into copies ("sheets"), with distinct values of a given argument on distinct sheets. It's helpful to compactify a sheet by adding a single point at infinity, "rolling the plane up" into a sphere (the "Riemann sphere").

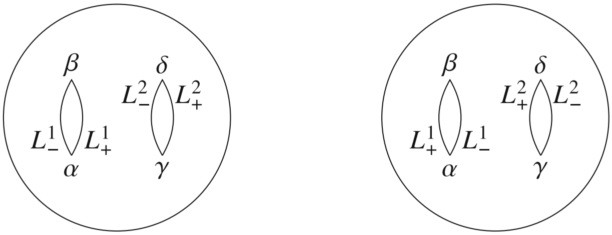
[29] An aside: This is already a point where the choice of the complex numbers as the domain is theoretically crucial. The inverted function makes sense if restricted to the real numbers, but the fundamental property of double periodicity only appears if the function ranges over all the complex numbers. Except in trivial cases, one of the periods must be a complex, non-real number and the other a real number. This illustrates a point that is often neglected in connection with the subject of ontology in mathematics: extending a domain can change which classifications of objects in the original domain count as "natural".

[30] This "length" is more formally the net effect of a complex velocity along the whole curve, where complex velocities combine differently than real velocities do (so that visually very different paths can all have the same "length" in this sense of same net effect). I'm grateful to Colin McClarty here.

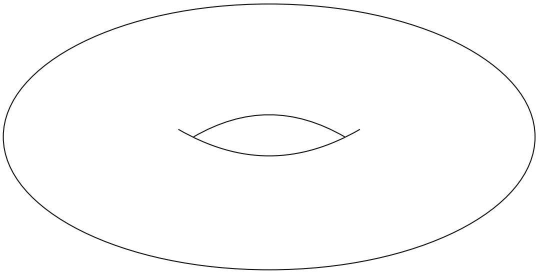
In the case of $z = \pm\sqrt{y}$ there will be two sheets, for $+\sqrt{y}$ and $-\sqrt{y}$. An integral with $\sqrt{(t-\alpha)(t-\beta)(t-\gamma)(t-\delta)}$ in the denominator needs two sheets, because of the $\sqrt{\ }$.[31]

There is a 0 on each sheet, but $0 = +\sqrt{0} = -\sqrt{0}$, so there should only be one 0. Also, there are continuous curves beginning on one sheet and ending on another. So the sheets must be "glued together". For reasons I won't go into, this has to be done in very specific ways.[32] For the elliptic integral, because the polynomial has zeros at $\alpha, \beta, \gamma, \delta$ there needs to be four points of contact when gluing the two spheres together.

Draw lines from α to β and γ to δ on each, and paste the spheres together at these cuts. α to β (L^1_- and L^1_+) and from γ to δ (L^2_- and L^2_+) to reflect that certain transitions are sensitive to the direction of a curve relative to the zeros.[33]



Connect the lines so that L^1_- on one sheet matches up with L^1_+ on the other and similarly each L^2_- on one sheet matches up with L^2_+ on the other. The result, topologically, is a torus/donut:



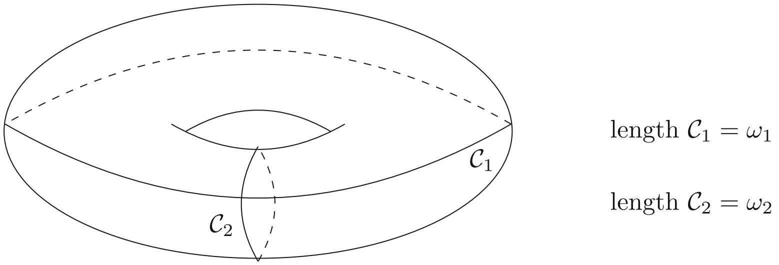
OK, so what about that elliptic integral? Recall that the integral will (very vaguely speaking) give the length of a path, sensitive only to the beginning and

[31] On the construction of the Riemann surface for $\frac{1}{p(z)}$, with $p(z)$ a polynomial, see Jones and Singerman (1987, pp. 157–67).

[32] Wilson (2006, pp. 312–19) discusses some of the relevant complexities in a philosophical context.

[33] I am of course hopscotching across many complications. For more detail see Jones and Singerman (1987, pp. 149–72 esp. p. 161) or Springer (1981, pp. 1–10).

the end points, except for curves that return onto themselves but don't divide the surface into inside/outside/boundary. On a torus (relative to a choice of coordinates) there is a particularly natural way to make an independent pair of such curves, corresponding to cycles $\mathcal{C}_1$ and $\mathcal{C}_2$:



Say that the value of the integral is ω_1 along the cycle $\mathcal{C}_1$, and ω_2 along $\mathcal{C}_2$. What is the value of the elliptic function $\Phi(k)$? Consider the inverse $\Phi^{-1}(z) = k$. Say that k is the length of a direct path from 0 to z_1. The same beginning and end are also the result of integrating the path from 0 to z_1 that goes along $\mathcal{C}_1$ m times, $\mathcal{C}_2$ n times and then to z_1. In terms of Φ, we get:

$\Phi(k) = \Phi(k + m\omega_1 + n\omega_2)$ for arbitrary integers m and n.

That is the explanation of why elliptic functions have precisely two periods. More than 150 years after its introduction, it remains a compelling diagnosis. It illustrates what Riemann meant when he spoke of establishing, “practically without computing” results that had required “tiresome calculations”, and of establishing properties of the function without reference to the means of expression.[34] Though Riemann's informal methodology was stated explicitly in an only somewhat formless way, his subsequent practice allows us to fill in many details of what he had in mind.

24.4.4 Riemann Surfaces, Fruitfulness and Informal Methodology

As the language of the quotes opening Sect. 24.4.1 indicates, the presentation of double periodicity via Riemann surfaces prompts, for many of us, a feeling that the phenomenon has been explained, and that in virtue of this explanation we understand what is going on.[35] It does happen that many people feel something

[34] The main reference point for Riemann's remarks were a specific cluster of results he had derived on hypergeometric series. I'm taking his words to be intended more broadly.

[35] No doubt some of the psychological force arises from the fact that this Riemann surface can be *visualised*. It's an issue worth study, and visualisation in mathematics has been the topic of excellent research in recent decades, but I'll set it aside in this paper. Giaquinto (2020) is an

compelling, something viscerally "right", about the explanation of double periodicity via Riemann surfaces, as the quotes in Sect. 24.4.1 reveal. In what ways can we see this as an objective evaluation of the mathematical facts themselves?

There is, of course, a lot to say, so I'll keep focus with the benchmark comparison from Sect. 24.3.1: why does behaviour of a function in $\mathbb{C}$ explain its behaviour in $\mathbb{R}$, rather than just provide a trick that conveys the illusion of understanding? At least part of an answer comes from a broader question of what is special about $\mathbb{C}$ for mathematical investigation and insight more broadly. To answer this you need to clarify the sorts of things that Gauss and Riemann gesture at in Sect. 24.3.2 when they write of beauty, roundness, harmony ... That preliminary gesturing has been given substance, not just by Gauss and Riemann's work, but also by the nearly two centuries of discovery that ensued.

This explanation of double periodicity presents the same situation: at least part of the robust value of the explanation is grounded in the status Riemann's framework overall has earned. Were Riemann's predictions correct? Did the Riemann approach, and in particular the concept of Riemann surface, prove fruitful, support further discoveries and subsequently verified conjectures, etc.? Not only is the answer yes, but it is hard to find superlatives adequate to express just how yes the answer is. It revolutionised, and remains indispensable to, not only complex analysis but a diverse range of other areas of mathematics, many not obviously connected to complex analysis.[36] From the point of view of its role in contemporary mathematics, there are compelling reasons to view "Riemann Surface" as a basic and fundamental concept. I'll return to this point in Sect. 24.6.2, after setting a point of contrast.

24.5 Contrast: Weierstrass and the $\wp$-Function

It will be illuminating to consider the contrasting virtues of the Weierstrass approach. In earlier writing (for example Tappenden, 2005b and Tappenden, 2006) my enthusiasm for Riemann was combined with a (rightly criticised) failure to do justice to the complementary power of Weierstrass's work.[37] Here I'll take steps to rectify this imbalance.

---

excellent survey of this work. The methodological significance of visualisation is complicated for reasons I discuss in Tappenden (2005b) (see especially Sect. 2.3).

[36] Speaking on the 100th anniversary of Riemann's thesis, Ahlfors wrote: "Very few mathematical papers have exercised an influence on the later development of mathematics which is comparable to the stimulus received from Riemann's dissertation. It contains the germ to a major part of the modern theory of analytic functions, it initiated the systematic study of topology, it revolutionized algebraic geometry, and it paved the way for Riemann's own approach to differential geometry." (Ahlfors, 1953, p. 3). The list could easily be extended: Lie groups, algebraic number theory, ...(Farkas and Kra, 1992, p. 1).

[37] See for example Fillion (2019). Jeremy Gray, Stephen Menn and Philip Kitcher were especially helpful in coaxing me to a more balanced point of view.

For Weierstrass, the basic objects of complex analysis are *analytic functions*: those that can be represented by a convergent power series $f(z) = \Sigma_{n=0}^{\infty} a_n (z-c)^n$ (with the a_i's and c complex numbers). Here is a series-based exploration of the double periodicity of elliptic functions, drawing on central Weierstrass concepts, methods and results.

Given two periods ω_1 and ω_2 with non-real ratio, we can define what is called the Weierstrass $\wp$—function for ω_1 and ω_2:

$$\wp(z) = \frac{1}{z^2} + \Sigma_{m,n \neq 0}[\frac{1}{(z - m\omega_1 - n\omega_2)^2} + \frac{1}{(m\omega_1 - n\omega_2)^2}]$$

This series can be seen to be doubly periodic "by inspection", since adding integer multiples of ω_1 or ω_2 just shifts everything down.[38] The series is uniformly convergent, so its derivative $\wp'$ exists. $\wp'$ is similarly obviously doubly periodic with the same periods as $\wp(z)$. Weierstrass produced from ω_1, ω_2 an integral whose inversion yields the $\wp$-function with periods ω_1, ω_2, so $\wp(z)$ is elliptic by the definition used here. Quite remarkably, it can be shown that *every* elliptic function $\Phi(z)$ with periods ω_1 and ω_2 can be given a simple representation in terms of the $\wp(z)$ and $\wp'(z)$ with those periods. There will be two rational functions (i.e. ratios of polynomials) R_1 and R_2 such that:

$$\Phi(z) = R_1(\wp(z)) + R_2(\wp(z))\wp'(z)$$

This representation preserves double periodicity. So: any elliptic function $\Phi(z)$ must be doubly periodic, because it admits of a double-periodicity preserving representation in terms of doubly periodic functions.

Is this a compelling account, and perhaps even an explanation? For those of us without a specialist's knowledge it lacks the immediacy of the Riemann surface explanation, but deeper immersion in the details could change that. What could support a case that this shows that double periodicity "didn't just happen", that we have an explanation and not a trick? Are there convincing reasons, beyond just the representation theorem, for regarding the $\wp$ function as non-arbitrary or in some sense "natural"? Does this function show up here by accident, or is there a reason?

As it happens, $\wp(z)$ is a remarkable function indeed, with many unexpected ramifications. I'll footnote sources for some basic points, and note a more profound connection in Sect. 24.6.3.[39] Furthermore, the $\wp$ function has turned out to be

[38] This conjures up a visual analogy with the periodicity of the sine function, different from the "go around the circle" one. Why is sine periodic with period 2π? Because adding 2π just shifts the sine curve to the right in such a way that the curve remains unchanged.

[39] cf. McKean and Moll (1997, section 2.8 (pp. 84–87)) or Stein and Shakarchi (2003, pp. 266–73) for clear explanations of why the series is the simplest and most reasonable one to pick, given what work the $\wp$ function is meant to do. For a variety of illustrations of the ubiquity of the $\wp$-function, see McKean and Moll (1997, pp. 84–104).

exceptionally fruitful, often in connections far from what a nineteenth century mathematician could envision in more than (at best) an indistinct way.[40]

24.6 Background: Riemann and Weierstrass Styles of Reasoning

24.6.1 Weierstrass Versus Riemann: Points of Debate

Say we ask naïvely: Is one of these an explanation and the other not? Or are both of them explanations, but one is not a *good* explanation? Or are both of them good, but one is *better*? With more than a century and a half of hindsight, and guided by our fruitfulness condition we can say that these questions are misleadingly simple: both are good—indeed profound—in different respects and for importantly different reasons. Two illustrations of many: The Riemann approach unifies many apparently profoundly distinct branches of mathematics,[41] while series-based approaches require minimal machinery.[42] What appears to be a consensus view is expressed by Catanese, commenting on "classical functions of a real variable are truly understood only after one extends their definition to the complex domain":

> ...the great appeal of [Complex] Function Theory rests on the ...variety of different methods and perspectives: polynomials, power series, analysis, and geometry, all of these illustrate several facets of the theory of holomorphic functions, complex differentiability, analyticity (local representation through a power series), conformality (Catanese, 2016, p. 196).[43]

This isn't to say that one can never make distinctions of better and worse, good and not-so-good, in connection with candidate mathematical explanations. One illustration (explored in Mancosu, 2001, pp. 108–112) is Weierstrass follower Pringsheim's explicitly expressed desire—in his 1925 *Vorlesungen über Funktionenslehre*—to "explain" phenomena in complex analysis. It is valuable work, but as Mancosu notes, it hasn't led anywhere, suggesting a negative assessment of the explanatory project.[44] Whatever comparisons could be made between Weierstrass'

[40] For example, avoiding integration proved useful for an extensive range of number types, such as p-adics (Roy, 2017, p. 183).

[41] Farkas and Kra (2001, p. XV), (Farkas and Kra, 1992, p. 1), (Napier and Ramachandran, 2011, p. vii).

[42] Walker (1996, p. xv). Walker also notes that the Eisenstein series-based treatment displays on its face the analogy with trigonometric functions. See also Weil (1976).

[43] See also Weyl (1995a) and Weyl (1995b) (a two part essay), developing the suggestion that algebraic and topological approaches are two fundamentally different and complementary modes of mathematical thinking, with Riemann and Weierstrass on complex analysis as the central reference point.

[44] This point is also made in Remmert (1991, pp. 351–2 and 431).

and Riemann's accounts, one may not unreasonably say that both are more explanatory than Pringsheim's.

The contrast of frameworks and styles did seem to be of great moment at the time, and we no longer see it with the same urgency largely because the issues that were raised came to be resolved, or in some cases reconceived, through mathematical investigation and methodological reflection. Here is one example: an objection raised by Weierstrass was the problem of extending complex variable theory from functions of *one* variable to functions of *several*. Among the reasons cited: the apparent intractability of the problem of generalizing the Cauchy-Riemann equations to several variables, and the apparent complexity of the task of generalizing Riemann surfaces to multiple dimensions.[45] Not just to many in the nineteenth century, but for decades later the case seemed compelling: even as late as 1966 Siegel describes the Weierstrass power series approach as stronger than Riemann's when the subject is functions of several variables.[46] But by 1996 we find an analyst stating a opinion roughly opposite to Siegel's (Laugwitz, 1999, pp. 148–52): "A century later we see that ... the general function theory of several variables must be viewed as a further development of Riemann's conception." Riemannian approaches to functions of several complex variables are now thriving, thanks to, among other innovations, Hörmander's work in the 1960's on the $\bar{\partial}$-equation, generalizing the Cauchy-Riemann equations, and the development of Riemann surfaces in n dimensions.[47] In this instance, the accepted view as to which approach was superior in just this one respect required over a century of research to reevaluate.

It will be illuminating to take up some of the objections raised to Riemann's methods, to see how they unfolded in the subsequent century and a half. I'll just take up two more topics briefly; the evolution of the understanding of this material was so rich and subtle that I can only cover a small part.

24.6.2 Rigour

Weierstrass acknowledged the fruitfulness of Riemann's framework while insisting that this is insufficient without an establishment of "systematic foundations".[48] Poincaré echoed this evaluation: "... the method of Riemann is above all one of

[45] cf. Weierstrass (1988, p. 115, p. 141) and Bottazzini and Gray (2013, ch. 9).

[46] Siegel (1973, p. 1); the German original appeared in 1966.

[47] On the $\bar{\partial}$-equation, I'm indebted to correspondence with Jon-Erik Fornæss. The emergence of one style of n-dimensional generalization of Riemann's approach is charted in Remmert (1998).

[48] "... function theory ... must be built on the foundation of algebraic truths, and that it is therefore not the right path when the "trancendant" ... is taken as the basis of simple and fundamental algebraic propositions. [Appeal to the "transcendant"] seems so attractive at first sight, in that through it Riemann was able to discover so many of the important properties of algebraic functions. (It is self-evident that, as long as he is working, the researcher must be allowed to follow every path he wishes; it is only a matter of systematic foundations.)" (*Werke* II p. 235).

discovery, that of Weierstrass is above all a method of demonstration." (Poincaré, 1900, p. 7). Sophus Lie put it this way:

> [Riemann's] astonishing mathematical instinct let him see immediately what his time didn't allow him to prove definitively by purely logical considerations, nonetheless these brilliant results are the best testament to the fruitfulness of his methods. (Lie and Scheffers, 1896, pp. v–vi)[49]

Indeed, Riemann made mistakes, and some of his arguments were obscure.[50]

After Riemann's early death at age 39 in 1866 he couldn't clarify and develop further, so that even on the 100th anniversary of Riemann's thesis, Ahlfors could famously speak of Riemann's "cryptic messages to the future" (Ahlfors, 1953, p. 3). But the "discovery/justification" distinction, and talk of "astonishing mathematical instinct" is misleading, to the extent that it presents discovery as an oracular revelation of ultimately inexplicable genius, the inscrutable voice of an external *daimon* whispering in Riemann's ear. There were principles and methods that Riemann grasped, however obscurely, which guided him to his "brilliant results" and a task for his successors was to articulate these principles in a more disciplined way (even if there might be other means to demonstrate the results, once discovered). When the work was not rigorous as it stood, it could usually be *made* rigorous in a way that retained its core insights.[51]

This rigorisation was also a creative process of (broadly speaking) interpretation in various directions. In Sects. 24.6.2.1 and 24.6.2.2 I'll consider aspects of the "orthodox" development of complex functions, but I'll briefly note two others.

---

[49] For one contemporary opinion, in connection with elliptic functions in particular, note Remmert's remark, concerning Abelian functions, which generalise elliptic functions: Though Weierstrass regarded the topic as crucial, "here Riemann's ideas were more fruitful" (Weierstraß, 1988, p. ix).

[50] Weierstrass himself noted several Riemannian lapses of rigour, for example Riemann's overly general formulation of the Dirichlet principle, as noted in Sect. 24.6.2.2. Another of Weierstrass's objections directly addressed §20 of Riemann's thesis, indicating a limitation on the generality of Riemann's approach. (Weierstrass says §19, but §20 is clearly intended.) Riemann states at the end of §20 that he is not *proving* that the class of ("monogenic") functions he is defining coincide with those that can be "expressed by operations on quantities", but that such a proof would be needed to view his approach as foundational to "a general theory of operations on quantities", and he certainly appears to believe that such a proof can be found. But though the equivalence holds for the most part, Weierstrass produced a class of counterexamples displaying that "the concept of a monogenic function of one complex variable does not coincide with the concept of a dependence that can be expressed by means of (arithmetical) operations on magnitudes" (Weierstrass, 1880, p. 79). Weierstrass was no doubt pleased to note, "the contrary has been stated by Riemann". (*ibid* p. 79 footnote) I am here indebted to Bottazzini and Gray (2013, p. 465), and to a communication from Bottazzini.

[51] For present purposes I'm presupposing that we can make sense of the idea that a precise set of concepts and techniques can to some degree or other accurately represent concepts/techniques that are only vaguely implicit in earlier work. There has been philosophical scrutiny of this, tracing back to Burge (1979) on Frege's view of partial grasp of concepts and Peacocke (1998) on the relation between the Leibniz/Newton treatment of calculus and later ones. Smith (2015) on the derivative is a good recent discussion.

Clebsch and his students treated Riemann's account of Abelian functions as a computational theory of algebraic curves and surfaces.[52] Developed by Max Noether and Clebsch student Alexander Brill, this flowed into the Italian tradition of algebraic geometry. Dedekind avoided appeals to geometric intuition by defining Riemann surfaces within a recognizably modern, "structural" approach to algebra in Dedekind and Weber (1882) and elsewhere. Subsequent development by Emmy Noether and her students led to the contemporary conception of abstract algebra.

24.6.2.1 Rigour: Intuition and Riemann Surfaces

The main objection to Riemann's approach from the Weierstrass side is the absence of rigour: At the time of Riemann's death, Riemann surfaces were not so much defined as they were described and graphically depicted.[53] There are two distinct claims about rigour in the Weierstrass passage quoted in footnote 48: Riemann's project as it stood lacked necessary "systematic foundations" and they "must be built on the foundation of algebraic truths". It is possible to address the first without the second, thereby introducing an expanded conception of rigour. The challenge for those attempting to develop Riemann's work in such a way as to preserve the Riemannian essence responsible for its fecundity as a means of discovery is to come up with a "conceptual" rather than "computational" style of *rigorous* proof.

Though it was a long process of refinement with many additional results required to set the stage, an adequate definition was finally arrived at, with the shape of the current conception visible in Hermann Weyl's *Die Idee der Riemannschen Fläche* (Weyl, 1997).[54] Weyl's account presents a further flowering of a Riemannian principle of mathematical ontology noted above: As Riemann viewed it as crucial to study functions without requiring reference to the linguistic means by which the functions were introduced and designated, Weyl lays a foundation for treating Riemann surfaces without reference to the functions through which they were originally introduced.[55]

---

[52] See, for example, Clebsch and Gordan (1866). A rich exploration of (*inter alia*) Clebsch's approach to Riemann is Gray (1989). An extensive, clear presentation of the Brill-Noether approach is Casas-Alvero (2019). Lê (2017) is a revealing immersion in Clebsch's computational conception of geometry, including the genus/deficiency relation (see esp. §4.1, §4.2) touched on in footnote 58 below.

[53] There were other objections, as noted in footnote 50. Another turned on "purity of method" considerations, as Mittag-Leffler (presumably channeling Weierstrass) wrote that even if Riemann's approach could be developed rigorously, it would "[introduce] elements into function theory that are in principle altogether foreign". Frostman (1966, pp. 54–5) cf. Tappenden (2006, p. 113) and Bottazzini and Gray (2013, p. 424).

[54] See Bottazzini and Gray (2013, pp. 612–620). A useful compact discussion of the historical events, including Weyl (1997) and its reception, is in Remmert (1998).

[55] Indeed, Weyl seems to suggest that Riemann was well aware of this possibility but chose to hold back from conveying "too strange ideas" to his contemporaries. Weyl (1997, p. VII) (Though

24.6.2.2 Rigour: Dirichlet Principle and Function Existence

Riemann's use of a function-existence principle called the Dirichlet principle posed a problem. It's a well-known story—I'll just recount key points.[56] As noted, Riemann treated functions as independent objects. Certain properties (zeros, number and type of singularities, ...) were established, then (ideally) the existence of a function with those properties would be proven. The Dirichlet principle justified such existence claims in key arguments. But the principle as Riemann stated it admitted counter-examples. Could the results be saved? Ultimately the Dirichlet principle in versions sufficient for Riemann's arguments was worked out. A nice example of the classic Lakatos (1976) pattern: proof $\Rightarrow$ counterexample $\Rightarrow$ refined proof.

Here I want to consider less the Dirichlet principle itself than one proposed diagnosis of its failure: that it appeals to an uncontrolled concept of function. This was a Weierstrassian objection,[57] but it also arose within an intramural dispute between the tradition of complex analysis that followed out an orthodox Riemannian style and Clebsch's research stream. The Clebsch approach was partly motivated by a desire for rigour, avoiding: "... all consideration of functions in general, which are always precarious, because the concept involves completely vague and unidentified possibilities" (Clebsch and Gordan, 1866, p. VI). This reservation continued to be cited as a defect by writers in the Clebsch tradition even nearly thirty years later: "[There are misgivings about] the operation with functions of indeterminate determination in the Riemann style. The function concept in such generality, incomprehensible and amorphous, no longer leads to verifiable conclusions." (Brill and Noether, 1894, p. 265).

Among the Riemannian responses to this was to maintain that the Clebsch strategy inverted the proper order of explanation. Thus Riemann's student Prym in a letter: "The attempt to base function theory on algebra is completely useless... algebra is a consequence of function theory rather than the other way around." (quoted in Neuenschwander, 1978/79, p. 61.) Today we would not say one needs to be taken as absolutely prior to the other—rather, they complement one another—but Prym's opinion is not just a matter of "what I like"; it rests on some subtle judgements of informal methodology. I'll consider just one example to illustrate: the definition of genus.

For Riemann, the genus is a topological property: in the simplest cases, it's the number of holes in a surface. Remarkably, it's equivalent to an algebraic property, reflecting the number of singularities short of the maximum for a surface of the given

perhaps this was among the things that prompted Weyl's mature 3rd edition reflection that his "enthusiastic preface betrayed the youth of the author" (Weyl, 1955, p. VII)).

[56] In addition to the technical details, the classic textbook presentation Courant (1950) begins with a brief but clear account of the physical motivations for the principle (pp. 1–3). For historical discussion, see Gray (2015) esp. 14.4, 16.4 and 18.3.

[57] Mittag-Leffler notes this in an 1875 letter, for example Frostman (1966, p. 54).

degree.[58] Klein used the contrast of definitions as a touchstone for his affirmation of the value of Riemann's approach:

> [An] objection to adopting Clebsch's presentation lies in the fact that, from Riemann's point of view, many points of the theory become far more simple and almost self-evident, whereas in Clebsch's theory they are not brought out in all their beauty. An example of this is presented by the idea of the deficiency [=genus] p. In Riemann's theory,... the invariability of p under any rational transformation is self-evident, while from the point of view of Clebsch this invariability must be proved by means of a long elimination, without affording the true geometrical insight into its meaning. (Klein, 1893, pp. 3–5)

Setting aside the comparative evaluation and just addressing the evaluation of Riemann's topological approach: Klein is right to emphasise the mathematical value and fruitfulness of understanding genus topologically and taking the topological invariance of genus as central.[59] Klein's informal methodology strives to further articulate the insight that Riemann is on to something vital in his characterisation of the genus.

But "function" did need to be clarified. This was among the driving forces for the development of set theory.[60] Hence, 100 years after Riemann's thesis, Carathéodory could frame his complex analysis textbook in the following way. After noting that Weierstrass developed an arithmetic treatment of unexcelled "rigour and beauty", he continues:

> During the last third of the 19th Century the followers of Riemann and those of Weierstrass formed two sharply separated schools of thought. However, in the 1870's Georg Cantor (1845–1918) created the Theory of Sets.... With the aid of Set Theory it was possible for the concepts and results of Cauchy's and Riemann's theories to be put on just as firm a basis as that on which Weierstrass' theory rests, ...and this led to the discovery of great new results in the Theory of Functions as well as of many simplifications in the exposition. (Carathéodory, 1954, p. vii)

This confers further substance on the title thesis of Gray (2009), in that nineteenth century analysis drove a refinement of our understanding of how a theory can be "rigorous". It also displayed a case that undercuts the "discovery/justification" distinction as manifested in the contrast of Riemann's "astonishing mathematical instinct/method of discovery" and Weierstrass's "method of demonstration/purely

---

[58] Clebsch called this the *Geschlecht*, usually translated "genus". (Cayley's term "deficiency" is sometimes used for genus defined in Clebsch's way, especially in older textbooks (for example Hilton, 1920, p. 113).) Popescu-Pampu (2016) is illuminating on the history of the variations on the genus concept.

[59] This is not to endorse the negative side of Klein's evaluation. Clebsch's non-topological approach to genus also marks out a central property as a birational invariant—in this case a central property of algebraic curves, which interacts systematically with the degree of the curve. For reasons of space I won't explore Weierstrass' important non-topological definition of genus via his "*Lückensatz*" (gap theorem). For the history, see Bottazzini and Gray (2013, §6.8.6) and Del Centina (2008). Edwards (2005, ch. 4) ("The Genus of an Algebraic Curve") explores the value of a further non-topological definition.

[60] cf. Jourdain's introduction to Cantor (1955). Some of the ways nineteenth century analysis shaped the development of set theory are explored in Ferreirós (1999).

logical means", as represented by the quotes opening Sect. 24.6.2. Riemann's work and its subsequent evolution suggest that it can be worthwhile to craft methods of justification that make proven means of discovery explicit, in the spirit of one of Frege's justifications of his foundations: "…if, by examining the simplest cases, we can bring to light what people have there done by instinct, and can extract from such procedures what is universally valid in them, may we not thus arrive at general methods for forming concepts and establishing principles which will be applicable also in more complicated cases?" (Frege, 1884, §2).

24.6.3 Weierstrass and Algebraic Addition Theorems

I've considered Weierstrassian objections that the Riemann tradition managed to overcome. Now I'll look at something of the highest importance to Weierstrass, about which the Riemann tradition says little: Weierstrass's characterisation of the elliptic functions as the most general class of functions with an algebraic addition theorem (AAT). In school we learn the addition formula for sine, which expresses $sin(x+y)$ as an algebraic function of sine and its derivative cosine applied to x and y:

$$sin(x+y) = sin(x)cos(y) + sin(y)cos(x)$$

We can mark out a more general characterisation of such patterns: A function $\phi(z)$ has an *algebraic addition theorem* if there is a nonzero polynomial $F(x_1, x_2, x_3)$ in three variables such that:

$$F(\phi(z_1+z_2), \phi(z_1), \phi(z_2)) = 0$$

Among the remarkable features of the $\wp$-function is that it has an addition formula; it follows without much fuss that all elliptic functions do.

$\wp(z_1+z_2) = \left(\frac{\wp'(z_1)-\wp'(z_2)}{\wp(z_1)-\wp(z_2)}\right)^2 - \wp(z_1) - \wp(z_2)$[61]

A result concerning addition theorems was a core component of Weierstrass's lectures in the mid-1860s and onwards: every analytic function that admits an algebraic addition formula is either an elliptic function, or a limiting case.[62]

[61] The addition theorem for $\wp$ directly entails an *algebraic* addition theorem for $\wp$, since $\wp'(z_1)$ and $\wp'(z_2)$ are algebraic functions of $\wp(z_1)$ and $\wp(z_2)$ cf. Akhiezer (1990, p. 45).

[62] Bottazzini and Gray (2013, §6.6.3) and Del Centina (2019). A clear textbook discussion of the mathematical details is Prasolov and Solovyev (1997, §2.9). In this section I'm also indebted to an unpublished manuscript by Mark Villarino, which cites the proof of Weierstrass's characterisation as indicating that "the 'cause' or 'explanation' of the existence of a period of the meromorphic function" is the combination of the AAT and the dispersion of points around essential singularities." (Villarino, 2022, p. 7).

(Specifically: Let $\phi(z)$ be an analytic function admitting an algebraic addition theorem. Then $\phi(z)$ is either: (i) an algebraic function of z (ii) an algebraic function of $e^{\frac{i\pi z}{\omega}}$ (ω some constant) or (iii) a doubly periodic function of z.) That is, Weierstrass proved that possession of an algebraic addition theorem was a *distinguishing characteristic* of the elliptic functions. Beginning with that basic characterisation, Weierstrass derives the $\wp$-function, and uses it as a stepping stone to the entire theory. Mittag-Leffler wrote in 1876 lectures drawing from Weierstrass' classes:

> [Is it] possible to find a characteristic property common to the doubly periodic functions and this particular sub-class of simply periodic functions, and which is the exclusive property of these functions and thus distinguishes them in contrast to all other analytic functions? Weierstrass found such a property in the addition theorem, and from this starting point he succeeded in developing the theory of elliptic functions to the highest degree of perfection that a mathematical theory may ever reach. (Mittag-Leffler, 1923, pp. 350–51)

Treating the addition theorem as the "characteristic property" represents a divergence between Weierstrass and Riemann that seems to rest on incommensurable values—there were research objectives that mattered a great deal to Weierstrass and his followers, which were not salient concerns for Riemann and his followers. This was not a capricious preference on Weierstrass's part: Given the problems Weierstrass wanted to address, emphasis on addition theorems *was* an exceptionally fruitful means of organisation, and it was reasonable to have the $\wp$-function play a starring role.

And more than 100 years later the set of considerations turned out to be fruitful in ways that Weierstrass couldn't have imagined. A particularly compelling example is in the theory of elliptic curves, where the addition theorems (and the $\wp$-function in particular) support the proof of the group law on cubic curves, illuminating a range of areas of mathematics from cryptography to Fermat's Last Theorem.[63]

24.7 Nineteenth Century Analysis as Philosophy of Mathematics: Reprise

The informal side of the pursuit of mathematical understanding is necessarily a dynamic process, because many of the reasons offered for particular theoretical choices and verdicts involve predictions and are hostages to the future. Is it possible

[63] For the use of the $\wp$-function to demonstrate the group law for elliptic curves see Koblitz (1993, §7) or Lang (1978, §§2–3). Griffiths and Harris (1994, p. 240) write of the group structure of cubic curves and the addition theorem for elliptic functions as arising from different "interpretations" of a basic equation containing the $\wp$-function. For connections to cryptography and number theory, see Washington (2008). Pastras (2020), as the name suggests, is an illustration of the long reach of the Weierstrass conception, the $\wp$-function in particular. Another illustration of the addition-theorem centred conception of elliptic functions in an applied context is the textbook (Hietarinta et al., 2016). See especially Appendix B. See also Nijhoff (2022) Appendix A for more explicit framing remarks.

to work out a general concept of function sufficiently sharp and disciplined for Riemann's analysis? Yes, though it took many decades before that was put to bed. Is Riemann's approach limited due to an inability to generalize from one to many variables? It turns out it isn't, but it took 100 years to work that out. In some cases, points of difference can turn on incommensurable values, and formless matters of 'styles of reasoning". Is the possession of an addition theorem a reasonable choice of basic property delineating a class of objects? On Weierstrass' way of doing things, absolutely, but for Riemann, not so much…

This in turn gives us a perspective on the Gauss and Riemann motivations for taking $\mathbb{C}$ as the domain for the study of the functions of analysis. They struggled to informally articulate the advantages through vague and often expressly aesthetic evaluations, but the research driven by the insights make it clear how much substance was behind the "cryptic messages to the future". Subsequent development bore this out: the vague ideas and proof techniques were generally rendered precise and rigorous. To the extent that the Gauss/Riemann remarks in §3 were bound up with tacit or explicit predictions, those predictions could hardly have been more emphatically confirmed. We now have a still partial, but more detailed sense of just what "beauty and roundness" gestured at, and of what substantial principles guided Riemann's "astonishing mathematical instinct".

Nineteenth century analysis is a particularly rich and vivid manifestation of a wider phenomenon: deep mathematical ideas can often outstrip the conceptual, technical and linguistic resources available to lay them out in a rigorous way. It is a long way from "beauty and roundness", or "hidden harmony and regularity" to where we are now. One of the things that drives mathematical investigation is the quest for the resources to articulate what is as yet only viewed through a glass darkly. What I'm calling informal methodology is part of the process of gaining intellectual mastery of the subject. People sought explanations and understanding, argued about what is fundamental and what is secondary, debated which arguments were/were not rigorous and why, and disputed what "rigour" was. Some scrutinized the relationship between the subject matter and the language used to describe it, some disputed which definitions capture the essence of an idea and which fasten on accidental properties … Activities that in other contexts would be called "philosophy".

Acknowledgments It is a great pleasure to dedicate this essay, with gratitude, to Jeremy Gray, as he has been a friend and intellectual mentor almost from the beginning. My happy encounter, as a graduate student, with Gray (1989) transformed my understanding of the potential interactions of history and philosophy of mathematics. It will be clear to anyone reading this paper what a debt it owes to Jeremy's writings, especially Gray (2015) and Bottazzini and Gray (2013). I've accumulated debts to so many people during the long evolution of this paper that a full list would be longer than any editor would allow, so this list is quite compressed. My apologies to anyone I have failed to mention. Early versions of the material in Sect. 24.4 were read at the University of Toronto, CUNY graduate center, Columbia University, UC Irvine, McGill and the University of Paris VII. I am grateful to Howard Stein for instructive written comments on that material. A later evolution, augmented with early versions of Sect. 24.5 was presented to the PoSe seminar at the University of Michigan, the 2015 APMP meetings in Paris, the 2017 MPMW at Notre Dame, and as a plenary address to the 2015 Canadian Mathematical Association meetings in Montréal. I am grateful to those audiences for questions and commentary, not just during the talks themselves but in subsequent discussions and correspondence over the years, especially the late Andrew Arana, Margaret Morrison, Philip Kitcher, Stephen Menn, Lydia Patton, Michael Hallett, Colin McClarty, Paolo Mancosu, Umberto Bottazzini, Marco Panza, Ivahn Smadja, Hourya Sinaceur, Karine Chemla, Bruno Belhoste, Renaud Chorlay and Emmylou Haffner. I've learned much about Riemann's philosophical dimensions from José Ferreirós. I am grateful to the referees of this paper for helpful comments. Thanks to Yu Zhongmiao for helpful feedback. I would have been lost without my Michigan mathematical colleagues, beginning with fondly remembered conversations with the late Juha Heinonen and Mario Bonk when these ideas were just taking form. On these topics I learned from James Milne, Mel Hochster, Al Taylor, Dan Burns and Jens-Eric Fornæss. Most recently I've learned much from email and conversation with Lizhen Ji. Finally I'd like to thank the editors of this volume collectively for conceiving and bringing this volume to completion.

References

Ahlfors, L. 1953. Development of the theory of conformal mapping and Riemann surfaces through a century. *Contributions to the theory of Riemann Surfaces, Annals of Mathematics Studies 30*: 3–13.

Akhiezer, N.I. 1990. *Elements of the Theory of Elliptic Functions*. Translations of Mathematical Monographs, vol. 79. Providence: American Mathematical Society. Translated from the second Russian edition by H. H. McFaden.

Avigad, J. 2006. Methodology and metaphysics in the development of Dedekind's theory of ideals. In *The Architecture of Modern Mathematics*, ed. J. Ferreirós and J. Gray, pp. 159–86. Oxford: Oxford University Press.

Bell, D. 2010. Poisson's remarkable calculation—a method or a trick? *Elemente der Mathematik 65*(1): 29–36.

Boros, G., and V. Moll. 2004. *Irresistible Integrals*. Cambridge: Cambridge University Press.

Bottazzini, U., and J. Gray. 2013. *Hidden Harmony - Geometric Fantasies. The Rise of Complex Function Theory*. New York: Springer.

Bottazzini, U., and R. Tazzioli. 1995. *Naturphilosophie* and its role in Riemann's mathematics. *Revue d'Histoire des Mathématiques 1*: 3–38.

Brill, A., and M. Noether. 1894. Die Entwickelung der Theorie der algebraischen Functionen in älterer und neuerer Zeit. *Jahresbericht der Deutschen Mathematiker-Vereinigung 3*: 107–566.

Burge, T. 1979. Individualism and the mental. *Midwest Studies in Philosophy 4*(1): 73–122.

Burgess, J.P. 2015. *Rigor and Structure*. Oxford: Oxford University Press.

Cantor, G. 1915–1955. *Contributions to the Founding of the Theory of Transfinite Numbers*. New York: Dover. 1955 Reprint of 1915 original Translated with an Introduction by Philip Jourdain.

Carathéodory, C. 1954. *Theory of Functions of a Complex Variable*, vol. 1. New York: Chelsea Publishing. Translated by F. Steinhardt.

Casas-Alvero, E. 2019. *Algebraic Curves, the Brill and Noether Way*. Universitext. Cham: Springer.

Catanese, F. 2016. Monodromy and normal forms. In *Karl Weierstraß (1815–1897)*, pp. 195–218. Wiesbaden: Springer.

Clebsch, A., and P. Gordan. 1866. *Theorie der Abelschen Functionen*. Leipzig: Teubner.

Courant, R. 1950. *Dirichlet's Principle, Conformal Mapping, and Minimal Surfaces*. New York: Interscience Publishers.

Dedekind, R. 1895. Über die Begründung der Idealtheorie. *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen*, Heft 1, pp. 106–113 Reprinted in *Werke* II p. 50–58.

Dedekind, R., and H. Weber. 1882. Theorie der algebraischen Funktionen einer Veränderlichen. *Journal für die Reine und Angewante Mathematik 92*: 181–290. Reprinted in Dedekind's *Werke* I pp. 248–349.

Del Centina, A. 2008. Weierstrass points and their impact in the study of algebraic curves: A historical account from the "Lückensatz" to the 1970s. *Annali dell'Universitá di Ferrara. Sezione VII. Scienze Matematiche 54*(1): 37–59.

Del Centina, A. 2019. The addition theorem in Weierstrass' theory of elliptic and Abelian functions. In *Serva di due padroni: Saggi di Storia della Matematica in onore di Umberto Bottazzini*, ed. A. Cogliati, pp. 129–69. EGEA spa.

Edwards, H.M. 2005. *Essays in Constructive Mathematics*. New York: Springer.

Eisenbud, D., and J. Harris. 1992. *Schemes: The Language of Modern Algebraic Geometry*. Boston: Wadsworth.

Farkas, H.M., and I. Kra. 1992. *Riemann Surfaces*. Graduate Texts in Mathematics, 2nd edn., vol. 71. New York: Verlag.

Farkas, H.M., and I. Kra. 2001. *Theta Constants, Riemann Surfaces and the Modular Group*. Providence: American Mathematical Society. An introduction with applications to uniformization theorems, partition identities and combinatorial number theory.

Ferreirós, J. 1999. *Labyrinth of Thought: A History of Set Theory and its Role in Modern Mathematics*. Boston: Birkhäuser Verlag.

Ferreirós, J. 2006. Riemann's Habilitationsvortrag at the crossroads of mathematics, physics and philosophy. In *The Architecture of Modern Mathematics. Essays in History and Philosophy*, ed. J. Ferreirós and J. Gray. Oxford: Oxford University Press.

Ferreirós, J. 2007. 'Ο Θεὸς 'Αριθμητίζει: The rise of pure mathematics as arithmetic with Gauss. In *The Shaping of Arithmetic After C. F. Gauss's* Disquisitiones Arithmeticae, pp. 235–268. Berlin: Springer.

Ferreirós, J., and J. Gray, eds. 2006. *The Riemannian Background to Frege's Philosophy*. Oxford: Oxford University Press.

Ferreirós, J., and E.H. Reck. 2020. Dedekind's mathematical structuralism: From Galois theory to numbers, sets, and functions. In *The Prehistory of Mathematical Structuralism*. Logic Computation Philosophy, pp. 59–87. New York: Oxford University Press.

Fillion, N. 2019. Conceptual and computational mathematics. *Philosophia Mathematica Series III 27*(2): 199–218.

Frege, G. 1953–1884. *The Foundations of Arithmetic*, 2nd ed. Evanston: Northwestern University Press. J. L. Austin, Trans. original published 1884, first edition 1950, Translations often modified with suggestions from Beaney 1997.

Freudenthal, H. 1975. Riemann. In *Dictionary of Scientific Biography*, ed. C. Gillespie, vol. 11, pp. 447–56. New York: Scribner's.

Frostman, O. 1966. Aus dem Briefwechsel von G. Mittag-Leffler. In *Festschrift zur Gedächtnisfeier für Karl Weierstraß 1815–1965*, ed. H. Behnke and K. Kopfermann, pp. 53–56. Wiesbaden: VS Verlag für Sozialwissenschaften.

Gauss, C. 1832. Theoria residuorum biquadraticorum: Commentatio seconda. *Göttingensche gelehrte Anzeigen 7*. Page reference to reprinting in *Werke II* pp. 93–148.

Gauss, C.F. 1880. *Briefwechsel zwischen Gauss und Bessel*. Leipzig: Englemann.

Gauss, C.F. 1927. *Werke XI*. Königlichen Gesellschaft der Wissenschaften.

Giaquinto, M. 2020. The epistemology of visual thinking in mathematics. In *The Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta (Spring 2020 ed.). Metaphysics Research Lab, Stanford University.

Gray, J. 1989. Algebraic geometry in the late nineteenth century. In *The History of Modern Mathematics*, ed. D. Rowe and J. McCleary, vol. I. Cambridge: Academic Press.

Gray, J. 2008. *Plato's Ghost: The Modernist Transformation of Mathematics*. Oxford: Oxford University Press.

Gray, J. 2009. Nineteenth century analysis as philosophy of mathematics. In *New Perspectives on Mathematical Practice*, ed. B. Van Kerkhove, pp. 138–49. Singapore: World Scientific.

Gray, J. 2015. *The Real and the Complex: A History of Analysis in the 19th Century*. Berlin: Springer.

Griffiths, P., and J. Harris. 1994. *Principles of Algebraic Geometry*. Wiley Classics Library. New York: Wiley. Reprint of the 1978 original.

Hietarinta, J., N. Joshi, and F.W. Nijhoff. 2016. *Discrete Systems and Integrability*. Cambridge Texts in Applied Mathematics. Cambridge: Cambridge University Press.

Hilbert, D. 1998. *Theory of Algebraic Number Fields*. Berlin: Springer. Originally published 1897. Ian T. Adamson (trans).

Hilton, H. 1920. *Plane Algebraic Curves*. Oxford: The Clarendon Press/Oxford University Press.

Hutcheson, F. 1973–1725. *An Inquiry Concerning Beauty, Order, Harmony, Design*. The Hague: Martinus Nijhoff. editor: Kivy, Peter.

Iwasawa, H. 2009. Gaussian integral puzzle. *The Mathematical Intelligencer 31*(3): 38–41.

Johnson, D.M. 1979. The problem of the invariance of dimension in the growth of modern topology. I. *Archive for History of Exact Sciences 20*(2): 97–188.

Jones, G.A., and D. Singerman. 1987. *Complex Functions*. Cambridge: Cambridge University Press. An algebraic and geometric viewpoint.

Kitcher, P. 1981. Mathematical rigor - who needs it? *Noûs 15*: 469–493.

Kitcher, P. 1984. *The Nature of Mathematical Knowledge*. Oxford: Oxford University Press.

Klein, F. 1893. *The Evanston Colloquium Lectures on Mathematics*. New York: McMillan. Alexander Ziwet (ed.).

Koblitz, N. 1993. *Introduction to Elliptic Curves and Modular Forms*, 2nd ed. New York: Springer.

Lakatos, I. 1976. *Proofs and Refutations, the Logic of Mathematical Discovery*. Cambridge: Cambridge University Press. Edited by John Worrall and Elie Zahar.

Lang, S. 1978. *Elliptic Curves: Diophantine Analysis*. Berlin: Springer.

Lange, M. 2010. What are mathematical coincidences and why do they matter? *Mind 119*: 307–40.

Laugwitz, D. (1999). *Bernhard Riemann: 1826–1866 Turning Points in the Conception of Mathmatics*. Boston: Birkhäuser. Abe Shenitzer (trans.).

Lê, F. (2017). Alfred Clebsch's "geometrical clothing" of the theory of the quintic equation. *Archive for History of Exact Sciences 71*(1): 39–70.

Lie, S., and G. Scheffers. 1896. *Geometrie der Berührungstransformationen*. Leipzig: Teubner. 1956 Reprint: Chelsea Publishers.

Mancosu, P. 2001. Mathematical explanation: Problems and prospects. *Topoi. An International Review of Philosophy 20*(1): 97–117.

Mancosu, P., ed. 2008. *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press.

Markushevich, A.I. 1987. *Theory of Functions of a Complex Variable*, vol. III. Englewood Cliffs: Prentice-Hall. Translated and Edited by Richard Silverman.

McKean, H., and V. Moll. 1997. *Elliptic Curves*. Cambridge: Cambridge University Press. Function theory, geometry, arithmetic.

Mittag-Leffler, G. 1876–1923. An introduction to the theory of elliptic functions. *Annals of Mathematics. Second Series 24*(4): 271–351. English Translation by Einar Hille of *En Metod att Komma i Analytisk Besittning af de Elliptiska Funktionerna* Stockholm 1876.

Napier, T., and M. Ramachandran. 2011. *An Introduction to Riemann Surfaces*. New York: Birkhäuser/Springer.

Needham, T. 1997. *Visual Complex Analysis*. New York: The Clarendon Press Oxford University Press.

Neuenschwander, E. 1978–1979. Der Nachlass von Casorati (1835–1890) in Pavia. *Archive for History Exact Sciences 19*(1): 1–89.

Nijhoff, F. 2022. Leeds math5492 lectures. http://www1.maths.leeds.ac.uk/~frank/math5492/AdvLectures.pdf. Accessed 22 April 2022.

Pastras, G. 2020. *The Weierstrass Elliptic Function and Applications in Classical and Quantum Mechanics—A Primer for Advanced Undergraduates*. SpringerBriefs in Physics. Cham: Springer.

Peacocke, C. 1998. Implicit conceptions, understanding and rationality. *Philosophical Issues 9*: 43–88.

Poincaré, H. 1900. L'oeuvre mathématique de Weierstrass. *Acta Mathematica 22*: 1–18.
Poincaré, H. 1985. *Papers on Fuchsian Functions*. New York: Springer. Translated from the French and with an introduction by John Stillwell.
Popescu-Pampu, P. 2016. *What Is the Genus?*. Lecture Notes in Mathematics, vol. 2162. Berlin: Springer.
Prasolov, V., and Y. Solovyev. 1997. *Elliptic Functions and Elliptic Integrals*. Memoirs of the American Mathematical Society. Providence: American Mathematical Society. Translated by D. Leites.
Remmert, R. 1991. *Theory of Complex Functions*. New York: Springer. Translated from the second German edition by Robert B. Burckel.
Remmert, R. 1998. From Riemann surfaces to complex spaces. In *Matériaux pour l'histoire des mathématiques au XXe siècle (Nice, 1996)*. Séminares et Congrès, vol. 3, pp. 203–241. Paris: Society of Mathematics France.
Riemann, B. 1851. Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse. Inauguraldissertation, Göttingen. Reprinted in Bernhard Riemann's *Gesammelte Mathematische Werke und Wissenschaftlich Nachlass* R. Dedekind and H. Weber with Nachträge ed. M. Noether and W. Wirtinger, 3rd edn R Narasimhan (ed) Springer, New York. 1990 1st ed. R. Dedekind, H Weber (eds.) Leipzig: Teubner 1876 pp.3–45. English Translation in Riemann 2004 pp.1–42. Page references to 3rd ed. of *Werke*.
Riemann, B. 1857a. Theorie der Abel'schen Functionen. *Journal für die reine und angewandte Mathematik 54*. Reprinted in Bernhard Riemann's *Gesammelte Mathematische Werke und Wissenschaftlich Nachlass* R. Dedekind and H. Weber with Nachträge ed. M. Noether and W. Wirtinger, 3rd edn R Narasimhan (ed) Springer, New York. 1990 1st ed. R. Dedekind, H Weber (eds.) Leipzig: Teubner 1876, pp.88–142. Page references to 3rd ed. of *Werke*.
Riemann, B. 1857b. Abstract of [1857c]. *Göttinger Nachrichten 1*. Bernhard Riemann's *Gesammelte Mathematische Werke und Wissenschaftlich Nachlass* R. Dedekind and H. Weber with Nachträge ed. M. Noether and W. Wirtinger, 3rd edn R Narasimhan (ed) Springer, New York. 1990 1st ed. R. Dedekind, H Weber (eds.) Leipzig: Teubner 1876 pp.84–5. Page references to 3rd ed. of *Werke*.
Riemann, B. 1857c. Beiträge zur Theorie der durch die Gauss'sche Reihe $f(\alpha, \beta, \gamma, x)$ darstellbaren Functionen. Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen *7*. Reprinted in Bernhard Riemann's *Gesammelte Mathematische Werke und Wissenschaftlich Nachlass* R. Dedekind and H. Weber with Nachträge ed. M. Noether and W. Wirtinger, 3rd edn R Narasimhan (ed) Springer, New York. 1990 1st ed. R. Dedekind, H Weber (eds.) Leipzig: Teubner 1876 pp.67–83. Page references to 3rd ed. of *Werke*English translation.
Riemann, B. 1876. *Gesammelte Mathematische Werke und Wissenschaftlich Nachlass*. Leipzig: Teubner. R. Dedekind and H. Weber eds. 2nd ed. with Nachträge ed. M. Noether and W. Wirtinger, 3rd ed. R Narasimhan (ed) Springer, New York. 1990 1st ed. R. Dedekind, H Weber (eds.) Leipzig: Teubner 1876 pp.3–45. English Translation of 2nd ed as *Collected Works* R. Baker, C. Christenson and H. Orde trans. 2004 Page references to 3rd ed. of *Werke*.
Riemann, B. 2004. *Collected Works*. Heber City UT: Kendrick Press. English Translation by R. Baker, C. Christenson and H. Orde of Riemann, B. *Gesammelte Mathematische Werke und Wissenschaftlich Nachlass* R. Dedekind and H. Weber eds. 2nd ed. with Nachträge.
Rowe, D.E. 2000. Episodes in the Berlin-Göttingen rivalry, 1870–1930. *The Mathematical Intelligencer 22*(1): 60–69.
Roy, R. 2017. *Elliptic and Modular Functions from Gauss to Dedekind to Hecke*. Cambridge: Cambridge University Press.
Scholz, E. 1982. Herbart's influence on Bernhard Riemann. *Historia Mathematica 9*: 413–440.
Siegel, C. 1973. *Topics in Complex Function Theory*, vol. III Abelian Functions and Modular Functions of Several Variables. New York: Wiley. E. Gottschling and M. Tretkoff (trans.).
Siegel, C.L. 1932. Über Riemanns Nachlaß zur analytischen Zahlentheorie. *Quellen und Studien zur Geschichte der Mathematik, Astronomie und Physik 2*: 45–80. Page references to reprinting in Bernhard Riemann's *Gesammelte Mathematische Werke und Wissenschaftlich Nachlass* 3rd ed. R Narasimhan (ed). Springer, New York. 1990 pp. 770–805

Sklar, L. 2000. *Theory and Truth: Philosophical Critique within Foundational Physics*. Oxford: Oxford University Press.

Skow, B. 2015. Are there genuine physical explanations of mathematical phenomena? *British Journal for the Philosophy of Science 66*(1): 69–93.

Smith, S.R. 2015. Incomplete understanding of concepts. *Mind 124*(496): 1163–1199.

Spivak, M. 1967. *Calculus*, 1st ed. New York: W. A. Benjamin.

Springer, G. 1981. *Introduction to Riemann Surfaces*. AMS Chelsea Publishing Series. Providence: American Mathematical Society.

Stein, E.M., and R. Shakarchi. 2003. *Complex Analysis*. Princeton Lectures in Analysis, vol. 2. Princeton: Princeton University Press.

Steiner, M. 1978. Mathematics, explanation and scientific knowledge. *Noûs 12*: 17–28.

Stillwell, J. 1989. *Mathematics and its History*. Berlin: Springer.

Tappenden, J. 1995. Extending knowledge and 'fruitful concepts': Fregean themes in the foundations of mathematics. *Noûs 29*: 427–467.

Tappenden, J. 2005a. The Caesar problem in its historical context: Mathematical background. *Dialectica 59*(fasc. 2): 237–264.

Tappenden, J. 2005b. Proof style and understanding in mathematics I: Visualization, unification and axiom choice. In *Visualization, Explanation and Reasoning Styles in Mathematics*, ed. P. Mancosu and K. Jørgensen. Berlin: Springer.

Tappenden, J. 2006. The Riemannian background to Frege's philosophy. In *The Architecture of Modern Mathematics: Essays in History and Philosophy*, ed. J. Ferreirós and J. Gray, pp. 97–132. Oxford: Oxford University Press.

Tappenden, J. 2008. Mathematical concepts: Fruitfulness and naturalness. In *The Philosophy of Mathematical Practice*, ed. P. Mancosu, pp. 276–301. Oxford: Oxford University Press.

Tappenden, J. 2012. Fruitfulness as a theme in the philosophy of mathematics. *Journal of Philosophy 109*(1–2): 204–219.

Tappenden, J. 2013. The mathematical and logical background to analytic philosophy. In *The Oxford Handbook of the History of Analytic Philosophy*. Oxford: Oxford University Press.

Tappenden, J. 2019. Infinitesimals, magnitudes and definition in Frege. In *Essays on Frege's Basic Laws of Arithmetic*, ed. by P. Ebert and M. Rossberg, pp. 235–63. Oxford: Oxford University Press.

Villarino, M. 2022. Algebraic addition theorems. https://arxiv.org/abs/1212.6471. Accessed March 2022.

Walker, P.L. 1996. *Elliptic Functions. A Constructive Approach*. Chichester: Wiley.

Washington, L.C. 2008. *Elliptic Curves*, 2nd ed. Boca Raton: Chapman & Hall/CRC. Number theory and cryptography.

Weierstraß, K. 1878–1988. *Einleitung in die Theorie der analytischen Funktionen*. Deutsche Mathematiker Vereinigung, Freiburg; Friedr. Vieweg & Sohn, Braunschweig. Vorlesung Berlin 1878. [Berlin lecture of 1878], Lecture notes taken and with supplementary material by Adolf Hurwitz, With a preface by R. Remmert, Edited and with a foreword by Peter Ullrich.

Weierstrass, K. 1880. Zur Funktionenlehre. *Monatsberichte Berlin 52*: 719–743. Page references to the reprinting in Weierstrass, K. *Abhandlungen aus der Funktionenlehre* Berlin: Teubner 1886, 67–101.

Weierstrass, K. 1886–1988. *Ausgewählte Kapitel aus der Funktionenlehre. Vorlesung, gehalten in Berlin 1886. Mit der akademischen Antrittsrede, Berlin 1857 uns drei weiteren Originalarbeiten von K. Weierstrass aus den Jahren 1870 bis 1880/86*. Liepzig: Teubner. Collection of Weierstrass lectures edited and annotated by R. Siegmund-Schultze, 1988.

Weil, A. 1976. *Elliptic Functions According to Eisenstein to Kronecker*. Berlin: Springer.

Weyl, H. 1913–1997. *Die Idee der Riemannschen Fläche*. Stuttgart: B. G. Teubner Verlagsgesellschaft mbH. Reprint of the 1913 German original, With essays by Reinhold Remmert, Michael Schneider, Stefan Hildebrandt, Klaus Hulek and Samuel Patterson, Edited and with a preface and a biography of Weyl by Remmert.

Weyl, H. 1955. *The Concept of a Riemann Surface*. Reading: Addison. translation by G. Maclane of 3rd revised edition of Weyl 1913.

Weyl, H. 1995a. Topology and abstract algebra as two roads of mathematical comprehension. I. *American Mathematical Monthly 102*(5): 453–460. Translated from the German by Abe Shenitzer.

Weyl, H. 1995b. Topology and abstract algebra as two roads of mathematical comprehension. II. *American Mathematical Monthly 102*(7): 646–651. Translated from the German by Abe Shenitzer.

Wilson, M. 2006. *Wandering Significance*. Oxford: Clarendon Press.

Chapter 25
What We Talk About When We Talk About Mathematics

Jeremy Avigad

Abstract This is an essay about the ways that philosophers talk about mathematics.

What do we talk about when we talk about mathematics? Numbers and functions, certainly. Algebraic structures sometimes. The structures can get pretty complicated. But these are really things that we talk about when we *do* mathematics. What do we talk about when we talk *about* mathematics, which has been around for a long time?

Philosophers have two ways of talking about mathematics. Some of them talk about what mathematicians say and do. When these philosophers talk about mathematics, they talk about definitions, theorems, and proofs, and sometimes calculations, questions, and conjectures. They also talk about methods, intuitions, and ideas, and maybe it is not so clear what those are. But even methods, intuitions, and ideas are found in what mathematicians say and do. So when these philosophers talk about mathematics, they are talking about mathematical talk.

The other way of answering the second question is to repeat the answer to the first question with more emphasis. When some philosophers talk about mathematics, they talk about numbers, functions, and algebraic structures, and about how our ordinary mathematical talk latches on to those things. When these philosophers talk about mathematics, they are talking about what mathematical talk is about.

Logicians make a big deal about the difference between *syntax* and *semantics*. When logicians talk about syntax, they talk about the rules of a language. When they talk about semantics, they talk about what a language means. So it's tempting to say that the first bunch of philosophers are interested in the syntax of mathematics and the second bunch of philosophers are interested in its semantics.

The distinction between syntax and semantics is useful in logic, but it is not very useful in philosophy. In fact, it causes a lot of problems. We ought to think long and hard about how we got stuck with these problems, and then we ought to figure out how to get out of the mess we are in.

J. Avigad (✉)
Carnegie Mellon University, Pittsburgh, PA, USA
e-mail: avigad@cmu.edu

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_25

25.1 The Problem

It helps to think about the ways we talk about mathematics. That's how we got axiomatic foundations like set theory and type theory. Foundations helped clarify the ways we talk about numbers and functions and algebraic structures. They also helped us think about important questions, like, is it o.k. to talk about functions we cannot compute? Is it o.k. to use the axiom of choice?

People came up with formal languages for mathematics before there were programming languages, or languages for databases and expert systems. Those are also formal languages, but formal languages for mathematics were there first, and we got them by thinking about the ways we talk about mathematics.

So theories of mathematical language are useful. What have philosophical theories of mathematical objects given us? Not much. Mathematicians did not decide that it was o.k. to use the axiom of choice because philosophers were able to tell them what it means. Philosophers can't even agree about what it means. That's not a problem, because you don't have to say what the axiom of choice means to do mathematics. You *do* have to decide whether to use the axiom of choice. If you are trying to do that, you ought to talk it over with people you know, especially if you want them to listen to you later on. It's probably better not to talk to philosophers.

I am not saying that semantics isn't useful. It's really useful in logic. There are a million proof systems for propositional logic, and what they all have in common is that they prove formulas that are always true, no matter how you interpret the variables. Without knowing what it means for a propositional formula to be true, you can't say what it means for a proof system to be correct, and then it's really hard to explain what all the good proof systems have in common.

It is also useful in computer science. The semantics of a programming language tells you what the programming language is supposed to do and what it means for a compiler to be correct. We use programming languages all the time. If we couldn't think about whether we are implementing them correctly, we would be in pretty bad shape.

Semantics is even useful in mathematics. On the face of it, a polynomial is an expression. It has terms, maybe a constant term, and a term of highest degree, and those are expressions too. But a polynomial can also be a function on the real numbers, which is a thing that the expression describes. A polynomial can also be an element in a polynomial ring. Polynomial rings give us ways of thinking about polynomials without worrying about whether $x+1$ and $1+x$ are the same. At some point, we have to say what the elements of a polynomial ring are, and one way is to do it is to say that they are expressions, maybe up to an equivalence relation. Knowing how to reason about expressions and how the expressions are related to what they express is generally helpful.

In all these situations, we have expressions that describe things and we have other ways of thinking about the things that the expressions describe. Semantics fits the pieces together.

When we talk about mathematical statements, how do we talk about the things that they describe? With mathematical statements, of course. We use mathematics to talk about things like numbers and groups and spaces. When we talk about them we are just doing mathematics. It's not like we have some other way of talking about them. Logicians and computer scientists have special-purpose languages, like the language of an algebraic structure or a programming language. But when they talk about those languages and what they mean, they use mathematics. There is no magical philosophical language that tells us what things *really* are, and we don't need one.

It's not just that worrying about what mathematical objects are isn't helpful. It causes a lot of problems. One of them is Benacerraf's problem, which goes like this. Scientists learn about the world by poking it and seeing what happens. They do experiments and measure things. But then how do they learn about numbers? You can't poke them and measure them because they aren't anywhere. If we can't learn about them in a scientific way, it's hard to say how they can be useful for science. Maybe you want to say that mathematics isn't part of science, but even so, if you are a responsible scientist and you use numbers, you ought to say how you know what you think you know about them.

This way of thinking about mathematics has to be wrong. Of course we can learn mathematics. That's one of the things we do in school and also when we get older. And of course mathematics is useful for building airplanes and bridges and for making our tax forms come out right. That's why we learn it. It's just that numbers and triangles are not like rocks and trees and sofas. We don't learn about them by bumping into them. They aren't even like atoms and magnetic fields. We learn about them in different ways.

I don't blame Paul Benacerraf for saying what was on his mind. Sometimes you have to talk about the things that are bothering you and get them out in the open. Then you can take a serious look at them and realize that you don't have to worry about them. The problem is, a lot of philosophers can't get over it.

There are many important questions about mathematics. Should we use computers to do proofs? Is mathematics on the right path? Is it getting too abstract? Is it getting too applied? How can we tell whether something is good mathematics? Does statistics count? Is AI going to change everything? Should we be worried? But now there is nobody left to talk about things like that. Mathematicians are too busy trying to prove their theorems and philosophers are too busy trying to figure out what numbers really are. Nobody wants to be a bad mathematician or a bad philosopher so they stick to what they are doing.

25.2 What Went Wrong

The philosophy of mathematics took a bad turn sometime in the twentieth century. The first half of the twentieth century was pretty good for philosophy of mathematics. It was bad for humanity, especially in Europe, but it was good for philosophy of mathematics. It was good for philosophy of science too.

The logical positivists got things going by telling everyone how science works. It is not as simple as they made it out to be, but one of the things they got more or less right is that the way we talk about things is important. Rudolf Carnap said a lot about "linguistic frameworks," but that just means ways of talking about things.

The logical positivists said that mathematics comes down to a choice of linguistic framework. This makes sense when you think about it. Doing mathematics means coming up with ways of thinking about things, which is pretty much the same as coming up with ways of talking about things. Having good ways of thinking about things means having good ways of thinking about the world. But mathematics is about the ways of thinking and not about the world. When we do science we choose a mathematical description a lot like the way we buy a car to get around. If the car doesn't work, we unload it on someone and get a better one.

But even though the logical positivists thought that mathematics comes down to making choices—they called them *pragmatic* choices—they also said that we shouldn't talk about the reasons for making those choices. They are outside the framework. That makes them metaphysical questions, which means that they are not scientific, which is bad.

If we are going to do mathematics, why shouldn't we talk about how we are doing it? If we have reasons for our choices, why can't they be scientific? If I want to decide what car to buy, I am sure as hell going to talk about it. I am going to think about all the things I want to do with the car. I am going to go to the library to look at all the car magazines and I am going to ask my friends for advice. It's hard to make decisions. It helps to talk about them.

Eventually W. V. O. Quine came along and said that the logical positivists were wrong. It's not just mathematics that is determined by language. All of science is determined by language, and there is nothing special about mathematics. It is all one big web of beliefs. Everything has to do with how we talk about things, and we had better make good choices about how we talk about things if we want science to work out.

But even though Quine thought we had to make choices, he also thought we shouldn't talk about them much. When we talk about science and do it right, we are just *doing* science. He also said that there isn't a principled distinction between talking about things and coming up with ways of talking about them. This was a jab at the logical positivists, who thought that this was exactly the difference between science and mathematics. Science is about things, and mathematics is about how we talk about them.

Philosophers like to talk about principled distinctions, but mathematics is different from science and it doesn't make sense to pretend they are the same. The logical positivists said that mathematics is different because it is analytic, which means that mathematicians define everything. Or stipulate everything. Definitions determine what words mean because that's what definitions do. Axioms are true because we decide that they are true, not because of the way the world is. Axioms are like definitions. They define the things they talk about.

Quine said that has to be wrong because when someone writes a dictionary and says that some word means something or other, they aren't supposed to make it up.

The definition is supposed to describe something that is already there, and the logical positivists didn't explain how the way mathematics got to be there is any different from the way that science got to be there. Then he backed up a little and allowed that sometimes definitions do other things. Some definitions clarify the meaning of words and other definitions are abbreviations. But even the definitions that are supposed to clarify are supposed to clarify things that are already there, and Quine didn't think that the abbreviations were all that interesting.

But that's the whole point. When you clarify something enough so that you know what the rules are, that's when you have mathematics. Any mathematician will tell you that coming up with good definitions is very hard to do. So Quine took the most interesting part of mathematics and made it sound too boring to talk about.

Mathematicians still think about important questions, but they are afraid that if they talk about them, they are doing philosophy, which is a waste of time. Philosophers think that if they try to be scientific about mathematics, then they are *not* doing philosophy, which, for them, is also a waste of time. So mostly we talk about things that don't matter, and when we talk about things that matter, we don't do it well. At least the mathematicians can go back to doing mathematics.

25.3 How to Fix Things

How can we avoid talking about things that don't matter? Sometimes it helps to look around and notice that nobody cares what we are saying. But that doesn't always work. Sometimes we tell ourselves that the reason nobody cares is that we are talking about things that are so deep and important that nobody else can appreciate them.

Another thing we can do is look at the history of mathematics. History is really interesting. When you read about how mathematics was done you see that people had very good ideas. They had to work hard to come up with the ideas. You can think about why they decided to talk about things the way they did and you can think about what makes the ideas they had so good.

It doesn't help to think about whether the numbers people used to talk about are the same as the numbers we talk about today and whether their words latched onto them in the same ways that ours do. It *is* interesting to think about how people used to talk about numbers and how we talk about numbers today and how our talk has changed. But that's not the same as thinking about how numbers have changed.

There is a guy I know who writes about mathematics. He has written about the history of analysis and the history of algebra and the history of geometry in the nineteenth century. He has written about where mathematics comes from and how it gets used in physics. He has also written about famous mathematicians and big ideas like modernism. Mostly he writes about what mathematicians thought and what they did. Sometimes he uses words like "ontology" and "epistemology." By that he just means the way people talk about things and think about them.

It would be nice if more mathematicians read about the history of mathematics. It would be nice if some philosophers read about it too, and even some people who aren't mathematicians or philosophers. Then we could all get together and talk about it. We could talk about mathematics and how it got to be the way it is. We could talk about why we like mathematics so much and what we like about it. We could even talk about how it might be different by the time our children and grandchildren are grown up and are doing mathematics on their own. It would be nice to talk. I am pretty sure we would like it.

Appendix

I am grateful to the editors for accepting this unconventional contribution to an otherwise scholarly collection. The title is an homage to Raymond Carver's short story, "What We Talk About When We Talk About Love." I originally set out to write an ordinary philosophical essay about the role of syntax and semantics in the philosophy of mathematics, but having chosen the title, it was hard to resist emulating Carver's narrative style. Doing so was liberating because it encouraged me to avoid overworn philosophical tropes and to formulate ideas as clearly and simply as I could.

The history of mathematics is a powerful philosophical tool, and thinking about what has changed and what has remained constant provides critical insights as to why we do mathematics the way we do. The guy in Sect. 25.3 who writes about mathematics is, of course, Jeremy Gray, who has always treated the history of mathematics as a history of ideas. His respect for the power of those ideas animates his work. I have learned a lot from him, and if this essay brings him a bit of enjoyment in return, it will have served its main purpose.

A secondary purpose was to explore the way that certain developments in twentieth century philosophy of mathematics have shaped the way we think about the subject. Carnap's influential "Empiricism, Semantics, and Ontology" (Carnap 1950) can be taken as an exemplar of the views attributed to the logical positivists in Sect. 25.2, and, of course, the counterpoint provided there is a summary of Quine's "Two Dogmas of Empiricism" (Quine 1951). Volumes have been written about the issues raised in these two publications, and readers can turn to the *Stanford Encyclopedia of Philosophy* (Creath 2021) for details and references. I also recommend Edmonds' recent book, *The Murder of Professor Schlick: The Rise and Fall of the Vienna Circle* (Edmonds 2020), for an engaging exposition of the historical context.

My goal here has not been to add to the debates, but, rather, to reflect on the way they have influenced the philosophy of mathematics. I find the things that Carnap and Quine have in common to be more striking than their differences, and I hope this essay makes it clear that I take their shared focus on the communicative and inferential norms of mathematics and the sciences to be an important philosophical advance.

But, curiously, this focus is not what drives the philosophy of mathematics today. This essay offers one possible explanation as to why not. It is easy to interpret Carnap's and Quine's portrayal of the relationship between philosophy and science as an implicit affirmation that the best thing that philosophers can do is to respectfully step aside while mathematicians and scientists do their work. It is not surprising that philosophers since then have resisted that conclusion and have instead turned their attention to puzzles in metaphysics and epistemology, topics that are comfortably within their wheelhouse.

This perceived dichotomy between thinking about mathematics and thinking about meaning, reference, and the nature of mathematical objects is unfortunate. There is a lot to be learned by paying attention to mathematics itself, and philosophers are well positioned to help us make sense of the norms, values, and goals of the practice. Mathematicians may be very good at doing mathematics, but that doesn't necessarily imply that they are good at thinking about what they do. Philosophers' training puts them in a position to assess the mathematical literature critically, analyze the conceptual and inferential structure, make sense of the implicit norms and expectations, study the means that mathematicians employ, and understand how they are suited to their goals. That requires familiarity with the relevant mathematics but not the same type of expertise. We still have a lot to learn about the nature of mathematics and its applications to science, industry, and policy.

Instead, an interminable focus on disconnected technical problems has had a devastating effect on philosophy of mathematics. A recent analysis of tenure-track positions advertised in *Jobs for Philosophers* in the 2021–2022 academic year doesn't even mention philosophy of mathematics in its categorization.[1] Digging into the data shows that the phrase "philosophy of mathematics" occurs in only three of the 201 advertisements, in each case listed among multiple areas of potential interest. Surely this is an indication that the field is no longer viewed as important. It is sad that a discipline that was so central to the philosophical tradition from ancient times to the middle of the twentieth century now barely registers a pulse.

But let me temper this doom and gloom with some more positive notes. First, colleagues assure me that the outlook for philosophy of mathematics is more optimistic in Europe, and I would not be surprised to learn that other communities have also managed to escape the gravity of the Anglo-American analytic tradition.

Second, whatever their long-term career prospects may be, a number of talented young people are throwing caution to the winds and finding ways of doing important work in the field. Meetings of the Association for the Philosophy of Mathematical Practice, of which Jeremy was a founding member, are lively and well attended. I only wish the organization would drop the phrase "philosophy of mathematical practice" in favor of "philosophy of mathematics." We should worry about philosophy of mathematics that *doesn't* have anything to do with

[1] https://philosopherscocoon.typepad.com/blog/2022/04/where-the-tt-jobs-werent-in-2021-22.html.

mathematical practice, and we should avoid depicting philosophy that does as anything less than the proper heir to a long philosophical tradition.

Third, there is still a lot to be done. Almost all of the philosophical papers I have written end with cheery exhortations to roll up our sleeves and get to work. See, for example, the last section of my "Reliability of Mathematical Inference" (Avigad 2021), which, incidentally, cites a number of the young philosophers alluded to in the previous paragraph.

Finally, there is considerable interest. At a time when it seems that every undergraduate is majoring in computer science, data science, or business, I still come across students from across the United States that are double-majoring in mathematics and philosophy. Rebelling against the segregation of science from the humanities, they are an encouraging reminder that there are still young people who find value in the scholarly traditions that have served us well for centuries. We would do well to support them.

At the end of the day, mathematicians are among the most philosophically inclined people on the planet. Dealing with creative flights of abstraction on a daily basis encourages constant reflection on the nature and meaning of their craft, so there is still room for a philosophy of mathematics that does the subject justice. All we need to do is take stock of where we are and where we want to be, and then figure out how to get from here to there.

References

Avigad, J. 2021. Reliability of mathematical inference. *Synthese 198*(8): 7377–7399.

Carnap, R. 1950. Empiricism, semantics, and ontology. *Revue Internationale de Philosophie 4*(11): 20–40.

Creath, R. 2021. Logical empiricism. In *The Stanford Encyclopedia of Philosophy*, ed. E.N. Zalta, Winter 2021 edn. Stanford: Metaphysics Research Lab, Stanford University.

Edmonds, D. 2020. *The Murder of Professor Schlick: The Rise and Fall of the Vienna Circle*. Princeton: Princeton University Press.

Quine, W.V.O. 1951. Two dogmas of empiricism. *Philosophical Review 60*(1): 20–43.

Part VII
The Making of a Historian of Mathematics

Chapter 26
History Is a Foreign Country: A Journey Through the History of Mathematics

Snezana Lawrence

Abstract This chapter has three strands that are interwoven together to trace a narrative showing how understanding of a contextual development of a mathematical concept or a technique is a necessary step in uncovering the historical development of the same. To generalise, three strands are a context and invention of an original mathematical technique; historian's own context and interpretation at a time of research and writing of a historical thesis; and finally, a dialogue between such interpretation and different audiences. In particular, the mathematical technique is that of Descriptive Geometry of Gaspard Monge; the research and interpretation that I made as I undertook during my PhD studies under the supervision of Jeremy Gray to whom this volume is dedicated is the second strand of the chapter; the lessons I learnt from communicating such interpretation to my audience of the time – Jeremy and my examiners, is the third strand.

This narrative is not intended to serve as a model of historical research and/or writing to be recreated by future PhD students or mathematical historians. It is but an example of what worked well to uncover some interesting new roles geometry played in nineteenth century England and France. It is a method that worked for me, but I hope there will be lessons for the new and aspiring historians of mathematics in this narrative too. It is also a personal tribute to my PhD supervisor, whose friendship and intelligence contributed not only to my academic success but to my continuing personal and professional development and wellbeing.

Theme: Contextualising mathematics in its social and material environments.

S. Lawrence (✉)
Department of Design Engineering and Mathematics, Middlesex University, The Burroughs, London, UK
e-mail: snezana@mathsisgoodforyou.com

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*, Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_26

26.1 Introduction

In the spring of 1991, I was completing my degree in architecture, and my little daughter was going to the first grade of primary school in Belgrade, the capital of then Yugoslavia. I was on a trip to Greece during Easter break. Whilst returning from this trip, we stopped at a restaurant on the border between Yugoslavia and Greece. As my friends and I we spoke over dinner, people next to our table asked us whether we heard of the news. They were from Netherlands and said that 'the war had started in Yugoslavia' that day. Of course, this was worrying, but we thought that they had misinterpreted the news they heard on the radio earlier in the day. The next ten years showed us that this piece of news was indeed true.

Months later, I found myself with my daughter in England. Finding work was difficult as UK was at the time having an economic crisis and architectural firms were making people redundant. They were not employing new architects and technicians at all, let alone those who just happened to drop in from another country and had no idea of British regulations or its building industry.

The only thing I could think of and the only thing I loved doing apart from architecture was teaching descriptive geometry which I did as a tutor for some years. Descriptive geometry was a popular academic subject in then Yugoslavia, and is still taught in many countries around the world, although not in France from which it originated (Belhoste 2003). As a first step towards achieving this goal, I thought I should find an English language textbook to pick technical terms related to descriptive geometry so that I could teach it well. This is how my introduction to the history of mathematics started.

For young readers, one must put a reminder here that this was all happening in the time before the internet. There was no Amazon, no online repositories of old books, and no internet forums where you could just ask any type of question like the one that was preoccupying me: 'why aren't there any books on descriptive geometry in England?'. In absence of these ways of finding answers, I went to my local, small-town library, travelled to see larger libraries in some larger towns near the one I lived in at the time, and then, one day by accident, stumbled upon a 'historian of mathematics' in a national newspaper. I perked up at first, as I didn't know one can be a historian of mathematics, and then thought that this historian, Professor David Singmaster, from the Southbank University in London, could perhaps, just perhaps, know how and why there weren't any new and more importantly old descriptive geometry books in England. If these books existed at all, they couldn't have all been sold or disappeared. But not a single book was to be found in any of the new or second-hand bookshops I visited, or any of the libraries I went through in the Southeast of Kent, England. I wrote a letter to David and told him simply and very briefly my story. I came to England from Yugoslavia; I wanted to teach descriptive geometry; I couldn't find any textbooks to get my language up to scratch; does he know why there aren't any old books on the subject?

To my surprise, the answer came quickly and was incredibly helpful and encouraging. David said that he had no idea, and that indeed as he was an American, he had heard of descriptive geometry. He said it was a well-known subject in the US; if I was there, he said, I could find some old books on the subject. But he had, he said, a colleague who would be able to say something on the history of the subject in England. He suggested I get in touch with Jeremy Gray from the Open University.

From this letter a few months passed. I looked Jeremy up in the local library and got his *Ideas of Space* (1979) first. From this book I learnt more of the history of non-Euclidean geometry which intrigued me. I followed this by reading, for the first time, an original text, Halsted's translation (1896) of Bolyai's appendix (1896) and literally felt like my mind had opened to a new world that I had no idea existed. This world of old mathematics texts took me to a different intellectual space to the one I existed in until then. It was the world in which I could learn, first-hand, from the people who were the first to realise some mathematical truths. Contrast that with the re-interpretation, summaries, and ambiguities I met with when learning about mathematical inventions! I was very excited about this. I realised that the old mathematics books would lead me to learning a little more about the history of mathematics, and that this can indeed, be an answer to my trying to learn how to teach descriptive geometry (and find a book in English language on it!). Later on, I also realised that this new interest in the history of mathematics in general had replaced the original interest or rather a desire to teach descriptive geometry. This realisation crystallised more clearly over the coming years.

To get back to the mid-1990s, I got in touch with Jeremy. I don't remember how I did that. I must have written a letter and explained all of this in my pretty broken English at the time. But making contact with Jeremy marked the beginning of my work in the history of mathematics.

26.2 The PhD Journey

I met Jeremy Gray and John Fauvel (1947–2001) one spring morning of 1994 at the Open University in Milton Keynes. We spoke about why there weren't any books on descriptive geometry in England and why the subject was never taught in the school system, unlike most of Europe, and unlike Eastern Europe and Yugoslavia where the subject still survived since the ninteenth century. Jeremy and John were very hospitable, took me around the university, and bought me a lunch. And then they said that it may be an interesting topic to investigate further. Would I consider doing some research on it? Of course I would – what more interesting could I imagine doing at the time? They suggested I could even register to do it as a PhD student. I was worried about that for two reasons: I didn't think my English was good enough to write a thesis, and I had absolutely no spare money for the university fees. We left it at that, and I said I would think about it.

The war in my country was raging by this time, and I realised I wasn't going back there any time soon. I had a job as a teaching assistant in a sixth form college, and was also teaching a couple of courses to adults: one was a course on computer aided design, and the other on foundation mathematics. I made some friends at this college, and one friend was very happy to drive me several times to the library of the University of Kent in Canterbury (from Deal where I lived at the time). I spent hours in this library trying to find and read some old mathematics books and some books on the history of mathematics. This friend and I used to walk around the shelves, picked the books, read them. We spoke about them. He persuaded me that my English was good enough and would get better if I read and wrote all the time. It was already 'getting better as we spoke' he exclaimed! He also said one day 'Let's look at that application form [for the Open University PhD programme] and everything it entails'. I am very indebted to this friend, who has sadly passed since. His name was Chris Cooper. He was truly a gentle-man. Without his support I would not have dared to write that application form, but I did, and I started my PhD studies.

26.3 Finding the Magic Word

So many diverse tribes and sticks have contributed to the formation of the English nation that it is not easy to draw a line between the native and the foreign elements. After all, the Jutes and Saxons and Angles were themselves immigrants, who came to this island in historical times; the main stock was transplanted, and is no more native to the soil than the branches which have been grafted into it from time to time. It seems a little arbitrary to fix on any definite date and designate the immigrants of the earlier times, component parts of the English race, while we speak of the later arrivals as aliens.

William Cunningham (1897). *Alien Immigrants to England*. Macmillan Co, p. 3.

Slowly I started finding my way around English libraries and educational system, but also on another more general existential level, finding my way around England. Through reading and writing notes on all the old mathematics books I was able to find, I was learning about the development of mathematics in England, and I was learning about descriptive geometry and why it wasn't accepted here. Finding answers to my questions about mathematics history began to make me also feel clearer about my predicament at the time. I realised this feeling of becoming to feel more settled was due in great part to learning the history of the new place in which I found myself. So, I pursued this path, as it took me towards the place that, although I didn't know it until then, started to feel like home.

I started PhD about three and a half years after I arrived in Britain, and the first year I was doing it I was a part-time student. Somehow Jeremy and John managed to get my fees reduced to an absolute minimum – I have no idea how they managed to do that. For the Christmas period of that year, Jeremy offered that I stay, with my daughter, at his sister-in-law's house in Oxford as a house-sitter. This would have given me an opportunity not only to explore the city, but also to see whether there is anything interesting in the Bodleian Library from the period I was interested in (late eighteenth, early nineteenth century). I jumped at the opportunity and was happily registered with the Bodleian Library where I spent time reading for a week or a little more over this period. I remember how quaint I found the admission system at the

Library – not only did I have to provide the usual documents so that they could see who I was, but I had to read aloud the text they gave me: a promise that I will not damage anything I see or read whilst in the library.

I was very lucky over that short period of time. Not only did I find new things in the Bodleian Library about descriptive geometry and its translations into English language – a few sparse and incomplete translations like the ones by Hall (1841), and Bradley (1861), but I also came across the work of William Cunningham (1849–1919), a Scottish historian. To be more precise, Cunningham was a Scottish historian of economics, and is credited by some as one of the inventors of economic history as a scholarly discipline. He wrote, for example, on *Growth of English Industry and Commerce in Modern times* (1882) and *Growth of English Industry and Commerce in Early and Middle Ages* (1890). Cunningham became a professor of economics at King's College London where he taught 1891–1897, later taught at Cambridge, and for a time also at Harvard. He was one of the founding fellows of the British Academy and was towards the end of his life appointed as Archdeacon of Ely, one of the largest, most beautiful, and rather important centres of Anglican Church in England (Fig. 26.1).

Cunningham's paper on descriptive geometry was one of his early papers in economic history, and I found it in the Bodleian. This paper, or a pamphlet, was published in Edinburgh in 1868, and set out to trace some *Notes on the History, Methods and Technological Importance of Descriptive Geometry, compiled with reference to Technical Education in France, Germany & Great Britain*. My research by then had already showed me that there were differences between how descriptive geometry was perceived, narrated on, and taught in different European countries in the early to mid-nineteenth century.

First of all, once a treatise on descriptive geometry was first published in France during the Revolution, in 1795 (Lawrence 2002; Sakarovitch 2005), it was soon

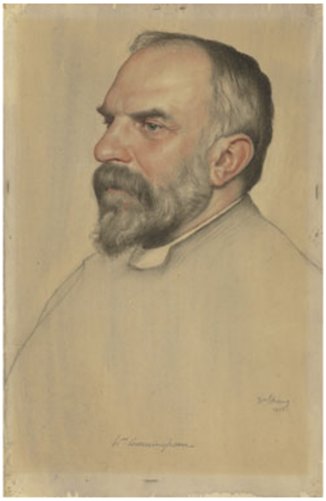


Fig. 26.1 William Cunningham, drawn by William Strang. (Creative Commons CC by National Galleries of Scotland)

realised not only in France, where its inventor Gaspard Monge (1746–1818) had a role in establishing the new system of education, but elsewhere (Barbin et al. 2019), that descriptive geometry could be used as a principal tool of graphical communication. It was undoubtedly the most coherent system of all available at the time in this respect, but various texts on it published in England, and in English language more broadly, seemed to have gone out of the way to find some faults with it (Lawrence 2002, 2019). So, Cunningham's text came as a real surprise to me at the time – surprise in the sense that someone else had also felt that there was such a difference in the approach to the technique between the French and the English.

There was a little nugget of genius in what Cunningham spotted, without ever having, to my best knowledge, done descriptive geometry himself. He didn't concentrate on politics (although as the time descriptive geometry was first taught England and France were intermittently at war – 1803–1815), but instead on a crucial difference between how space was communicated by using descriptive geometry in England and elsewhere through the analysis of the language used to describe procedures in the textbooks on the subject. He identified, most importantly, that the terms that the English books used, 'plan and elevation' were not actually native to the original technique of Monge. His view was that

> *it is impossible to express the co-ordinate relation of the two planes of projection in such terms as Plan and Elevation, which involve the special ideas – Horizontal and Vertical.*
> William Cunningham (1868) p. 25.

Instead, Cunningham suggested the use of the word 'rabatting' based on the original French verb, *rabattre* which means to turn, to turn down, to pull (Fig. 26.2).

Having found the simple explanation between these different perceptions that I couldn't, until then, articulate so well, I was very happy to report on my finding next time when I met Jeremy for a tutorial. 'Hang on' Jeremy said, 'I've never heard of that word before. When did you say this word was used, what was the publication date of this paper?' And so, we looked at the *Oxford English Dictionary*, the 'definitive record of the English language' which gives the instances of the use of each word as given in print for the first time, with examples. And we noticed that the date of the first use of this word, *rabatting*, was in fact noted wrongly, some 14 years after the paper Cunningham wrote it in. I wrote a letter to the Editors of the OED, and got a reply reporting that they will amend this error in all the subsequent editions of the *Dictionary*.

Finding this little mistake about this rarely used word in English language, became an important point of persuasion for the Open University to award me a grant to start doing my PhD full-time. And from there on, I was registered as a full-time student (Fig. 26.3).

26.4 The New Worlds That Research Leads You to

> Unlike the vague lineaments of times ahead, the fixed past has been sketched by countless chroniclers. Its vestiges in landscape and memory reflect innumerable details of what we and our predecessors have done and felt. The richly elaborated past seems more familiar

than the geographically remote, in some respects even more than our own nearby present; the here and now lacks the felt density and completeness of what time has filtered and ordered.

David Lowenthal (1985). *The Past is a Foreign Country*. Cambridge University Press, p. 3.

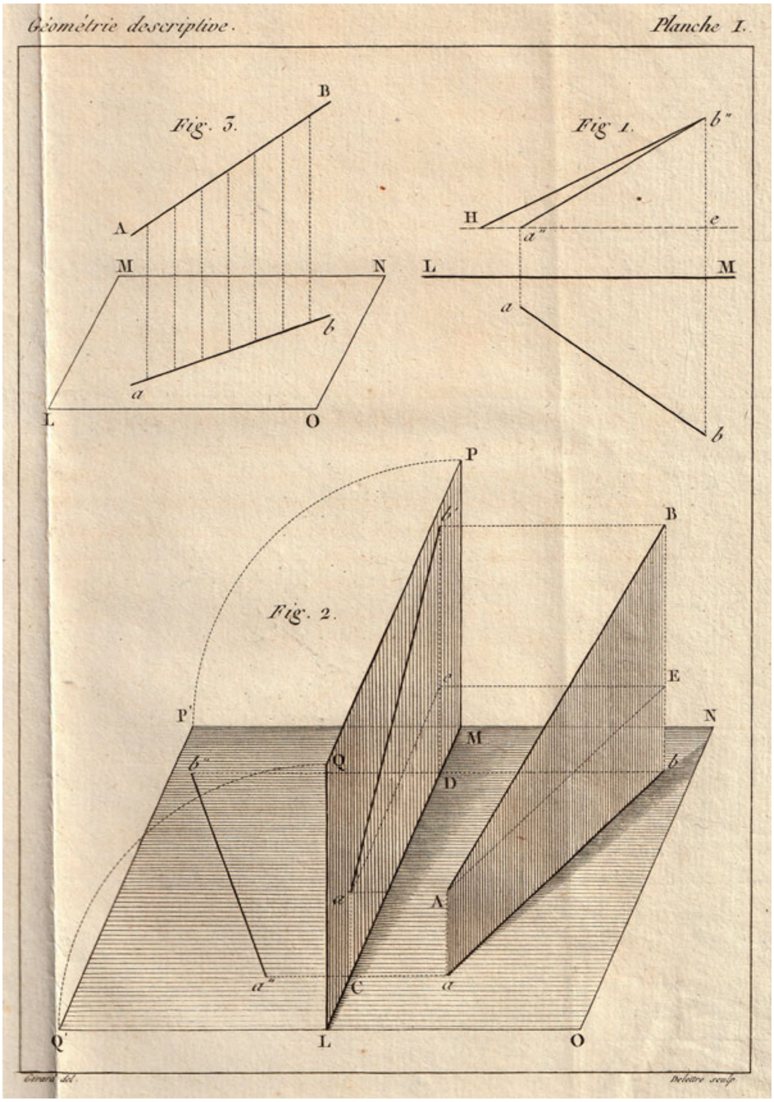


Fig. 26.2 (Bottom of the plate) from Monge's *Géométrie Descriptive* (1795), showing how a line in space, AB, can be defined by its projections onto horizontal and vertical planes, and how, in order to extract measurements, the vertical plane can be considered revolved to lie flat in the horizontal plane. This describes the process of rabatting, or bringing the plane in which a particular length lies into the projection plane, in order to get the real length of a segment in question

> Now this is quite incorrect. For, assuming that an elevation means a projection on a vertical plane, how can we, with any propriety, talk of projecting the line A B upon a plane in which it actually lies—that is to say, its "plan-"projecting" plane ? No, the operation referred to is correctly described as "rabatting the plan-projecting plane of "the line," or simply "rabatting the line." By the use of this term the student sees at once that he can rabatt a line A B in space either upon the horizontal or vertical plane at pleasure, whereas, if he calls the former operation "making "an elevation," he is little likely ever to perform the latter; for the simple reason that he has no name to give to it.

Fig. 26.3 Cunningham's use of the term 'rabatting' and phrase 'to rabatt' is its first occurrence in the English language and directly derived from the French. (Cunningham 1868, p. 25)

At some point after I began to work on my thesis, I started also to see that what was available on the surface, did not always agree with the snippets of the internal narratives of the subjects of my research as I came across them. In particular, this related to the original works or the unpublished papers of authors such as Peter Nicholson (1765–1844), Scottish architect and mathematician, who worked with people like Sydney Smirke (1797–1877) the English architect and one of the designers of the British Museum in London. Apart from the interest and contribution Nicholson made to introducing and modifying descriptive geometry into a graphical communication in English publications, his social network and correspondence between some of the people in that network, showed a deeper underlying interest in the stone-masons' craft. Their language and the objects they were interested in, as well as the institutions they belonged to or helped establish, began to feel quite different to the external history. I started unearthing new things I haven't heard of before (Lawrence 2002, pp. 69–96). I began to look further through their correspondence and their social networks, in order to try and see what was *really* going on with these people almost 200 years ago. A lesson for a new student in the history of mathematics is to see whether they can find any correspondence of the people they are most interested in. Letters are a treasure trove for historians: they are primary source for the researcher to find directly 'from the horse's mouth' what people were living through, at the time.

In this process there is, of course, a danger that you begin building a 'museum' to house all your nuggets of understanding and feelings that you collect about the past you are investigating. In the process you then may become little like a tourist who buys a souvenir but doesn't really remember the real place they visited. Lowenthal (1985) rightly pointed out that there is some escapism in the history in that respect. Certainly, the history was becoming clearer to me than my own life and situation at the time. In my private life, everything seemed to be in a flux and I had no idea where things would lead eventually. But in an increased clarity about the history of descriptive geometry in England in the nineteenth century, every single little thing that didn't 'fit' was therefore becoming a nuisance that had to be investigated. Not that past had to be 'clean' from these imperfections of the things that I didn't understand, but rather I realised there was a huge gap between the official history, the recorded history, and what people were really involved in and what mattered to them most.

By this stage on my PhD journey, I realised that there was an underlying animosity between England and France at the beginning of the nineteenth century. Napoleonic wars (1803–1815) should certainly have given me a hint of what was going on, but in all seriousness, I could not believe that this would have influenced mathematicians in terms of doing mathematics. I did, however, begin to realise that the overall, public, general picture and the publications that came out of this period may have been influenced by the perceptions or even more importantly by the *expectations* of certain perceptions and that for sure, these were partially influenced by the political and social life of the time. But underneath there seemed to have existed an underlying structure of communication, even friendship, and to certain extent a set of undeclared connections that didn't seem to relate to anything I knew about.

26.5 National Lore and an International Movement

> Q. Why was you made a Fellow-Craft?
> A. For the sake of the letter G.
> Q. What does that G denote?
> A. Geometry or the Fifth Science.
> Pritchard (1730) *Masonry Dissected*, London, p. 12

Thus came the secondary stage of my PhD research. At this point I began to know more about the connections certain people I came across in my research so far had, in the most unexpected ways. These connections all in some way related to how descriptive geometry was used or translated between England and France, and some were in fact part of the movement of establishing the emerging architectural and engineering professions in both countries across the Channel. What I wanted to know was how would, for example, certain individuals be able to get hold of a copy of Monge's book from the Revolutionary period, and even study it and make their own inventions related to it (Lawrence 2002, 115–128) when England and France were most of the time in a war, in this period?

These kinds of questions led me to explore geometry in a wider context and to see the role it had in the, for the want of a better word, its 'underground' existence in eighteenth and nineteenth century, and in particular its role in the civil society movement of Freemasonry. Within this context, geometry, and especially descriptive geometry as it was originally related to the stone-masons' knowledge and craft – as did Freemasonry itself – was viewed as something almost as a

Fig. 26.4 An engraving published in 1734, showing a 'Master-Mason rare, Whose mystic Portrait does declare, The Secrets of Free Masonry...'. It shows the main symbols of Freemasonry and their applicability to initiates' life philosophy. (United Grand Lodge of England collection)

sacred knowledge in increasingly secularised societies in both France and England. In France of course, in the aftermath of Revolution, even Freemasonry became different to that of England, so that French Freemasons did not to need to believe in the 'Supreme Being' or any deity for that matter. But both French, and English 'branches' of Freemasonry, played an important role in establishing new institutions of education (Cumming 1954), And in England, the University College London, was the first university in the country to be based on religious un-denominationalism, and was, at the time, even considered a Freemasonic project as it was opened by the then Grand Master, Duke of Sussex (Augustus Frederick, 1773–1843) in full Freemasonic regalia in 1829.

It intrigued me to see that Freemasonry seemed to have become an international force towards the latter part of the eighteenth century and at the same time drew ontological links with architecture as a profession and geometry at the core knowledge of architects. I spoke to Jeremy about this, and the only thing he could recommend was to go into the Freemasons' Hall in London, the first Grand Lodge of the world, and try to see what they had in their archive. So I did. I managed to get access to the library and archive of the United Grand Lodge in London, and over painstaking work over some months had dug up traces of a network of connections in architecture and engineering no one had known about until then (Lawrence 2002). I also found that Monge himself was an enthusiastic Freemason, as well as his much less known and influential counterpart in Britain, Peter Nicholson (Lawrence 2002; pp. 101–153) (Fig. 26.4).

But here we come to the point where Jeremy would have happily left me to it rather than followed as a supervisor on my new research trek. Although I could see that a number of scholars globally at approximately the same time were coming to similar conclusions and wrote a number of books on the subject, most notably Jacobs (1983) and who followed with more on the same theme post my PhD (Jacobs 2005, 2006), I could see how it could look like in our then contemporary context. This too can't be forgotten when describing a historical research or the contemporary scene from which a historian works. At the time, in the late 1980s and early 1990s, there was a political affair in the UK that negatively impacted the image of Freemasonry. For the sake of keeping mathematics in the spotlight as well as the historical rather than contemporary Freemasonry, the full description of this affair is avoided in this chapter, but suffice it to say that in the UK at least, this became a very contentious issue and the whole organisation began to also be seen in the light of what was going on at the time. And so, it wasn't a surprise to me when Jeremy, on a number of occasions around this time, had started our pre-dinner tutorials at his home with this metaphor:

> Mathematics is a swamp in which there are many stagnant pools, and through which there wind many narrow, twisting paths. It is the teacher's job to escort the students along the narrow, twisting paths and one by one to push them into the swamp.

According to Jeremy (but after an unsuccessful search for the original quote, hence not evidenced further) this quote is taken from Abram Samoilovitch Besicovitch (1891–1970). In retrospect I think this metaphor came to Jeremy as I entered these waters that he considered probably a little murky for his taste. But we did come out at the other end, together, with the finished thesis. I had managed to understand the networks, developments, schools, establishment of professions of architecture and engineering, and the role Freemasonry played in all this, in both England and France in the nineteenth century. More importantly I had also learnt that there, indeed, were only a handful of books on descriptive geometry written in Britain, somewhere in the middle of the nineteenth century. The technique never took hold in England, and it never became properly understood. In its place were introduced techniques similar to descriptive geometry, but not quite it – like the one invented by Peter Nicholson (Lawrence 2002, 2019). Whilst its application in France (and other places in which it was taken up), see in particular Barbin et al. (2019), included applications to other branches of mathematics, analysis for example, in England it became for ever known as a little more than a geometrical drawing technique (Lawrence 2019).

26.6 The Inner Circle

I completed my PhD in 2002. I submitted it a year earlier, but in that year many things had happened. John Fauvel who originally also helped with getting through the administration for my full-time grant to materialise, sadly died. A number of other things happened by then in relation to my status too. I was granted an indefinite leave to remain in Britain; the war in Yugoslavia was finished; my daughter and I had contemplated going back but decided that this – Yugoslavia and our life there – had become a part of our past. We had settled in England, had got a little house on the south coast, and both felt very much British. In fact, we became officially British in the early 2000s.

In Britain however, the positions for historians of mathematics had not materialised in the meantime bar the one at the Open University and a couple in Scotland. Rather than thinking of what else to do, I plunged myself into education and enjoyed an enormously happy and successful decade of teaching mathematics to secondary school pupils. Sometimes the students would ask me what history had to do with mathematics, but in time I became better at teaching, and they became more and more enchanted with mathematics through its history. Jeremy followed me all the way as a supporter and an observer, but mostly a very good friend. Whenever I needed a reference in these times, he had been right there by me. In time, the number of projects in the history of mathematics and their role in education multiplied in my work. I had various projects on integrating history in teaching of mathematics at primary and secondary level in the late 2000s. I eventually moved into the teacher training, and after a decade of working there, I moved into Higher Education, where I have happily found a new, academic home, since 2009. Currently I work at Middlesex University in the Design Engineering and Mathematics Department and try to include history of mathematics in all the modules on which I teach. One of these modules itself is partially dedicated to history of mathematics, and partially to communication of mathematical concepts. Internationally I am perhaps even more recognised for my work in this area, and have been voted in as a Chair of the International Study Group in the History and Pedagogy of Mathematics (an affiliate group of the International Mathematics Union) in 2020 (I finish my mandate in 2024).

In the words of a colleague who once told me why I should make sure I completed my PhD, especially in the times when I wasn't sure I could sustain this financially, I always remembered that the PhD is indeed a necessary step to join the 'inner circle'. It is rare that you will be considered a professional historian of mathematics (or indeed academic in any discipline) without it. I'm happy I followed this advice.

26.7 Some Suggestions for New Researchers

The journey I have undertaken was unique to me and no one else can have exactly the same experience whilst becoming a historian of mathematics, and therefore won't have the same outcomes. However, there are some steps that can be identified for those starting up in the history of mathematics that I hope would be useful for the novices: and by this, I mean to those who are new to the history of mathematics.

Firstly, the original question sustained my interest and research throughout the period of starting with, and completing my PhD. There had been many a difficult moment during that time and many opportunities to give it all up and start something else, but it was that original question that I really wanted an answer to, not only for myself, that sustained me. This was because the question, and its answer, said something about how mathematics is made, discovered, communicated, disseminated, and learnt. So the conclusion from my experience is that the core, central research question, if you are at the point when you are contemplating a PhD work, should be the one that is crucially important to *you*, and that you believe would be crucially important for others to hear about it too. Of course that is probably a definition of a PhD: you should believe answering your research question can contribute, in a significant way, to the scholarship if you were able to answer it. In my case, the answer turned out to be something else than what I thought it was originally, and even my original question had slightly changed, but nevertheless the intensity of my desire succoured me through difficulties that many doctorate students have: of persisting with PhD studies, writing it all up, and not getting a better paid job in which you wouldn't have the time to complete what you had started!

Secondly, the history of mathematics is a field that is heavily reliant on your knowing both mathematics and history. Neither is more important, and in some cases, you will be inspired by history, and in some by mathematics. What I learnt to be even more important than the two disciplines in a way, is how you weave these two strands into your history of mathematics narrative. You don't have to be a creative mathematician, and you do not have to have, currently, a training of a historian. But you do need to know the branch of mathematics you are researching pretty well in order to understand the nuances of different approaches or material culture you come across. In my case, this entailed me being able to understand before I came across the proof, that there were crucial differences between the ways space was understood if it was to be communicated by the technique I was interested in (descriptive geometry) in two different countries. If you are a student thinking of beginning this type of journey, you also need to read about various methods of history. Perhaps the best one I could recommend as a starter is an old gem, the little book by Collingwood (1946).

Thirdly, the material culture and context are important for the history of mathematics not only because they show you how people lived and thought in the time of your interest, but because they are the remnants of the space they created for themselves during their lives. The material remnants of past times literally have the power to take you to your research subjects' metaphorical, and sometimes, their real homes. These artefacts and their reconstructed 'totality' – how they were used, how they were made, what they were made for, and why they mattered to the people – was what surrounded *your* people in *their* time and space. It is this material culture that can transport you to their world, to a place they created and made with love. In the case of my research, this material culture was heavily outside of the scope of my original question, but understanding it was crucial to answering the original question. In other words, even if the research takes you to a place you never thought it would, you should try to venture and understand it, especially if it will enlighten, and illuminate in some way, your original question.

26.8 An Epilogue: The History Is a Foreign Country

Oh call back yesterday . . . bid time return.
Shakespeare, Richard II, 3.2.

The opening phrase of the novel *The Go-Between*, by L. P. Hartley 1953, is 'The past is a foreign country'. It is a novel in which the childhood memories are revisited by a man, who in his childhood tried to help two lovers by passing messages between them. The love affair had disastrous consequences because of the context and social norms of the times, and only with the reflection undertaken with the passage of time, the main character understands all the forces that contributed to this outcome. Without the love affair, or the going-in between, I often thought how I turned to history because of a disaster of the war and of losing my country. Trying to find something personal in a landscape and culture that was completely foreign to me as I landed in England, certainly became interwoven with my original very pragmatic desire to find employment by doing something meaningful that connected not only with my skills and abilities but with what I brought with me from the previous, happier times of my life. I had a great luck to stumble upon Jeremy. It happened that, additionally to all the benefits of finding what I could do, settling down, and finding good employment, I also had a great friend in this quest for meaning. My supervisor Jeremy Gray and his whole family without exception, accepted me, and my search, without question, without doubt, with gentleness of spirit and great intelligence and kindness, and made me realise that home really is where your friends are.

References

Barbin, Évelyne, Marta Menghini, and Klaus Volkert, eds. 2019. *Descriptive Geometry, the Spread of a Polytechnic Art: The Legacy of Gaspard Monge*, International Studies in the History of Mathematics and its Teaching (ISHMT). New York City: Springer.

Belhoste, Bruno. 2003. *La Formation d'une technocratie : l'École polytechnique et ses élèves de la Révolution au Second Empire*. Paris: Belin.

Bolyai, John. 1896. *The Science Absolute of Space, Independent of the Truth or Falsity of Euclid's Axiom XI* (which can never be decided a priori). Translated from the original of János Bolyai, 1832, "Appendix scientiam spatii absolute veram exhibens; a veritate aut falsitate axiomatis XI Euclidei, a priori haud unquam", by George B. Halsted. The Neomon Press.

Bradley, Thomas. 1861. *Elements of Geometrical Drawing or, Practical Geometry, Plane and Solid Including Both Orthographic and Perspective Projection*. London: Chapman and Hall.

Collingwood, R. George. 1946. *The Idea of History*. Oxford: Oxford University Press.

Cumming, Ian. 1954. Freemasonry and Education in 18th century France. *History of Education Journal* 5 (4): 118–123.

Cunningham, A. William. 1868. *Notes on the History, Methods and Technological Importance of Descriptive Geometry, Compiled with Reference to Technical Education in France, Germany, and Great Britain*. Edinburgh: Edmonston and Douglas.

———. 1882. *Growth of English Industry and Commerce in Modern Times*. Cambridge: Cambridge University Press.

———. 1890. *Growth of English Industry and Commerce in Early and Middle Ages*. Cambridge: Cambridge University Press.

———. 1897. *Alien Immigrants to England*. London: Macmillan.
Gray, Jeremy. 1979. *Ideas of Space: Euclidean, non-Euclidean, and Relativistic*. Oxford: Oxford University Press.
Hall, T. Grainger. 1841. *The Elements of Descriptive Geometry: Chiefly Intended for Students in Engineering*. London: J. W. Parker.
Hartley, L. Poles. 1953. *The Go-Between*. London: Hamish Hamilton.
Jacobs, Margaret. 1983. *The Radical Enlightenment: Pantheists, Freemasons and Republicans*. London: George Allen & Unwin.
———. 2005. *The Origins of Freemasonry: Facts and Fictions*. Pennsylvania: University of Pennsylvania Press.
———. 2006. *Strangers Nowhere in the World: The Rise of Cosmopolitanism in Early Modern Europe*. Pennsylvania: University of Pennsylvania Press.
Lawrence, Snezana. 2002. Geometry of Architecture and Freemasonry in 19th century England. Thesis submitted for the award of PhD in History of Mathematics, Open University, UK.
———. 2019. Descriptive Geometry in England: Lost in Translation. In *Descriptive Geometry, the Spread of a Polytechnic Art: The Legacy of Gaspard Monge*, International Studies in the History of Mathematics and its Teaching (ISHMT), ed. Évelyne Barbin, Marta Menghini, and Klaus Volkert, 313–335. New York City: Springer.
Lowenthal, David. 1985. *The Past is a Foreign Country*. Cambridge: Cambridge University Press.
Monge, Gaspard. 1795. Géométrie descriptive. In *Les Séances des écoles normales recueillies par des sténographes et revues par des professeurs*. Paris.
Pritchard, Samuel. 1730. *Masonry Dissected*. London.
Sakarovitch, Joël. 2005. Gaspard Monge, Géométrie descriptive, first edition (1795). In *Landmark Writings in Western Mathematics 1640–1940*, ed. I. Grattan-Guinness, 225–241. Amsterdam: Elsevier.

Chapter 27
Reflections

Jeremy Gray

Abstract The accidental story of how I became a historian of mathematics, from the University of Warwick and back via the Open University, and met some of the colleagues and friends I acquired along the way.

I didn't set out to become a historian of mathematics. I graduated from Oxford in 1969 and went to Warwick to do an MSc. and a Ph.D. in mathematics. The MSc. went well enough, but almost immediately on embarking on the Ph.D. things fell apart between me and my supervisor. Inspired by the general folly of the times (it was 1970, man) I tried to go it alone—it wasn't clear actually who else could have supervised me, but nor did I push back hard enough—and predictably I dropped out. It was the first big failure of my academic life, and very depressing. I moved to London in 1972, got a part-time job at what was then called South Bank Polytechnic—the upside of the early 1970s was that there were jobs—and wondered what to do next.

My MSc., plus a willingness to look at the history of mathematics, enabled me to get a one-year job as a production assistant at the BBC to work on programmes being made for the upcoming course on the subject at the Open University. This was at Alexandra Palace, the BBC's first TV studios, and just up a hill from where I lived. My contract ended in 1974, and I moved to the Open University as a course assistant on that course. The OU was the last of new universities, and it showed. The central staff were people who had not been caught up in earlier appointments to the new universities of the 1960s, and they were a mixed bag. The production model was also strange. The idea was there would be a small core of staff who would be permanent, and would decide there should be a course on, let's say, history of mathematics. One of them would perhaps write a detailed outline, and a course assistant would be hired to produce the full version; this job was not permanent, but the course would be. By the time I joined it was already clear that this model didn't

J. Gray (✉)
Open University, School of Mathematics and Statistics, Milton-Keynes, UK
e-mail: jeremy.gray@open.ac.uk

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7_27

work; we were never able to produce courses as planned and on schedule. But for the course assistants, this was life-saving. Already there weren't so many jobs out there, so to be the person writing a course that had been commissioned by someone else who couldn't do it on their own was a way to stay in academia. And these course assistants were recent Ph.D.s who knew about research, and whose career depended on doing more. Finally, the university did the decent thing and gave these people tenure, and in 1978 we all became lecturers.

The OU's first course on the history of mathematics was already running, its best material being on ancient mathematics, written at short notice by Bartel van der Waerden, and the early history of the calculus by Henk Bos. In the late 1970s, I had the opportunity to produce an entirely new course with John Fauvel, who had also passed through Warwick, and that led to the source book *The History of Mathematics: a Reader* (1981), which sold surprisingly well for more than a decade. Sadly, John died in 2001. He could make anything interesting, or maybe he had a magpie's eye for the interesting bits, and I learned from him to appreciate the importance of elementary mathematics and how it is integrated into the social life around it. It's another approach to the history of mathematics, much less concerned with elites, but in his hands often revelatory. Various pressures on the OU led to the end of the course, and since then, in one of my long-running collaborations, June Barrow-Green, Robin Wilson, and I have been rewriting and updating it as a two-volume book, keeping its reliance on the study of sources in translation, and we make a good team. One volume came out in 2019 and the second and final volume came out at the end of 2022.

There'd been a second-year lecture course at Warwick on hyperbolic geometry in the early 1970s, and like all the graduate students I'd help tutor it. The lecturer had included historical information, which he took from Bonola's book on non-Euclidean geometry (Bonola, 1912) and I thought I'd look at it. Up to page 80 it's full of famous names getting nowhere: one fallacious argument in defence of the parallel postulate after another. After page 80, in addition to the names of János Bolyai, Lobachevskii, and Gauss, are people I'd never heard of, almost all doing well, and it seemed to me that the magic ingredient must be described and presented as such somewhere in the book. Well, it's not; Bonola didn't seem to appreciate hyperbolic trigonometry, and it just slithers in unannounced. I checked what historical literature I could find, and realised that no-one had put the story together the way I wanted to, so I had a question and an original answer. I wrote it up, with some helpful editorial advice it was published, and I started to become a historian of mathematics.

In fact, it wasn't my first published article. That was on Johann Lambert, a fascinating figure we ought to know more about. I'd met Laura Tilling at a conference of the newly founded British Society for the History of Mathematics (BSHM), where we commiserated over what happens if you give a talk with Tom Whiteside in the audience and have to face his absolute confidence that the evidence for one of your claims is not as you have described it. I had given a talk on Lambert, and Laura and I teamed up to write a piece on him (Gray and Tilling, 1978). This was much harder for me; Lambert's eighteenth century German and the German script severely taxed the one year's German I'd done at school.

The BSHM had been set up in 1971, and it was run by a number of elderly academics, as they seemed to me then, with professional positions in the colleges of London University or the various Polytechnics, and it contained people with a genuine interest in the history of mathematics as well as those who hoped it would be useful in mathematics education. None of them had much experience of research in the subject, and so with two exceptions standards were low. Of these, Whiteside was one and Ivor Grattan-Guinness was the other. Whiteside seemed to have Newton covered, and anyway I wanted to look at more modern topics, related to the present-day mathematics syllabus. Ivor was a product of the philosophers of science in the LSE and he had an enormous ability to find sources. He was working at the time on his monumental three-volume work *Convolutions in French Mathematics* (1990), but he was smarting under the heavy criticism the old guard at the *Archive for the History of Exact Sciences* had levelled at him over his article ‘Did Cauchy plagiarise Bolzano?’ that the journal had recently published. He said Cauchy had, but the short answer was and still is ‘No’, and a flavour of the criticism was Freudenthal’s remark “Of course you must read between the lines, but first you must read the lines themselves”. I would have quit on the spot, but Ivor bravely battled through, although his generally poor relations with professional mathematicians remained an obstacle all his life. As a result, the BSHM was a society of people who were amateur historians, prone to thinking that something they’d read in some old book or journal could be written up as a piece of historical research. I fitted in nicely. The only difference was that I wrote about topics of greater mathematical difficulty, as I began to see where the discovery of non-Euclidean or hyperbolic geometry had led.

In 1979, it seemed to me that I should try again to get a Ph.D., this time in history of mathematics. I was writing *Ideas of Space* at the time, and a flatmate suggested that I take it to Warwick and see if they’d call it a Ph.D. thesis. Very wisely, the Mathematics Department there said they wouldn’t, because they hadn’t supervised it, but they would take me back to write one if I would agree to accept the help of David Fowler and Ian Stewart, even though they were not strictly speaking historians of mathematics (naturally, I did), and the OU said I could go ahead as long as my work for them did not suffer. This all meant that I never did a formal course in history of science, let alone mathematics, and I don’t think there was any possibility of enrolling anywhere in the U.K. on a graduate course in the history of mathematics. But Ian and David were very helpful, as was Henk Bos later, and I have no regrets at having had to find my own way. Overall, the help I had at Warwick more than cancelled out my original unhappy experience there.

I knew from my link to Warwick that Bill Thurston was making Fuchsian and Kleinian functions big topics again. It was soon clear that both Klein and Poincaré had been involved in the nineteenth century, and that non-Euclidean geometry had been crucial to their discoveries. But who was Fuchs, and what had he contributed? His work took me back to Riemann and the hypergeometric equation, and therefore to Gauss. I started to fill in the gaps in my knowledge, helped by a meeting early on with Wilhelm Magnus in New York, who put me on to Schlesinger’s papers and his intimidatingly large *Handbuch* on differential equations. Periodically I worried that my entire thesis was in there somewhere. I was also helped on that trip by Tom

Hawkins at Boston University and Garrett Birkhoff at Harvard. I began to realise that nineteenth century mathematics is full of intriguing connections between what can seem to be separate aspects, and that recognising and exploiting these connections has been a driving force in the subject.

I went to Paris in December 1980, partly for a holiday with my girlfriend Sue (later my wife) who was studying there, partly for research, all set to say that I'd looked for Poincaré's unpublished entry to the Académie's Prize Competition in 1880 and had failed to find it, although it marks his first use of non-Euclidean geometry and opened the road to his theory of Fuchsian and Kleinian functions. The morning I went to the Académie des Sciences I learned that they would be closing at lunchtime for the long holiday break. I asked if they had any Poincaré documents, and immediately they took an ornate box off the shelves in the room I was in, and there, halfway down, were the papers I wanted. I asked for copies to be made, and when I got back to London I wrote to Jean Dieudonné to tell him what I'd found. With the help of Sue's mother I wrote a paper about the discovery in French, then I finished my thesis, and graduated in 1982. Much later, Scott Walter and I published them as (Poincaré, 1997).

With my Ph.D. behind me and some study leave stored up, I could travel, and I was lucky enough to go to Brandeis for two terms in 1983–1984. They paid me generously to do some elementary teaching, which left me ample time to sit in on David Eisenbud's algebraic geometry class, and drop in on classes in Harvard by Joe Harris and Serge Lang. It was all too hard for me, but in any case I only wanted to acquire a rough and ready sense of what schemes are and what they are for. I had a vague idea that I might write about the history of algebraic geometry around 1900. But the good relations I had with David Eisenbud lasted, and we should shortly finish a project that had its hazy beginnings all those years ago. While there, I met Roger Cooke. I suspect most of us are, as I am, indebted to him for his knowledge of Russian and his knowledge of mathematics in Russia, which became difficult to acquire with the decline of history of mathematics there after the collapse of the Soviet Union.

When I returned, I set about turning my thesis into a book, which came out in 1986, the same year that Judith Field and I published our book on Desargues. We surely met through the BSHM, and she had the idea that with her good French and my understanding of the mathematics we could do something worthwhile. Since then, Jan Hogendijk has made some valuable criticisms, and in the last couple of years Jean-Yves Briend and Marie Anglade have subjected the *Brouillon Projet* to a very thorough and illuminating analysis, while the works of Andrea del Centina (see e.g. his 2020) and Kirsti Andersen (see e.g. her 2007) have elaborated a rich picture of perspective, projective geometry, and related issues in which Desargues played an important part. I'm glad to have had something to do with it, but no regrets about returning to the nineteenth century.

I believe it was in the early 1980s when I was first invited to what was then an annual gathering of historians of mathematics at Oberwolfach, the German Mathematical Society's wonderful conference centre in the Black Forest, with its splendid runs of journals. If I remember correctly, one of speakers on that occasion took out of his bag what he proclaimed was the skull of Swedenborg. This was not what meetings of the BSHM had prepared me for! On the other hand, for the first time I met significant numbers of historians with high professional standards. Without attempting to be complete, there was Kirsti Andersen and Henk Bos, Adolf Youschkevich, Hans Wussing, Ivor Grattan-Guinness, and a number of (west) German historians including Christoph Scriba and Eberhard Knobloch. If not at that Tagung then at one soon after, Scriba produced one of a large number of large red books that represented what they intended to publish concerning Leibniz's work on algebra, which was everything; there was a heated discussion about the merits of this, all in the days before the internet. A surprising number of the talks were in English, but the randomised seating allocation at meal times was good for my German.

Also at the meeting were a number of younger historians, some of whom quickly became good friends: Umberto Bottazzini, Jesper Lützen, and Erhard Scholz among them (it turned out that Erhard had spent the MSc. year at Warwick when I was there, but we only barely remembered each other). He was in some sense a student of Henk's, Jesper was a student of Kirsti's, and I think Umberto may have been more independent. We were united in being the rising generation in those political times, and in trying to do work that was simply better than much of what the previous generation had done, at least as far as our common field, the nineteenth century, was concerned. Henk was very supportive, and it took me some time to appreciate what he was saying, because he always looked behind a question to see where it was coming from, what presumptions it concealed, and why it might be worth asking.

The language issue is an obstacle to all this work. I benefitted from being able to write in my native language, and ever since I've tried to redress the balance by polishing other people's articles before they are published, provided they start off in reasonably good English. The downside is that not very much high level mathematics was written in English in the nineteenth century, the rich archives are in France, Germany, Italy, and most problematically in Russia. For my friends, and many other European historians, the balance is reversed: rich, relatively accessible archives, but the knowledge that publishing in English is the best way to a readership. My friend David Rowe has drawn in recent years on the advantage of being an English speaker in Germany to produce increasingly rich pictures of mathematics in Göttingen; (see e.g. Rowe, 2018). If you worked on the nineteenth century, the archival problem was to some extent reduced by the fact that there was simply no good account in any language: Boyer's (1968) but much revised was superficial and Kline actively disliked modern pure mathematics (his account of nineteenth century matters in his (1972) is, however, quite thorough and reliable, if more of a reference work than a critical history). Accordingly, just putting together an account based on what was in the journals and not simply the folklore of the working mathematician seemed to most of us to be worth doing.

Missing from older histories of modern mathematics was much attention to the methods by which the results had been achieved. I think most of us were starting to realise that how results were found was as much part of the story as the accumulation of the results themselves. It's a difficult balance to strike, and it led on to tricky questions over what to do with arguments that are not proofs by modern standards: what made them convincing at the time, and how should they be presented? It also became easier to ask why certain results were wanted at all, quite independently of any present-day significance they may have.

Thus Erhard Scholz wrote his book (1980) on the origin of the idea of a manifold in the period from Riemann to Poincaré, which impressed me greatly, Umberto Bottazzini wrote his book (1986) on real and complex analysis—it was the complex part that was truly innovative—and Jesper Lützen produced his biography of Liouville (1990), which is exceptional in being grounded in all of Liouville's notebooks (a resource that does not seem to exist for other mathematicians, unfortunately). We are often reminded not to write 'Whig history', history that leads progressively to us; Grattan-Guinness came to hold views about this by making a distinction between history and heritage that I think said more about his strained relationship with mathematicians than anything else. I think there was a move to tell it as it really was so far as the sources would permit, and yes, that often meant what was in the nineteenth century mathematics journals. But sometimes that was done exceptionally well, the papers and books by Tom Hawkins—e.g. his (2000)—being another good example that I learned a lot from. But by now, the OU was making everyone work hard. I wrote a book or two in the next 20 years, and a number of articles that kept me reading and writing and therefore visible. They are of uneven merit, but most of them have something in them worth saying.

In the 1980s, attempts to ground the history of mathematics in social history lost some of their charm with the change in the ambient politics of the time, although good work was still done; perhaps the massive volumes currently being edited by Joe Dauben and David Rowe will revitalise things. In their place came a string of well-researched biographies: Jesper Lützen on Liouville as already mentioned, Bruno Belhoste on Cauchy (1991), Tony Crilly on Cayley (2006), Joe Dauben on Cantor (1979), the books by Anne Koblitz (1983) and Roger Cooke (1984) on Kovalevskaya, Detlef Laugwitz on Riemann (1999), Vladimir Maz'ya and Tatyana Shaposhnikova on Hadamard (1998), Karen Parshall on Sylvester (2006), Arild Stubhaug on Abel (2000), Lie (2002), and Mittag-Leffler (2010), and only recently Uta Merzbach (2018) on Dirichlet. They all strike a different balance between the life and the work, and make different appeals to different people, but they all are substantial (and not 'Whig') works of history, and they all have tried to be readable. Lützen's solution has a lot to recommend it: a first half on the life, a second half on the works that goes into much more mathematical detail.

Quite generally, in the 1980s there was a move to ask what it was that historians of mathematics were trying to do and what, indeed, they should try to do. Were there any big themes, or sweeping subjects that would make more valuable topics to research? Certainly there were substantial historical works focussed on particular themes: Hawkins' work on Lie theory from Lie via Killing to Cartan is a case in point. Scholz took up the study of Hermann Weyl. He gave an ICM address on Weyl's geometry in 1994 that emphasised his Fichtean roots (Scholz, 1995), and then turned to his cosmology, and ever since he has been in a productive relationship with physicists. Karine Chemla was beginning to suggest a way of reading Chinese mathematics that questioned what constituted a proof (see her essays in her 2012): obviously I wasn't going to take up Chinese mathematics, but we became good friends—we first met at a conference in Luminy in the early 1980s—and I have ever since regarded her as a force for fresh thinking and conceptual rigour that can drive up standards across our field. Several of my favourite books raised the idea of major shifts in mathematics, Bos's *Redefining Geometrical Exactness* (2000), for example. By now there were also several competing French camps, all of which tended to emphasise methodological rigour (defining a corpus, etc.). Soon, it was no longer possible to write good history of mathematics in advance of a serious wrestle with the sources; the folklore could no longer be accepted, and there were numerous readable books and papers out there trying to replace it. In this spirit, Bottazzini and I decided to rescue complex analysis from the shadow of real analysis, and eventually published our big book in 2013, to the surprise of our colleagues, I suspect.

In 1989 I persuaded the London Mathematical Society to launch a series of books on the history of mathematics. By then various Americans had the same effect on the American Mathematical Society, and the two series soon merged. They operated with very different attitudes to money, but I was very glad to be involved as an editor and to have a hand in some books of lasting merit. This, and my later work as an Editor-in-Chief with my good friend Jed Buchwald of *Archive for History of Exact Sciences*, which sprang out of my term at the Dibner Institute MIT in 1996, has been my way of making up for small number of Ph.D. students I was able to attract. Small here means two. One was June Barrow-Green, about whom more below, the other Snezana Lawrence, who came from Serbia during the Bosnian war of 1992–1995, wrote a fine Ph.D. thesis on the Freemasons in England which deserves to be published, and went on to become an influential figure in mathematics education and a good friend. It may well have been on that trip to the Dibner that I also met the indefatigable John Stillwell, who has translated so many things and also written several books that illuminate important topics that wrongly get overlooked, not least his translations of papers by Poincaré (1985) and (2010). I also met John McCleary then, who encouraged me to stay interested in differential geometry. I owe more to these two than it is easy to say, and I'm very grateful for their support.

There came a point in my life some 20 years ago where I didn't expect a nineteenth century journal or book to say anything major that I didn't already know about or come up with a reference to a paper that would truly surprise me. Of course I had a shallow grasp of most things, and I had deliberately avoided anything to do with partial differential equations, or advanced number theory; I learned a lot, both in content and approach, from Goldstein, Schappacher and Schwermer's *The Shaping of Arithmetic* (2007). Still, I had the feeling it was now possible to map out the salient features of the nineteenth century mathematical landscape. What then, I asked myself, was I meant to do with all this knowledge? Certainly I should deepen my understanding of it, but to be a better historian meant asking better historical questions and coming up with better historical answers.

The idea of modernism in mathematics appealed to me. In 1989, the late Herbert Mehrtens had published his *Moderne Sprache Mathematik* that launched the idea of mathematical modernism and coupled it to an exploration of broader social and intellectual issues not too far, in some directions, from the Nazi time and all packaged in the Foucauldian analysis of the day. I had no liking for Foucauldian tales of power, I felt that for all its originality Mehrtens's book still relied too much on the received wisdom about Hilbert, Klein, Poincaré, and Brouwer (although the attention paid to Hausdorff was new to me), and dwelt too much on their explicitly philosophical views to the exclusion of the range of the mathematical work. It seemed to me that wherever one looked in the years just before 1900 there were advocates of a new way of doing mathematics—in algebra, analysis, geometry, logic, right across the board—and that this point of view had pretty much won in many more places than Germany and France to boot. Quite some scholarship on specific cases supported this point of view, and quite a bit also enabled me to balance it more subtly than Mehrtens' division into Moderne and Gegenmoderne, and on the back of that to sketch out how the new modern mathematics renegotiated its relations with science and even philosophy. At all events, that's what I attempted to do in *Plato's Ghost*, which Princeton U.P. published in 2008.

Unfortunately, Mehrtens never commented on my book, came out at left the field for his chosen field of cultural criticism. I can see now that he may have thought I gave him too little credit. His intellectual world was not mine, and that is also part of the story. I wish too that I had made more of the implications of modernism as I interpreted it for the contemporary domains of applied mathematics. It seems to me more than ever that a full study of how modernist mathematics related to physics would be a good test of any modernism thesis, but such a book has yet to be written. I've recently written a little more about this in the memorial volume to Herbert (see Gray, 2023). On the other hand, I learned recently that *Plato's Ghost* inspired a breakaway movement in literary studies from literature and science to literature and mathematics, which is oddly gratifying.

Many of the historians I have most respect for have based new analyses on unpublished sources. I began by trying to make better sense of published sources, and slowly felt the need for a language to describe what I had read. I didn't want to write *Mathematical Reviews* for the past, and I began to think of what kind of knowledge mathematics might be. How was it acquired, secured, conveyed? What was it, more precisely, that a mathematician knew? What mattered, and why? What do mathematicians mean when they say there are good and less good ways to prove things? These questions were much debated around 1900, and by various historians and philosophers around 2000. I had the luxury of being able to ignore modern philosophy of mathematics, although I read quite a bit of it, and I'd picked up some sort of education in philosophy at Oxford by attending many philosophy seminars (especially the magnificent weekly seminar run by Freddie Ayer that every serious philosophy student attended). This was a year or two before Oxford brought in a degree in Mathematics and Philosophy, but I'm glad I missed it. I would have learned more logic and less mathematics than would have been good for me. Gradually, words like 'epistemology', 'ontology', and 'semantics' crept into my work. I began to see how to fit my questions about the nature of old mathematical arguments into debates about fruitful proofs, purity of method, depth (e.g my 2015), and other themes.

These questions were the purview of philosophers of mathematics and logic out there who did not think highly of structuralism or questions about mathematics that see only set theory as capturing what they wanted to get at. Probably my first acquaintance here was my friend Jamie Tappenden, whom I met in Boston when I visited the Dibner Institute in 1996. His subtle analysis of how Frege can be seen as influenced by Riemann appealed to me. Somewhere along the line I met José Ferreirós, who combines interests in logic and history and philosophy of mathematics to an extraordinary degree, and we also became good friends, and in the push to establish the Association for the Philosophy of Mathematical Practice (APMP) I was happy to be called in as a sufficiently useful historian. At the early meetings I also met Paolo Mancosu, whose own work expertly bridges close readings of many mathematical works with novel questions of a logical and philosophical kind (see e.g. the many essays collected in his 2010). The book he edited, *The Philosophy of Mathematical Practice*, is the envy of us all for managing to contain elementary and more advanced essays in a coherent whole. In this group I met another good friend, Jeremy Avigad, who combines expertise in formal mathematics and machine-produced proofs with an ability to reflect on what mathematicians actually do. Or, rather, have actually done, it being very difficult to follow mathematicians at work, if only because they don't have laboratories to run, and nobody likes to be watched filling up a wastebasket with bad ideas. I also got to know and become friends with Colin McLarty, the category theorist and logician who has taught me a lot about Emmy Noether.

Independently, and somewhat differently, my friend Moritz Epple had taken to approaching mathematics through his theory of epistemological configurations, which he adapted from the work of Hans-Jörg Rheinberger's (1997)—see (Epple, 2011). I got to know him through Oberwolfach, and through his supervisor, David Rowe, who also supervised Leo Corry; his thesis later became his *Modern Algebra and the Rise of Mathematical Structures*, a rare and valuable book on twentieth century mathematics. Leo Corry went on in his (2004) to transform our views of Hilbert by using volumes of his lectures notes to document his interest in applied mathematics, which has implications for any theory of modernism in mathematics. Moritz Epple's trajectory has taken him from an early interest in Wittgenstein to his much-admired, philosophically acute book on knot theory (1999), and the immense labour he expended editing the works of Hausdorff in ten volumes. Hausdorff is key figure in any account of mathematical modernism, and I suspect all of us know and appreciate Hausdorff's work because of the work done by Moritz Epple, Erhard Scholz, Walter Purkert, and, of course, Egbert Brieskorn (among many others). The upshot of all of this was that a number of us found ways to ways to write about the nature and significance of mathematical discoveries that complement the more traditional approach, which was also becoming deeper and more insightful.

As Mehrtens had noted, Poincaré was a counter-modern of some stature. So much of what I had written about led up to Poincaré, and the work of others pushed me further towards him. The first of my former students and now good friend June Barrow-Green turned her Ph.D. thesis into her book (1997) on Poincaré and the three-body problem, one of the few books by a historian that I believe mathematicians are entirely comfortable with. It contains her thorough rewriting of the story based on her discovery of the original manuscript that Poincaré had submitted for King Oscar the Second's prize, with its major mistake that Mittag-Leffler had tried to cover up. There was good work already done by others on various aspects of Poincaré's work: Arthur Miller (1981) and Scott Walter (2009) (on special relativity); Dieudonné (1989), Karanbir Sarkaria (1999), Stillwell (1985 and 2010), and others on his topology (see Ji and Wang, 2020 and 2022); numerous philosophers, including Gerhard Heinzmann (2008), on the four volumes of Poincaré's essays; and above all the continuing efforts of the people in the Archive Henri Poincaré in Nancy to produce editions of Poincaré's correspondence. Finally, the gravitational pull proved irresistible. I had helpful visits to Nancy that taught me a lot about the institutional side of Poincaré's life, and Princeton U.P. produced my *Henri Poincaré: a Scientific Biography* in 2013, just missing the hundredth anniversary of his death.

High level books should be no more difficult than they have to be, but below the snow line the problem of connecting to an audience is a real one. As I have already said, 2001 was a time for me to move on, but I didn't know how. An OU colleague suggested I approach Warwick. David Fowler, by then a respected historian of Greek mathematics, had recently taken early retirement on medical grounds, and in 2002 Warwick agreed to have me teach a course one term a year provided that it pulled in large numbers (David's course on Newton had been capped at twenty to ensure each student presented a paper). I leapt at this, but what to teach? The history of non-Euclidean geometry, of course. I bashed out a full set of notes, expanding the subject to include some projective geometry, Felix Klein's view of geometry, Poincaré's non-Euclidean geometry, and Hilbert's axiomatisation of elementary geometry. The students liked it in adequate numbers, and after 4 years, during which I discovered the pleasures of rewriting a course from one year to the next in the light of its reception—something the OU cannot provide—it was time to pick a new topic. I wasn't happy with the reception of my big book with Umberto Bottazzini, so I distilled from it a course on real and complex analysis in the nineteenth century. Four years later I produced a course on nineteenth century algebra, which has some new things to say about Klein, his book on the icosahedron of 1884, and Galois Theory. How much mathematics descends from Gauss and Galois! I concluded my time at Warwick with a course on the history of ordinary and partial differential equations. I found it quite shocking that there was no single substantial book on the history of differential equations and the real reason we have the calculus (which is not to bewilder students with epsilons and deltas), although Craig Fraser and Tom Archibald's papers gave me a good start on several topics.

All these courses became books published in the Springer SUMS series. They are a compromise intended to help keep history of mathematics alive at a difficult time for it in the U.K. and perhaps the USA. The one on geometry is used at the OU in its MSc. course. A course in thirty lectures with assessment material and advice on how to tackle it is not the ideal place to publish original research, there is always the problem of keeping the material accessible, although Warwick students can be very bright, but I believe the topic is popular with students and part of what stops universities from teaching it more often is the problem of how to assess it. Anyway, the deed is done, and perhaps inevitably the level of mathematical difficulty increased from one course to the next. Despite John Fauvel's example, I continue to believe that the history of mathematics from the time of Newton and Leibniz to yesterday is easier to explain to an audience with at least some mathematics (it should be a third-year course, not a first-year one), such as undergraduates and higher-level school teachers, people with mathematics degrees. Or, sadly less likely, historians of science. Each group has its merits and poses its challenges.

I had hoped that my scientific biography of Poincaré would reach historians of science, and perhaps it did. But the relations between historians of mathematics and historians of science, specifically physics, are complicated and vary from country to country. How a closer relationship with historians of science might work in future is not clear, but one possibility is this. Now that we know so much more about the early calculus, thanks to the ability of Niccolò Guicciardini, George Smith, and others to get beyond the imposing eight volumes of Newton's *Mathematical Papers* that Tom Whiteside had completed, and we can read what has finally come out of the immense Leibniz project thanks to Eberhard Knobloch, Richard Arthur and David Rabouin, somebody really should write a good up-to-date book about it. It could continue into the eighteenth century more easily than before, thanks to the very positive development of the Euler Archive, a fine example of collaborative work on the web, with its extensive array of translations. We may then also get beyond Truesdell's comprehensive vision of rational mechanics, and rediscover Lambert.

For myself, prompted by Niccolò Guicciardini, like several of us, to think about anachronism, I have recently been back to Bonola in my (2021), and seen properly for the first time where he went wrong. Wedded to an anachronistic view of geometry, he put hyperbolic trigonometry in the wrong place and messed up his account of Lobachevskii as a result. It's odd to think that had Bonola's book been better I might never have got started.

Since then, I have swung back to hard mathematics. Retirement from the OU and Warwick has only served to revive two long-running, much interrupted collaborations, one with Mario Micallef at Warwick (on Douglas, minimal surfaces, and the first Fields Medal, which has turned out to require an extensive prehistory about partial differential equations) and the other with David Eisenbud at MSRI Berkeley (on Macaulay and the theory of polynomial rings after Max Noether). Happily, my co-authors understand the mathematics much better than I do, and generate fresh questions. I wish I'd worked with mathematicians more, but that's not so easy to arrange.

In conclusion, I was lucky to have enough French and German to read papers in mathematics, and to graduate at a time when the history of mathematics was being revived, lucky to get a job teaching the history of mathematics, and luckier still to land initially on non-Euclidean geometry, a topic of perennial interest at many levels, leading as it does to work by Gauss, Riemann, Poincaré, and many others, and numerous mathematical, philosophical, and historical questions. I look back at some things I wrote and wish that I had taken more time, and pushed a little deeper, but it took me a long time to work out why history of mathematics matters and how perhaps it can be done better. For a long time I just followed my nose, tracing the implications of the discovery of non-Euclidean geometry, happy to find that others had written in depth on related topics. It takes time to have read enough to build up a store of information that assembles into a worthwhile argument with some chance of being tested. I have also worked either as a coauthor, editor, or simply as a colleague alongside several other historians, and together we have contributed to the production of a much more detailed set of accounts of mathematics in the nineteenth century, richer in mathematical detail, in social context, and in historical analysis. This makes it easier to see what still needs to be done, but that is the paradox of research.

I have not been able to mention here everyone who helped me, worked on the same topics and enriched the history of mathematics with different perspectives, wrote books and articles I felt able to rely on, or even those who disagreed with me. (I do think there should be more disagreement in the subject; if it is well intended and not tendentious it would be productive.) I owe a special thanks, of course, to the editors of this book, especially Lizhen Ji, not only for all his work pulling this book together but for a spirited email correspondence that I hope will long continue. I apologise to anyone who feels I have left them out, but this brief autobiography is surely long enough already.

I owe a lot, in ways that extend well beyond the academic, to my colleagues in Europe and America, who were people I could learn from, and who became good friends. Those friendships are, in the end, among the most rewarding part of my life, outside of the wonders of living with Sue and sharing our lives with our daughters Martha and Eleanor. I thank them all.

References

Andersen, K. 2007. *The Geometry of an Art: The History of the Mathematical Theory of Perspective from Alberti to Monge*. Berlin: Springer.

Anglade, M., and J.Y. Briend. 2017. La notion d'involution dans le *Brouillon Project de Girard Desargues*, *Archive for History of Exact Sciences 71*: 543–588.

Barrow-Green, J.E. 1997. *Poincaré and the Three-Body Problem*. American and London Mathematical Societies, HMath 11.

Barrow-Green, J., J.J. Gray, and R. Wilson, 2019/2022. *The History of Mathematics: A Source-based Approach*, vol. 2. Providence: American Mathematical Society.

Belhoste, B. 1991. *Augustin-Louis Cauchy: A Biography*. Berlin: Springer.

Bonola, R. 1912. *Non-Euclidean Geometry*. Chicago: Open Court Publications. Transl. H.S. Carslaw, Repr. Dover 1955.

Bos, H.J.M. 2000. *Redefining Geometrical Exactness: Descartes' Transformation of the Early Modern Concept of Construction*. Berlin: Springer.

Bottazzini, U. 1986. *The Higher Calculus. A History of Real and Complex Analysis from Euler to Weierstrass*. Berlin: Springer.

Bottazzini, U., and J.J. Gray. 2013. *Hidden Harmonies — Geometric Fantasies: The Rise of Complex Function Theory*. Berlin: Springer.

Boyer, C.B. 1968. *A History of Mathematics*. Hoboken: Wiley.

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

Cooke, R. 1984. *The Mathematics of Sonya Kovalevskaya*. Berlin: Springer.

Corry L. 1996/2004. *Modern Algebra and the Rise of Mathematical Structures*. Birkhäuser, 2nd revised edn. Berlin: Springer.

Corry, L. 2004. *David Hilbert and the Axiomatization of Physics (1898–1918): From* Grundlagen der Geometrie *to* Grundlagen der Physik. Alphen aan den Rijn: Kluwer.

Crilly, T. 2006. *Arthur Cayley: Mathematician Laureate of the Victorian Age*. Baltimore: Johns Hopkins University Press.

Dauben, J.W. 1979. *Georg Cantor: His Mathematics and Philosophy of the Infinite*. Cambridge: Harvard University Press.

Dauben, J.W., and D.E. Rowe. (eds.) to appear. *A Cultural History of Mathematics*, vol. 6. London: Bloomsbury.

del Centina, A. 2020. Pascal's mystic hexagram, and a conjectural restoration of his lost treatise on conic sections. *Archive for the History of Exact Sciences 74*: 469–521.

Dieudonné, J. 1989. *A History of Algebraic and Differential Topology, 1900–1960*. Basel: Birkhäuser.

Epple, M. 1999. *Die Entstehung der Knotentheorie*. Wiesbaden: Vieweg.

Epple, M. 2011. Between timelessness and historiality: on the dynamics of the epistemic objects of mathematics. *Isis 102*: 481–493.

Fauvel, J., and J.J. Gray (eds.). 1987. *The History of Mathematics: A Reader*. New York: Macmillan. In association with the Open University.

Goldstein, C., N. Schappacher, and J. Schwermer (eds.). 2007. *The Shaping of Arithmetic After C. F. Gauss's* Disquisitiones Arithmeticae. Berlin: Springer.

Grattan-Guinness, I. 1970. Did Cauchy plagiarise Bolzano? *Archive for the History of Exact Sciences 6*: 372–400.

Grattan-Guinness, I. 1990. *Convolutions in French Mathematics, 1800–1840*, vol 3. Basel: Birkhäuser.

Gray, J.J. 1979/1989. *Ideas of Space, Euclidean, non-Euclidean and Relativistic*, 2nd edn. Oxford: Oxford University Press.

Gray, J.J. 1981. Les trois suppléments de Poincaré, etc. *Comptes rendus de l'Académie des Sciences 293*, 87–90.

Gray, J.J. 1986. *Linear differential equations and group theory from Riemann to Poincaré*, 2nd edn. Basel: Birkhäuser. With three new appendices and other additional material, 2000.

Gray, J.J. 2008. *Plato's Ghost: The Modernist Transformation of Mathematics*. Princeton: Princeton University Press.

Gray, J.J. 2011. *Worlds out of Nothing; A Course on the History of Geometry in the 19th Century*, 2nd edn. Berlin: Springer.

Gray, J.J. 2013. *Henri Poincaré: A Scientific Biography*. Princeton: Princeton University Press.

Gray J.J. 2015. Depth – a Gaussian tradition in Mathematics. *Philosophia Mathematica 23*(2): 177–195.

Gray, J.J. 2018. *A History of Abstract Algebra: From Algebraic Equations to Modern Algebra*. Berlin: Springer.

Gray, J.J. 2021. Anachronism: Bonola and non-Euclidean geometry. In *Anachronism in the History of Mathematics*, ed. N. Guicciardini, pp. 281–306. Berlin: Springer.

Gray, J.J. 2023/to appear. Poincaré and counter-modernism. Special issue on Herbert Mehrtens and his work *Science in Context*.

Gray, J.J., and J.V. Field. 1986. *The Geometrical Work of Girard Desargues*. Berlin: Springer.

Gray, J.J., and L. Tilling. 1978. Johann Heinrich Lambert, mathematician and scientist, 1728–1777. *Historia Mathematica 5*: 13–41.

Hawkins, T. 2000. *The Emergence of the Theory of Lie Groups: An Essay in the History of Mathematics, 1869–1926*. Berlin: Springer.

Heinzmann, G. 2008. Poincaré wittgensteinien? In *Wittgenstein. État des lieux*, ed. E. Rigal, pp. 274–289. Paris: Vrin.

Ji, L., and Wang, C. 2020. Poincaré's stated motivations for topology. *Archive for History of Exact Sciences*, 74: 381–400.

Ji, L. and Wang, C. 2022. Poincaré's works leading to the Poincaré Conjecture. *Archive for History of Exact Sciences 76*: 223–260.

Kline, M. 1972. *Mathematical Thought from Ancient to Modern Times*. Oxford: Oxford University Press.

Koblitz, A.H. 1983. *A Convergence of Lives. Sofia Kovalevskaia: Scientist, Writer, Revolutionary*. Basel: Birkhäuser.

Laugwitz, D. 1999. *Bernhard Riemann 1826–1866: Turning Points in the Conception of Mathematics* (A. Shenitzer. transl.). Basel: Birhkäuser.

Lützen, J. 1990. *Joseph Liouville, 1809–1882. Master of Pure and Applied Mathematics*. Berlin: Springer.

Mancosu, P. (ed.) 2008. *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press.

Mancosu, P. 2010. *The Adventure of Reason: Interplay between Philosophy of Mathematics and Mathematical Logic, 1900–1940*. Oxford: Oxford University Press.

Maz'ya, V., and T. Shaposhnikova. 1998. *Jacques Hadamard, A Universal Mathematician*. Providence, RI: American Mathematical Society/London Mathematical Society.

Mehrtens, H. 1989. *Moderne Sprache Mathematik*. Berlin: Suhrkamp.

Merzbach, U. 2018. *Dirichlet: A Mathematical Biography*. Basel: Birkhäuser.

Miller, A.I. 1981. *Albert Einstein's Special Theory of Relativity: Emergence (1905) and Early Interpretation (1905–1911)*. Boston: Addison-Wesley.

Parshall, K.H. 2006. *James Joseph Sylvester: Jewish Mathematician in a Victorian World*. Baltimore: Johns Hopkins University Press.

Poincaré, H. 1985. *Papers on Fuchsian Functions*, edited and translated, with an introduction by J. Stillwell. Berlin: Springer.

Poincaré, H. 1997. *Trois Suppléments sur la Découverte des Fonctions Fuchsiennes*, ed. J.J. Gray, and S. Walter. Berlin/Blanchard: Akademie Verlag/Blanchard.

Poincaré, H. 2010. *Papers on Topology: Analysis Situs and Its Five Supplements*, edited and translated, with an introduction by J. Stillwell. American and London Mathematical Societies, HMath 37.

Rheinberger, H.-J. 1997. *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Stanford: Stanford University Press.

Rowe, D.E. 2018. *A Richer Picture of Mathematics: The Göttingen Tradition and Beyond*. Berlin: Springer.

Sarkaria, K.S. 1999. The topological work of Henri Poincaré. In *History of Topology*, ed. I.M. James, pp. 123–167. Amsterdam: North-Holland.

Scholz, E. 1980. *Geschichte des Mannigfaltigkeitsbegriffs von Riemann bis Poincaré*. Basel: Birkhäuser.

Scholz , E. 1995. Hermann Weyl's "Purely Infinitesimal Geometry". *Proceedings of the International Congress of Mathematicians 2*: 1592–1603.

Schlesinger, L. 1897. *Handbuch der Theorie der Linearen Differentialgleichungen*, vol. 3. Leipzig: Teubner.

Stubhaug, A. 2000. *Niels Henrik Abel and his Times: Called Too Soon by Flames Afar*. Berlin: Springer.

Stubhaug, A. 2002. *The Mathematician Sophus Lie: It Was the Audacity of My Thinking*. Berlin: Springer.

Stubhaug, A. 2010. *Gösta Mittag-Leffler: A Man of Conviction*. Berlin: Springer.

Walter, S. 2009. Henri Poincaré, theoretical physics, and relativity theory in Paris. In *Mathematics Meets Physics*, ed. K.-H. Schlote, and M. Schneider. Frankfurt: Harri Deutsch Verlag.
Whiteside, D.T. (ed.). 1967–1981. *The Mathematical Papers of Isaac Newton*, vol. 8. Cambridge: Cambridge University Press.

Appendix A
Photos of the Venerated

Jeremy Gray in Scotland

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*,
Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7

Top: Jeremy Gray and Kirsti Andersen. Bottom: Jeremy Gray talking (2014)

Jeremy Gray in Xi'An (2010). Top: with R. Chorlay, R. Siegmund-Schultze, M. Schneider, K. Chemla

Jeremy Gray and Sue Lawrence with grandson Rufus (2021)

Jeremy Gray with granddaughter Benny (2022)

From left to right: N. Guicciardini, L. Corry, K. Chemla, J. Barrow-Green, J. Gray, M. Epple, E. Scholz, U. Bottazzini, J. Lützen (valediction conference 2014)

Index of Names

K. Chemla et al. (eds.), *The Richness of the History of Mathematics*, Archimedes 66, https://doi.org/10.1007/978-3-031-40855-7