CMS Books in Mathematics

Glen Van Brummelen
Michael Kinyon
Editors

Mathematics and the Historian's Craft

The Kenneth O. May Lectures

Canadian Mathematical Society
Société mathématique du Canada

CMS Books in Mathematics
Ouvrages de mathématiques de la SMC

1 Herman/Kučera/Šimša Equations and Inequalities
2 Arnold Abelian Groups and Representations of Finite Partially Ordered Sets
3 Borwein/Lewis Convex Analysis and Nonlinear Optimization
4 Levin/Lubinsky Orthogonal Polynomials for Exponential Weights
5 Kane Reflection Groups and Invariant Theory
6 Phillips Two Millennia of Mathematics
7 Deutsch Best Approximations in Inner Product Spaces
8 Fabian et al. Functional Analysis and Infinite-Dimensional Geometry
9 Křížek/Luca/Somer 17 Lectures on Fermat Numbers
10 Borwein Computational Excursions in Analysis and Number Theory
11 Reed/Sales (Editors) Recent Advances in Algorithms and Combinatorics
12 Herman/Kučera/Šimša Counting and Configurations
13 Nazareth Differentiable Optimization and Equation Solving
14 Phillips Interpolation and Approximation by Polynomials
15 Ben-Israel/Greville Generalized Inverses, Second Edition
16 Zhao Dynamical Systems in Population Biology
17 Göpfert et al. Variational Methods in Partially Ordered Spaces
18 Akivis/Goldberg Differential Geometry of Varieties with Degenerate Gauss Maps
19 Mikhalev/Shpilrain/Yu Combinatorial Methods
20 Borwein/Zhu Techniques of Variational Analysis
21 Van Brummelen/Kinyon Mathematics and the Historian's Craft: The Kenneth O. May Lectures

Glen Van Brummelen Michael Kinyon
Editors

Mathematics and the Historian's Craft

The Kenneth O. May Lectures

With 91 Figures

Springer

Glen Van Brummelen
Bennington College
Bennington, VT 05201
USA
gvanbrum@bennington.edu

Michael Kinyon
Department of Mathematical Sciences
Indiana University South Bend
South Bend, IN 46634-7111
USA
mkinyon@iusb.edu

Mathematics Subject Classification (2000): 01-06, 01A85

Library of Congress Control Number: 2005923503

ISBN-10: 0-387-25284-3 Printed on acid-free paper.
ISBN-13: 978-0387-25284-1

Printed in the United States of America. (EB)

9 8 7 6 5 4 3 2 1

springeronline.com

To Miriam May, in memory of Ken

Preface

1974 was a turning point for the history and philosophy of mathematics in North America. After years of planning, the first issue of the new journal *Historia Mathematica* was printed. While academic journals are born and die all the time, it was soon clear that *Historia Mathematica* would be a major factor in shaping an emerging discipline; shortly, it became a backbone for a global network of professional historians of mathematics. In the same year, the Canadian Society for History and Philosophy of Mathematics (CSHPM) was founded, adopting *Historia Mathematica* as its official journal. (In the 1990s, the CSHPM recognized its broader mission by naming *Philosophia Mathematica* as its official philosophical journal, rechristening *Historia Mathematica* as its historical journal.) Initially consisting almost entirely of Canadian members, the CSHPM has become in practice the North American society for the scholarly pursuit of history and philosophy of mathematics. The joint establishment of society and journal codified and legitimized the field, commencing what has become a renaissance of activity for the past 30 years.

These initiatives were begun by, and received much stimulus from, one man: Kenneth O. May, of the Institute for History and Philosophy of Science and Technology at the University of Toronto. May was a brilliant researcher, but he recognized that the viability of the fledging discipline required administrative leadership as well. In the introduction that follows, Amy Shell-Gellasch, CSHPM archivist, describes May's life and some of his achievements. Central to May's vision of the history of mathematics was the dichotomy between the role of the historian and the use that a mathematician might find for history. Mathematical practitioners, for reasons of pedagogy or in order to contextualize their own work, tend to focus on finding the antecedents for current mathematical theories in a search for how particular sub-disciplines and results came to be as they are today. On the other hand, historians of mathematics eschew the current state of affairs, and are more interested in questions that bear on the changing nature of the discipline itself. How, for instance, have the standards of acceptable mathematical practice differed through time and across cultures? What role do institutions and organizations play in the development of the subject? Does mathematics naturally align itself with the sciences or the humanities, or is it its own creature, and do these distinctions matter? The lead article in this volume, by Ivor Grattan-Guinness, is a strong statement on what makes history of mathematics unique, and reflects well May's own vision for our field.

May passed away, too early, in 1977. However, his legacy lives on partly through our thriving community; the continued prosperity of the CSHPM, *Historia Mathematica*, and *Philosophia Mathematica* are a resounding testament to that. In 2002, on the 25th anniversary of his passing, the CSHPM held a special meeting in May's honour. One of our actions at this meeting was to re-christen the keynote addresses at our annual general meetings as the "Kenneth O. May Lectures". Each of our annual meetings is a special occasion: while also providing a forum for presentations on all aspects of history and philosophy of mathematics, each meeting focuses on a specific theme, with activity revolving around an invited keynote address by a scholar of international repute. The diversity of these sessions over the years, witnessed in the table below, is a clear testament to the breadth and significance of the CSHPM's activities.

Since 1988 the CSHPM has preserved a record of the scholarly activities of the annual general meeting through the production of a volume of Proceedings, to which all speakers are invited to contribute. These Proceedings, distributed internally to Society members, are by now a repository of a great deal of valuable research. Some of these works have appeared elsewhere but many which deserve wider exposure have not; this volume represents our first attempt to correct this state of affairs. By printing the Kenneth May Lectures since 1990, we hope not only to choose some of the finest work presented at CSHPM meetings but also to present ourselves to the broader scholarly community. This volume represents by example who we are, how we approach the disciplines of history and philosophy of mathematics, and what we find important about our scholarly mission.

Many things happen over fifteen years. The editors attempted to reach all May lecturers since 1988, but were not wholly successful. Also, some of their lectures appeared later in formal scholarly journals (which the Proceedings is not), and some of these later versions incorporated improvements. In these cases we have chosen to reprint the polished final articles rather than the original lectures. One implication of this is that the bibliographic standards vary from article to article, reflecting the different sources in which the articles appeared. We are grateful to the following organizations that granted us permission to reprint articles free of charge from the pages of their journals and books: the Association for Symbolic Logic, the Canadian Mathematical Society (CMS), the Mathematical Association of America, and *Philosophia Mathematica*.

As editors of this volume, we have received a great deal of support from many people. The CSHPM, both its executive and its members, has been pivotal in working with us over the past year to produce the best possible

public imprint for the Society. The authors of the papers in this volume and archivist Amy Shell-Gellasch have combined to produce a truly admirable body of work. The editors of the CSHPM Proceedings over the years, listed below, have moved mountains to produce these volumes. Jonathan and Peter Borwein, editors of the CMS Books in Mathematics, provided highly valued encouragement and advice. Ina Lindemann, Mark Spencer, and Anne Meagher of Springer Verlag helped tremendously in bringing this volume to fruition. Thanks also go to Dennis Richter for technical support. Our families have sacrificed in their own ways, putting up with late dinners and with occasionally absent parents; we thank them especially for their patience. Finally, our greatest gratitude is due to the man to whom this volume is dedicated. Ken, your vision lives and prospers in the 21st century. Without your insight and formative efforts, the CSHPM might not be here today. Thank you.

Glen Van Brummelen and Michael Kinyon

A note on the title. Ken May considered the practice of the history of mathematics to be a unique melding of the crafts of mathematician and historian. This entails sensitivity both to the mathematical content of the subject, and to the various contexts in which it can be understood. Our daily work is constantly informed by our attempts to achieve this delicate balance. In Ken's words:

"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."[1]

[1]Kenneth O. May, "What is good history and who should do it?", *Historia Mathematica* 2 (1975), 453.

Annual Meeting Themes & Kenneth O. May Lecturers Since 1990

2003: Maritime Mathematics (Halifax, NS)
– Jim Bennett, *Geometry, Instruments and Navigation: Agendas for Research, 1500-1800*

2002: In Memory of Kenneth May (Toronto, ON)
– Ivor Grattan-Guinness, *History or Heritage? Historians and Mathematicians on the History of Mathematics*

2001: French Mathematics (Québec, PQ)
– Jean Dhombres, *The Applied Mathematics Origins of Lebesgue Integration Theory and Why it was Read as Pure Mathematics During the First Years of the 20th Century*

2000: History of Mathematics at the Dawn of a New Millennium (Hamilton, ON)
– Rüdiger Thiele, *Hilbert and his 24 Problems*

1999: Joint meeting with the British Society for History of Mathematics (Toronto, ON)

1998: Late 19th-Century Mathematics (Ottawa, ON)
– Volker Peckhaus, *19th-Century Logic: Between Philosophy and Mathematics*

1997: Science and Mathematics (St. John's, NF)
– Rüdiger Thiele, *The Mathematics and Science of Leonhard Euler*

1996: Ancient Mathematics (St. Catharines, ON)
– Alexander Jones, *Greek Applied Mathematics*

1995: Mathematics Circa 1900 (Montreal, PQ)
– Joseph W. Dauben, *Cantor and the Epistemology of Set Theory*

1994: History of Mathematics in the United States and Canada (Calgary, AB)
– Thomas Archibald (co-author Louis Charbonneau), *Mathematics in Eastern British North America in the Nineteenth Century: Some Preliminary Remarks*
– Karen Hunger Parshall, *The Emergence of the American Mathematical Research Community 1876-1900*

1993: Philosophy of Mathematics (Ottawa, ON)
– Stuart Shanker, *Turing and the Origins of Artificial Intelligence*

1992: Ethnomathematics (Charlottetown, PEI)
– Michael Closs, *The Ancient Maya: Mathematics and Mathematicians*

1991: Women in Mathematics (Kingston, ON)
– Ann Hibner Koblitz, *Women in Mathematics: Historical and Cross-Cultural Perspectives*

1990: History and Pedagogy (Victoria, BC)
– Judith Grabiner, *Was Newton's Calculus a Dead End? A New Look at the Calculus of Colin Maclaurin*

CSHPM/SCHPM Presidents

1974 – Charles V. Jones
1975, 1976 – Viktors Linis
1977, 1978 – J. L. Berggren
1979, 1980 – G. de B. Robinson
1981, 1982 – Wesley Stevens
1983, 1984, 1985 – Edward J. Barbeau
1986 – Marshall Walker
1987 – Louis Charbonneau
1988, 1989 – J. L. Berggren
1990, 1991 – Craig Fraser
1992, 1993, 1994, 1995 – Thomas Archibald
1996, 1997 – Robert Thomas
1998, 1999 – James J. Tattersall
2000, 2001 – Glen Van Brummelen
2002, 2003 – J. L. Berggren
2004, 2005 – Robert Bradley

CSHPM/SCHPM Proceedings Editors

1988, 1989 – Tasoula Berggren
1990 – Francine Abeles, Victor Katz, Robert Thomas
1991 – Hardy Grant, Israel Kleiner, Abe Shenitzer
1992-1999 – James J. Tattersall
2000, 2001 – Michael Kinyon
2002-present – Antonella Cupillari

Copyright Permissions

Thomas Archibald and Louis Charbonneau. Mathematics in Canada before 1945: A preliminary survey, in Peter Fillmore, ed., *Mathematics in Canada*, vol. I, Ottawa, ON: Canadian Mathematical Society, pp. 1-90. The article appears in both English and French; only the English version (pp. 1-43) is reprinted here.

Judith V. Grabiner. Was Newton's calculus a dead end? The continental influence of Maclaurin's treatise of fluxions, *American Mathematical Monthly* **104** (5) (1997), 393-410.

Ivor Grattan-Guinness. History or heritage? An important distinction in mathematics and for mathematics education, *American Mathematical Monthly* **111** (1) (2004), 1-12.

Ann Hibner Koblitz. Mathematics and gender: Some cross-cultural observations, in Gila Hanna, ed., *Towards Gender Equity in Mathematics Education*, Dordrecht: Kluwer, 1996, pp. 93-109.

Volker Peckhaus. 19th century logic between philosophy and mathematics, *Bulletin of Symbolic Logic* **5** (4) (1999), 433-450. Copyright held by the Association for Symbolic Logic.

Stuart Shanker. Turing and the origins of AI, *Philosophia Mathematica* **3** (1) (1995), 52-85.

Contents

List of Contributors

Thomas Archibald
Dept. of Mathematics and Statistics,
Acadia University, Wolfville, N. S.
B4P 2R6 Canada
tom.archibald@acadia.ca

Jim Bennett
Museum of the History of Science,
University of Oxford,
Broad Street, Oxford. OX1 3AZ
England
jim.bennett@mhs.ox.ac.uk

Louis Charbonneau
Département de Mathématiques,
Université de Québec a Montréal,
Montréal, QC H2X 3Y7 Canada
charbonneau.louis@uqam.ca

Joseph W. Dauben
Department of History,
Herbert H. Lehman College, CUNY,
250 Bedford Park Blvd West,
Bronx, NY 10469-1589 USA

Ph.D. Program in History,
The Graduate Center,
City University of New York,
365 Fifth Ave,
New York, NY 10016-4309 USA
jdauben@att.net

Judith V. Grabiner
Pitzer College,
1050 North Mills Avenue,
Claremont, CA 91711 USA
jgrabiner@pitzer.edu

Ivor Grattan-Guinness
Middlesex University at Enfield,
Middlesex EN3 4SF England
eggigg@ghcom.net

Alexander Jones
Department of Classics,
University of Toronto,
97 St. George Street,
Toronto, ON M5S 2E8 Canada
alexander.jones@utoronto.ca

Ann Hibner Koblitz
Arizona State University
P.O. Box 873404,
Tempe, AZ 85287-3404 USA
koblitz@asu.edu

Karen Hunger Parshall
Departments of History and
Mathematics, University of Virginia,
P. O. Box 400137, Charlottesville,
VA 22904-4137 USA
khp3k@virginia.edu

Volker Peckhaus
Universität Paderborn, Fakultät für
Kulturwissenschaften – Philosophie,
Warburger Str. 100, D–33098
Paderborn, Germany
volker.peckhaus@upb.de

Stuart Shanker
Departments of Philosophy and
Psychology,
Atkinson College, York University,
North York, ON M3J 1P3 Canada
shanker@yorku.ca

Rüdiger Thiele
Karl-Sudhoff-Institut
für Geschichte der Medizin und der
Naturwissenschaften,
University of Leipzig, D-04109
Leipzig, Germany
thieler@server3.medizin.uni-leipzig.de

Kenneth O. May (1915-1977)
Originally appeared in *Historia Mathematica* **5** (1) (1978), 2.
Reprinted with permission of Elsevier Science.

Introduction: The Birth and Growth of a Community

Amy Shell-Gellasch

CSHPM/SCHPM Archivist

The Canadian Society for History and Philosophy of Mathematics, or Société d'Histoire et de Philosophie des Mathématiques, affectionately known as CSHPM/SCHPM, is a society of those interested in the history or philosophy of mathematics. Our constitution states that "the aim of the society is to promote throughout Canada discussion, research, teaching and publishing in the history and the philosophy of mathematics. Any person with interest in the history or in the philosophy of mathematics is eligible to become a member." Those statements are clearly obvious and necessary; however, they do not convey the depth, breadth or quality of the society and its members.

Currently the society has over two hundred members in nineteen countries, including Brazil, Sweden, Bangladesh and Japan. Though most of our members are academics, some do not work in academia but are simply consumers of the subject, either personally, as educators, or through professional interest from other disciplines. The diversity of the CSHPM is also its strength: our different motives and perspectives combine to produce richer portraits of the history of mathematics than we could achieve individually.

Our primary goal is to provide our members with the means to both present and receive current research in the field. This is done primarily through our Annual Meetings and the resulting *Proceedings*, as well as through our official historical journal *Historia Mathematica* and philosophical journal *Philosophia Mathematica*. Our semi-annual newsletters allow members to keep abreast of events in the field as well as interact with one another. Occasionally we hold joint meetings with our sister organization, the British Society for the History of Mathematics, as well as with the Canadian Mathematical Society. Our underlying goal, possibly the more important of the two, is to establish a community of scholars, practitioners and consumers of the history and philosophy of mathematics, with all the qualities and interactions that the word "community" implies.

The groundwork for establishing that community was laid in 1972. In that year Kenneth May sent letters to several colleagues inquiring into the desire among practitioners to organize a society in the history and philosophy of mathematics. The responses that May received show an enthusiastic reception to the idea. In May 1973 the first meeting of the new organization took place at the Learned Societies Congress (often shortened to the “Learneds”) at Queen’s University in Kingston, Ontario. The society was officially formed the following summer when the society’s constitution was approved.

Kenneth Ownsworth May (1915-1977), at the time of the founding of the society, was at the Institute for History and Philosophy of Science (IHPS) at the University of Toronto. In addition he was the editor of the journal *Historia Mathematica*, which he officially launched in 1974, having been in newsletter form for the previous two years. May was an accomplished mathematician, historian and educator. He studied mathematics and economics at the University of California at Berkeley, receiving his A. B. in 1936, and his M. A. in 1937. En route to his doctorate under Griffith Evans, his life took many turns. At the recommendation of Evans, May became a fellow of the Institute of Current World Affairs in 1937, studying economic, social, and political conditions in Europe. He traveled to England and Russia to conduct his research. The next year May married and resigned his fellowship since his position and its funding were unsure. He and his wife then studied at the Sorbonne in Paris and became active in the workers movement. By 1939, he returned to Berkeley to resume work on his doctorate in mathematics with applications to social theory. However, in 1940 he was dismissed from his teaching duties at the University due to his involvement in the Communist Party. In 1942 May ran unsuccessfully for State Treasurer of California on the Communist ticket, nevertheless gaining 44% of the vote.

Just before finishing his thesis, May’s life changed again. With Russia allied with the U.S. for the war, May sought to join the service; however, married men were not accepted in the service at that time. When his wife filed for divorce in mid-1942, May was able to enlist. May joined the 87th Mountain Infantry (10th Mountain Division), and served in the Aleutians (1943) and in Italy (1945). In 1944 he remarried, and after the war he and his second wife stayed in Italy, where May taught mathematics at the Army University Study Center. He returned to California in 1946 and defended his thesis, “On the Mathematical Theory of Employment”, under Evans. May then accepted an assistant professorship at Carleton College in Northfield, Minnesota. During the late 1940s he published and presented his research in mathematics and industrial theory with titles such as his 1947 “The Aggregate Effect of Technological Changes in a Two-Industry Model”. Throughout the 1950s, May’s research focused on election theory, in which he published extensively. In the 1960s May’s interests were directed towards the history of mathematics. In 1966 he moved to the University of Toronto’s Institute for the History and Philosophy of Science, where he promoted the history of mathematics and science through his involvement in the IHPS, his founding of *Historia Mathematica*, and the founding of the CSHPM in 1973.

At the last moment, May did not attend the 1973 Queen’s University meeting at which the society was founded. Charles V. Jones of York University chaired the meeting. At that meeting, the name of the society was agreed upon and *Historia Mathematica* was selected as its official journal. Jones was elected President, along with Thomas Settle as Vice President and J. Lennart Berggren as Secretary-Treasurer. At this meeting a modest set of papers in the history and philosophy of mathematics was presented. During 1973 and early 1974, Jones and Settle drafted the original bylaws by which the society still operates (with only slight modifications).

The following year the first official meeting of the society was held at the Learneds, with sixty charter members. A more extensive program of papers was presented at this meeting, including invited papers from all three of the executive members. These were the first annual guest presentations, of which this volume contains a sampling. After May's death in 1977 the Kenneth O. May fund was established, which helps to bring noted historians to the annual meetings as guest speakers.

Over the years, the society has traditionally held its annual meeting at the Learneds Congress (now the Congress of the Humanities and Social Sciences) every spring. From time to time we sponsor joint sessions with the Canadian Society for the History and Philosophy of Science, the first in 1974. In 1996 reciprocal memberships between the two organizations were introduced. Our most recent meeting with the Canadian Mathematical Society occurred in 2000 in Hamilton; another is planned for the year 2005. On a grander scale, at about the same time CSHPM and the British Society for History of Mathematics (BSHM) became sister organizations, with joint meetings held in 1997 in Oxford, 1999 in Toronto and most recently, 2004 in Cambridge.

As our membership continues to grow and diversify, so does interest in history and philosophy of mathematics from the mathematical community at large. In the past few years, interest in using the history of mathematics in teaching to motivate learning at both the school and collegiate levels, and an interest in the subject in its own right, has increased dramatically. To facilitate this new interest from those outside of the specialty, the History of Mathematics Special Interest Group of the Mathematical Association of America (HOM SIGMAA) was formed in 2002. Initial discussions leading to the formation of this new organization occurred during the annual CSHPM meeting in Quebec, 2001. Two members of the society drafted the constitution of this new group. Though HOM SIGMAA and CSHPM are not officially affiliated, they maintain a close informal working relationship. The CSHPM focuses on scholarly activity in the history and philosophy of mathematics, and HOM SIGMAA focuses primarily on the pedagogical aspects of the history of mathematics. The goal is for a symbiotic relationship that will promote not competition but complementary pursuits. Currently, all the HOM SIGMAA executive members are also CSHPM members.

This volume represents the next major project undertaken by the CSHPM. Since 1973, a wide variety of original work in the history and philosophy of mathematics has been presented at our annual meetings. That work has been recorded since 1988 in our internally produced annual *Proceedings*. After sixteen years, it is time to present some of that material to a wider audience. Though all the papers presented in the *Proceedings* deserve wider attention, this volume will showcase the papers presented by the keynote speakers at the Annual Meetings. These papers are of a consistently high quality by known experts in the field.

On behalf of the Society, I would like to thank Michael Kinyon and Glen Van Brummelen for their time and energy in seeing this project to completion. I would also like to thank all the speakers over the years who have shared their research and interest in the history of mathematics with us at our Annual Meetings. Of course these meetings, and the Proceedings that result, could not happen without those who devote many hours to making sure that the meetings run smoothly and that the *Proceedings* are published. We hope to share our love of the history of mathematics with you through this sampling of CSHPM activities.

1

History or Heritage? An Important Distinction in Mathematics and for Mathematics Education*

Ivor Grattan-Guinness

Middlesex University at Enfield

To the fond memory of John Fauvel (1947–2001)

1.1 Interest and Disagreements

During recent decades there has been a remarkable increase in work in the history of mathematics, including its relevance to mathematics education. But at times considerable differences of opinion arise, not only about its significance but even concerning *legitimacy*–that is, whether or not an historical interpretation counts as history at all. In this paper I consider the latter issue, and also note some consequences for education.

The disagreements are general, in that they may arise for any branch of mathematics in any period or culture; so they need a general resolution. I offer one in the form of a distinction in the ways of interpreting a piece of mathematics of the past. Take such a mathematical notion N; it could be anything from one notation through a definition, proof, proof–method or algorithm to a theorem, a wide-ranging theory, a whole branch of mathematics, and ways of teaching it. By its 'history', which becomes a technical term, one considers the development of N during a particular period: its launch and early forms, its impact, and applications in and/or outside mathematics, and so on. It addresses the question 'What happened in the past?' by offering descriptions. Maybe some kinds of explanation will also be attempted to answer the companion question 'Why did it happen?'.

History should also regard as important two companion questions, namely 'What did not happen in the past?' and 'Why not?'. The reasons may involve the other side of this distinction, which I call 'heritage'. There one is largely concerned with the effect of N upon later work, during any relevant period including that of its launch. Some modernised versions of N are likely to be

*First published in the *American Mathematical Monthly* **111** (1) (2004) 1–12.

taken, for heritage is largely concerned with the question 'How did we get here?', that is, to some current version of the context in question.

The distinction between history and heritage is often sensed by people who study some mathematics of the past, and feel that there are fundamentally different ways of doing so. Hence the disagreements can arise; one man's reading is another man's anachronism, and his reading is the first one's irrelevance. The discords often exhibit the differences between the approaches to history usually adopted by historians and those often taken by mathematicians.

The claim put forward here is that *both history and heritage are legitimate ways of handling the mathematics of the past; but muddling the two together or asserting that one is subordinate to the other, is not.* Many consequences flow from this stance, which will be treated in sections 3 and 4; first let us take a simple and well-known example, from the distant past.

1.2 Pythagoras's Theorem, Euclid Style

One of the best-known theorems in Euclid's *Elements* (fourth century B.C.E.) concerns the sides of a right-angled triangle ABC in Figure 1.

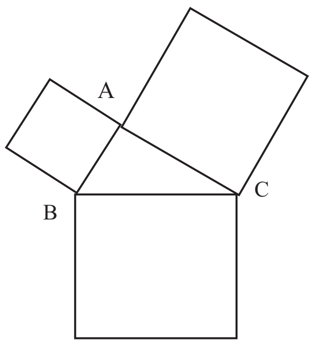


Fig. 1.1.

We recognise it as saying of the sides AB, AC, and BC that

$$AB^2 + AC^2 = BC^2; \tag{1.1}$$

but Euclid actually says something quite different [11, Book 1, Proposition 47J]: in right-angled triangles the square on the side subtending the right angle is equal to the squares on the sides containing the right angle. There is an attached diagram of which Figure 1 is part, and the differences between it and (1) are basic. Not only is (1) algebraic whereas the figure is geometric: the diagram shows the squares outside the triangle, which (1) does not convey.

Were any of the squares to lie over the triangle, then both (1) and the theorem would still be true: but the complicated proof, not shown in the figure, could not be effected. The algebraic character of (1) emerges further when, as was and is commonly done, the letters 'a', 'b', and 'c' are used for the sides: for algebra is the branch of mathematics in which special words and especially symbols are used to a significant extent to represent constants, unknowns, variables, and operations.

Another important difference concerns the word 'on'. Euclid never used the phrase 'side squared', for in his geometrical Books he never multiplied geometric magnitudes together, either in the statement of theorems or (more importantly) in any proof. For example, he did not draw upon side-squaring when proving Pythagoras's theorem, either in the complicated proof just mentioned, which relies upon congruence, or in a more elegant one for the more general theorem about rectangles with the same ratio of sides set upon the sides of the triangle, where the proof deploys similar triangles and ratio theory [11, Book 6, Proposition 31]. Thus 'BC^2' is already a transgression from his geometry (and the frequent use in diagrams of small letters such as 'a^2' even more so). Instead Euclid constructed a square *on* a given line–indeed, in the proposition immediately preceding Pythagoras's theorem [11. Book 1, Proposition 46].

The issue is more profound than it may seem. Both here and everywhere else in the *Elements* Euclid works with *lines* rather than *lengths*, the latter being lines upon which some arithmetical measure has been imposed. Euclid presented geometry without *arithmetic* in the sense just explained; numbers are also present, but for other purposes, such as saying that this line is twice that line, or that the ratio of two lines is the same ratio as $5 : 7$. In the same way he worked with planar regions but not (measured) areas, with solids but not volumes, with angles but not in degrees. By contrast, which is sometimes overlooked, in the arithmetical Books 7–9 multiplication of integers themselves occurs as usual [15].

These remarks concern the history of Euclid. When one moves to its heritage, then a quite different situation arises, in which (1) and many other such equations are prominent. For the *Elements* played a major role in the development of common algebra among some of its Arabic initiators, and a still greater one when Europe at last woke up during the twelfth century and began to elaborate that algebra with symbols introduced both for unknown quantities and for operations. *Both* (1) and Pythagoras's theorem as shown by the figure are legitimate readings of Euclid, but are quite different from each other.

The *Elements* is a particularly interesting historical example, because common algebra as in (1) became the dominating reading of Euclid (including in mathematics education) to such an extent that during the nineteenth century it also became the normal historical interpretation; apparently Euclid had been a 'geometric algebraist', talking geometry but really practising common algebra. A supporter of this reading was T. L. Heath, whose English edition and translation, first published in 1908, is still the most widely used, usually now in the second edition [11]. Greek specialists tell me that his translation is very reliable both to the language and to the mathematics; in particular, for Pythagoras's theorem and all other contexts he says there 'square on the side', not 'square of the side' as many earlier translations had rendered (the word 'apo' can admit both 'on' and 'of' as translations) but which can easily lead to the algebraic 'side squared'. Nevertheless, Heath added to his translation many algebraic versions of the propositions without seeming to notice the differences entailed.

While some historians of that time did not follow the algebraic interpretation of Euclid—for example, the Dutchman E. J. Dijksterhuis [26, chap. 5]—the standard view came under severe challenge only from the 1960s onwards. In particular, in the mid 1970s the historian Sabatei Unguru attacked it strongly, to the opposition of some mathematicians interested in history. Unguru's charges of anachronism and ahistory are largely vindicated: his mathematician opponents were inheritors [20].

We shall take another Euclid example in section 8. First, though, let us explore some general consequences of the distinction.

1.3 Some Principal Differences between History and Heritage

The distinction between the history of a notion N and its heritage obviously involves its respective pre- and post-histories; but much more is at hand, for history has to use post-history also. To see this, let us consider the advice, which is quite often put forward for history of all kinds, about a way of being 'history-minded' about N (say, Pythagoras's theorem in Euclid); namely, forget everything that has happened since N was formed, and read Euclid with the eyes with which he wrote it. But this advice *begs the question at hand.* For in order to forget everything E that has happened since N, then one has to know E already; however, to do that one needs to be able to distinguish E from the history and pre-history of N; but this is the task to be attempted.

Thus the distinction between history and heritage rests in part upon the ways in which notions later than N are to be used. When they are determined to be later notions, the view urged here is this: by all means bring them to bear, and deploy them to understand the heritage from N, but *avoid* feeding them back to appraise its history (such-and-such did not happen). Further, when considering periods intermediate between that of N and some later ones such as now, apply the distinction carefully. Thus, in our example the equation (1) is not only part of the history of René Descartes and the heritage of Euclid but also belongs to the heritage from, among others, the algebraist François Viète in the sixteenth century, whose work also belongs to the history of Descartes. Note also that history is usually *a history of heritages*; it is a tale of mathematicians taking and modifying notions from the past (often pretty recent) without enquiring about the history of those notions.

Various other matters can be explored; a more detailed discussion, largely focussed upon history, is given in a companion paper [16]. The following table summarises the main features of handling past notions N in the two different ways suggested. An apparent contradiction between the third and fourth rows needs to be addressed. When the historian reconstructs past muddles, he will conflate notions that we now know to be different, a feature that the inheritor will stress. But the difference that the reconstruction exposes is that between past ignorance of the distinction, which is different from our (and the inheritor's) present knowledge of it.

Feature	History	Heritage
Motivation(s) to N	Important issue; maybe hard to find (for example, for Euclid's *Elements*)	Probably only of minor interest
Types of influence	Can be negative as well as positive; both should be noted	Likely to draw only upon the positive cases
Relationships of N to earlier and to later notions	Major issue; differences stressed as much as similarities, maybe more	Important issue: similarities stressed more than differences
Handling unclarities evident in N	Reconstruct them, and as clearly as possible	Recognise them, but clean them up
Successful developments	Very important: but also study failures, delays, missed opportunities, and late arrivals	Likely to be the main concern
Role of chronology	Usually important; can be hard to establish	Beyond broad details, not likely to matter so much
Historical consequences	May try to reconstruct the *foresight* (hopes, and so on) for N held by the historical figures	May try to construct *hindsight* and historical perspective of the developments after N

Determinism?	Preferably not claimed: the actual developments were so-and-so, but not necessarily so-and-so	May carry a determinist flavour; we *had* to get here (but see the history column!)
Foundations of a theory	Dig down to them, and build upon a swamp	Lay them down and build up from them, like on solid ground
Level of importance or popularity of N	Can vary over time, independently of content; should be noted (and maybe explained)	Not normally considered; current importance assigned

1.4 Changing Habits

A further type of issue, which is not susceptible to tabular expression, concerns the use of notions that have become standard and therefore are now used habitually. Such habits may well help in determining the heritage from N: but historical anachronism can easily arise, which needs to be controlled. I note three important examples.

First, after an interesting history of its own from the 1870s [9], Georg Cantor's set theory has been part of our mathematical furniture for just over a century; so for the mathematics of this period its use may well be faithful. Now collections of things have been handled in mathematics since at least Greek antiquity; but the earlier theory of so doing was part-whole theory, where (say) British women form part of the class of women, membership is not distinguished from inclusion, and an object is not distinguished from its unit class. The differences between part-whole and set theories are considerable, both technically and philosophically, and the historian needs to mark them carefully. By contrast, the inheritor can deploy set theory with little chance of deception.

Second, while the influence of Euclid was great in Western mathematics, his stress on axioms and common notions was rarely imitated (though to some extent Newton's *Principia* is an example). The axiomatisation of mathematical theories became more prominent only during the late nineteenth century, especially in connection with the axioms of Euclidian and non-Euclidian geometries, and the emergence of abstract algebras [8]. Both developments attracted the attention of David Hilbert, and led him to launch the wide-ranging use of axiomatisation during the first half of the twentieth century, an attitude that has now become pretty standard: a clear path of heritage can be traced up to present-day practises. But the historian should be careful when looking at the structure of earlier mathematical theories, for axiomatisation may well not be prominent beyond specifying basic principles or laws.

Cantor's set theory is a good example: while it too was developed during the late nineteenth century, he showed little interest in the axioms that it may require.

Third, vector and matrix theory have become standard fare in mathematics, though (especially in the second case) only from the l930s onwards and after rather scrappy historical developments in various contexts during the nineteenth century. Once again, care should be exercised in applying them to earlier work. For example, much of the mechanics developed by figures such as L. Euler, J. L. Lagrange, and P. S. Laplace can be rewritten in vectorial and matricial forms, but historical understanding will not profit. For none of these figures knew that their theories could be developed in terms of strings or arrays of scalar elements; they worked instead in terms of collections of simultaneous linear or differential equations, or quadratic and bilinear forms [14, chaps. 5–6]. The introduction of vectors or matrices is not merely a matter of changing notation; new theories are involved. It is of course nice to save such space, for one thing; but if the historian does deploy these theories, then a chronological health warning should be appended.

By contrast to all these cautions to the historians, the inheritors can execute all these reformulations of theory quite legitimately; indeed, much nice heritage mathematics may emerge. Further, some history of mathematics produced after the initial period under study might be created; for, as was mentioned in section 1, mathematicians normally read the past in a heritage spirit.

As an example, take Lagrange and others in mechanics. A major problem, which he formulated in the 1770s, was to prove mathematically that the planetary system was stable. (Previous figures such as Newton and Euler had relied on God to watch out for danger; that is, a religion influenced mathematics.) In terms of matrix theory, Lagrange's brilliant theory sought proof of the reality of all the eigenvalues and eigenvectors (to use modern terms) of a certain matrix. But he had no such theory, and worked with the corresponding quadratic forms; so did Laplace, who adapted his results to some extent; neither man found a watertight proof. The next major contribution came in 1829 from (surprisingly) A. L. Cauchy, and in 1829 he did formulate 'tableaux' of scalar entries in his own work on this problem [17]. Thus matrix theory may—indeed, should—be used to describe Cauchy's contribution, and thus to help us to grasp an important part of his heritage from his predecessors. And we also have a nice example of the 'What did not happen?' question; for Cauchy never realised the significance of his achievement and rarely used it later, so that unfortunately he was not an influential founder of the spectral theory of matrices.

1.5 Some Philosophical Background

It is obvious that this talk of earlier and later notions, the development of theories, and so on, is not confined to mathematics: such features occur also in the histories of other sciences (including technology, engineering, and medicine), and indeed elsewhere (for example and a nice one, practices to be adopted and avoided in the so-called authentic performance of older music). The main general principles that underlie the foregoing discussion are as follows.

First, history is *unavoidable*, whether one likes it or not. A mathematician who presents his theory without concern with history is not thereby immune from it. For example, an enthusiast for axiomatics mentioned in section 3 will lay out his theory in a very formal way without reference to predecessors or precedents; but they will be there, including previous formal theories laid out by preceding axiomatists without reference to their own predecessors or precedents. Thus the question of whether or not one can use history in mathematics is miscast: it is rather the question of whether it is done consciously or not. Indeed, independent of the content of this paper, it is useful to have some general historical idea of a topic of interest, whatever it may be.

Second, knowledge and ignorance go together. This symbiosis has not received the general philosophical attention that it deserves. In particular and of special significance for mathematics, there is knowledge of ignorance, especially when one formulates a problem. When, for example, J. P. G. Dirichlet studied the convergence problem of Fourier series in the late 1820s, he knew that he did not know sufficient conditions on a function to establish convergence to it: finding some was precisely his problem. Having done so, he knew that he did not know whether or not they could be weakened, thereby setting the next problem in this chain (to which the first answer was the Lipschitz condition, by the way). One can also have ignorance of ignorance, or unawareness, where people do not know that they do not know something because the required connections between notions have not yet been laid down. Thus Dirichlet did not know that he did not know how his proof bore upon the specification of function spaces, because that notion did not emerge until the late nineteenth century [21].

Third, and following from the preceding line of thought, knowledge of all kinds is stratified into theory, metatheory, ... For mathematics this means not only metamathematics of the technical kind that Hilbert launched, but also informal kinds. In particular, the history of notion N is one kind, its heritage is another, manners of its possible teaching a third, heuristic strategies to explain its significance a fourth, and there may well be others. The relationship between knowledge and ignorance just outlined lie in the metatheory of the notions involved. Similarly, metatheory requires metametatheory as its own forum for discussion, and so on upwards as far as is needed. An example of metametatheory is the history of the history of mathematics, an interesting story recently recorded in detail in [10]; the comments on Heath in section 2 form an example of it; and this paper itself is a self-referring example, with its heritage (if any) awaited!

The recognition of history and heritage as metatheoretic also releases both historians and inheritors from the need to like what they find in the past that they study. Why should they? After all, they were not there (as a rule). The point seems obvious enough; after all, one can be a good historian (or inheritor) of, say, military history without being a militarist. Yet not infrequently historians and inheritors become overly attached to their objects and figures of study, in any kind of history, and feel that they have to defend what they find. While of course such attachment can be felt if it arises naturally, no compunction to it should even be encouraged.

The generality of stratification is an insight forged in connection with symbolic logic in the early 1930s, thanks principally to Kurt Gödel and Alfred Tarski. In logic the distinction of (object-level) logic itself from metalogic is especially tricky but thereby all the more important; as was known already in Greek times, failure to make a distinction of some kind admits nasty paradoxes. Gradually stratification spread into other disciplines, especially mathematics and some types of philosophy. One follower, inspired by Tarski in the mid 1930s, was Karl Popper. Several parts of his philosophy of fallibilism are metaphilosophical; for example, his preference for indeterminism over determinism [19]. Of particular relevance to this paper is his essay 'On the Sources of Knowledge and Ignorance' [18, introduction], for it contains an insight largely missing from other kinds of philosophy; that ignorance is nice, for it is the site (in metatheory) of our problems when construed as knowledge of ignorance. In most other philosophies ignorance is a disease to be cured by the acquisition of knowledge however that acquisition is claimed to occur (see [25, chaps. 1–6] for the various forms of this view maintained within the sceptical tradition of philosophy). So far explicit use of stratification has not been widely canvassed among prevalent philosophies of history (which are well surveyed in [23]); but it seems worthy of further elaboration.

1.6 General Remarks about History in Mathematics Education

In recent decades a considerable and international increase has developed in the use of history in mathematics education, in order to temper and challenge the normal picture of mathematics as a human-free zone, all answers but no questions, all solutions but no problems. Several edited or authored books and special issues of journals have appeared containing material of various kinds: textbooks significantly informed by the relevant history; summary histories of particular developments; surveys of the lives and works of important historical figures; international and/or multicultural comparisons of the development of (more or less) the same theories; translations of original texts with commentary; and suggested strategies for using history in teaching practice, both in specific contexts and in general. The emphasis often falls upon motivation and context, on showing that mathematics is after all human activity despite appearances, and moreover that much of it is not Western in origin. The range of concerns is well captured in a recent volume [12].

Most attention seems to have fallen on teaching at school and college level, but the university level has also been addressed. Much more work has been done on pure mathematics than on applied or applicable mathematics, or on probability and statistics; a redress of balance would be most welcome. I do not attempt to review this literature here, but I consider the place and utility of the distinction between history and heritage in mathematics education in general.

As with researchers in history mentioned in section 1, there is an evident sense of the distinction in this kind of educational literature, or at least an intuition that the mathematics of the past can be used in different ways. Where is mathematics education to be found between history and heritage? My answer is that *that is exactly where it should be found*, so that it can profit from *both* sides. In particular, if notion N is to be taught, then both its history and its heritage can be used. Euclid's *Elements* is a good example, where the inherited use of algebra has been well used quite frequently. In addition, the historical Euclid deserves attention, with its geometry presented without arithmetic with lines rather than lengths, and the beautiful theory of ratios used in both his geometry and his arithmetic.

1.7 History-Satire and the Calculus

In the paper [13] I introduced long ago the term 'history-satire' to characterise a way in which history and also heritage can be used in mathematics education. Under it the broad features of the historical record are respected and used; but usually many detours and complications occur that, while they attract the historian, will impede teaching and so should be set aside or at most treated only in passing. The 'genetic method' of Otto Toeplitz, which he introduced initially in the late 1920s in connection with teaching the differential and integral calculus, is similar in sentiment [24]. More recently the Mathematical Association of America published a novel and important textbook in real-variable mathematical analysis by David Bressoud, in which he gives prominent places to the main developments, especially of the nineteenth century, such as Fourier series [5].

As Bressoud duly notes, a major innovation of the century was the founding of analysis in the 1820s by Cauchy. His approach was based upon a newly sophisticated theory of limits, not with limit left as an intuitive notion. Undoubtedly it was much superior to the preceding versions in the organisation of the subject and statements and proofs of the theorems; however the loss in heuristics was heavy, and both his colleagues and students objected forcefully to it [14, chaps. 10–11 passim, and 20.8].

For an explicit example, here is a use of history-satire that I found helpful in my own teaching. In a remarkable analogy, Cauchy adapted his real-variable analysis to complex variables and their functions and thereby introduced a major new subject into mathematics. But it seems a strange subject when first learnt: it uses the corresponding expressions as in real-variable analysis, but there are no curves, tangents to them, or areas underneath them to think about or look at. Among the many theorems that Cauchy proved, a main one is now named after him: namely, that the integral of a single-valued and differentiable function with a continuous derivative around and inside a closed contour C is zero. To students, including me long ago, it seemed to be a peculiar result: and a quick and doubtless valid proof using the Cauchy-Riemann equations and Green's theorem did not assuage the perplexity.

Cauchy developed his theory fitfully from the 1810s to the 1840s [22], and this version of his theorem is the last one, with the complex plane available as the site for C. I found that an earlier stage of his theory helped in understanding the theorem. In his treatment of the real-variable integral of $f(x)$ over the range $x_0 \leq x \leq X$ (I use his symbols) he formed the area sum S for a partition of values of x over the range, took successive subpartitions and formed the corresponding sums, and defined the integral as the limiting value of the sequence if it existed at all. This manner of defining the integral has long been standard, and his version is still worth reading and teaching [6, lecture 21].

Soon afterwards Cauchy deployed his analogy. He defined the integral of a finite-valued and continuous complex-variable function '$f(x + y\sqrt{-1})$' by forming the expression corresponding to S for $f(x)$ but with $x+y\sqrt{-1}$ taking a sequence of values between the limits $A =$ '$x_0+y_0\sqrt{-1}$' and $B =$ '$X+Y\sqrt{-1}$' for which both x and y were continuous functions of a parametric variable t. Then, drawing upon integration by parts and the calculus of variations, he proved that the value of this integral between A and B '*is independent of the nature of the functions*' involved [7, sec. 3]. The closed contour theorem then follows by taking the integral along one sequence of values between A and B and back along another sequence under which the required conditions obtain; the two integrals for the two sequences cancel out, so that the value of the integral over C is zero.

Working through the theorem this way certainly took more than a few lines; but the understanding increased substantially, especially as the definition of the real-variable integral had already been taught elsewhere. My account follows Cauchy historically to the extent of deliberately avoiding diagrams, for both types of integral: at that stage in his career he regarded geometric notions as unrigorous and so wished to avoid them. The status of geometry makes a nice point to debate in the classroom, and in fact I increased the measure of satire by using diagrams myself. I also ignored several special cases and other details of the theory as Cauchy was then developing it. But I raised questions such as whether or not Cauchy assumed the derivative of f to be continuous (yes, but implicitly); and I also taught his 1825 version of the residue theorem, noting that, contrary to most later practice, he allowed $x + y\sqrt{-1}$ to go through, and not just round, a pole of $f(x + y\sqrt{-1})$ [6, sec. 8].

The considerations of this section have used the calculus and mathematical analysis because these case studies happen to come from it. But genetic approaches and history-satire can be applied to any mathematical notion or level of teaching.

1.8 The Proposals of Bashmakova

The relationships between knowledge and ignorance outlined in section 5 deserve serious consideration, including the niceness of ignorance as the source of problems (big or small) to tackle. One important area of education where these relationships are prominent is the design of a course syllabus and the manner and order of teaching the topics proposed, when in effect the designer is considering the stage at which the pupils or students should cease to be ignorant of some specific notions.

Let us take an example, examining the historiography proposed in recent years by the Russian historian I. G. Bashmakova, for two of her books have recently been translated into American and published by the Mathematical Association of America for their utility in mathematics education. While dealing with the history of common algebra, her position is put forward in a general way, most explicitly in a joint paper with I. M. Vandaloukis [4]. For them, there are two main stages in handling an historical text. 'First the text should be "translated" into the [sic] contemporary mathematical language, i.e. an adequate model for it should be constructed. This is absolutely necessary in order to *understand* the text, to reveal its mathematical content' (p. 251). In the next stage 'it is necessary to embed the considered work in the context of science of its day' (p. 252).

The authors state that the second stage is 'more difficult' than the first: in my view it might well be impossible, since the first stage will have put so much heritage in place that the historical context could be masked. They state the aim of heritage very clearly: 'the mathematicians of every new age reconsider the previous material and restate it in new terms, thus making it readily available and applicable for the contemporary scientist' (p. 250).

The examples given in Bashmakova's writings seem to exhibit the conflation of history and heritage, without the stress on the distinction between them that was argued in section 7. For example, she takes Proposition 4 of Book 2 of Euclid's *Elements* to express the quadratic identity

$$(a+b)^2 = a^2 + 2ab + b^2 \tag{1.2}$$

as a legitimate prime reading [1, p. 88], [3, p. 165].

In a more recent short history of algebra, coauthored with G. Smirnova, (2) is held to be 'equivalent' to the diagram [3, chap. 2]; and throughout this book the modern notations dominate, although the older terms and symbols are also presented in some detail [3, chaps. 4–5]. The dominance of heritage

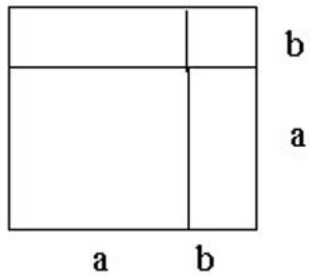


Fig. 1.2.

is clear in a general historiographical appendix, where in specifying the term 'geometric algebra' the authors characterise algebra as an historical category drawn from 'the class of problems associated with algebra today' (p. 164). For them, therefore, Euclid's Book 2 is concerned with algebraic identities such as (2) (see especially Bashmakova in [2]): indeed, her most recent stance is to impose algebraic readings onto ancient arithmetic and geometry for all cultures (see [3, pp. 163–172], where Bashmakova and Smirnova vote for the mathematicians and against Unguru in the disagreement noted earlier in section 2).

The preference for modern notations in the book fits its primarily educational purpose well, exposing an important chain of heritage influences. But the quoted general statements of *historical interpretation* seem to involve heritage mistaken as history. For me, in Book 2 Euclid presents theorems relating subregions of planar rectilinear constructions involving rectangles, squares, and triangles (as in the cited example, where the relative locations of the subsquares and subrectangles are lost in the ubiquitous sign '+'); the algebraic content is empty, as also in all his other geometry Books. By contrast, algebra looms very large in the post-Grecian heritage from Euclid's geometry. Both readings are valuable to mathematics education, though better presented as distinct sources. Indeed, like Euclid himself the history of the theory of polynomial equations is especially suitable for historical satire.

1.9 Concluding Remark

In this paper, and in more detail in its companion [16], I assert that the history of mathematics differs fundamentally from heritage studies in the use of the mathematics of the past, and that both are beneficial in mathematics education when informed by the mathematics of the past. The majority of the examples presented come from fairly modern periods. This is no accident, for they constitute my specialist areas; thus the examples as such have no particular significance. Indeed, since the distinction between history and heritage is held to be general, then indefinitely many more examples could be presented; the reader is invited to construct some of his or her own. A rich resource comes from considering the many ways in which notions are changed, especially when they are (major) theorems or theories. These include the alteration of known results by extension, generalisation, and/or abstraction; reaction to counterexample; the exposure as axioms or as procedures of assumptions previously taken for granted; the adaptation of algorithms; the introduction, or maybe removal, of connections between branches (such as geometry with or without arithmetic); classifications into kinds of objects in a theory; switches between axiom, theorem, and definition; and new applications, both within mathematics and to other disciplines.

More attention has been paid in this paper to issues concerning history and historiography than to heritage and heritage studies; but no value judgement is involved, for, as stated in section 1, neither activity is subordinate to the other one. A companion paper concentrating on good and bad practices in heritage work could be written. The two activities are distinct but they interact in fruitful ways, each posing questions for the other to address.

Acknowledgment. This paper is based upon a plenary lecture delivered to the joint annual meeting of the Mathematical Association of America and the American Mathematical Society that was held in Baltimore in January 2003. Thanks are offered to the former organisation for the invitation.

References

1. I. G. Bashmakova, Diophantine equations and the evolution of algebra, *Transl. Amer. Math. Soc.* **147** (2) (1990) 85-100.
2. ———, *Diophantus and Diophantine Equations* (trans. A. Shenitzer), Mathematical Association of America, Washington, D.C., 1997.
3. I. G. Bashmakova and G. Smirnova, *The Beginning and Evolution of Algebra* (trans. A. Shenitzer), Mathematical Association of America, Washington, D.C., 2000.
4. I. G. Bashmakova and I.M. Vandaloukis, On the justification of the method of historiographical interpretation, in *Trends in the Historiography of Science*, K. Gavroglu et al., eds., Kluwer. Dordrecht, Boston and London, 1994, pp. 249–264.
5. D. M. Bressoud, *A Radical Approach to Real Analysis*, Mathematical Association of America. Washington, D.C., 1994.
6. A. L. Cauchy, *Résumé des Lecons Données a l'Ecole Polytechnique sur le Calcul Infinitesimal*, vol. 1 [and only], de Bure, Paris, 1823; also in *Oeuvres Complètes*, ser. 2. vol. 4. Gauthier-Villars, Paris, 1898, pp. 5–261.
7. ———, *Memoire sur les intégrales définies, prise entre des limites imaginaires*, de Bure, Paris, 1825; also in *Oeuvres Complètés*, ser. 2, vol. 15, Gauthier-Villars, Paris, 1974, pp. 41–189.
8. J. Cavaillès, *Methode axiomatique et formalisme*, 3 vols., Hermann, Paris, 1938.
9. J. W. Dauben, *Georg Cantor*, Harvard University Press, Cambridge, Mass., 1979; reprinted by Princeton University Press, Princeton, 1990.

10. J. W. Dauben and C. J. Scriba. eds., *Writing the History of Mathematics: Its Historical Development*, Birkhäuser, Basel, 2002.
11. Euclid, *Elements*: edition used: *The Thirteen Books of Euclid's Elements*, 2nd ed.. 3 vol.,. ed. and trans. T. L. Heath). Cambridge University Press, Cambridge, 1926: reprinted by Dover, New York, 1956: 1st ed. 1908.
12. J. Fauvel and J. van Maanen, eds., *History in Mathematics Education. The ICME Study,* Kluwer, Dordrecht, 2000.
13. I. Grattan-Guinness, Not from nowhere. History and philosophy behind mathematical education, *Int. J. Math. Edu. in Science and Technology* **4** (1973) 421–453.
14. ________, *Convolutions in French Mathematics, 1800–1840*, 3 vols., Birkhäuser, Basel, and Deutscher Verlag der Wissenschaften, Berlin, 1990.
15. ________, Numbers, magnitudes, ratios and proportions in Euclid's *Elements*: How did he handle them? *Historia Mathematica* **23** (1996) 355–375; printing correction in **24** (1997) 213.
16. ________, The mathematics of the past. Distinguishing its history from our heritage, *Historia Mathematica* **31** (2004) 163–185.
17. T. W. Hawkins, Cauchy and the spectral theory of matrices, *Historia Mathematica* **2** (1975) 1–29.
18. K. R. Popper, *Conjectures and Refutations*, Routledge and Kegan Paul, London, 1963.
19. ________, *The Open Universe. An Argument for Indeterminism*, Hutchinson, London, 1982.
20. D. Rowe, New trends and old images in the history of mathematics, in *Vita Mathematica. Historical Research and Integration with Teaching*, R. Calinger, ed., Mathematical Association of America, Washington, D.C., 1996, pp. 3–16.
21. R. Siegmund-Schulze, Die Anfänge der Functionalanalysis, *Archive for History of Exact Sciences* **26** (1982) 13–71.
22. F. Smithies, *Cauchy and the Creation of Complex Function Theory*, Cambridge University Press, Cambridge, 1997.
23. M. Stanford, *An Introduction to the Philosophy of History*, Blackwells, Oxford, 1997.
24. O. Toeplitz, *The Calculus. A Genetic Approach*, University of Chicago Press, Chicago, 1963.
25. P. Unger, *Ignorance. A Case for Scepticism,* Clarendon Press, Oxford, 1975.
26. K. van Berkel, *Dijksterhuis. Een biografie*, Bert Bakker, Amsterdam, 1996.

2

Ptolemy's Mathematical Models and their Meaning

Alexander Jones

Department of Classics, University of Toronto

In the middle decades of the second century of our era, a Greek-speaking Egyptian living in the town of Canopus close to Alexandria carried out a massive scientific program centring on the writing of about a dozen books on astronomy, astrology, optics, harmonics, and cartography.[1] Unlike his near contemporary Galen, Ptolemy evidently did not lead the sort of career, and certainly did not have the self-trumpeting personality, that would procure notoriety among one's contemporaries, and so we know scarcely anything about his life. But his works were well enough appreciated, in spite of their severe style and uncompromising technicality, so that the great part of them were preserved, almost the sole remnants of their kind of scientific writing from antiquity. Though ranging widely in subject matter, these books revolve around two great themes: mathematical modelling of phenomena, and methods of visual representation of physical reality. In the following, I wish to consider what Ptolemy thought the relationship was between his models and the physical nature that he was describing.

To begin, let us look briefly at what his predecessors made of this question. The explanations of phenomena offered by Greek physical science varied greatly, but they were often framed in terms of two broad principles: first, that change in matter can be reduced to the operations of a small number of fundamental qualities, typically hot, cold, wet, and dry; and secondly, that the phenomena can be modelled by mathematical objects. These principles tended to be regarded as mutually exclusive, so that certain phenomena were referred to qualitative and others to quantitative explanation.

What decided which kind of rationale was appropriate in any particular situation? The historical reality was surely that people stuck to whichever

[1]For the biographical data on Ptolemy see Toomer 1987. (Given the informal nature of the present paper, I have thought it appropriate to furnish the text with references only to translations of the pertinent works and to a few particularly helpful works of modern scholarship. The translations of passages quoted in this paper are my own.)

kind seemed to work best for the subject matter. Certain areas of experience lent themselves more obviously to mathematical modelling than others; in particular, the motions of the heavenly bodies, the esthetics of musical intervals, and the visual perception of shape, distance, and motion were sensed to have a quantitative regularity not shared by physical change in materials such as heating, melting, or burning, which on the contrary seemed to have a fairly direct connection with transference of the qualities hot, cold, wet, and dry.

Aristotle's cosmology, with its inner globe of more or less stratified earth, water, air, and fire enclosed in an outer spherical shell of ether, was in part motivated by this polarity, and in return gave it an *a priori* rationalization. The matter of the heavens–the part of the cosmos where the stars, planets, sun, and moon dwell–is of a different kind from the four mundane elements, subject to a different natural motion (circular revolution as opposed to motion towards the cosmic centre), and not subject to any other kind of change. Aristotle's ether has the power to force change in other things, but considered by itself, its only property is eternally regular circular motion. Hence an Aristotelian astronomy has everything to do with mathematics, and nothing to do with elementary qualities. Earth, air, fire, and water, on the other hand, can be forced by an external agent to move in any direction or to change properties, and, moreover, these processes vary unpredictably in degree and duration. This is why, even if the continual changes among the four elements–including life itself–can be traced back through a chain of cause and effect to the physical action of the heavenly bodies (most importantly the sun's annual revolution alternating between north and south), terrestrial phenomena are not as regular and periodic as the celestial revolutions:

> We see that when the sun comes closer, coming-into-being takes place, and when it recedes, ceasing-to-be takes place, and each happens in equal time. . . . But it often happens that things cease to be in a shorter time because of the mixture of things with one another; for since their matter is not uniform and not the same everywhere, necessarily their comings into being too are not uniform, and some are faster and some slower. . . . (Aristotle, *De Gen. et Corr.* 336b16)

Aristotle's cosmology thus explains why we can have a mathematical astronomy. It does not, however, account for the possibility of mathematical sciences dealing with special aspects of the world of the four elements, although Aristotle recognized that possibility, since he classified optics and harmonics, along with astronomy, as sciences embedding mathematics, or indeed as *branches* of mathematics (*Physics* 194a6). Here and there in the Aristotelian corpus we encounter *obiter dicta* confirming that Aristotle recognized that mundane phenomena could be subject to mathematical constraints, for example in the following passage where he speculates on a possible analogy between harmonic theory and colour theory:

> We have to discuss the other colours [besides white and black], distinguishing the number of ways that they can arise. Now white and black can be placed side by side in such a way that each one cannot be seen because of its tiny size, but the product of the two becomes visible in this way. This cannot appear either as white or as black. But since it must have *some* colour, and it cannot be either of these, it must be a mixture and some different form of colour. In this way one can suppose that there are more colours besides white and black, and that they are numerous in accordance with ratio. For they can lie next to each other in the ratio three to two, and three to four, and in ratios of other numbers; and others can be wholly in no ratio, but incommensurable by some excess and defect. And it is possible that these things subsist in the same manner as (musical) concords; for the colours that are in numbers that form good ratios, just like concords in the other context, would seem to be the most pleasant of colours, for example sea-purple, red, and a few others like these, for the same reason that there are just a few concords, while those that are not in such numbers are the other colours. (*De Sensu* 439b19)

But an analogy is not an explanation, and we are left in the dark as to why simple ratios of whole numbers should have a special status in a world of geometrically continuous matter and change. Similarly one is left wondering why vision follows straight lines if it is in fact a process of continuous change in nonuniform matter.

For a working scientist of the Hellenistic or Roman periods in search of a broad rationalizing framework in which to set his own theorizing, Aristotle's cosmology and conception of matter were not the only ones on offer. In the first place, Epicurus revivified atomism into an elaborate, strictly materialistic physics in which all matter and change are reduced to the chance motions of eternal atoms, endowed with a minimum of properties (shape and size), in an infinite void. Epicurus has sometimes been portrayed as a prophet of science; in reality he was no friend to the sciences of his time. He endeavoured to show how the phenomena for which the astronomers sought unique explanations could result from numerous different physical situations, any of which might be temporarily valid at some time and place within his boundless universe; his theory that vision occurs by means of films of atoms that continually peel off bodies and fly off in all directions would not have stood up long to the scrutiny of a practitioner of geometrical optics; and in general he contemned any inquiry into nature that was not subordinated to his ethical goals, freeing humanity from avoidable pain and fear.

The physics of the Stoics was closer to Aristotle's. We find again an insistence that matter is geometrically continuous and reducible to variable mixtures of a restricted number of fundamental stuffs, which at one level of analysis prove to be the familiar earth, air, fire, and water. The Stoic cosmos is finite and spherical, but there is no outer shell of special unchanging matter for the heavenly bodies, and the cosmos in its present differentiated state has a finite span of life. On the other hand, although Stoic physics is strictly materialist, its cosmos is orderly and rational. The organization of the cosmos and its parts is effected by pneuma, a vital mixture of fire and air that extends in varying degrees throughout the cosmos and that has the power to "tense" the bodies with which it is intermixed. In place of a reductionist explanation of the mathematical behaviour of phenomena, we encounter a deistic appeal to the will of the cosmic mind. (It should, however, be kept in mind that our sources for Stoic physics are less satisfactory and more controversial than those for Epicurean physics, and in any case Stoicism was considerably more open to innovation than Epicureanism with its *ipse dixit* deference to its founder's pronouncements.)

Alongside these more or less coherent systems there existed a looser tradition of physical speculation, which we call "Peripatetic" because its most prominent known advocates, in particular Theophrastus and Strato, were close associates and followers of Aristotle. This was an eclectic approach, grounded in observation and analogy, and again materialistic. Properties of matter and processes of change are explained in fairly mechanical terms, for example by supposing that materials can be composed of particles that can be packed loosely or tightly, but the particles lack the permanence of true atoms and are less denuded of innate characteristics. Aristotle's fifth element seems to have won no following; the heavens were instead supposed to be composed mostly or entirely of fire. This fire might be endowed with special properties, perhaps, but the divide between the celestial and mundane spheres was inevitably blurred.[2]

Such were the main lines of physical thought evolving during the century following Aristotle's death. It was also at this time that the earliest surviving works that treat physical problems using mathematical models were written. These include works on astronomy by Autolycus, Euclid, and Aristarchus, works on statics by Archimedes, and works on optics and harmonics by (or at least ascribed to) Euclid. What is striking about these works is not only the attempt to deduce phenomena through explicit axiom and theorem structures, but also the fact that these works seem deliberately to evade physical interpretation of the axioms.

One would dearly like to know what developments the subsequent three and a half centuries brought. The state of evidence is far from encouraging. Thus, of the numerous books written by undoubtedly the most important mathematical scientist of this period, Hipparchus, we possess only one, and with scarce gratitude and less justice we tend to dismiss that work as atypical and uninteresting. Among the philosophers, Posidonius stands out as a writer who undoubtedly exerted a considerable influence on physical thought. One recognizes in some of the reports of his lost writings the tincture of Peripatetic

[2]The rejection of Aristotle's fifth element is ably discussed by Falcon 2001, 121-183.

physics in his Stoicism for which he was later criticized; it is harder to discern a serious engagement with the mathematical sciences.

For that, we must turn to Theon of Smyrna, a Platonist philosopher of far lesser distinction than Posidonius, but one with the accidental merit that a large part of one of his books, *The Mathematics Useful for Reading Plato*, has come down to us.[3] Theon's mathematics embraces harmonics and astronomy, and the long astronomical section is of particular interest here. Theon exposes, with geometrical demonstrations, the epicyclic and eccentric models as assemblages of circular paths in the plane; but he insists that these circles are not mere abstract conceptions but stand for spheres of ether such that, for example, an epicycle is a rotating sphere nested in the gap between two concentric spherical shells which revolve together, bearing the epicycle with them. Theon was a mere generation older than Ptolemy, but this is enough to establish that the revival of Aristotle's etherial spheres and their adaptation to non-homocentric models was not due to Ptolemy, though it may have been fairly new science in his time.

It makes sense in several ways to begin considering Ptolemy's attitude to mathematical models in the context of his astronomy. This was the science closest to his heart, the only one on which he is known to have written a multiplicity of books. His central treatise on astronomical modelling, known to us as the *Almagest*, preceded most of the others, yet it followed upon a quarter-century of personal observation and analysis.[4] It is also a monumental piece of reasoning, much more complex and at the same time more structurally unified than his other large works.

The models with which the *Almagest* is concerned are kinematic geometrical constructions built up from circular motions representing the paths travelled by the heavenly bodies (the sun, moon, planets, and fixed stars). Most of the bulk of the *Almagest*, and most of its mathematics (in the usual sense of the word), is devoted to determining the radii, rotational velocities, and orientations of the components of each model. These parts, taken in isolation, leave open the question whether the circles in the diagrams stand for some sort of physical bodies in motion, or whether they are just abstract analytical components of a complex motion which the heavenly bodies perform due to undetermined physical causes.[5] We can at least dismiss a third option, that they are mere computational devices with no necessary relation to what the heavenly bodies really do, but by which one can reproduce the phenomena seen by a terrestrial observer; Ptolemy's treatment of parallax and eclipses depends on the assumption that his lunar and solar models correctly describe the distances of the sun and moon from the earth as well as their directions from the observer.

[3] The most reliable translation is Dupuis 1892.

[4] Toomer 1984.

[5] On the question of Ptolemy's realism in the *Almagest* and *Planetary Hypotheses* see Lloyd 1978.

However, the broader deductive structure of the *Almagest* decisively commits Ptolemy to believing that his circles stand somehow for material bodies, even if it is not made explicit precisely how they do so. This may be seen by examining how Ptolemy arrives at each model *before* turning to the deduction of its numerical parameters. I take as an illustration Ptolemy's model for the moon, since Ptolemy's presentation of this model in Books 4 and 5 of the *Almagest* is particularly explicit about the stages by which the model is worked out.

Ptolemy starts out in *Almagest* 4.5 with a working hypothesis, which he warns us will later be disproved, that the moon has a "single and invariant" anomaly, that is, that its apparent progress along the ecliptic has a periodic variation that always repeats exactly. He asserts that two models identically produce this phenomenon. In one model (Fig. 1), the centre of an epicycle E travels eastward along a circular deferent concentric with the earth T with uniform angular velocity (relative to an arbitrary stationary radius from the earth's centre), while the moon M travels along the epicycle uniformly in the opposite direction (relative to the radius from the earth's centre to the epicycle's centre). The angular velocity of the moon on its epicycle is slightly less than that of the epicycle's centre on the deferent. In the other model (Fig. 2), the centre of an eccentric circular orbit C revolves with a slow uniform westward motion along a circle concentric with the earth T, while the moon M travels along the orbit with a uniform eastward motion (relative to the radius from the earth's centre to the eccentre's centre). The two models, as Ptolemy proves, are kinematically interchangeable; that is, any set of positions in space of the moon for specific dates generated by the one model can be generated identically by the other. Moreover, Ptolemy knows already that they are both incorrect, because the lunar anomaly is not simply periodic. Nevertheless, Ptolemy selects the epicyclic model as the basis for a preliminary lunar theory in which the numerical parameters are determined by analysis of observations of lunar eclipses. It is noteworthy that Ptolemy makes a point of showing that the same parameters result from several different sets of observation reports, thus establishing that the preliminary model is computationally valid for all eclipses (and by extension, all oppositions).

The motivation for Ptolemy's selection of the epicyclic model only becomes fully evident when he shows in *Almagest* 5.2 how it is defective. He finds that the "equation," or difference, between the moon's observed position and its mean position (that is, the direction to the epicycle's centre according to the model) is in general greater than the model predicts, with the discrepancy vanishing when the moon is at 0° or 180° elongation from the sun and maximum when it is at ±90° elongation. In an epicyclic model the equation is explained by the planet's motion on the epicycle, so that the new phenomenon (essentially equivalent to "evection" in later lunar theory) would amount

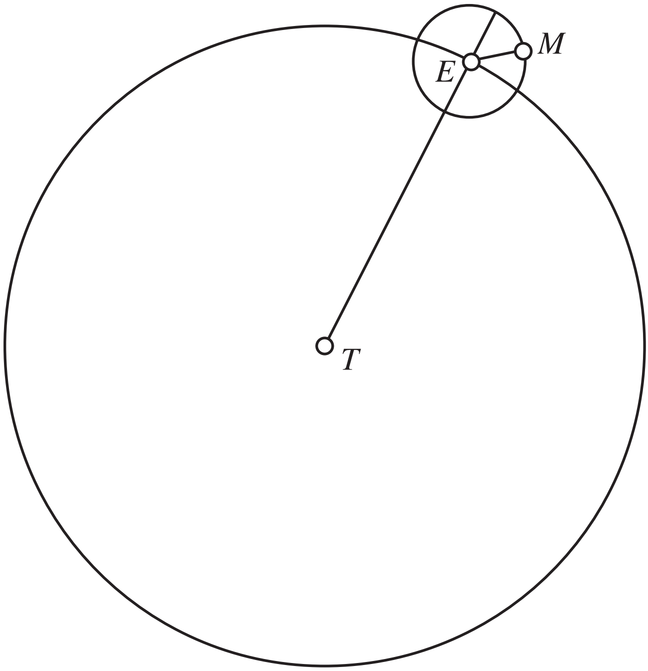


Fig. 2.1. Ptolemy's simple epicyclic model for the moon.

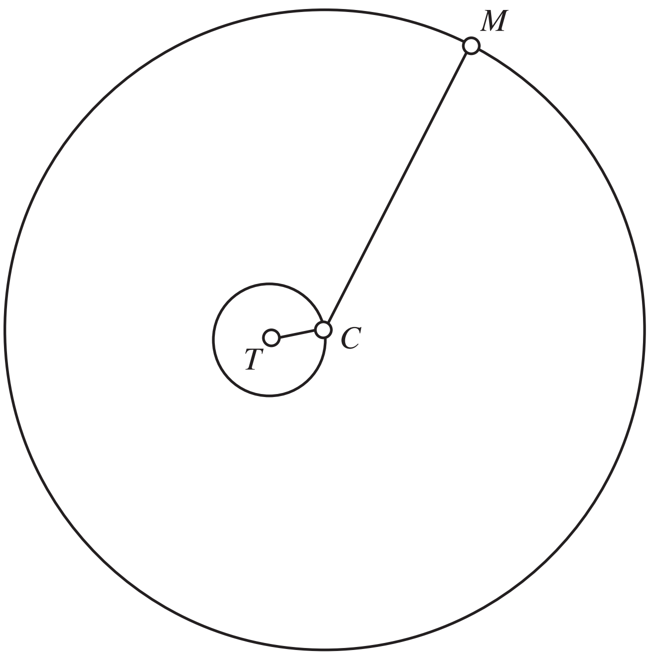


Fig. 2.2. Eccentre model equivalent to the model of Fig. 1.

to an apparent enlargement of the epicycle.[6] Ptolemy accounts for this easily (Fig. 3) by replacing the concentric deferent in the preliminary model with an eccentric deferent, the centre C of which revolves around the earth T at a rate such that the epicycle's centre E (which is still revolving uniformly as seen from T) is furthest from the earth whenever the mean moon and the mean sun are aligned or diametrically opposite. Now it is true that if Ptolemy had employed an eccentric orbit to effect the anomaly in Book 4, he could have corrected the model in Book 5 by adding an epicycle (or even a second, independent eccentricity), but the relation of the components to the phenomena would have been much less intuitive. And in any case Ptolemy fine-tunes the model, for closer agreement with observations, by stipulating that the moon's motion on its epicycle is uniform as measured relative to a revolving radius, not drawn from the centre of the deferent C or from T, but from a distinct point D such that T is always the midpoint of C and D. This could not, I believe, be translated in any straightforward way into a model in which the primary anomaly is produced by an eccentre.

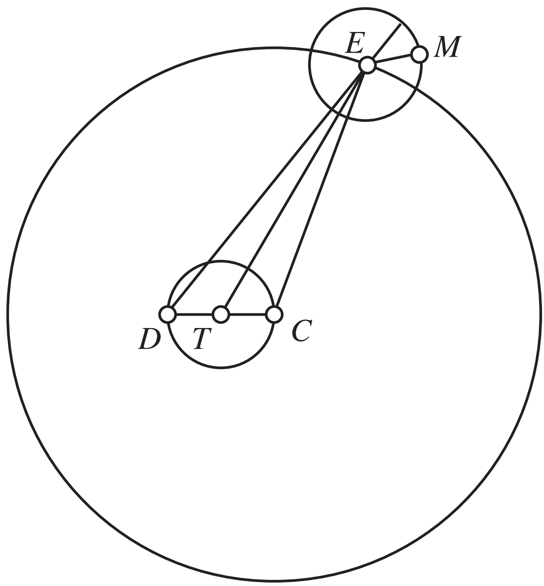


Fig. 2.3. Ptolemy's eccentre-and-epicycle model for the moon.

Thus all the stages from the selection of a basic model type to the final model are motivated in Ptolemy's exposition by the requirements of agreement with observations, simplicity, and a clear one-to-one correspondence of the

[6]For the relationship between Ptolemy's so-called second anomaly of the moon and the component called "evection" in modern lunar theory, see Neugebauer 1975 v. 3, 1108–1109.

elements of the model to the basic facts about the moon's motion. A similar account could be given for Ptolemy's deduction of the models for the sun and the five planets. Ptolemy makes no appeal in these parts of the *Almagest* to physical constraints arising from the corporeal nature of the model. But this is because those constraints have already been taken into account at a still earlier stage, the decision to build all the models out of uniform circular motions, which is made once and for all in *Almagest* 3.3, just before the first discussion of the sun's model. Here Ptolemy writes:

> The next task being to exhibit also the apparent anomaly of the sun, the assumption must first be made that the shiftings of the planets [including the sun and moon] in the trailing direction of the heavens [i.e., westward] are uniform, just like the movement of the whole [heavens] in the leading direction [i.e., the daily eastward rotation of the heavens], and they are circular by nature, that is, the straight lines that are imagined as leading the heavenly bodies or their circles in their revolutions sweep out in all cases equal angles in equal times with respect to the centres of each one's revolutions, while the apparent anomalies pertaining to them are produced by the positions and arrangements of the circles *on their spheres*, by means of which they make their motions, and nothing in nature really occurs that is foreign to their *eternity* in connection with the imagined irregularity of the phenomena.

This is one of only a handful of references in the *Almagest* to the circles in the models as being on the surfaces of spheres; when he does this, it is always in a matter-of-fact way, implying that the reader will already be familiar with the conception. In this particular passage Ptolemy uses language connecting the idea of uniform circular motion with physical nature and eternity, so that ether, though not explicitly named, is inevitably called to mind.

And this brings us back to Ptolemy's very first chapter, *Almagest* 1.1. Here he defines the science of which his subject matter is a part, which he calls "mathematics" (the *Almagest*'s real title is *Mathematical Composition*), as the study of shapes and spatial movements in all kinds of bodies, whether eternal and etherial or perpetually changing and composed of the four elements. Mathematics offers "sure and unshakeable knowledge," and when concerned with the etherial heavens, this knowledge is as eternal as its objects. In other words, the conviction that the heavens are composed of etherial bodies, which are by their composition both eternal and subject to no kind of change except circular revolution, guarantees the legitimacy and truth of the kind of reasoning that the *Almagest* embodies. It is noteworthy that, while practically every other theoretical hypothesis in the *Almagest* is justified by some empirical argument, the hypothesis of the etherial nature of the heavens is given axiomatically at the beginning.

His claim to be arriving at "sure and unshakeable knowledge" in the *Almagest* turns out in practice to have certain limitations. Numerical parameters are, by his confession, knowable only approximately, and in particular the rates of rotation become more precisely known as we accumulate a longer temporal span of observation reports. Ptolemy does not say outright whether he believes that his specific model structures (exclusive of their numerical parameters) are certainly valid. What the *Almagest* does affirm through its broad plan is that Ptolemy's models suffice to explain all the known phenomena of the heavenly bodies, including eclipses, planetary retrogradations, and visibility conditions. But the impression of finality is moderated, not only by the way that Ptolemy recounts his discovery of the moon's evection (might there be other phenomena waiting to be noticed?), but also by his passing reference to alterations he had only lately made in his models for Mercury and Saturn. In a famous passage towards the end of the work (*Almagest* 13.2), he justifies the complexity of his models for the latitudinal motion of the planets by affirming that the principle of simplicity in models should not be allowed to override the necessity to account for the phenomena, since what seems complex to us with our experience of the imperfections of mechanisms built from the four mundane elements may be simple to essences that are eternal and free from hindrance. Implicit in this is his confidence that his models really are the simplest that can be brought into agreement with observation.

Ptolemy's reticence regarding any but the most fundamental properties of ether and regarding the way in which the geometrical objects that constitute the *Almagest* models are supposed to be instantiated in etherial "spheres" in the actual heavens may be due partly to a reluctance to digress from the core subject matter of the book, but another reason may be that he had not yet given these topics much thought (just as he tells us in *Almagest* 2.13 that he has not yet worked out the list of geographical locations that he eventually delivered in the *Geography*). In a much later work, the *Planetary Hypotheses*, Ptolemy has considerably more to say about the spheres.[7]

The *Planetary Hypotheses* is ostensibly an exposition of the *Almagest* models, with some revisions, described in a manner that will be helpful for people who wish to make demonstration models or planetaria, with the parts either manually adjustable to their positions at any date or driven by a mechanism. After a first book in which Ptolemy sets out the parameters of all the models individually and proposes a scheme for nesting the models one inside the next, from the moon's model outwards to those of Saturn and finally the fixed stars, he turns in Book 2 to a consideration of the models as three-dimensional corporeal objects, that is, the "spheres" alluded to in the *Almagest*. Here Ptolemy engages in an extended discussion of his notion of how etherial matter works.

[7]The original Greek text of the *Planetary Hypotheses* is extant only for the first part of Book 1, for which see Heiberg 1907, 70–106; there is no reliable modern translation from the Greek. The whole of Book 1 in the medieval Arabic translation is edited and translated into French in Morelon 1993. For the Arabic text of Book 2, one currently depends on the German translation of L. Nix in Heiberg 1907, 111–145 and the facsimile of a manuscript in Goldstein 1967. Murschel 1995 is an excellent synopsis of the work.

Etherial bodies, he says, are subject to no external influence or alteration. To each independent motion in the kinematic models there corresponds a rotating etherial body incited into motion by the power of the visible heavenly body that it bears. These visible bodies (i.e., the sun, moon, planets, and stars) are the same in composition as the matter that surrounds them. They differ, however, in that they issue rays that have a power to penetrate other bodies, analogous to our intellects and vision. Moreover, their ability to set their spheres in motion is analogous to the power of our minds to cause our bodies to move; but in the celestial case the movement is utterly effortless.

Aristotle's cosmology had been strongly influenced by Eudoxus' astronomical models, in which the motions of the heavenly bodies were produced by combinations of circular motions all concentric with the centre of the cosmos; hence he could ascribe to ether a "natural motion" always perpendicular to any radius from the cosmic centre (in contrast to the natural motion of the four elements, which is always rectilinearly centripetal). For Ptolemy this cannot do, but he proposes a novel principle, that what Aristotle had characterized as natural rectilinear motion is in fact only natural to a body that has been removed from its "natural place." The etherial bodies, being already in their natural place, are subject to no tendency to migrate up or down, but are free to stand still or rotate effortlessly. "Mathematics" (i.e., deductive mathematical astronomy in the style of the *Almagest*) allows for two possibilities for the shapes of the etherial bodies. On the one hand, they can be spherical shells and solid spheres; but if so, they do not have to be imagined as being driven in their rotations by their axes, as in a mundane machine. Indeed, the entire polar regions of the spheres seem to Ptolemy to be superfluous to their motions, so that he is prepared to restrict the mobile bodies to equatorial slices of spheres and spherical shells, so-called "tambourines" and "rings," which are presumably sandwiched between regions of stationary ether. Since distinct etherial bodies can slide freely against each other, there is no need to imagine "unwinding" spheres that cancel out the revolutions of outer spheres, such as Aristotle had imposed on his mechanistic interpretation of Eudoxus in *Metaphysics Λ*.

When we come to the detailed description of each heavenly body's physical model, we find that the basic conception is similar to Theon of Smyrna's, but extended to include eccentric as well as epicyclic motions. Fig. 4 (a cross-section through the plane of the moon's orbit) shows how Ptolemy conceives of the arrangement of etherial bodies that bring about the moon's motion, on the assumption that the bodies are complete spheres or spherical shells. The entire apparatus must be thought of as being spun about the poles of the celestial equator with the daily rotation of the heavens. The outermost shell A rotates around the poles of the ecliptic with the slow motion of the moon's nodes. Within this is a shell B rotating around the poles of the inclined plane of the moon's orbit at the rate that, in the model of *Almagest* Book 5, the centre of the moon's eccentre revolves around the earth relative to the nodes. Cut out of shell B (and actually dividing it into two noncontiguous parts) is an eccentric shell C that has embedded within it the solid epicyclic sphere D. C and D together revolve uniformly as seen from the centre of the cosmos T with the rate that the centre of the epicycle revolves around the earth in *Almagest* 5. Finally, the epicyclic sphere rotates, carrying embedded close to its surface the moon M itself, producing the primary anomaly. This physical model is wholly consistent with the *Almagest* model, except that Ptolemy abandons the special radius with respect to which the moon's regular revolution on the epicycle is reckoned, instead stipulating that the moon's revolution is uniform relative to the radius from the centre of the cosmos. At the beginning of the *Planetary Hypotheses*, Ptolemy writes that the models as set out in this work incorporate some revisions to the *Almagest* models based on newer analysis of observations, but also that he is making some minor simplifications purely for the sake of an easier construction of demonstration models; one is left uncertain which kind of change is being made here in the lunar model.

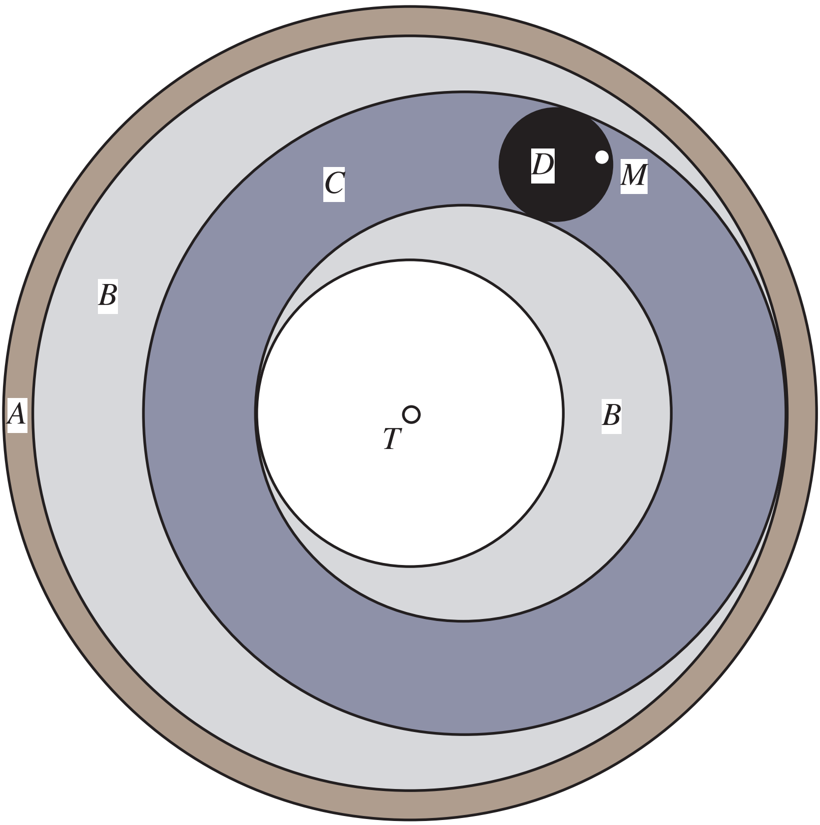


Fig. 2.4. Cross-section of Ptolemy's etherial-spheres model for the moon.

To the extent that the *Planetary Hypotheses* is intended as a description of the reality of the heavens (as opposed to its professed purpose of giving designs for didactic three-dimensional illustrations that we can construct), the pronouncements in the book are more equivocal than those of the *Almagest*. Ptolemy is quite sure of the etherial composition of the heavens, and also quite sure of the fundamental geometrical structures of the celestial motions; but the specific way that these geometrical structures are embodied in the etherial matter is open to alternatives (complete spheres or equatorial slices, spinning driven by cosmic souls or by planetary rays or by the axes). Ptolemy generally tells us which way he is inclined to choose, but these topics are not the province of the "unshakeable knowledge" of mathematics.

With the *Tetrabiblos* (a work written after the *Almagest* but probably well before the *Planetary Hypotheses*), Ptolemy turns from pure contemplation of the celestial realm of ether to an investigation of the action of the heavens upon the world of the four elements.[8] The fundamental assumption, comparable in its role in Ptolemy's astrology to the hypothesis of the uniform circular motion of ether in his astronomy, is propounded in *Tetrabiblos* 1.2:

> The fact would appear utterly obvious to everyone through even a few considerations that some power is given forth and reaches from the etherial and eternal nature to all the region around the earth, which is in all respects subject to change, with the first elements below the moon, i.e., fire and air, surrounded and directed by the movements in the ether and surrounding and directing all the rest, i.e., earth and water and the plants and animals within them.

There are, however, important differences between these fundamental hypotheses. In both, the etherial matter is simply a given. But the property of uniform circular motion in the *Almagest* is justified on *a priori* grounds (circular rotation being the only kind of eternally unchanging motion that can be conceived), whereas the property of exerting a power of change on the four elements is argued directly from empirical facts. Ptolemy backs it up with a series of examples of situations where laymen know perfectly well, and act on the knowledge, that the motions of the heavenly bodies affect (or at least predict) mundane phenomena such as seasons and weather, floods and tides, and the generation of plants and animals. Secondly, and more significantly for our topic, uniform circular motion is a *mathematical* behaviour, which leads immediately to the modelling of the *Almagest*, while power to change the elements is by its nature not mathematical, since it operates with qualities such as hot and cold, wet and dry. And this creates a problem: how can cause-and-effect relations operating at the qualitative level, and largely within the "irregular" sublunary part of the cosmos, be well described by the predictive mathematical models of astrology?

[8] Robbins 1940.

It must be confessed that Ptolemy evades this problem. Essentially Ptolemy relies on the orderliness of the heavens to justify the mathematical structure of the predictive schemes of his astrology, but appeals to the disorderliness and complexity of the mundane environment to explain why astrological predictions, even when made according to the most correct principles, are not certain to be borne out. Moreover, the schemes that Ptolemy sets out to rationalise are in great part the rather chaotic traditional practices of the astrology of his time. Though he allows himself to reform or suppress some of this tradition in accordance with his physics, he can only go so far in that direction since his claim that astrology is a valid science depends heavily on the assumption that the traditional practices really work. One can sense his delight in finding here and there some apparent pattern in the jumble, for example when he finds harmonically significant ratios embedded in the “aspects” (astrologically significant linkages of zodiacal signs forming sides of triangles, squares, or hexagons), or when–indulging in a topos beloved of authors–he recovers from an old, neglected, and nearly illegible manuscript a gloriously complicated rationale for the seemingly nonsensical but empirically verified “Egyptian” system of terms (divisions of zodiacal signs associated with individual planets). Elsewhere Ptolemy almost seems to give up trying to explain, and lapses into catalogues of astrological associations scarcely distinguishable from the manuals of astrologers who were less sophisticated from a scientific point of view.

Optics, which in antiquity meant the study of visual perception, was a more fruitful subject for the interplay between mathematical and physical modelling. As in astronomy, there existed a range of well-established phenomena that lent themselves to explanation in terms of a geometrical model, in this case the “visual ray,” diagrammed as a straight line extending from the viewer’s eye to a point on an object. The hypothesis is that when a visual ray exists between the eye and a point on a body, that point is seen. The eye (or mind) always perceives the seen point as being in the direction in which the ray sets out from the eye, even if the ray is reflected or refracted at the interface between two bodies. This directional information provides the eye with indications of the shape, position, and movement of bodies; on the other hand, the ray conveys to the eye either limited knowledge or no knowledge at all about the distances to the point it perceives.

But what are these lines really? The classic exposition of Greek geometrical optics, repeatedly cited or paraphrased by later authors, was Euclid’s *Optics*. This treatise does not explain the physical nature of the visual rays but does specify that they are discrete, with spaces between the individual rays that grow wider as the rays fan out towards more distant objects; moreover, some of the explanations of visual phenomena appear to assume that the rays are somehow attached to the eye (so that as the eye moves, the rays move accordingly). The gaps between the rays provide an explanation of the fact that objects are seen less clearly, or not seen at all, as they become more distant. But the gaps also lent themselves to a physical interpretation of the rays that is found in Peripatetic texts approximately contemporary with Euclid. According to this interpretation the eye emits, through pores in its surface, exquisitely thin and straight rods of matter (typically composed of fire) that extend with unimaginable swiftness until they encounter a body, at which point their progress may be impeded (in which case a visual perception of the body occurs) or reflected (if the surface is smooth enough) or refracted (if the body is porous).[9]

In his *Geography* Ptolemy treats the problem of planar map projections as essentially one of optics: how can one devise an appropriate framework of lines representing parallels and meridians to give the illusion of a part of a spherical surface?[10] Ptolemy takes for granted many elementary perspective consequences of the hypothesis of rectilinear visual rays. In one passage concerning the relation between the appropriate size of a map and the expected distance of the spectator from it, Ptolemy invokes the Euclidean gaps, which suggests that at this stage in his career (not long after the *Almagest* and *Tetrabiblos*), if he had any opinion at all about the physical nature of vision, it was not far removed from the Peripatetic notion of discrete material emanations.

When he came to write his *Optics* (a work that I suspect was among his last writings), Ptolemy had changed his mind.[11] He now speaks of the eye as emitting an entity conventionally translated as the "visual flux," a cone comprising a geometrical continuum of rectilinear rays that are stronger or weaker in perceptive power both to the extent that they have to extend a shorter or longer distance from eye to object, and to the extent that they are nearer to or further from the central axis of the cone.[12] It is to this weakening of the rays, rather than any supposed gaps between them, that fuzzy vision of distant or peripheral objects is due. Thus Ptolemy's geometrical treatment of visual phenomena thus preserves the parts of the Euclidean scheme that depend on the rectilinearity of the visual rays (namely, perspective phenomena, reflections, and refractions) but replaces the somewhat clumsy Euclidean handling of visual resolution with a more flexible and powerful hypothesis.

Unfortunately the entire first book of the *Optics* is lost, and with it Ptolemy's discussion of the physical makeup of the visual flux. Obviously he cannot have thought of it as a body, at least not the kind that displaces other bodies that formerly occupied its space, which is the only kind of body

[9] The Peripatetic texts and their possible relation to the Euclidean model are discussed in Jones 1994.

[10] Berggren and Jones 2000.

[11] The *Optics* survives, lacking its beginning and end, only in a medieval Latin translation of an Arabic translation, a circumstance that causes great difficulties of interpretation. The French translation in Lejeune 1989 and the English one in Smith 1996 are both useful, though under the circumstances neither can claim to represent Ptolemy's meaning exactly throughout.

[12] The Latin term rendered as "visual flux" is *uisus*, which almost certainly represents the same Greek word *opsis* that, in Euclidean optics, refers to the single visual rays; but Ptolemy used a different word when he meant an individual line of sight.

that is envisioned in Aristotle's or in Peripatetic physics. It seems likely that Ptolemy resorted to ideas from Stoic physics, which allowed for having distinct elements occupy the same space as if in layers. In this manner the Stoics could hypothesise that the entire cosmos was pervaded and regulated by *pneuma*. Ptolemy may have suggested that the eye issues the visual flux as an overlapping layer of matter in the space between eye and object; or perhaps more likely, he could have attributed to the eye a faculty of radiating a tensing power, creating the flux by means of the *pneuma* already present in the intervening space. We recall that in the *Planetary Hypotheses* he asserted a kinship between the motive power of the heavenly bodies and the analogous power in living things. As it happens, one of the very few references to Ptolemy's *Optics* in other authors that appear to pertain to its lost first book is a sentence in a work on physical topics by the eleventh-century Byzantine writer Simeon Seth: "Ptolemy says in his *Optics* that the visual *pneuma* is etherial and composed of the fifth element." This "visual *pneuma*" is probably the substance of the cone of the visual flux, and so we have a remarkable fusion of Aristotelian and Stoic element theory. The sixth-century philosopher Simplicius gives us a further clue when he writes:

> It should be noted that Ptolemy in his book *On the Elements* and in his *Optics*, and the great Plotinus, and Xenarchus in his *Difficulties Addressing the Fifth Element*, assert that motion in a straight line belongs to the elements when they are still in a place that is not natural to them, but (such motion) no longer belongs to them when they have assumed their natural place.... Manifestly they do not move when they are completely in their natural state, but, as the aforesaid men, i.e., Ptolemy, Xenarchus, and Plotinus, say, when they are in their natural state and in their proper places the elements either stand still or move in a circle.

This is precisely the notion that we have seen Ptolemy putting forward in the *Planetary Hypotheses*—which Simplicius does not cite here. It is obvious why Ptolemy would have repeated it in a (no longer extant) work on the elements; but what relevance can it have had in the *Optics*? I suspect that Ptolemy was invoking it here for a purpose converse to his purpose in the *Planetary Hypotheses*. There, the point was that etherial bodies in the heavens can spin freely and effortlessly even if their revolution is not concentric with the centre of the cosmos; in the *Optics*, perhaps Ptolemy claimed that the etherial matter of the visual flux, connected with our sight and thus *displaced* from its natural place in the heavens, travels in straight lines. A final observation worth making is that the Xenarchus cited by Simplicius as sharing this idea was active about the late first century B.C., so that here we may be able to identify the source of a principle that, in Ptolemy's hands, simultaneously accounts for the geometrical properties of the models of astronomy and of optics.[13]

I have saved for last what may have been Ptolemy's first major effort at mathematical modelling, the *Harmonics*.[14] The subject of this work calls for some explanation. Ancient Greek music was essentially melodic unison melody, occasionally employing singing or playing at the octave or the sounding of simultaneous distinct notes as an effect, but free of harmony in the modern sense. There existed numerous systems of relative pitches (i.e., scales) in which melodies could be composed, none of which involved a sequence of intervals quite like the diatonic scales on which most modern Western music is based. The science of harmonics, as Ptolemy presents it, investigates models that explain why certain intervals and combinations of intervals are esthetically pleasing and hence exist as constituents of the music actually produced in Ptolemy's time.

Unlike the kinematic models of the *Almagest* and the visual rays of the *Optics*, the models of the *Harmonics* are not geometrical but arithmetical. The model for any interval between musical pitches is a ratio of whole numbers, the question at issue being what rules determine the whole-number ratios that correspond to the intervals of existing musical scales. Ptolemy credits the ratio model to the Pythagoreans, though he disagrees with what he sees as their tendency to develop *a priori* modelling principles that are not referred to empirical evidence in an appropriate manner. In the course of criticizing the Pythagoreans (and the more fundamentally wrong-headed Aristoxeneans) and evolving his own models, Ptolemy makes more explicit pronouncements about the interplay between *a priori* and empirical reasoning in science than in any of his other works.

Ptolemy's harmonic models are built up from three kinds of esthetically satisfying intervals: (a) *homophones*, i.e., intervals between notes that sound nearly alike, being identical in pitch or separated by one or more octaves, modelled by ratios always of the type $m : 1$, e.g., $1 : 1$ or $2 : 1$ or $4 : 1$; (b) *concords*, i.e., intervals between notes that sound different but akin, and that form the more stable larger intervals in scales, modelled by ratios of the type $m : n$ such that m is often but not always equal to $n + 1$, e.g., $3 : 2$ or $4 : 3$ or $8 : 3$; and (c) the smaller *melodic* intervals between consecutive notes of a scale, which are almost always modelled by ratios of the type $(n + 1) : n$, e.g., $9 : 8$.

The ratios are observable through the devices or instruments that make the notes. This is clearest in cases where the difference between notes follows from a difference between lengths in an instrument. For example, in wind

[13]The "fragments" of the lost part of Ptolemy's *Optics* (there are only four known) are collected in Lejeune 1989, 271. On Xenarchus, see Falcon 2001, 272, s.v. "Senarco."

[14]Barker 1989, 270-391 provides the best of the existing translations. West 1992 is a splendidly lucid introduction to all aspects of Greek music.

instruments one can measure the length of the pipe, say from the reed of an *aulos* (conventionally rendered by tin-eared classicists as "flute," but actually a double reed like an oboe or shawm) to one of the finger-holes. For his harmonic demonstrations, Ptolemy prescribes instruments involving tensed strings, since these allow the maximum control and precision in the tunings and measurements. Thus it is by dividing a tensed string with a bridge into two parts in the ratio 4 : 3 that Ptolemy establishes the association of this ratio with the *tetrachord*, the principal fixed interval in the Greek scales (in modern terminology, a "fourth").

But Ptolemy knows that length is not the only factor contributing to pitch. Thickness and density, among other characteristics of the bodies that produce the notes, are other variables that determine pitch; for this reason, before allowing us to try out ratios on a tensed string, Ptolemy instructs us to conduct a careful check of each part of the string to ensure that equal short lengths sound equal notes. Hence it is not at all easy to give a physical interpretation to the numbers in the modelling ratios that fully *explains* the musical intervals. Somehow a multiplicity of quantitative properties of a sounding body, some of them more straightforwardly measurable than others, give rise to a single abstract magnitude in the air in which the sound subsists.

In the chapters where he discusses the nature of sound and musical tone (*Harmonics* 1.3-4), Ptolemy does not try to explain the nature of sound more deeply than his initial definition that it is "a modification (*pathos*) of air when it is struck" (*Harmonics* 1.1), except for the conclusion that differences in pitch ("sharpness" and "heaviness") are a form of quantity. He does, however, restrict the scope of harmonic science to the study of sequences of discrete sounds, each of which has a constant pitch, so that one may speak of stable relations or "ratios" between the notes. The special status of *whole-number* ratios enters the discussion circuitously, by way of the review of the Pythagorean model, and although Ptolemy uses divisions of a tensed string to provide empirical justification that the homophones and concords are modelled by ratios of small whole numbers, he provides no *a priori* justification of this fact.

But patience is rewarded. When Ptolemy has completed his set task of deducing a more or less complete set of models to describe the systems of tuning current in his time (*Harmonics* 3.2), he embarks on a new project of describing how harmonic theory illuminates our understanding of aspects of the cosmos that have no direct connection with sound, namely the behaviour of human beings and of the heavens. It turns out that harmonics is not really a science concerning sound at all. It is a science that discovers far deeper and more general truths about our world, exploiting one specific part of it that happens to be exceptionally well adapted to the interplay between sensory observation and rational deduction that, for Ptolemy, constitutes scientific method. The true subject of harmonics is *harmonia*, "the form of rational causation (i.e., causation arising from reason and intellect) that concerns good ratios of motions," and this is necessarily present in all things that can move themselves, and above all in the most rational self-movers, namely, people and celestial spheres.

What this means is that the special status of whole-number ratios is a manifestation of the Good (in the Platonic sense) that the intellect apprehends and puts into action. One way that our intellects do this is by constructing musical instruments to produce sounds that fit the ideal ratios (since, after all, the sounds spontaneously produced by natural objects would not be recognized as music). Because of the close correspondence between measurable quantities in the instruments and the notes that we hear (which we can compare but not measure), we can discover the laws governing the order that our souls impose on this external matter. But these same laws are also recognizable, Ptolemy maintains, in the arrangement, motions, and powers of the heavenly bodies, which we discover through astronomy and astrology, and they must exist in our own characters, virtues, and emotions, where the quantitative relations are not apparent to our senses.[15]

These closing chapters of the *Harmonics* have received faint praise from modern readers, and it is undoubtedly true that the identification of detailed correspondences between the elements of his theory of musical tunings and an assortment of ethical, astronomical and astrological concepts is not Ptolemy's *forte*. But there can be no doubt that the principle motivating this *péché de jeunesse* was close to Ptolemy's heart, the conviction that the mathematical behaviour that we find here and there in the cosmos is structure imposed for the sake of the Good by minds upon a world that would otherwise be governed by disorder.

Bibliography

Barker, A. 1989. *Greek Musical Writings.* Vol. 2. *Harmonic and Acoustic Theory.* Cambridge.

Berggren, J. L., and A. Jones. 2000. *Ptolemy's* Geography: *An Annotated Translation of the Theoretical Chapters.* Princeton.

Dupuis, J. 1892. *Théon de Smyrne philosophe platonicien: Exposition des connaissances mathématiques utiles pour la lecture de Platon.* Paris.

Falcon, A. 2001. *Corpi e movimenti. Il* De caelo *di Aristotele e la sua fortuna nel mondo antico.* Napoli.

Goldstein, B. R. 1967. *The Arabic Version of Ptolemy's* Planetary Hypotheses. Transactions of the American Philosophical Society 57.4. Philadelphia.

Heiberg, J. L. 1907. *Claudii Ptolemaei Opera quae exstant omnia.* Vol. 2. *Opera astronomica minora.* Leipzig.

[15]See Swerdlow 2004 for a discussion of this part of the *Harmonics*, with emphasis on Ptolemy's celestial harmonics.

Jones, A. 1994. "Peripatetic and Euclidean Theories of the Visual Ray." *Physis* 31, 47-76.

Lejeune, A. 1989. *L'*Optique *de Claude Ptolémée dans la version latine d'après l'arabe de l'émir Eugène de Sicile.* 2nd edition. Collection de travaux de l'Académie internationale d'histoire des sciences 31. Leiden.

Lloyd, G. E. R. 1978. "Saving the Appearances." *Classical Quarterly* 28, 202–222. Reprinted with preface in G. E. R. Lloyd, *Methods and Problems in Greek Science*, Cambridge, 1991, 248-277.

Morelon, R. "La version arabe du *Livre des hypothèses* de Ptolémée." *Mélanges de l'Institut dominicain d'études orientales du Caire* 21 (1993) 7–85.

Murschel, A. 1995. "The Structure and Function of Ptolemy's *Physical Hypotheses of Planetary Motion.*" *Journal for the History of Astronomy* 26, 33–61.

Neugebauer, O. 1975. *A History of Ancient Mathematical Astronomy.* 3 vols. Berlin.

Robbins, F. E. 1940. *Ptolemy: Tetrabiblos.* Loeb Classical Library 435. Cambridge, Mass.

Smith, A. M. 1996. *Ptolemy's Theory of Visual Perception: An English Translation of the Optics with Introduction and Commentary.* Transactions of the American Philosophical Society 86.2. Philadelphia.

Swerdow, N. M. 2004. "Ptolemy's *Harmonics* and the 'Tones of the Universe' in the *Canobic Inscription.*" *Studies in the History of the Exact Sciences in Honour of David Pingree*, ed. C. Burnett, J. P. Hogendijk, K. Plofker, and M. Yano. Islamic Philosophy Theology and Science: Texts and Studies 54. Leiden. 137–180.

Toomer, G. J. 1984. *Ptolemy's Almagest.* London.

__________ . 1987. "Ptolemy." *Dictionary of Scientific Biography* 15, 207–224.

West, M. L. 1992. *Ancient Greek Music.* Oxford.

3

Mathematics, Instruments and Navigation, 1600-1800

Jim Bennett

Museum of the History of Science, University of Oxford

This article will consider four episodes in the history of navigation that are not part of the customary story. They will be set within a broad overview of developments in the period 1600 to 1800. Will this topic qualify as history of mathematics, as it must if it is to fall within the rubric of the Kenneth May Lecture? On our contemporary understanding of the mathematical discipline that may seem doubtful, and we may have to expand how we view mathematics and alter our assumptions about the identity of the mathematician, if we are to admit navigation in the seventeenth and eighteenth centuries. The expanded view we will need sits more comfortably in the sixteenth century, but the attitudes it contains are still sufficiently strong to illuminate the seventeenth and to have at least some relevance to the eighteenth. However, increasingly there were other visions of mathematics in play; there were challenges to the mathematical legitimacy of professional groups and a general shift away from the identity characteristic of the earlier period.

That earlier identity saw mathematics — mainly geometry — as engaged in a world of action and as central to success in that world. It was regarded as a vital tool for practising in a range of professional arenas. Navigation is our focus on this occasion, but the same was true of astronomy, surveying, architecture, warfare, engineering, and so on. The mathematician was one whose mode of life engaged with these worlds and whose skills and originality were valued there. While this was the arena of mathematical practice, mathematicians were not engaged with natural philosophy, as they would become later. They were not concerned with causal explanations of how the world operates or with what might be its material constitution. That mathematical culture — mathematics as action — was dominant in the sixteenth century, remained strongly present in the seventeenth, and was subordinate in the eighteenth. Taken as a whole, the four episodes of this article construct a narrative of decline.

What characterises the period around 1600 in the generally received history of navigation? Emphasis is placed on the effective and commonly adopted technique for finding latitude, based on altitude measurements of either the stars (notably the Pole Star) or the sun. This was sufficiently successful to alter navigational practice, at least in the open ocean, and to displace the bearing and distance technique (based on the magnetic compass, the log for measuring speed, the sand-glass for time, the traverse board for recording and the chart for plotting position). 'Bearing and distance' was replaced by latitude sailing, where a course would be set significantly to the east or west of the target destination and followed until the relevant latitude was achieved. This latitude would then be maintained while sailing west or east to landfall.[1]

As with other assumptions we will encounter, this one is only partly true, because without a corresponding technique for finding longitude and in the absence of permanently clear skies, the seaman still required dead reckoning based on a bearing and distance technique. Latitude sailing and dead reckoning were complementary and were used together, and this would be the case for a couple of centuries to come.

It may be that the apparent solution for latitude has thrown our attention too strongly on the complementary problem of longitude. We might be better advised to be guided by a contemporary opinion of the inadequacies in navigational practice, and to take Edward Wright's *Certain Errors in Navigation*, published in 1599, as a starting point.[2] The errors in question were:

- problems with the magnetic compass, particularly in relation to magnetic variation;
- problems with the design and use of the cross-staff, one of the instruments used for measuring the altitudes required in latitude sailing;
- the inadequacies of the plane chart; and
- the inadequacies of astronomical tables used for navigation; in particular, tables of declinations of stars and of the sun.

Wright does not nominate the longitude. At this stage there was no direct method for finding longitude; the hope of finding a predictable relationship between magnetic variation and longitude, based on a geometric 'theoric' (as it was called) after the manner of the astronomers, had failed. Wright was at the centre of the recognition of that failure through his promotion of Simon Stevin's *Haven-Finding Art* and of William Gilbert's *De Magnete*.[3] Both

[1] For general accounts, see E.G.R. Taylor, *The Haven-Finding Art: a History of Navigation from Odysseus to Captain Cook* (London, 1971); D.W. Waters, *The Art of Navigation in England in Elizabethan and Early Stuart Times* (London, 1978); J.B. Hewson, *A History of the Practice of Navigation* (Glasgow, 1983).

[2] E. Wright, *Certaine Errors in Navigation, Arising Either of the Ordinarie Erroneous Making or Vsing of the Sea Chart, Compasse, Crosse Staffe, and Tables of Declination of the Sunne, and Fixed Starres Detected and Corrected by E. W.* (London, 1599)

[3] S. Stevin, trans. E. Wright, *The Hayen-Finding Art, or, the Way to Find any Hauen or Place at Sea, by the Latitude and Variation* (London 1599); W. Gilbert, *De magnete, magneticisque corporibus, et de magno magnete tellure* (London, 1600).

works assume that the magnetic variation is totally dependent on local contingencies and is not subject to any global pattern or theoric. Wright does not cover longitude because he is dealing with the actual practice of navigation in the absence of a longitude method, and with the inadequacies of that practice.

In fact Wright's book reflects much of the agenda for the development of navigation and navigational theory in the seventeenth century. This is not to say that he was influential in setting that agenda, but rather that the set of problems he identified were in fact addressed in the seventeenth century, which reflects well on his ability to identify the central errors in the practice of his time. Each of my four episodes begins with one of the errors that concerned Edward Wright in 1599.

Magnetic Variation and Planetary Motion

Wright was at the centre of the English interest in a set of questions concerning both the possible application of variation to position finding at sea, and the need to manage variation so as to use the steering compass as accurately as possible. One outcome of this interest was the discovery of secular changes in variation, announced by Henry Gellibrand in 1635.[4]

Rather than ending any hope of position finding by variation, as might have been expected now that it was known that variation in a given location changes continuously, the discovery of this secular dimension to variation stimulated interest in finding some predictive account of the changes in variation — some more complex theoric that included a time variable. This may be less surprising than meets the eye. A wholly contingent variation, depending simply on the irregularities of the local terrain and to be discovered only by measurement and mapping, is a daunting (if not boring) prospect. However, if this pattern changes with time, there may be an underlying pattern that might be discovered and linked to some physical hypothesis. The prospect at least holds some interest.

The almanac publisher and teacher of navigation Henry Bond began his work on variation following Gellibrand's announcement, and from 1636 he published predictions that variation, then some degrees to the east, would reduce to zero in 1657 and then increase to the west.[5] He stated that his prediction was based on an account of the earth's magnetism that would yield a longitude method. The prediction turned out to be true, and Christopher Wren was sufficiently impressed that he stated in his inaugural address as Gresham Professor of Astronomy in 1657 that the study of variation may well

[4] H. Gellibrand, *A Discourse Mathematical on the Variation of the Magneticall Needle* (London, 1635).

[5] H. Bond, ed., *The Seaman's Kalender* (London, 1636).

yield a method for longitude 'than which, former Industry hath hardly left any Thing more glorious to be aim'd at in Art.'[6]

This much is familiar, but I want to take as my first noteworthy episode Wren's other interest in variation pursued at the same time. Using a very long magnetic needle, he was making a detailed and systematic study of changes in variation — not just the long-term movements, such as the gradual decrease to zero in 1657, but also the cycles of change that lay within these. Since he was aiming to detect an *annual* cycle within the overall pattern of change ('I hope to discover the Annual Motion of Variation & Anomalies in it'[7]), he must have thought he was dealing with a link between physical characteristics of the earth, or local phenomena that were amenable to experimental philosophy, and astronomy. This, of course, sounds very Newtonian, since it has the character of what has been called the 'Newtonian synthesis.' A fundamental causal agency, located in the earth and in celestial bodies, has both local consequences we can investigate on earth and celestial consequences we can measure in the heavens. These sets of consequences are essentially the same, but they manifest themselves at different distances.

Wren pointed out that Kepler himself acknowledged his debt to William Gilbert for the physical explanation Kepler gives of planetary motion based on an interaction between magnetic bodies. Wren said that this elliptical astronomy required 'Perfection,' i.e., further elaboration of both its geometry and its causal explanatory account, and he believed that this refinement would come from a study of the magnetic cycles of the earth, which would be reflected in its orbital motion and thus in astronomical measurements. Wren says that the study of the earth's magnetism is 'a Kind of Terrestrial Astronomy, an art that tells us the Motions of our own Star we dwell on.'[8]

Wren's programme of magnetic dynamics gave way to an account based on a different influence at a distance, but still involving an attractive force — now a single, central force, governed by a distance law and combined with a principle of rectilinear inertia. It is worth remembering that Wren was involved, along with Hooke and Halley, in formulating the principles of this programme for planetary dynamics.[9] What seems to have been overlooked in the standard history is his earlier projected synthesis. While it had a very similar scope and ambition, it depended on a different set of physical observations and speculations; it emerged from a sustained tradition of work on magnetism, which had been driven by its importance for the practice and development of navigation.

---

[6] C. Wren, Jnr, *Parentalia: or, Memoirs of the Family of the Wrens* (London, 1750), p. 206.

[7] J.A. Bennett, 'A Study of *Parentalia*, with Two Unpublished Letters of Sir Christopher Wren', *Annals of Science*, 30 (1973), 129-47, see p. 147.

[8] Wren, *op. cit.*, p. 206.

[9] J.A. Bennett, 'Hooke and Wren and the System of the World', *British Journal for the History of Science*, 8 (1975), 32-61.

From Plane to Mercator Sailing

Returning to Wright's *Errors* for the starting point of my second episode, we find him exercised by the continued use of the plane chart, i.e., one where there was no systematic accommodation of the curvature of the earth by means of a geometrical projection. In 1569 Gerard Mercator had published a chart based on a projection that was admirably suited for use at sea, especially for plotting compass bearings since they were projected as straight lines, but Mercator had not revealed the geometry on which this was based and, until that was available, there was no way that his invention could be generally used.

Wright explained the mathematics behind the Mercator chart in *Certain Errors*, and the following generation of English mathematicians, based at Gresham College in London, translated this mathematics into a practical technique for seamen. They did so by providing two things:

- protocols for each of the different types of calculations a navigator might face in plotting a position, and
- instruments that made these protocols achievable through routines of manipulation adapted to each problem.

The mathematicians involved in this work were Henry Briggs, one of Wright's associates, the first Professor of Geometry at Gresham College, and Edmund Gunter, the third Professor of Astronomy. Briggs contributed to the development of John Napier's logarithms, while Gunter was the foremost designer of mathematical instruments in the England of his day. The problem with sailing by the Mercator chart was that it inevitably involved some entanglement with trigonometry. A consequence of keeping the rhumb lines or compass bearings straight for the convenience of the navigator was that the scale of the chart was a function of the secant of the latitude.

This was beyond the regular seaman, but Gunter devised instruments that reduced operating with trigonometrical functions to a series of manipulations with instruments, and he explained both the instruments themselves and the protocols for their use in his book *De sectore et radio* of 1623.[10] The 'sector' of the title was a calculating instrument that had been proposed or modified by a number of mathematicians, including Galileo; Gunter's contribution was both to include trigonometrical functions on the lines of his sector, and to translate the calculations required for 'Mercator sailing' into routines that could be carried out with this instrument.

The basic principles of the sector relied on the proportional characteristics of similar triangles, but underlying the 'radius' — the second calculating instrument in Gunter's book of 1623 — was logarithms. The logarithms of trigonometrical functions were provided as lines on Gunter's radius, or 'rule' as it became commonly known, and were added or subtracted (that is, the

[10] E. Gunter, *De sectore et radio* (London, 1623).

trigonometrical functions were multiplied or divided) using a pair of dividers. It was not long before, instead of a single rule and dividers, some instruments had pairs of logarithmic scales sliding along each other. The 'sliding Gunter' of the seventeenth century is the origin of the ubiquitous logarithmic slide rule.

This critical development in the history of calculation arises from the contemporary problems of navigation and the established mathematical practice of rendering techniques accessible to navigators through the designing of appropriate instruments. No other set of circumstances in the period would have promoted such an outcome: only in navigation was there a pressing need for a complex and totally unfamiliar technique involving the management of large numbers to be rendered accessible to a large population of relatively unsophisticated practitioners. This episode is an important moment in the history of practical mathematics, but it is also significant for the history of navigation. The general development of routines of observation and calculation on board ship, the growing discipline in the management of these routines, and the gradual adoption and mastery of the Mercator chart are processes as important to progress in navigation as the more glamorous business of latitude finding and searching for longitude.

The 'radius' of Gunter's title *De sectore et radio* was so called because Gunter originally advised that the logarithmic scales should be marked on the radius of a cross-staff. This makes the link between the genesis of the logarithmic rule and contemporary problems in navigation even more emphatic.

Altitude Measurement by Backstaff and Octant

The cross-staff brings us back to Wright's *Certain Errors* and the beginning of my third episode. Wright was exercised by several problems with the use of the cross-staff: the user had to look in two directions simultaneously; if he was sighting the sun, he was looking at the sun; and — the point that Wright emphasises — the centre of the angle required is at the centre of the observer's eyeball, and this inaccessibility results in further inaccuracy.[11]

So far as solar sights were concerned, for finding latitude from the meridian altitude of the sun, an effective and popular solution came in the form of an alternative instrument, the backstaff (see figure 1). In the backstaff the quadrant is divided into two portions: usually a 65-degree arc, and one of 25 degrees drawn to a much larger radius. A near-sight, held to the eye, can be moved along the 25-degree arc, while the user views the horizon through this sight and a far-sight (called the horizon vane) at the common centre of the two arcs. A shadow vane can be set to any position on the 65-degree arc. In practice this is done so that the reading on the scale is some 10 or 15 degrees

[11] W.F.J. Mörzer Bruyns, *The Cross-Staff: History and Development of a Navigational Instrument* (Amsterdam, 1994).

Fig. 3.1. Backstaff by J. Gilbert, London, 1745. ©Museum of the History of Science, Oxford.

less than the expected altitude. The two scales diverge from a common zero, so that adding the two readings gives the angle subtended at the horizon vane by the shadow vane and the near-sight. In making his measurement, the navigator has his back to the sun and, keeping the horizon in view and the shadow of the shadow vane on the horizon vane, moves the near-sight down the 25-degree arc until the maximum angle is reached, and to continue he would have to begin to move the sight back. The sum of the two scale readings is the meridian altitude of the sun. To find the altitude of the equator, account must be taken of the solar declination, i.e. the angle of the sun above or below the equator, which varies by the time of year. For this the observer must have a table linking date and solar declination; such a table is sometimes found engraved on the instrument. The latitude is then the complement of the altitude of the equator. Wright's concerns have been met: the user looks in a single direction (that of the horizon) and looks away from the sun, while the angle being measured is external to his eye.

Backstaves are relatively common, which shows that they were much used, for generally speaking the survival rate for wooden instruments such as these is very low. They were robustly made, intended to cope with the rigours of seaborne life. Nonetheless, despite this unpretentious working context, if we look closely at individual examples we find that they have quite sophisticated features, that they are ingenious in their design, and that they have surprising, even impossible ambitions for accuracy.

The general design maximises the potential for precision by magnifying the measuring scale differentially without enlarging the whole instrument. The backstaff would be impossible to manage if an entire quadrant were made to the radius of the 25-degree arc. Since the 65-degree arc is used only for setting the shadow vane to a particular reading, it need only be divided to single degrees, but enlarging the arc where the measurement is taken allows it to carry a much more closely divided scale. In fact the scale used is imported from astronomical instruments — used, for example by Tycho Brahe and Johannes Hevelius — namely, the diagonal or transversal scale. The scale is commonly divided to degrees, then to 30 minutes, 10 minutes and 5 minutes. But diagonal lines between the 10-minute divisions and crossed by 10 equally-spaced arcs concentric with the main scale, carry the sub-division down to one minute — an unrealistic ambition, given the nature of the observation.

Two features demonstrate further refinements of the backstaff design. While the 65-degree scale is marked on the face of the arc, it is usual for the 5-degree divisions to be repeated on the rim. The rim divisions never quite coincide with the scale on the face, being consistently offset slightly. At first this may seem to be carelessness, but it is found on every instrument, so must be deliberate. It reflects the concern that the observer needs to register the position of the centre of the sun, but it is the limb of the sun that determines the extent of the shadow. The apparent discrepancy in the positions of the rim divisions is meant to give the user the option to correct for the semi-diameter of the sun. Of course there are other uncertainties at play, with penumbra effects and so on, so that this feature only indicates a concern over the problem; it is not a full solution.

The problem is completely solved by a second refinement: a convex lens can be fitted to the 65-degree arc as an alternative to the shadow vane. This focuses the light of the sun onto the horizon vane. It is often said that this was for use in hazy conditions, when sunlight was insufficiently strong to cast a clear shadow, but given the concern about recording the position of the sun's centre, the lens might have been preferable under a range of conditions. The vanes so rarely survive that it is difficult to know to what extent lenses were used.

Despite its unpretentious origins, emerging from the world of mathematical practitioners, instrument makers and navigators rather than university mathematicians, the backstaff has considerable sophistication and surprising ambition in its concern for precision. It was replaced eventually in the eighteenth century by the octant, also known as the 'Hadley quadrant' after one of its designers, the mathematician, optical experimenter and Fellow of the Royal Society John Hadley. His designs were published in the *Philosophical Transactions* after he had presented them to a meeting of the Royal Society in 1731, a very different context for the introduction of a navigational instrument. Hadley began, not by describing the instrument, but with what he calls a 'Principle in Catoptrics,' on which he develops a geometrical construction related to two successive reflections and elaborates five corollaries.[12] This was

[12] J. Hadley, 'The Description of a New Instrument for Taking Angles,' *Philosophical Transactions*, 37 (1731), 147-57.

a presentation tuned to his audience and to Hadley's standing in the Royal Society.

When Hadley moved to the practical realisation of his ideas, he offered two designs. One was much like a design later attributed to Newton, on the basis of a drawing found in Edmond Halley's papers. Hadley spends less time on his second instrument, but it was much closer to the arrangement that was commonly adopted. Sea trials were requested at the Royal Society, by James Bradley and Halley, and accordingly Hadley had Jonathan Sisson make a brass octant, which was tried at sea and the results reported to the Society.

This development took place in the context of the Royal Society. Beyond this institutional context, the instrument makers already had a successful and familiar design in the backstaff, and they did not receive the new instrument in a passive way, but assimilated to it some of the features of the backstaff. For example, they dispensed with Hadley's telescopic sight, substituting the pinhole sight with which they and their customers were familiar. They added the diagonal scale from the backstaff. In fact, the scales on early octants are very similar to the backstaff scale, which is detrimental to the putative accuracy of the observation. A scale on an octant that is physically similar to one on a backstaff must be numbered at twice the rate, because the index arm which carries the mirror moves through only half the angle moved by the reflected ray. Therefore on an octant, the standard backstaff scale will measure to only two minutes of arc.

In fact this is only an apparent loss, since the whole measurement with the octant is a more accurate procedure than with the backstaff. Nonetheless, these adaptations show that the design was not simply adopted by the makers — it was modified, through the incorporation of features of the existing instrumentation for measuring angles. In fact, the next development in the octant, namely, the adoption of the vernier instead of the diagonal scale, again taken from astronomical instruments, seems to have been an initiative of the makers; it is not present in Hadley's designs.

During this period — from, say, 1730 to 1750 — it was by no means clear that the octant would come to dominate altitude measurements at sea. There were a number of proposed designs, both for reflecting instruments and for developments in the backstaff. One feature common to the latter proposals was the focusing lens (known as a 'Flamsteed glass'); in the contest between the two classes of instrument, the possibility of a lens for focusing sunlight on a hazy day was thought to be an important competitive feature of the backstaff and of neo-backstaves. Another feature was that some of the neo-backstaves incorporated an artificial horizon, so that they could be used when the real horizon was obscured either by land or cloud. John Elton's quadrant, for example, was a developed form of backstaff, also made by Sisson, published in the same volume of the *Philosophical Transactions* as the one containing

Hadley's quadrant.[13] The inclusion of an index arm, which is horizontal when the measurement is taken, allowed Elton to add a longitudinal bubble level as an artificial horizon, and to attach a vernier, in order to dispense with the diagonal scale. Other neo-backstaves, such as the similar designs promoted by the makers George Adams and Benjamin Cole in 1748, included the focusing lens and the vernier. In 1733 Hadley himself described to the Royal Society what he called a 'Quadrant for taking a Meridional Altitude at Sea, when the Horizon is not visible.'[14] It is completely different from his octant, focusing the sun's image onto a target and having a bubble-level artificial horizon. So, even Hadley was part of the neo-backstaff discussion.

My third episode, then, shows again that the familiar account, in this case the evident superiority of Hadley's quadrant and its rapid and untroubled assimilation into the instrumentation of navigation, is inadequate. The reality was more complex. The makers and practitioners already had a stable, sophisticated and successful instrument in the backstaff, and aspects of this technology were incorporated into a fluid and active exchange between mathematicians and makers in this period of evolution in the instrumental techniques for finding latitude.

Astronomy and Longitude

The message from my first three episodes, that the past is more complex and interesting than we might have imagined and that this complexity and interest come to light when you look more closely, applies especially to my final topic, the longitude. In this case it is astonishing that one can imagine that the past can have as simple a story as has recently been told.[15] While it is true that a popular book must be less detailed and technical than a scholarly monograph, here it is the morality of the tale, rather than its technical content, that has been simplified to an implausible extent.

One of the more prominent and enduring seventeenth-century responses to the inadequacy of astronomical tables - the fourth source of error identified by Wright - was the foundation of the Royal Observatory at Greenwich, whose immediate aim was an astronomical solution to the problem of finding longitude at sea. The context that brought Wright's final *Error* to the notice of the political establishment in England takes us back to Henry Bond. In 1676 Bond had published a book with the striking and apparently authoritative title *The Longitude Found. Examined by Six Commissioners Appointed*

[13] J. Elton, 'A Descrition of a New Quadrant for Taking Altitudes Without an Horizon, either at Sea or Land,' *Philosophical Transactions*, 37 (1731), 273-9.

[14] J. Hadley, 'A Spirit Level to be Fixed to a Quadrant for Taking a Meridional Altitude at Sea, when the Horizon is not Visible,' *Philosophical Transactions*, 38 (1733), 167-72.

[15] D. Sobel, *Longitude: the True Story of a Lone Genius who Solved the Greatest Scientific Problem of his Time* (London, 1996).

by the King's Majesty.[16] The title was accurate in one respect — the method had been examined and recommended by a Royal Commission — but had a technique for finding longitude at sea really been discovered, as Bond had been asserting on his own behalf for many years?

Soon afterwards a book by Peter Blackborrow appeared with the title *The Longitude Not Found: or, an Answer to a Treatise by Henry Bond.*[17] Despite the ruling of the distinguished commissioners, who included the President of the Royal Society, a former Professor of Astronomy in Oxford now Bishop of Salisbury, and the current Professor of Geometry in Gresham College, at least one commentator — the unknown Blackborrow, who scarcely seemed qualified to judge — remained unconvinced. Hindsight seems to hand the verdict to him. Bond's scheme, as revealed to the Royal Commissioners, was in the general class of the longitude-by-variation solutions, though he brought measurements of magnetic inclination, or dip, into the equation. His theory involved magnetic poles rotating in the atmosphere, lagging behind the motion of the earth and, for this reason, moving in a circle displaced from the geographical poles.

What was in Bond's favour at the commission? Its members had some impressive evidence that the theory worked, namely, the prediction of zero variation in London in 1657. Hazardous prediction and subsequent confirmation are said to be the touchstone of the scientific method, but modern scientists might have difficulty subscribing to Bond's theory of rotating poles in the atmosphere, despite its empirical success. We know, in fact, that the Royal Commission itself harboured serious doubts. Robert Hooke confided to his diary that he, probably in collusion with other commissioners, 'Found it ignorant and groundless and fals but resolved to speak favourably of it.'[18] He, or they, may have adopted this less than candid policy as part of a more complex stratagem within the fractured contemporary arguments over the longitude, and it is here that Wright's concern about astronomical tables is relevant. While the Royal Commission was active, a proposal was made at court by a French associate of the King's mistress for a lunar method for finding longitude, and Charles referred the matter to substantially the same commissioners as were considering Bond's solution. They may have wanted to dispatch this foreign intrusion by supporting Bond's ineffectual theory. It was in response to this interest in the lunar method, and the absence of the necessary astronomical data, that Charles was moved to establish the observatory at Greenwich.

Was the otherwise obscure Blackborrow right? Did he see through an impossible theory? Had he detected a conspiracy of vested interest that had

[16] H. Bond, *The Longitude Found: or, a Treatise Shewing an Easie ... Way ... to Find the Longitude* (London, 1676).

[17] P. Blackborrow, *The Longitude Not Found: or an Answer to a Treatise by ... Henry Bond* (London, 1678).

[18] H.W. Robinson and W. Adams, *The Diary of Robert Hooke* (London, 1935), p. 97.

resulted in a favourable account from a commission whose members did not believe their own report? No. In fact, Blackborrow's main concern was to attack the Copernican theory, and Bond had assumed a moving earth. Once we begin to look beneath the initial appearance, a story of simple virtue is rarely sufficient.

In the eighteenth century we encounter an even thicker longitude plot, one that has attracted interest and spawned controversy and division from its inception. The spin that has been popular recently offers a single, linear tale with a consistent moral: the just and heroic struggle of a virtuous individual, John Harrison, against the calumnies of powerful vested interests. A humble but determined man, with right on his side, took on the prejudice of a self-serving establishment and, after a prolonged campaign based on integrity and justice, won through to recognition and reward.

In fact this story is not at all new: it was told at the time, as part of the campaign waged for the reward. In its final stages it gained the eloquent championship of Edmund Burke, when in Parliament in 1773, he berated the Government's response to Harrison's claims:

> Where, Sir, is the dignity, where is the sense, where even the justice of the representative of a great, powerful, enlightened, and maritime nation, when a petition of a man is laid before them, claiming not a favour, but justice; claiming that reward which law would give him, and to see it refused — upon what principle? Why, a man of 83 is to make new watches; and he is not only to make them, but to make new voyages to the Indies to try them. Good God, Sir, can this be a British House of Commons?[19]

What could the Prime Minister, Lord North, do but make excuses about delays and arrange for a direct subvention from Parliament, outside the provision of the Longitude Act?

Here again, just as in the case of Henry Bond a century earlier, opinion was strongly divided, even among those not directly involved. Toward the end of the century Harrison was famous as the winner of the fabulous reward and as an outstanding watchmaker, but it would have been difficult to have found any informed commentator who thought that, in any effectual sense, he had solved the longitude problem. At the height of the debate surrounding Harrison's watch, one writer in 1765 considered the very idea that this longitude solution had been proved by a voyage to the West Indies 'such an insult upon common sense as cannot be read without indignation.'[20] After this stage in Harrison's campaign had resulted in the award of £10,000, another writer concluded in 1770 that the longitude was 'still a secret, and likely to continue so, for tho many thousands of pounds have been paid for the pretended discovery thereof,

[19] *The Parliamentary History of England*, vol. xvii, 1773 (London, 1813), columns 841-3.

[20] *The Gentleman's Magazine*, 35 (1765), p. 34.

we remain just as wise as we were before the discovery, except the ill success of it happens to teach us so much wit as to take better care of our money for the future.'[21] Clearly very different judgements were possible.

We might wonder why it is that in our own cynical and sophisticated age, a story of virtue versus jealousy and greed, of ingenuity versus ignorance and prejudice, and of humility versus arrogance and disdain has struck a chord. Despite all the worldly wisdom we use to assess stories today, the analysis we expect from expert reporters, the weight we give to different explanations, and the account we take of individual motives, we are prepared to accept an idealistic story located safely in the past. Do we really regard our predecessors in the eighteenth century so differently from ourselves that their story is one of virtue and villainy? Can we really forget that historians are inclined to consider Georgian England as an 'age of jobbery'? We might then believe that the fabulous longitude prize of £20,000 could be won by getting the answer right, and by getting that right answer, so to speak, properly 'marked'. What we have seen here is that a closer look often reveals a more complex and qualified story, but one that is richer and more interesting.

[21] W. Emerson, *The Mathematical Principles of Geography, Navigation and Dialling* (London, 1770), p. 172. See J. Betts, 'Arnold and Earnshaw: the Practical Solution,' in W.J.H. Andrewes, ed., *The Quest of Longitude* (Cambridge, Massachusetts, 1996), pp. 312-28. For a general account, see J. Bennett, 'The Travels and Trials of Mr Harrison's Timekeeper', in M.-N. Bourguet, C. Licoppe and H.O. Sibum, eds, *Instruments, Travel and Science: Itineraries of Precision from the Seventeenth to the Twentieth Century* (London, 2002), pp. 75-95.

4

Was Newton's Calculus a Dead End? The Continental Influence of Maclaurin's *Treatise of Fluxions**

Judith V. Grabiner

Pitzer College

4.1 Introduction

Eighteenth-century Scotland was an internationally-recognized center of knowledge, "a modern Athens in the eyes of an enlightened world." [74, p. 40] [81] The importance of science, of the city of Edinburgh, and of the universities in the Scottish Enlightenment has often been recounted. Yet a key figure, Colin Maclaurin (1698–1746), has not been highly rated. It has become a commonplace not only that Maclaurin did little to advance the calculus, but that he did much to retard mathematics in Britain—although he had (fortunately) no influence on the Continent. Standard histories have viewed Maclaurin's major mathematical work, the two-volume *Treatise of Fluxions* of 1742, as an unread monument to ancient geometry and as a roadblock to progress in analysis. Nowadays, few people read the *Treatise of Fluxions*. Much of the literature on the history of the calculus in the eighteenth and nineteenth centuries implies that few people read it in 1742 either, and that it marked the end—the dead end—of the Newtonian tradition in calculus. [9, p. 235], [49, p. 429], [10, p. 187], [11, pp. 228–9], [43, pp. 246–7], [42, p. 78], [64, p. 144]

But can this all be true? Could nobody on the Continent have cared to read the major work of the leading mathematician in eighteenth-century Scotland? Or, if the work was read, could it truly have been "of little use for the researcher" [42, p. 78] and have had "no influence on the development of mathematics"? [64, p. 144]

We will show that Maclaurin's *Treatise of Fluxions* did develop important ideas and techniques and that it did influence the mainstream of mathematics. The Newtonian tradition in calculus did not come to an end in Maclaurin's Britain. Instead, Maclaurin's *Treatise* served to transmit Newtonian ideas in calculus, improved and expanded, to the Continent. We will look at what these ideas were, what Maclaurin did with them, and what happened to this work afterwards. Then, we will ask what by then should be an interesting

*First published in the *American Mathematical Monthly* **104** (5) (1997), 393–410.

question: why has Maclaurin's role been so consistently underrated? These questions will involve general matters of history and historical writing as well as the development of mathematics, and will illustrate the inseparability of the external and internal approaches in understanding the history of science.

4.2 The Standard Picture

Let us begin by reviewing the standard story about Maclaurin and his *Treatise of Fluxions*. The calculus was invented independently by Newton and Leibniz in the late seventeenth century. Newton and Leibniz developed general concepts—differential and integral for Leibniz, fluxion and fluent for Newton—and devised notation that made it easy to use these concepts. Also, they found and proved what we now call the Fundamental Theorem of Calculus, which related the two main concepts. Last but not least, they successfully applied their ideas and techniques to a wide range of important problems. [9, p. 299] It was not until the nineteenth century, however, that the basic concepts were given a rigorous foundation.

In 1734 George Berkeley, later Bishop of Cloyne, attacked the logical validity of the calculus as part of his general assault on Newtonianism. [12, p. 213] Berkeley's criticisms of the rigor of the calculus were witty, unkind, and—with respect to the mathematical practices he was criticizing—essentially correct. [6, v. 4, pp. 65–102] [38, pp. 33–34] [82, pp. 332–338] Maclaurin's *Treatise* was supposedly intended to refute Berkeley by showing that Newton's calculus was rigorous because it could be reduced to the methods of Greek geometry. [10, pp. 181–2, 187] [9, pp. 233, 235] Maclaurin himself said in this preface that he began the book to answer Berkeley's attack, [63, p. i] and also to rebut Berkeley's accusation that mathematicians were hostile to religion. [78, p. 50]

The majority of Maclaurin's treatise is contained in its first Book, which is called "The Elements of the Method of Fluxions, Demonstrated after the Manner of the Ancient Geometricians." That title certainly sounds as though it looks backward to the Greeks, not forward to modern analysis. And the text is full of words—lots of words. So much time is spent on preliminaries that it is not until page 162 that he can show that the fluxion of ay is a times the fluxion of y. Florian Cajori, whose writings have helped spread the standard story, compared Maclaurin to the German poet Klopstock who, Cajori said, was praised by all, read by none. [10, p. 1881] While British mathematicians, bogged down with geometric baggage, studied and revered the work and notation of Newton and argued with Berkeley over foundations, Continental mathematicians went onward and upward analytically with the calculus of Leibniz. The powerful analytic results and techniques in eighteenth-century Continental mathematics were all that mathematicians like Cauchy, Riemann, and Weierstrass needed for their nineteenth-century analysis with its even greater power, together with its improved rigor and generality. [9, ch. 7] [49, p. 948] This story became so well known that it was cited by the literary critic Matthew Arnold, who wrote, "The man of genius [Newton] was continued by... completely powerless and obscure followers.... The man of intelligence [Leibniz] was continued by successors like Bernoulli, Euler, Lagrange, and Laplace —the greatest names in modern mathematics." [1, p. 54; cited by [61, p. 151]

Now since I myself have contributed to the standard story, especially in delineating the links among Euler, Lagrange, and Cauchy. [38, chs. 3–61] I have a good deal of sympathy for it, but I now think that it must be modified. Maclaurin's *Treatise of Fluxions* is an important link between the calculus of Newton and Continental analysis, and Maclaurin contributed to key developments in the mathematics of his contemporaries. Let us examine the evidence for this statement.

4.3 The Nature of Maclaurin's *Treatise of Fluxions*

Why—the standard story notwithstanding—might Maclaurin's *Treatise of Fluxions* have been able to transmit Newtonian calculus, improved and expanded, to the Continent? First, because the *Treatise of Fluxions* is not just one "Book," but two. While Book I is largely, though not entirely, geometric, Book II has a different agenda. Its title is "On the *Computations* in the Method of Fluxions." [my italics] Maclaurin began Book II by championing the power of symbolic notation in mathematics. [63, pp. 575–576] He explained, as Leibniz before him and Lagrange after him would agree, that the usefulness of symbolic notation arises from its generality. So, Maclaurin continued, it is important to demonstrate the rules of fluxions once again, this time from a more algebraic point of view. Maclaurin's appreciation of the algorithmic power of algebraic and calculus notation expresses a common eighteenth-century theme, one developed further by Euler and Lagrange in their pursuit of pure analysis detached from any kind of geometric intuition. To be sure, Maclaurin, unlike Euler and Lagrange, did not wish to detach the calculus from geometry. Nonetheless, Maclaurin's second Book in fact, as well as in rhetoric, has an algorithmic character, and most of its results may be read independently of their geometric underpinnings, even if Maclaurin did not so intend. (In his Preface to Book I, he even urged readers to look at Book II before the harder parts of Book I.) [63, p iii] The *Treatise of Fluxions*, then, was not foreign to the Continental point of view, and may have been written in part with a Continental audience in mind.

Nor was this algebraic character a secret open only to the reader of English. There was a French translation in 1749 by the Jesuit R. P. Pézénas, including an extensive table of contents. [62] Lagrange, among others, seems to have used this French edition (since he cited it by the French title [58, p. 17] though he cited other English works in English [58, p. 18]). Pézénas' translation, moreover, was neither isolated nor idiosyncratic, but part of the activity of a network of Jesuits interested in mathematics and mathematical physics, especially work in English, with Maclaurin one of the authors of interest to them. [84, pp. 33, 221, 278, 517, 655] For instance, Pézénas himself translated other English works, including those by Desaguliers, Gardiner's logarithmic tables, and Seth Ward's *Young Mathematician's Guide* [83, pp. 571–2] Thus there was a well-worn path connecting English-language work with interested Continental readers. Furthermore, the two-fold character of the Treaties of Fluxions was noted, with special praise for Book II's treatment of series, by Silvestre-François Lacroix in the historical introduction to the second edition of his highly influential three-volume calculus textbook. [52, p. xxvii] Unfortunately, though, recognition of the two-fold character has been absent from the literature almost completely from Lacroix's time until the recent work by Sageng and Guicciardini. [42] [78] We shall address the reasons for this neglect in due course.

4.4 The Social Context: The Scottish Enlightenment

Another reason for doubting the standard picture comes from the social context of Maclaurin's career. Eighteenth-century Scotland, Maclaurin's home, was anything but an intellectual backwater. It was full of first-rate thinkers who energetically pursued science and philosophy and whose work was known and respected throughout Europe. One would expect Scotland's leading mathematician to share these connections and this international renown, and he did.

Although Scotland had been deprived of its independent national government by the Act of Union of 1707, it still retained, besides its independent legal system and its prevailing religion, its own educational system. The strength and energy of Scottish higher education in Maclaurin's time is owed in large part to the Scottish ruling classes, landowners and merchants alike, who saw science, mathematics, and philosophy as keys to what they called the "improvement" of their yet underdeveloped nation. [65, p. 254] [80, pp. 7–8, 10–11] [17, pp. 127, 132–3] Eighteenth-century Scotland, with one-tenth the population of England, had four major universities to England's two. [80, p. 116] Maclaurin, when he wrote the *Treatise of Fluxions*, was Professor of Mathematics at the University of Edinburgh. Edinburgh was about to become the heart of the Scottish Enlightenment, and Maclaurin until his death in 1746 was a leading figure in that city's cultural life.

Mathematics played a major role in the Scottish university curriculum. This was in part for engineers; Scottish military engineers were highly in demand even on the Continent. [17, p. 125] Maclaurin himself was actively interested in the applications of mathematics, and just before his untimely death had planned to write a book on the subject. [36] [68, p. xix] In addition, mathematics and Newtonian physics were part of the course of study for prospective clergyman. [80, p. 20] The influential "Moderate" party in the

Church of Scotland appreciated the Newtonian reconciliation of science and religion. [16, pp. 53, 57]

Maclaurin's position in Edinburgh's cultural life was not just that of a technically competent mathematician. For instance, he was part of the Rankenian society, which met at Ranken's Tavern in Edinburgh to discuss such things as the philosophy of Bishop Berkeley; the society introduced Berkeley's philosophy to the Scottish university curriculum. [24, p. 222] [17, p. 133] [65, p. 197] Maclaurin and his physician friend Alexander Monro were the founders and moving spirits of the Edinburgh Philosophical Society. [65, p. 198] With Newton's encouragement, Maclaurin had become the chief spokesman in Scotland for the new Newtonian physics. His posthumously published book, *An Account of Sir Isaac Newton's Philosophical Discoveries*, was based on material Maclaurin used in his classes at Edinburgh, and the book was of great interest to philosophers. [24, p. 137] That book became well known on the Continent. It was translated into French almost as soon as it appeared, by Louis-Anne Lavirotte in 1749, and the first part appeared in Italian in Venice in 1762.

Another branch of Scottish science, namely medicine, also had many links with the Continent and was highly regarded there. Medical students went back and forth between Scotland, Holland, and France. [17, p. 135] [80, p. 7]

The best-known figures of eighteenth-century Scotland had major interactions with, and influence upon, Continental science and philosophy. [39] [81] Let it suffice to mention the names of four: the philosopher David Hume, who was a student at Edinburgh in Maclaurin's time; the geologist James Hutton, who attended and admired Maclaurin's lectures; [34, pp. 577–8] and, a bit after Maclaurin's time but still subject to his influence on Scottish higher education, the chemist Joseph Black and the economic and political philosopher Adam Smith. Maclaurin himself had twice won prizes from the Académie des Sciences in Paris, once in 1724 for a memoir on percussion, and then in 1740 (dividing the prize with Daniel Bernoulli, P. Antoine Cavalleri, and Leonhard Euler) for a memoir on the tides. [79, p. 611] [39, pp. 400–401]

Scotland in the eighteenth century nurtured first-rate intellectual work on mathematics, philosophy, science, medicine, and engineering, and did it all as part of a general European culture. [39, p. 412] [81, passim] The *Treatise of Fluxions* was the major mathematical work of a Scottish mathematician of considerable reputation on the Continent, a major work philosophically attuned to the enormously influential Newtonian physics and the Continentally popular algebraic symbolism. Such a work would certainly be of interest to Continental thinkers. Social considerations may not suffice to determine mathematical ideas, but they certainly affect the mathematician's ability to make a living, to get research support, and to promote contact and communication with other mathematicians and scientists at home and abroad. And so it was with Maclaurin.

4.5 Maclaurin's Continental Reputation

An even better reason for not accepting the traditional view of Maclaurin is that his work demonstrably was read in the eighteenth century, and was read by the big names of Continental mathematics. He had a Continental acquaintance through travel and correspondence. Even before the *Treatise of Fluxions*, his reputation had been enhanced by his Académie prizes and by his books on geometry. He was thus a respected member of an international network of mathematicians with interests in a wide range of subjects, and the publication of the *Treatise of Fluxions* was eagerly anticipated on the Continent.

The *Treatise of Fluxions* of 1742 was Maclaurin's major work on analysis, incorporating and somewhat dwarfing what he had done earlier. It contains an exposition of the calculus, with old results explained and many new results introduced and proved. Maclaurin seems to have included almost everything he had done in analysis and its applications to Newtonian physics. In particular, the findings of his Paris prize paper on the tides were included and expanded. His other papers, the posthumous and relatively elementary *Algebra*, and his works on geometry as such—though highly regarded—do not concern us here, but his Continental reputation was enhanced by these as well.

Let us turn now to some specific evidence for the Continental reputation of Maclaurin's major work. In 1741, Euler wrote to Clairaut that, though he had not yet seen the Paris prize papers on the tides, "from Mr. Maclaurin I expect only excellent ideas." [47, p. 87] Euler added that he had heard from England (presumably from his correspondent James Stirling) that Maclaurin was bringing out a book on "differential calculus," and asked Clairaut to keep him posted about this. In turn, Clairaut asked Maclaurin later in 1741 about his plans for the book, [66, p. 348] which Clairaut wanted to see before publishing his own work on the shape of the earth. [47, p. 110] Euler did get the *Treatise of Fluxions*, and read enough of it quickly to praise it in a letter to Goldbach in 1743. [48, p. 179] Jean d'Alembert, in his *Traité de dynamique* of 1743, [22, sec. 37, n.] praised the rigor brought to calculus by the *Treatise of Fluxions*. D'Alembert's most recent biographer, Thomas Hankins, argues that Maclaurin's *Treatise*, appearing at this time, helped persuade d'Alembert that gravity could best be described as a continuous acceleration rather than a series of infinitesimal leaps. [44, p. 167] D'Alembert's general approach to the foundations of the calculus in terms of limits clearly was influenced by Newton's and Maclaurin's championing of limits over infinitesimals, in particular by Maclaurin's clear description of limits in one of the parts of his *Treatise of Fluxions* that explicitly responds to Berkeley's objections (and which incidentally may be the first explicit description of the tangent as the limit of secant lines; see Section 7). [44, p. 23] [63, pp. 422–3] Lagrange in his *Analytical Mechanics* [55, p. 243] said that Maclaurin, in the *Treatise of Fluxions*, was the first to treat Newton's laws of motion in the language of the calculus in a coordinate system fixed in space. Though C. Truesdell [80, pp. 250–3] has shown that Lagrange was wrong because Johann Bernoulli and Euler were ahead of Maclaurin on this, the fact that Lagrange believed this is one more piece of evidence for the Continental reputation of Maclaurin as mathematician and physicist.

4.6 Maclaurin's Mathematics and Its Importance

The previous points show that Maclaurin could have been influential, but not that he was. Five examples will reveal both the nature of Maclaurin's techniques and the scope of his influence: a special case of the Fundamental Theorem of Calculus; Maclaurin's treatment of maxima and minima for functions of one variable; the attraction of spheroids; what is now called the Euler-Maclaurin summation formula; and elliptic integrals.

a. Key Methods in the Calculus. Two methods were central to the study of real-variable calculus in the eighteenth and nineteenth centuries. One of these is studying real-valued functions by means of power-series representations. This tradition is normally thought first to flower with Euler; it is then most closely associated with Lagrange, and, later for complex variables, with Weierstrass. The second such method is that of basing the foundations of the calculus on the algebra of inequalities—what we now call delta-epsilon proof techniques—and using algebraic inequalities to prove the major results of the calculus; this tradition is most closely associated with the work of Cauchy in the 1820's. I have traced these traditions back to Lagrange and Euler in my work on the origins of Cauchy's calculus. [38, chs. 3–6] It is surprising, at least if one accepts the standard picture of the history of the calculus, that both of these methods—studying functions by power series, basing foundations on inequalities—were materially advanced by Maclaurin in the *Treatise of Fluxions*. It is especially striking that the importance of Maclaurin's work on series-work based, it is well to remember, on Newton's use of infinite series—was recognized and praised in 1810 by Lacroix, who also linked it with the series—based calculus of Lagrange. [52, p. xxxiii]

Maclaurin skillfully used algebraic inequalities in his proof of a special case of the Fundamental Theorem of Calculus. He showed, for a particular function, that if one takes the fluxion of the area under the curve whose equation is $y = f(x)$, one gets the function $f(x)$. In his proof, Maclaurin adapted the intuition underlying Newton's argument for this fact in *De Analysi* [69]—that the rate of change of the area under a curve is measured by the height of the curve—but Maclaurin's proof is more rigorous. Although Maclaurin's argument proceeds algebraically, the concepts involved resemble those of the Greek "method of exhaustion" (more precisely termed by Dijksterhuis "indirect passage to the limit"). [26, p. 130] A key step in this Greek work is first to assume that two equal areas or expressions for areas are unequal, and then to argue to a contradiction by using inequalities that hold among various rectilinear areas. Newton in the *Principia* had based proofs of new results about areas and curves on methods akin to those of the Greeks. Maclaurin carried this much further. It was Maclaurin's "conservative" allegiance to Archimedean *geometric* methods that led him to buttress the *kinematic* intuition of Newton's calculus with *algebraic* inequality proofs.

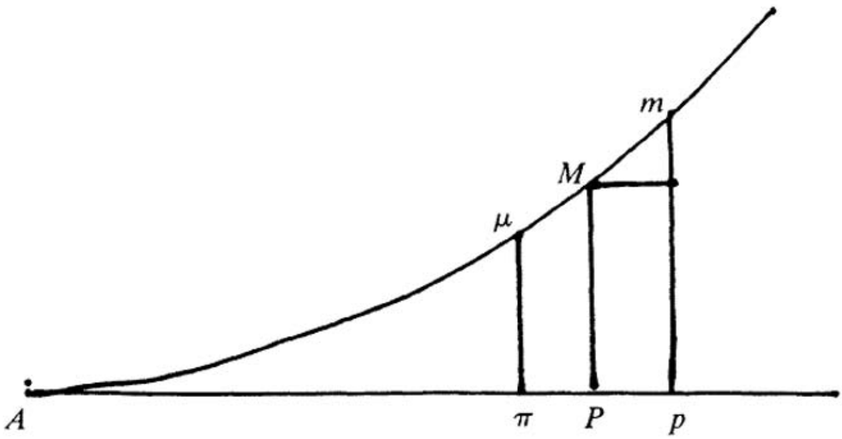


Fig. 4.1.

What Maclaurin proved in the example under discussion is that, if the area under a curve up to x is given by x^n, the ordinate of the curve must be $y = nx^{n-1}$, which is known to be the fluxion of f. [63, pp. 752–754] Maclaurin's diagram for this is much like the one Newton gave in the *De Analysi.* [69, pp. 3–4] Maclaurin began by saying that, since x and y increase together, the following inequality holds between the areas shown:

$$x^n - (x-h)^n < yh < (x+h)^n - x^n. \tag{4.1}$$

(Maclaurin gave this inequality verbally; I have supplied the "<" signs; also, I use "h" for the increment where Maclaurin used "o") Now Maclaurin recalled an algebraic identity he had proved earlier: [63, p. 583; inequality notation added]

$$\text{If } E < F, \text{then } nF^{n-1}(E-F) < E^n - F^n < nE^{n-1}(E-F). \tag{4.2}$$

(It may strike the modern reader that, since nx^{n-1} is the derivative of x^n, this second inequality is a special case of the mean-value theorem for derivatives. I shall return to this point later.)

Now, letting $x - h$ play the role of F and x play the role of E, $E - F$ is h and the first inequality in (2) yields

$$n(x-h)^{n-1}h < x^n - (x-h)^n.$$

Similarly, if $F = x$ and $E = x + h$, then $E - F = h$ and the second inequality in (2) becomes

$$(x+h)^n - x^n < n(x+h)^n h.$$

Combining these with inequality (1) about the areas, Maclaurin obtained

$$n(x-h)^{n-1}h < yh < n(x+h)^{n-1}h.$$

Dividing by h produces

$$n(x-h)^n < y < n(x+h)^{n-1}. \tag{4.3}$$

Recall that, given that the area was x^n, Maclaurin was seeking an expression for y, the fluxion of that area. A modern reader, having reached the inequality (3), might stop, perhaps saying "let h go to zero, so that y becomes nx^{n-1}," or perhaps justifying the conclusion by appealing to the delta-epsilon characterization of limit. What Maclaurin did instead was what Archimedes might have done, a double *reductio ad absurdum*. But what Archimedes might have done geometrically and verbally, Maclaurin did algebraically. He assumed first that y is *not* equal to nx^{n-1}. Then, he said, it must be equal to $nx^{n-1} + r$ for some r. First, he considered the case when this r was positive. This will lead to a contradiction if h is chosen so that $y = n(x+h)^{n-1}$, since, he observed, inequality (3) will be violated when $h = (x^{n-1} + r/n)^{1/(n-1)}$. Similarly, he calculated the h that produces a contradiction when r is assumed to be negative. Thus there can be no such r, and $y = nx^{n-1}$. [63, p. 753]

Maclaurin introduced this proof by saying something surprising for a *Treatise of Fluxions*: that the use of the inequalities makes the demonstration of the value of y "independent of the notion of a fluxion." [63, p. 752] (Of course one would need the notion of fluxion to interpret y as the fluxion of the area function x^n, but the proof itself is algebraic.) This proof was presumably part of his agenda in writing the more algebraic Book II of the *Treatise* for an audience on the Continent, where fluxions were suspect as involving the idea of motion. Later Lagrange, in seeking his purely algebraic foundation for the calculus, explicitly said he wanted to free the calculus from fluxions and what he called the "foreign idea" of motion. It is thus striking that Lagrange's *Théorie des fonctions analytiques* (1797) gives a more general version of the kind of argument Maclaurin had given, applying to any increasing function that satisfies the geometric inequality expressed in (1). In place of the algebraic inequality (2), Lagrange used the mean-value theorem. [58, pp. 238–9] [38, pp. 156–158] The similarity of the two arguments does not prove influence, of course, but it certainly demonstrates that Maclaurin's work, which we know Lagrange read (e.g., [58, p. 17]), uses the algebra of inequalities in a way consistent with that used by Lagrange and his successors.

Maclaurin's argument exemplifies the way his *Treatise* reconciles the old and the new. The double reductio ad absurdum reflects his Archimedean agenda. Treating the area as generated by a moving vertical line, and then searching for the relationship between the area and its fluxion, are Newtonian. Maclaurin did not have a general proof of the Fundamental Theorem in this argument, but relied on an inequality based on the specific properties of a specific function. Nonetheless, he had the precise bounding inequalities for the area function used later by Lagrange, and he used an algebraic inequality proof in a manner that would not disgrace a nineteenth-century analyst.

Inequality-based arguments in the calculus as used by Lagrange and Cauchy owe a lot to the eighteenth-century study of algebraic approximations, and it once seemed to me that this was their origin. But the algebra of inequalities as used in Continental analysis, especially in d'Alembert's pioneering treatment of the tangent as the limit of secants in the article "Différentiel" in the *Encyclopédie*, [19] must owe something also to Maclaurin's translation of Archimedean geometry into algebraic dress to justify results in calculus. Throughout the eighteenth century, practitioners of the limit tradition on the Continent use inequalities; a clear line of influence connects Maclaurin's admirer d'Alembert, Simon L'Huilier (who was a foreign member of the Royal Society), the textbook treatment of limits by Lacroix, and, finally, Cauchy. [38, pp. 80–87]

Now let us turn to some of Maclaurin's work on series. There is, of course, the Maclaurin series, that is, the Taylor series expanded around zero. This result Maclaurin himself credited to Taylor, and it was known earlier to Newton and Gregory. It was called the Maclaurin series by John F. W. Herschel, Charles Babbage, and George Peacock in 1816 [51, pp. 620–21] and by Cauchy in 1823. [14, p. 257] Since it was obvious that Maclaurin had not invented it, the attribution shows appreciation by these later mathematicians for the way Maclaurin used the series to study functions. A key application is Maclaurin's characterization of maxima, minima, and points of inflection of an infinitely differentiable function by means of its successive derivatives. When the first derivative at a point is zero, there is a maximum if the second derivative is negative there, a minimum if it is positive. If the second derivative is also zero, one looks at higher derivatives to tell whether the point is a maximum, minimum, or point of inflection. These results can be proved by looking at the Taylor series of the function near the point in question, and arguing on the basis of the inequalities expressed in the definition of maximum and minimum. For instance (in modern [Lagrangian] notation), if $f(x)$ is a maximum, then

$$\begin{aligned} f(x) &> f(x+h) = f(x) + hf'(x) + \tfrac{h^2}{2!}f''(x) + \cdots, \text{and} \\ f(x) &> f(x-h) = f(x) - hf'(x) + \tfrac{h^2}{2!}f''(x) - \cdots \end{aligned} \tag{4.4}$$

if h is small. If the derivatives are bounded, and if h is taken sufficiently small so that the term in h dominates the rest, the inequalities (4) can both hold only if $f'(x) = 0$. If $f'(x) = 0$, then the h^2 term dominates, and the inequalities (4) hold only if $f''(x)$ is negative. And so on.

I have traced Cauchy's use of this technique back to Lagrange, and from Lagrange back to Euler. [38, pp. 117–118] [37, pp. 157–159] [58, pp. 235–61] [29, Secs. 253–254] But this technique is explicitly worked out in Maclaurin's *Treatise of Fluxions*. Indeed, it appears twice: once in geometric dress in Book I, Chapter IX, and then more algebraically in Book II. [63, pp. 694–696] Euler, in the version he gave in his 1755 textbook, [20] does not refer to Maclaurin on this point, but then he makes few references in that book at all. Still we might suspect, especially knowing that Stirling told Euler in a letter of 16 April 1738 [91] that Maclaurin had some interesting results on series, that Euler would have been particularly interested in looking at Maclaurin's applications of the Taylor series. Certainly Lacroix's praise for Maclaurin's work on series must have taken this set of results into account. [52, p. xxvii] Even more important, Lagrange, in unpublished lectures on the calculus from Turin in the 1750's, after giving a very elementary treatment of maxima and minima, referred to volume II of Maclaurin's *Treatise of Fluxions* as the chief source for more information on the subject. [7, p. 154] Since Lagrange did not mention Euler in this connection at all, Lagrange could well have not even have seen the *Institutiones calculi differentialis* of 1755 when he made this reference. This Taylor-series approach to maxima and minima (with the Lagrange remainder supplied for the Taylor series) plays a major role in the work of Lagrange, and later in the work of Cauchy. It is because Maclaurin thought of maxima and minima, and of convexity and concavity, in Archimedean geometrical terms that he was led to look at the relevant inequalities, just as the geometry of Archimedes helped Maclaurin formulate some of the inequalities he used to prove his special case of the Fundamental Theorem of Calculus.

b. Ellipsoids. We now turn to work in applied mathematics that constitutes one of Maclaurin's great claims to fame: the gravitational attraction of ellipsoids and the related problem of the shape of the earth. Maclaurin is still often regarded as the creator of the subject of attraction of ellipsoids. [85, pp. 175, 374] In the eighteenth century, the topic attracted serious work from d'Alembert, A.-C. Clairaut, Euler, Laplace, Lagrange, Legendre, Poisson, and Gauss. In the twentieth century, Subramanyan Chandrasekhar (later Nobel laureate in physics) devoted an entire chapter of his classic *Ellipsoidal Figures of Equilibrium* to the study of Maclaurin spheroids (figures that arise when homogeneous bodies rotate with uniform angular velocity), the conditions of stability of these spheroids and their harmonic modes of oscillation, and their status as limiting cases of more general figures of equilibrium. Such spheroids are part of the modern study of classical dynamics in the work of scientists like Chandrasekhar, Laurence Rossner, Carl Rosenkilde, and Norman Lebovitz. [15, pp. 77–100] Already in 1740 Maclaurin had given a "rigorously exact, geometrical theory" of homogeneous ellipsoids subject to inverse-square gravitational forces, and had shown that an oblate spheroid is a possible figure of equilibrium under Newtonian mutual gravitation, a result with obvious relevance for the shape of the earth. [39, p. 172] [86, p. xix] [85, p. 374]

Of particular importance was Maclaurin's decisive influence on Clairaut. Maclaurin and Clairaut corresponded extensively, and Clairaut's seminal 1743 book *La Figure de la Terre* [18] frequently, explicitly, and substantively cites his debts to Maclaurin's work. [39, pp. 590–597] A key result, that the attractions of two confocal ellipsoids at a point external to both are proportional to their masses and are in the same direction, was attributed to Maclaurin by d'Alembert, an attribution repeated by Laplace, Lagrange, and Legendre, then by Gauss, who went back to Maclaurin's original paper, and finally by Lord Kelvin, who called it "Maclaurin's splendid theorem." [15, p. 38] [85, pp. 145, 409] Lagrange began his own memoir on the attraction of ellipsoids by praising Maclaurin's treatment in the prize paper of 1740 as a masterwork of geometry, comparing the beauty and ingenuity of Maclaurin's work with that of Archimedes, [57, p. 619] though Lagrange, typically, then treated the problem analytically. Maclaurin's eighteenth- and nineteenth-century successors also credit him with some of the key methods used in studying the equilibrium of fluids, such as the method of balancing columns. [39, p. 597] Maclaurin's work on the attraction of ellipsoids shows how his geometric insights fruitfully influenced a subject that later became an analytic one.

c. The Euler-Maclaurin Formula. The Euler-Maclaurin formula expresses the value of definite integrals by means of infinite series whose coefficients involve what are now called the Bernoulli numbers. The formula shows how to use integrals to find the partial sums of series. Maclaurin's version, in modern notation, is:

$$\sum_{h=0}^{\infty} F(a+h) = \int_0^a F(x)dx + \tfrac{1}{2}F(a) + \tfrac{1}{2}F'(a)$$

$$-\tfrac{1}{720}F'''(a) + \tfrac{1}{30240}F^{(}v)(a) - \cdots$$

$$[35, \mathrm{pp.}84 - -86]$$

James Stirling in 1738, congratulating Euler on his publication of that formula, told Euler that Maclaurin had already made it public in the first part of the *Treatise of Fluxions*, which was printed and circulating in Great Britain in 1737. [47, p. 88n] [91, p. 178] (On this early publication, see also [63, pp. iii, 691n]). P. L. Griffiths has argued that this simultaneous discovery rests on De Moivre's work on summing reciprocals, which also involves the so-called Bernoulli numbers. [40] [41, pp. 16–17] [25, p. 19] In any case, Euler and Maclaurin derived the Euler-Maclaurin formula in essentially the same way, from a similar geometric diagram and then by integrating various Taylor series and performing appropriate substitutions to find the coefficients. [31] [32] [33] Maclaurin's approach is no more Archimedean or geometric than Euler's; they are similar and independent. [63, pp. 289–293, 672–675] [35, pp. 84–93] [67] In subsequent work, Euler went on to extend and apply the formula further to many other series, especially in his *Introductio in analysin infinitorum* of 1748 and *Institutiones calculi differentialis* of 1755. [35, p. 127] But Maclaurin, like

Euler, had applied the formula to solve many problems. [63, pp. 676–693] For instance, Maclaurin used it to sum powers of arithmetic progressions and to derive Stirling's formula for factorials. He also derived what is now called the Newton-Cotes numerical integration formula, and obtained what is now called Simpson's rule as a special case. It is possible that his work helped stimulate Euler's later, fuller investigations of these important ideas.

In 1772, Lagrange generalized the Euler-Maclaurin formula, which he obtained as a consequence of his new calculus of operators. [53] [35, pp. 169, 261] In 1834, Jacobi provided the formula with its remainder term, [46, pp. 263, 265] in the same paper in which he first introduced what are now called the Bernoulli polynomials. Jacobi, who called the result simply the Maclaurin summation formula, cited it directly from the *Treatise of Fluxions*. [46, p. 263] Later, Karl Pearson used the formula as an important tool in his statistical work, especially in analyzing frequency curves. [72, pp. 217, 262]

The Euler-Maclaurin formula, then, is an important result in the mainstream of mathematics, with many applications, for which Maclaurin, both in the eighteenth century and later on, has rightly shared the credit.

d. Elliptic Integrals. Some integrals (Maclaurin used the Newtonian term "fluents") are algebraic functions, Maclaurin observed. Others are not, but some of these can be reduced to finding circular arcs, others to finding logarithms. By analogy, Maclaurin suggested, perhaps a large class of integrals could be studied by being reduced to finding the length of an elliptical or hyperbolic arc. [63, p. 652] By means of clever geometric transformations, Maclaurin was able to reduce the integral that represented the length of a hyperbolic arc to a "nice" form. Then, by algebraic manipulation, he could reduce some previously intractable integrals to that same form. His work was translated into analysis by d'Alembert and then generalized by Euler. [13, p. 846] [23] [27, p. 526] [28, p. 258] In 1764, Euler found a much more elegant, general, and analytic version of this approach, and worked out many more examples, but cited the work of Maclaurin and d'Alembert as the source of his investigation. A.-M. Legendre, the key figure in the eighteenth-century history of elliptic integrals, credited Euler with seeing that, by the aid of a good notation, arcs of ellipses and other transcendental curves could be as generally used in integration as circular and logarithmic arcs. [45, p. 139] Legendre was, of course, right that "elliptic integrals" encompass a wide range of examples; this was exactly Maclaurin's point. Thus, although his successors accomplished more, Maclaurin helped initiate a very important investigation and was the first to appreciate its generality. Maclaurin's geometric insight, applied to a problem in analysis, again brought him to a discovery.

4.7 Other Examples of Maclaurin's Mathematical Influence

The foregoing examples provide evidence of direct influence of the *Treatise of Fluxions* on Continental mathematics. There is much more. For instance, Lacroix, in his treatment of integrals by the method of partial fractions, called it "the method of Maclaurin, followed by Euler." [52, Vol. II, p. 10] [63, pp. 634–644] Of interest too is Maclaurin's clear understanding of the use of limits in founding the calculus, especially in the light of his likely influence on d'Alembert's treatment of the foundations of the calculus by means of limits in the *Encyclopédie*, which in turn influenced the subsequent use of limits by L'Huilier, Lacroix, and Cauchy, [38, chapter 3] (and on Lagrange's acceptance of the limit approach in his early work in the 1750's). [7] Although the largest part of Maclaurin's reply to Berkeley was the extensive proof of results in calculus using Greek methods, he was willing to explain important concepts using limits also. In particular, Maclaurin wrote, "As the tangent of an arch [arc] is the right line that limits the position of all the secants that can pass through the point of contact.. . though strictly speaking it be no secant; so a ratio may limit the variable ratios of the increments, though it cannot be said to be the ratio of any real increments." [63, p. 423] Maclaurin's statement answers Berkeley's chief objection—that the increment in a function's value is first treated as non-zero, then as zero, when one calculates the limit of the ratio of increments or finds the tangent to a curve. Maclaurin's statement is in the tradition of Newton's *Principia* (Book I, Scholium to Lemma XI), but is in a form much closer to the later work of d'Alembert on secants and tangents. [20] Maclaurin pointed out that most of the propositions of the calculus that he could prove by means of geometry "may be *briefly* demonstrated by this method [of limits]." [63, p. 87, my italics]

In addition, Maclaurin had considerable influence in Britain, on mathematicians like John Landen (whose work on series was praised by Lagrange), Robert Woodhouse (who sparked the new British interest in Continental work about 1800), and on Edward Waring and Thomas Simpson, whose names are attached to results well known today. [42] Going beyond the calculus, Maclaurin's purely geometric treatises were read and used by French geometers of the stature of Chasles and Poncelet. [90, p. 145] Thus, though Maclaurin may not have been the towering figure Euler was, he was clearly a significant and respected mathematician, and the *Treatise of Fluxions* was far more than an unread tome whose weight served solely to crush Bishop Berkeley.

4.8 Why a *Treatise of Fluxions*?

The *Treatise of Fluxions* was not really intended as a reply to Berkeley. Maclaurin could have refuted Berkeley with a pamphlet. It was not a student handbook either; this work is far from elementary. Nor was it merely written to glory in Greek geometry. Maclaurin wrote several works on geometry per se. But he was no antiquarian. Instead, the *Treatise of Fluxions* was the major outlet for Maclaurin's solution of significant research problems in the field we now call analysis. Geometry, as the examples I gave illustrate, was for Maclaurin a source of motivation, of insight, and of problem-solving power, as well as being his model of rigor.

For Maclaurin, rigor was not an end in itself, or a goal pursued for purely philosophical reasons. It was motivated by his research goals in analysis. For instance, Maclaurin developed his theory of maxima, minima, points of inflection, convexity and concavity, orders of contact, etc., because he wanted to study curves of all types, including those that cross over themselves, loop around and are tangent to themselves, and so on. He needed a sophisticated theory to characterize the special points of such curves. Again, in problems as different as studying the attraction of ellipsoids and evaluating integrals approximately, he needed to use infinite series and know how close he was to their sum. Thus, rigor, to Maclaurin, was not merely a tool to defend Newton's calculus against Berkeley—though it was that—nor just a response to the needs of a professor to present his students a finished subject—though it may have been that as well. In many examples, Maclaurin's rigor serves the needs of his research.

Moreover, the *Treatise of Fluxions* contains a wealth of applications of fluxions, from standard physical problems such as curves of quickest descent to mathematical problems like the summation of power series—in the context of which, incidentally, Maclaurin gave what may be the earliest clear definition of the sum of an infinite series: "There are progressions of fractions which may be continued at pleasure, and yet the sum of the terms be always less than a certain finite number. If the difference betwixt their sum and this number decrease in such a manner, that by continuing the progression it may become less than any fraction how small soever that can be assigned, this number is *the limit of the sum of the progression*, and is what is understood by the value of the progression when it is supposed to be continued indefinitely." [63, p. 289] Thus, though eighteenth-century Continental mathematicians did not care passionately about foundations, [38, pp. 18–24] they could still appreciate the *Treatise of Fluxions* because they could mine it for results and techniques.

4.9 Why the Traditional View?

If the reader is convinced by now that the traditional view is wrong, that Maclaurin's *Treatise* did not mark the end of the Newtonian tradition, and that not all of modern analysis stems solely from the work of Leibniz and his school, the question arises, how did that traditional view come to be, and why it has been so persistent?

Perhaps the traditional view could be explained as follows. Consider the approach to mathematics associated with Descartes: symbolic power, not debates over foundations; problem-solving power, not axioms or long proofs. The Cartesian approach to mathematics is clearly reflected in the work and in the rhetoric of Leibniz, Johann Bernoulli, Euler, Lagrange—especially in the historical prefaces to his influential works—and even Cauchy. These men, the giants of their time, are linked in a continuous chain of teachers, close colleagues, and students. Some topics, like partial differential equations and the calculus of variations, were developed mostly on the Continent. Moreover, the Newton-Leibniz controversy helped drive English and Continental mathematicians apart. Thus the Continental tradition can be viewed as self-contained, and the outsider sees no need for eighteenth-century Continental mathematicians to struggle through 750 pages of a *Treatise of Fluxions*, which is at best in the Newtonian notation and at worst in the language of Greek geometry. Lagrange's well-known boast that his *Analytical Mechanics* [55] had (and needed) no diagrams, thus opposing analysis to geometry at the latter's expense, reinforced these tendencies and enshrined them in historical discourse. But the explanation we have just given does not suffice to explain the strength, and persistence into the twentieth century, of the standard interpretation. The traditional view of Maclaurin's lack of importance has been reinforced by some other historiographical tendencies that deserve our critical attention.

The traditional picture of Maclaurin's *Treatise of Fluxions* radically separates his work on foundations, which it regards as geometric, sterile, and antiquarian, from his important individual results, which often are mentioned in histories of mathematics but are treated in isolation from the purpose of the *Treatise*, in isolation from one another, and in isolation from Maclaurin's overall approach to mathematics. Strangely, both externalist and internalist historians, each for different reasons, have reinforced this picture.

For instance, in the English-speaking world, viewing the *Treatise* as only about Maclaurin's foundation for the calculus, and thus as a dead end, has been perpetuated by the "decline of science in England" school of the history of eighteenth-century science, stemming from such early nineteenth-century figures as John Playfair, and, especially, Charles Babbage. [77] [2] [4] Babbage felt strongly about this because he was a founder of the Cambridge Analytical Society, which fought to introduce Continental analysis into Cambridge in the early nineteenth century. This group had an incentive to exaggerate the superiority of Continental mathematics and downgrade the British, as is exemplified by their oft-quoted remark that the principles of "pure d-ism" should replace what they called the "dot-age" of the University. [5, ch. 7] [10, p. 274] The pun, playing on the Leibnizian and Newtonian notation in calculus, may be found in [2, p. 26]. These views continued to be used in the attempt by Babbage and others to reform the Royal Society and to increase public support for British science.

It is both amusing and symptomatic of the misunderstanding of Maclaurin's influence that Lacroix's one-volume treatise on the calculus of 1802, [50] translated into English by the Cambridge Analytical Society with added notes on the method of series of Lagrange, [51] was treated by them, and has been considered since, as a purely "Continental" work. But Lacroix's short treatise was based on the concept of limit, which was Newtonian, elaborated by Maclaurin, adapted by d'Alembert and L'Huilier, and finally systematized by Lacroix. [38, pp. 81–86] Moreover, the translators' notes by Babbage, Herschel, and Peacock supplement the text by studying functions by their Taylor series, thus using the approach that Lacroix himself, in his multi-volume treatise of 1810, had attributed to Maclaurin. This is, of course, not to deny the overwhelming importance of the contributions of Euler and Lagrange, both to the mathematics taught by the Analytical Society and to that included by Lacroix in his 1802 book, nor to deny the Analytical Society's emphasis on a more abstract and formal concept of function. But all the same, Babbage, Herschel, and Peacock were teaching some of Maclaurin's ideas without realizing this.

In any case, the views expressed by Babbage and others have strongly influenced Cambridge-oriented writers like W. W. Rouse Ball, who said that the history of eighteenth-century English mathematics "leads nowhere." [5, p. 98] H. W. Turnbull, though he wrote sympathetically about Maclaurin's mathematics on one occasion, [88] blamed Maclaurin on another occasion for the decline: "When Maclaurin produced a great geometrical work on fluxions, the scale was so heavily loaded that it diverted England from Continental habits of thought. During the remainder of the century, British mathematics were relatively undistinguished." [89, p. 115]

Historians of Scottish thought, working from their central concerns, have also unintentionally contributed to the standard picture. George Elder Davie, arguing from social context to a judgment of Maclaurin's mathematics, held that the Scots, unlike the English, had an anti-specialist intellectual tradition, based in philosophy, and emphasizing "cultural and liberal values." Wishing to place Maclaurin in this context, Davie stressed what he called Maclaurin's "mathematical Hellenism," [24, p. 112] and was thus led to circumscribe the achievement of the *Treatise of Fluxions* as having based the calculus "on the Euclidean foundations provided by [Robert] Simson," [24, p. 111] who had made the study of the writings of the classical Greek geometers the "national norm" in Scotland. The "Maclaurin is a geometer" interpretation among Scottish historians has been further reinforced by a debate in 1838 over who would fill the Edinburgh chair in mathematics. Phillip Kelland, a candidate from Cambridge, was seen as the champion of Continental analysis, while the partisans of Duncan Gregory argued for a more geometrical approach. Wishing to enlist the entire Scottish geometric tradition on the side of Gregory, Sir William Hamilton wrote, "The great Scottish mathematicians, . . . even Maclaurin, were decidedly averse from the application of the mechanical procedures of algebra." [24, p. 155] Though Kelland eventually won the chair, the dispute helped spread the view that Maclaurin had been hostile to analysis. More recently, Richard Olson has characterized Scottish mathematics after Maclaurin as having been conditioned by Scottish common-sense philosophy to be geometric in the extreme. [70, pp. 4, 15] [71, p. 29] But in emphasizing Maclaurin's influence on this development, Olson, like Davie, has overstated the degree to which Maclaurin's approach was geometric.

By contrast, consider internalist historians. The treatment of Maclaurin's results as isolated reflects what Herbert Butterfield called the Whig approach to history, viewing the development of eighteenth-century mathematics as a linear progression toward what we value today, the collection of results and techniques which make up classical analysis. Thus, mathematicians writing about the history of this period, from Moritz Cantor in the nineteenth century to Hermann Goldstine and Morris Kline in the twentieth, tell us what Maclaurin did with specific results, some named after him, for which they have mined the *Treatise of Fluxions*. [13, pp. 655–63] [35, pp. 126ff, 167–8] [49, pp. 522–3, 452, 442] They either neglect the apparently fruitless work on foundations, or, viewing it as geometric, see it as a step backward. It is of course true that many Continental mathematicians used Maclaurin's results without accepting the geometrical and Newtonian insights that Maclaurin used to produce them. But without those points of view, Maclaurin would not have produced those results.

Both externalist and internalist historians, then, have treated Maclaurin's work in the same way: as a throwback to the Greeks, with a few good results that happen to be in there somewhat like currants in a scone. Further, the fact that Maclaurin's book, especially its first hundred pages, is very hard to read, especially for readers schooled in modern analysis, has encouraged historians who focus on foundations to read only the introductory parts. The fact that there is so much material has encouraged those interested in results to look only at the sections of interest to them. And the fact that the first volume is so overwhelmingly geometric serves to reinforce the traditional picture once again whenever anybody opens the *Treatise*. The recent Ph.D. dissertation by Erik Sageng [78] is the first example of a modern scholarly study of Maclaurin's *Treatise* in any depth. The standard picture has not yet been seriously challenged in print.

4.10 Some Final Reflections

Maclaurin's work had Continental influence, but with an important exception—his geometric foundation for the calculus. Mastering this is a major effort, and I know of no evidence that any eighteenth-century Continental mathematician actually did so. Lagrange perhaps came the closest. In the introduction to his *Théorie des fonctions analytiques*, Lagrange could say only, Maclaurin did a good job basing calculus on Greek geometry, so it can be done, but it is very hard. [58, p. 17] In an unpublished draft of this introduction, Lagrange said more pointedly: "I appeal to the evidence of all those with the courage to read the learned treatise of Maclaurin and with enough knowledge to understand it: have they, finally, had their doubts cleared up and their spirit satisfied?" [73, p. 30]

Something else may have blunted people's views of the mathematical quality of Maclaurin's *Treatise*. The way the book is constructed partly reflects the Scottish intellectual milieu. The Enlightenment in Britain, compared with that on the Continent, was marked less by violent contrast and breaks with the past than by a spirit of bridging and evolution. [75, pp. 7–8, 15] Similarly, Scottish reformers operated less by revolution than by the refurbishment of existing institutions. [16, p. 8] These trends are consistent with the two-fold character of the *Treatise of Fluxions*: a synthesis of the old and the new, of geometry and algebra, of foundations and of new results, a refurbishment of Newtonian fluxions to deal with more modern problems. This contrasts with the explicitly revolutionary philosophy of mathematics of Descartes and Leibniz, and thus with the spirit of the *mathématicien* of the eighteenth century on the Continent.

Of course Scotland was not unmarked by the conflicts of the century. During the Jacobite rebellion in 1745, Maclaurin took a major role in fortifying Edinburgh against the forces of Bonnie Prince Charlie. When the city was surrendered to the rebels, Maclaurin fled to York. Before his return, he became ill, and apparently never really recovered. He briefly resumed teaching, but died in 1746 at the relatively young age of forty-eight. Nonetheless, the Newtonian tradition in the calculus was not a dead end. Maclaurin in his lifetime, and his *Treatise of Fluxions* throughout the century, transmitted an expanded and improved Newtonian calculus to Continental analysts. And Maclaurin's geometric insight helped him advance analytic subjects.

We conclude with the words of an eighteenth-century Continental mathematician whose achievements owe much to Maclaurin's work. [39, pp. 172, 412–425, 590–597] The quotation [66, p. 350] illustrates Maclaurin's role in transmitting the Newtonian tradition to the Continent, the respect in which he was held, and the eighteenth-century social context essential to understanding the fate of his work. In 1741, Alexis-Claude Clairaut wrote to Colin Maclaurin, "If Edinburgh is, as you say, one of the farthest corners of the world, you are bringing it closer by the number of beautiful discoveries you have made."

Acknowledgement. I thank the Department of History and Philosophy of Science of the University of Leeds, England, for its hospitality while I was doing much of this research, and the Mathematics Department of the University of Edinburgh, where I finished it. I also thank Professor G. N. Cantor for material as well as intellectual assistance, and Professors J. R. R. Christie and M. J. S. Hodge for stimulating and valuable conversations.

References

1. Arnold, Matthew, The Literary Importance of Academies, in Matthew Arnold, *Essays in Criticism*, Macmillan, London, 1865, 42–79. Cited in [61].
2. Babbage, Charles, *Passages from the Life of a Philosopher*, Longman, London, 1864.
3. [Babbage, Charles], "Preface" to *Memoirs of the Analytical Society Cambridge*, J. Smith, 1813. Attributed to Babbage by Anthony Hyman, *Charles Babbage: Pioneer of the Computer*, Princeton University Press, Princeton, 1982.
4. Babbage, Charles, *Reflections on the Decline of Science in England, and Some of its Causes*, B. Fellowes, London, 1830.
5. Ball, W. W. Rouse, *A History of the Study of Mathematics at Cambridge*, Cambridge University Press, Cambridge, 1889.
6. Berkeley, George, *The Analyst, or a Discourse Addressed to an Infidel Mathematician*, in A. A. Luce and T. R. Jessop, eds., *The Works of George Berkeley*, vol. 4, T. Nelson, London, 1951, 65–102.
7. Borgato, Maria Teresa, and Luigi Pepe, Lagrange a Torino (1750–1759) e le sue lezioni inedite nelle R. Scuole di Artiglieria, *Bollettino di Storia delle Scienze Matematiche*, 1987, 7: 3–180.
8. Bourbaki, Nicolas, *Elements d'histoire des mathématiques*, Paris Hermann, Paris, 1960.
9. Boyer, Carl, *The History of the Calculus and Its Conceptual Development*, Dover, New York, 1959.
10. Cajori, Florian, *A History of the Conceptions of Limits and Fluxions in Great Britain from Newton to Woodhouse*, Open Court, Chicago and London, 1919.
11. Cajori, Florian, *A History of Mathematics*, 2d. ed., Macmillan, New York, 1922.
12. Cantor, G. N., Anti-Newton, in J. Fauvel et al., eds., *Let Newton Be!*, Oxford University Press, Oxford, 1988, pp. 203–222.
13. Cantor, Moritz, *Vorlesungen über Geschichte der Mathematik*, vol. 3, Teubner, Leipzig, 1898.
14. Cauchy, A..L., *Résumé des leçons données a l'école royale polytechnique sur le calcul infinitésimal*, in *Oeuvres completes*, Ser. 2, vol. 4, Gauthier-Villars, Paris, 1899.
15. Chandrasekhar, S., *Ellipsoidal Figures of Equilibrium*, Yale, New Haven, 1969.
16. Chitnis, Anand, *The Scottish Enlightenment: A Social History*, Croom Helm, London, 1976.
17. Christie, John R. R., The Origins and Development of the Scottish Scientific Community, 1680–1760, *History of Science* 12 (1974), 122–141.
18. Clairaut, A.-C., *Théorie de la figure de la terre*, Duraud, Paris, 1743.
19. d'Alembert, Jean, "Différentiel," in [21].
20. d'Alembert, Jean, and de la Chapelle, "Limite," in [21].
21. d'Alembert, Jean, et al., eds., *Dictionnaire encyclopédique des mathématiques*, Hotel de Thou, Paris, 1789, which collects the mathematical articles from the Diderot-d'Alembert *Encyclopédie*.
22. d'Alembert, Jean, *Traité de dynamique*, David l'Ainé, Paris, 1743.
23. d'Alembert, Jean, Récherches sur le calcul integral, *Histoire de l'Académie de Berlin* (1746), 182–224.
24. Davie, George Elder, *The Democratic Intellect: Scotland and her Universities in the Nineteenth Century*, The University Press, Edinburgh, 1966.

25. De Moivre, Abraham, *Miscellanea analytica*, Tonson and Watts, London, 1730.
26. Dijksterhuis, E. J., *Archimedes*, 2d. ed., Tr. C. Dikshoorn, Princeton University Press, Princeton, 1987.
27. Enneper, Alfred, *Elliptische Functionen: Theorie und Geschichte*, 2d. ed., Nebert, Halle, 1896.
28. Euler, Leonhard, De reductione formularum integralium ad rectificationem ellipsis ac hyperbolae, *Nov. Comm. Petrop.* 10 (1764), 30–50, in L. Euler, *Opera Omnia*, Teubner, Leipzig, Berlin, Zurich, 1911–, Ser. I, vol. 20, 256–301.
29. Euler, Leonhard, *Institutiones calculi differentialis*, 1755, sections 253–255; in *Opera*, Ser. I, Vol. XI.
30. Euler, Leonhard, *Introductio in analysin infinitorum*, Lausanne, 1748; in *Opera*, Ser. I, vols. 8–9.
31. Euler, Leonhard, Inventio summae cuiusque seriei ex data termino generali, *Comm. Petrop.* 8 (1741). 9–22; in *Opera*, Ser. I, vol. 14, 108–123.
32. Euler, Leonhard, Methodus generalis summandi progressiones, *Comentarii Acad. Imper. Petrop.* 6 (1738), 68–97; in *Opera*, Ser. II, vol. 22.
33. Euler, Leonhard, Methodus universalis serierum convergentium summas quam proxime inveniendi, *Comm. Petrop.* 8 (1741), 3–9; in *Opera*, Ser. I, vol. 14, 101–107.
34. Eyles, V. A., Hutton, *Dictionary of Scientific Biography*, Scribner's, New York, 1972, vol. 6, 577–589.
35. Goldstine, Herman, *A History of Numerical Analysis from the 16th through the 19th Century*, Springer-Verlag, New York, Heidelberg, Berlin, 1977.
36. Grabiner, Judith V., A Mathematician Among the Molasses Barrels: MacLaurin's Unpublished Memoir on Volumes, *Proceedings of the Edinburgh Mathematical Society* 39 (1996), 193–240.
37. Grabiner, Judith V., *The Calculus as Algebra: J.-L. Lagrange, 1736–1813*, Garland Publishing, Boston, 1990.
38. Grabiner, Judith V., *The Origins of Cauchy's Rigorous Calculus*, M.I.T. Press, Cambridge, Mass., 1981.
39. Greenberg, John L., *The Problem of the Earth's Shape from Newton to Clairaut*, Cambridge University Press, Cambridge, 1995.
40. Griffiths, P. L., Private communication.
41. Griffiths, P. L., The British Influence on Euler's Early Mathematical Discoveries, preprint.
42. Guicciardini, Niccolo, *The Development of Newtonian Calculus in Britain, 1700–1800*, Cambridge University Press, Cambridge, 1989.
43. Hall, A. Rupert, *Philosophers at War: The Quarrel between Newton and Leibniz*, Cambridge University Press, Cambridge, 1980.
44. Hankins, Thomas, *Jean d'Alembert: Science and the Enlightenment*, Clarendon Press, Oxford, 1970.
45. Itard, Jean, Legendre, *Dictionary of Scientific Biography*, vol. 8, 135–143.
46. Jacobi, C. G. J., De usu legitimo formulae summatoriae Maclaurinianae, *Journ. f. reine u. angew. Math.* 18 (1834), 263–272. Also in *Gesammelte Werke*, vol. 6, 1891, pp. 64–75.
47. Juskevic, A. P., and R. Taton, eds., *Leonhard Euleri Commercium Epistolicum*, Birkhauser, Basel, 1980; in Leonhard Euler, *Opera*, Ser. 4, vol. 5.
48. Juskevič, A. P., and Winter, E., eds. *Leonhard Euler und Christian Goldbach: Briefwechsel, 1729–1764*, Akademie-Verlag, Berlin, 1965.

49. Kline, Morris, *Mathematical Thought from Ancient to Modern Times*, Oxford, New York, 1972.
50. Lacroix, S. F., *Traité élémentaire de calcul différentiel et de calcul intégral*, Duprat, Paris, 1802. Translated as [51].
51. Lacroix, S.-F., *An Elementary Treatise on the Differential and Integral Calculus*, translated, with an Appendix and Notes, by C. Babbage, J. F. W. Herschel, and O. Peacock, J. Deighton and Sons, Cambridge, 1816.
52. Lacroix, S.-F., *Traité du calcul différentiel et du calcul integral*, 3 vols., 2d. ed., Courcier, Paris, 1810–1819. Vol. 1, 1810.
53. Lagrange, J.-L., Sur une nouvelle espece de calcul rélatif a la differentiation et a l'intégration des quantités variables, *Nouvelles Memoires de l'académie de Berlin*, 1772, 185–221; in *Oeuvres*, vol. 3, 439–476.
54. Lagrange, J.-L., *Leçons sur le calcul des fonctions*, new ed., Courcier, Paris, 1806; in *Oeuvres*, vol. 10.
55. Lagrange, J.-L., *Mechanique analytique*, 2d. ed., 2 vols., Courcier, Paris, 1811–1815; in *Oeuvres*, vols. 11–12.
56. Lagrange, J.-L., Note sur la métaphysique du calcul infinitesimal, *Miscellanea Taurinensia* 2 (1760–61), 17–18; in *Oeuvres*, vol. 7, 597–599.
57. Lagrange, J.-L., Sur l'attraction des sphéroides elliptiques, *Memoires de l'académie de Berlin*, 1773, 121–148. Reprinted in *Oeuvres de Lagrange*, vol. III, 619ff.
58. Lagrange, Joseph-Louis, *Théorie des fonctions analytiques*, Imprimérie de Ia République, Paris, An V [1797]; compare the second edition, Courcier, Paris, 1813, reprinted in *Oeuvres de Lagrange*, pub. M. J.-A. Serret, 14 volumes, Gauthier-Villars, Paris, 1867–1892, reprinted again, Georg Oms Verlag, Hildesheim and New York, 1973, vol. 9.
59. Legendre, Adrien-Marie, Mémoires sur les integrations par arcs d'ellipse et sur la comparaison de ces arcs, *Mémoires de l'Academie des sciences*, 1786, 616, 644–673.
60. Legendre, Adrien-Marie, *Traite des fonctions elliptiques et des intégrales eulériennes, avec des tables pour en faciliter le calcul numerique*, 3 vols., Paris, 1825–1828.
61. Loria, Gino, The Achievements of Great Britain in the Realm of Mathematics, *Mathematical Gazette* 8 (1915), 12–19.
62. Maclaurin, Colin, *Traité de fluxions*, Traduit de l'anglois par le R. P. Pézénas, 2 vols., Jombert, Paris, 1749.
63. Maclaurin, Colin, *A Treatise of Fluxions in Two Books*, Ruddimans, Edinburgh, 1742.
64. Mahoney, Michael, Review of [42], *Science* 250 (1990), 144.
65. McElroy, Davis, *Scotland's Age of Improvement*, Washington State University Press, Pullman, 1969.
66. Mills, Stella, *The Collected Letters of Colin Maclaurin*, Shiva Publishing, Nantwich, 1982.
67. Mills, Stella, The Independent Derivations by Leonhard Euler and Colin Maclaurin of the Euler-MacLaurin Summation Formula, *Archive for History of Exact Science* 33 (1985), 1–13.
68. Murdoch, Patrick, An Account of the Life and Writings of the Author, in Colin Maclaurin, *Account of Sir Isaac Newton's Philosophical Discoveries*, For the Author's Children, London, 1748, i-xx; reprinted, Johnson Reprint Corp., New York, 1968.

69. Newton, Isaac, *Of Analysis by Equations of an Infinite Number of Terms*, J. Stewart, London, 1745, in D. T. Whiteside, ed., *Mathematical Works of Isaac Newton*, vol. 1, Johnson Reprint, New York and London, 1964, 3–25.
70. Olson, Richard, *Scottish Philosophy and British Physics, 1750–1880*, Princeton University Press, Princeton, 1975.
71. Olson, Richard, Scottish Philosophy and Mathematics, 1750–1830, *Journal of the History of Ideas* 32 (1971), 29–44.
72. Pearson, Karl, *The History of Statistics in the Seventeenth and Eighteenth Centuries* [written 1921–1933]. Ed. E. S. Pearson, Charles Griffin & Co., London and High Wycombe, 1976.
73. Pepe, Luigi, Tre 'prime edizioni' ed un' introduzione inedita della Fonctions analytiqucs di Lagrange, *Boll. Stor. Sci. Mat.* 6 (1986), 17–44.
74. Phillipson, Nicholas, The Scottish Enlightenment, in [76], pp. 19–40.
75. Porter, Roy, The Enlightenment in England, in [76], pp. 1–18.
76. Porter, Roy, and Mikulas Teich, eds., *The Enlightenment in National Context*, Cambridge University Press, Cambridge, 1981.
77. Playfair, John, Traité de Mechanique Celeste, *Edinburgh Review* 22 (1808), 249–84.
78. Sageng, Erik Lars, *Colin Maclaurin and the Foundations of the Method of Fluxions*, unpublished Ph.D. Dissertation, Princeton University, 1989.
79. Scott, J. F., Maclaurin, *Dictionary of Scientific Biography*, vol. 8, 609–612.
80. Shapin, Stephen, and Arnold Thackray, Prosopography as a Research Tool in History of Science: The British Scientific Community, 1700–1900, *History of Science* 12 (1974), 95–121.
81. Stewart, M. A., ed., *Studies in the Philosophy of the Scottish Enlightenment*, Clarendon Press, Oxford, 1990.
82. Struik, D. J., *A Source Book in Mathematics, 1200–1800*, Harvard University Press, Cambridge, Mass., 1969.
83. Taton, Juliette, Pezenas, *Dictionary of Scientific Biography*, vol. 10, Scribner's, New York, 1974, 571–2.
84. Taton, René, ed., *Enseignement et diffusion des sciences en France au XVIII^e siècle*, Hermann, Paris, 1964.
85. Todhunter, Isaac, *A History of the Mathematical Theories of Attraction and the Figure of the Earth, from the Time of Newton to That of Laplace*, Macmillan, London, 1873.
86. Truesdell, C., Rational Fluid Mechanics, 1687–1765, introduction to Euler *Opera*, Ser. 2, vol. 12.
87. Truesdell, C., The Rational Mechanics of Flexible or Elastic Bodies, 1638–1788, in Euler, *Opera*, Ser. 2, vol. 11.
88. Turnbull, H. W., *Bicentenary of the Death of Colin Maclaurin (1698–1746)*, The University Press, Aberdeen, 1951.
89. Turnbull, H. W., *The Great Mathematicians*, Methuen, London, 1929.
90. Tweedie, Charles, A Study of the Life and Writings of Colin Maclaurin, *Mathematical Gazette* 8 (1915), 132–151.
91. Tweedie, Charles, *James Stirling*, Clarendon Press, Oxford, 1922.

5

The Mathematics and Science of Leonhard Euler (1707–1783)

Rüdiger Thiele

Karl-Sudhoff-Institut für Geschichte der Medizin und der Naturwissenschaften, University of Leipzig

> Laplace, qui pourtant n'avait pas pris, dans ses écrits, pour modèle le célèbre géomètre de Bâle, ne cessait de répéter aux jeunes mathématiciens ces paroles mémorables que nous avons entendues de sa propre bouche: *Lisez Euler, lisez Euler, c'est notre maître à tous.*[1]
> Guglielmo Libri-Carucci dalla Sommaja (1803-1869)

On 23 October 1783 a memorial session of the Imperial Academy of Sciences in St. Petersburg, Russia took place. Professor Nicolas (Nikolaus) Fuss (1755–1826), one of Euler's assistants during Euler's long period of blindness, delivered his famous Éloge de Monsieur Léonard Euler. He began with the following words:

> Représenter le cours de la vie d'un grand homme qui a illustré son Siècle en éclairant le monde, c'est faire l'éloge de l'esprit humain.[2]

At this time French was in common use not only in Paris, but as the lingua franca in Berlin and in St. Petersburg as well. One may summarize Fuss' words in this way: the task of giving a survey of Euler's life and work means giving a survey of the human intellect.

In 1783, in the year of Euler's death and 17 years before the turn of century, one of the most influential figures of the German Enlightenment, Georg Christoph Lichtenberg (1742–1799), devoted a paper to such a survey of the human intellect but to a smaller one. He proudly imagined how coming centuries might view his century. He set up an inventory of the scientists Isaac Newton (1642–1727), Gottfried Wilhelm Leibniz (1646–1716) and Leonhard

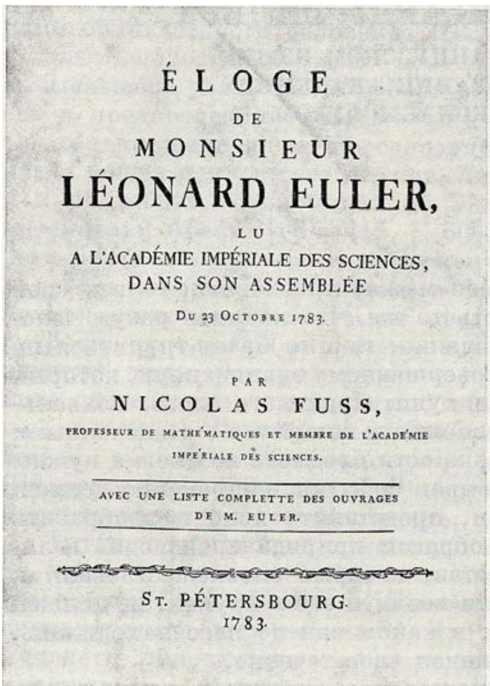
[1] Journal des Savants, 1846, p. 51: "Although Laplace in his writings did not regard that famous Basel mathematician as a model, he never ceased repeating these memorable words to young mathematicians which we [G. Libri-Carucci] heard from Laplace himself: Read Euler, read Euler, he is the master of us all."

[2] Fuss 1783, p. 1; German translation in: EO I, 1, p. XLIII.

Leonhard Euler
April 15, 1707 – September 9, 1783
Engraving after the plaster cast by J. Rachette (1781)
Courtesy Universitätsbibliothek Basel.

Euler (1707–1783) as well as the monarchs Peter the Great (1672–1725), Frederick the Great (1712–1786) and Catherine the Great (1729–1796). Concerning the year 1783 Lichtenberg mentioned a “huge new state” (referring to England’s acceptance of America’s Declaration of Independence in 1776) and

Fig. 5.1. Nicolas Fuss. Silhouette by F. Anting (1784). Due to a recommendation by Daniel Bernoulli, the Swiss Fuss became Euler's assistant in 1772, later professor of mathematics, in February 1783 member of the Petersburg Academy, and in 1784 Euler's grandson-in-law. Courtesy Sudhoff-Institut, Universität Leipzig.



ÉLOGE

DE

MONSIEUR

LÉONARD EULER,

LU

A L'ACADÉMIE IMPÉRIALE DES SCIENCES, DANS SON ASSEMBLÉE

DU 23 OCTOBRE 1783.

PAR

NICOLAS FUSS,

PROFESSEUR DE MATHÉMATIQUES ET MEMBRE DE L'ACADÉMIE IMPÉRIALE DES SCIENCES.

AVEC UNE LISTE COMPLETTE DES OUVRAGES DE M. EULER.

ST. PÉTERSBOURG.

1783.

Fig. 5.2. N. Fuss's eulogy on Euler delivered at a meeting of the Petersburg Academy in October 23, 1783. Courtesy Universitätsbibliothek Basel.

an air balloon (Montgolfière).[3] In the three memoirs of Catherine,[4] however, Euler, like any other scientist in Russia, is not mentioned, symbolic of the two cultures in the sense of Charles Percy Snow (1905–1980).[5] Snow argued that in the major branches of Western culture (above all science and the humanities) little or nothing is known about the other one, and that between the two camps communication is difficult if not impossible. Even the first edition of the Encyclopaedia Britannica took no notice of Euler in 1771.[6]

A century after Euler, the German historian of mathematics Hermann Hankel (1839–1873) properly remarked that Euler represented the scientific consciousness of the 18th century in the best way.[7] Recently, Clifford A. Truesdell (1919–2000) estimated that Euler wrote about one third of the mathematical works in the 18th century (including mathematical physics). But in general culture, the age of reason is regarded more as the age of John Locke (1632–1704) or Voltaire (1694–1678), rather than as the age of the natural sciences. A visit to any bookshop confirms this view. The author of a best-selling book on the Enlightenment[8] published in the promising Cambridge series "New approaches to European history" managed not to mention Euler in the entire book. However, it was the progress of the natural sciences that influenced more and more the rise of industry and, as a consequence, determined the conditions of human living. In this culture Euler has been appreciated; in his lifetime (in 1765) Euler was awarded with a prize (£300) by the British Parliament for his elaboration of two lunar theories, among them one with important consequences for navigation (1753, E 187); and in modern times (in 1935) the International Astronomical Union honoured him with a lunar crater, the diameter of which is 27 km wide, named after him (latitude 23.3 N and longitude 53.4 W).

---

[3] Lichtenberg, 1972, 3, 62–63. "Und was ich [the 18th century] gesehen habe? O genug. Ich habe Peter den Ersten gesehen und Katharina und Friedrich ... und Leibniz und Newton und Euler ... Bist du [reader] damit zufrieden? Gut. Aber sieh hier noch ein paar Kleinigkeiten: Hier habe ich einen neuen ungeheuren Staat [USA], ... und siehe endlich habe ich in meinem 83sten Jahr ein Luftschiff [Montgolfière] gemacht."

[4] Katharina II, 1986.

[5] Snow 1959. *The Two Cultures and the Scientific Revolution* (1959) and its sequel *The Two Cultures and the Scientific Revolution: A Second Look* (1964). The first heading is the title of C. P. Snow's Rede Lecture at Cambridge on the gap between science, and literature and religion.

[6] Edinburgh, vol. 2. 1771.

[7] Hankel 1982. "Euler, der das wissenschaftliche Bewußtsein in der Mitte des vorigen Jahrhunderts am vollständigsten vertritt, definiert [als Funktion . . .] ", p. 64. Reprint of a speech delivered in Tübingen in 1870.

[8] Outram 1996.

Youth

Precisely three towns determine the life of Leonhard Euler: Basel, a town of Roman origin; Berlin, founded in the 13th century; and St. Petersburg, founded in 1703.[9] He was born in Basel on 15 April 1707, some days before the Act of Union between England and Scotland constituted one Parliament in the island. Basel was an independent city of the empire and a centre of learning.

The Euler family is first mentioned in 1287 near Lindau, a town in southern Germany on Lake Constance. The German name Euler [oyler] sounds like Eule [oyle], the German word for owl, but the name refers to a wet meadow called an Au (diminutive: Äule [oyle]) in German, strictly speaking to a possessor of an Äule called an Äuler [oyler]. This Au may be found in the names of many smaller German towns, as in Nass*au* and Dess*au*.

Euler's father, Paul Euler (1670–1745), was a Protestant minister married to another minister's daughter, Margarete Brucker (1677–1761). The Euler family was rather poor. Most of Paul Euler's ancestors were comb-makers. In 1708 Paul Euler became pastor at the village of Riehen, two or three miles away from Basel, and his family soon moved there.

At Riehen[10] Euler grew up with his parents and later with two sisters in two rooms of the parsonage. Nevertheless, he was surrounded by educated people. Leonhard received his first instruction from his father. Paul Euler taught mathematics to his son using the widespread *Coss* (1525) by Christoff Rudolff (1500?–1549?) in Michael Stifel's (1487?–1567) edition of 1553 — a well-known book printed in many editions but difficult to read. Paul Euler himself was also a mathematician, having been a pupil of the famous James (Jakob) Bernoulli (1654–1705), and James's successor was his younger brother John (Johann) Bernoulli (1667–1748).

About 1713 — the time when Newfoundland was ceded to England — Leonhard Euler moved back to Basel to attend a grammar school, in which mathematics was not taught. Therefore Euler had private lessons from a Calvinist priest Johannes Burckhardt (1691–1743), who is known as a supporter of John Bernoulli in Bernoulli's clash with Brook Taylor (1685–1731). Burkhardt seems to have played an important role in Euler's mathematical education. Daniel Bernoulli (1700–1782), son of John Bernoulli, referred to Burckhardt as "magni Euleri praeceptor in mathematicis" (the teacher of the great Euler in mathematics).[11] Incidentally, the treatment of children in the 18th century is not generally regarded as having been lenient and kindly.

---

[9] In 1710 St. Petersburg had 8,000 inhabitants; in 1725 already 70,000; and in 1796 a quarter of a million.

[10] In the local Swiss dialect Rieche (pronounced ri:chϵ, ch as in the Scottish pronunciation of Loch).

[11] Fuss 1843, 2. Letter from Daniel Bernoulli to Euler (September 4, 1743), 529–537. "Vor etlichen Tagen ist der grosse Burcard [Burckhardt], Magni Euleri praeceptor in mathematicis, gestorben", p. 535.

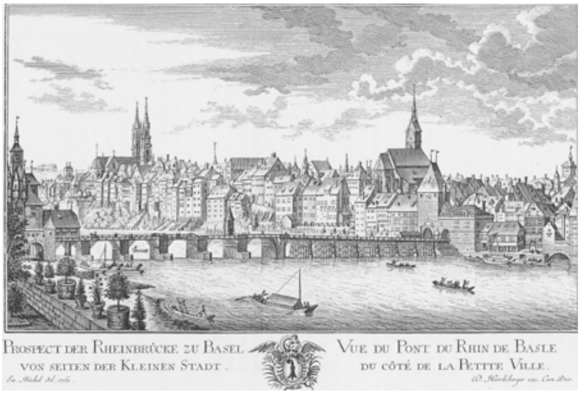


Fig. 5.3. View of the city of Basel. Etching by W. Herrliberger (1761) after a drawing by E. Büchel (1759). From the right: the end of the quarter Schifflände, the Church of St. Martin, the Old University (on the bank of the river Rhine), the Gothic Münster. Courtesy Universitätsbibliothek Basel.

Fig. 5.4. View of the Church of St. Martin in modern times, taken from the left side of the bridge shown in fig. 3. Photo R. Thiele.

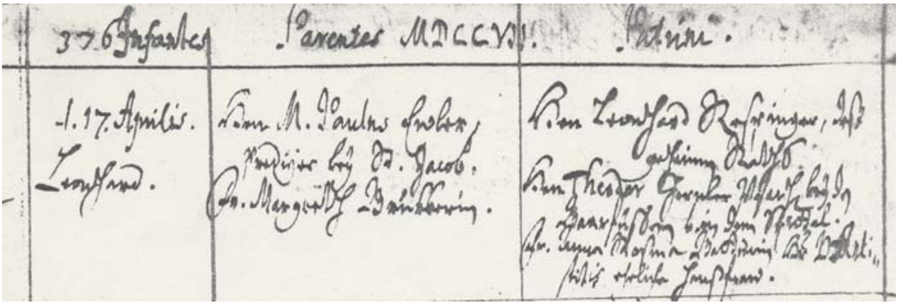


Fig. 5.5. Parish register of St. Martin in Basel for the year MDCCVII (1707). Entry of Leonhard Euler's christening ceremony on page 376, no. 1. Columns from left: date and infant to be baptized, parents, godfathers (including Leonhard Respinger). Courtesy Staatsarchiv Basel.

Pupils were frequently flogged by teachers, and occasionally an enraged father would appear in the classroom and respond in kind.

Study

In October 1720, at the not unusual age of 13, Leonhard entered the University of Basel. As a citizen of Basel he was allowed to do so whereas country folk could not enter the University before 1798. The city's university was founded in 1460 by Pope Pius II (1405-1464, pope since 1458) and was the first in Switzerland. The famous Dutch scholar Erasmus (1469?-1536) taught at this university in the 16th century, making the city a centre of humanism. In 1723 Euler and John II (Johann II), son of John Bernoulli, took their Master's degrees. Paul Euler wanted his son to follow him into church service; hence, Leonhard started to study theology.

The Calvinist clergy — even its most pious members — strongly advocated "law and order," mainly for political reasons. This created an ambiguous religious situation. Euler grew up in a spirit of submission to religious discipline. He upheld this spirit of true piety and religious discipline, and later on in Berlin and St. Petersburg he advocated parish affairs in this spirit to the very last. He remained a devout Calvinist all his life, and he conducted family prayers for his whole household, usually finishing off with a sermon.[12]

The University of Basel was very small in those days (19 professors and about 100 students); however, it was a mathematical centre of Europe under John I Bernoulli. At this small university it was inevitable that Euler and Bernoulli would meet. Indeed, soon Euler's mathematical abilities earned him the esteem of John I Bernoulli, who advised him to study mathematics. Finally Leonhard's father relented, and Leonhard left the Faculty of Theology and

[12] Cf. R. Thiele 2005a; M. Raith 1983, 459-470; F.G. Hartweg 1979.

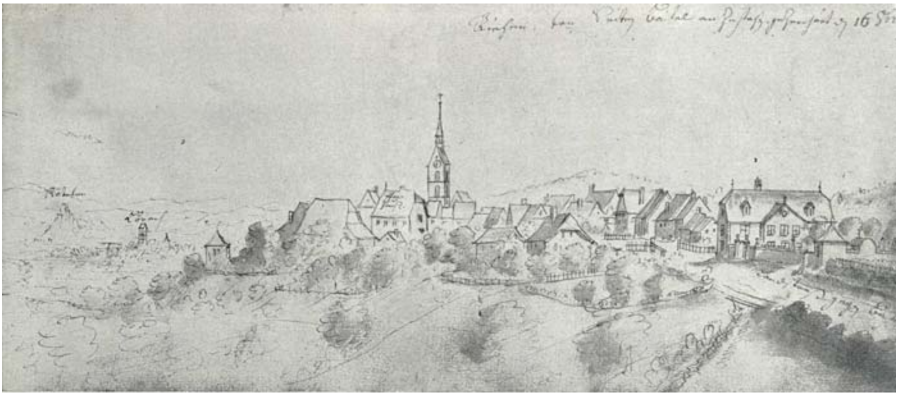


Fig. 5.6. View at Riehen, a village near Basel where Euler spent his childhood. The village church is St. Martin where Euler's father was a Protestant minister. Courtesy Universitätsbibliothek Basel.

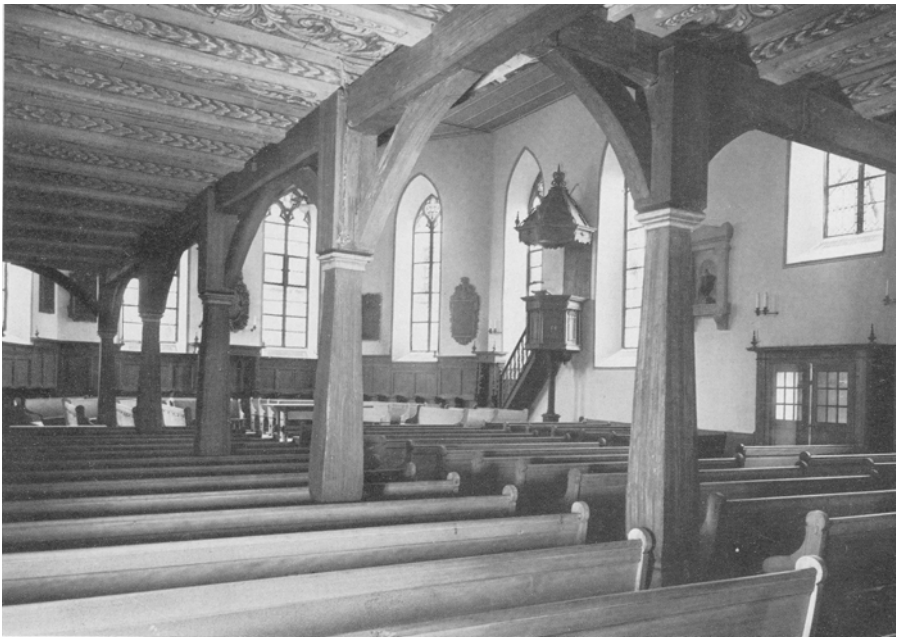
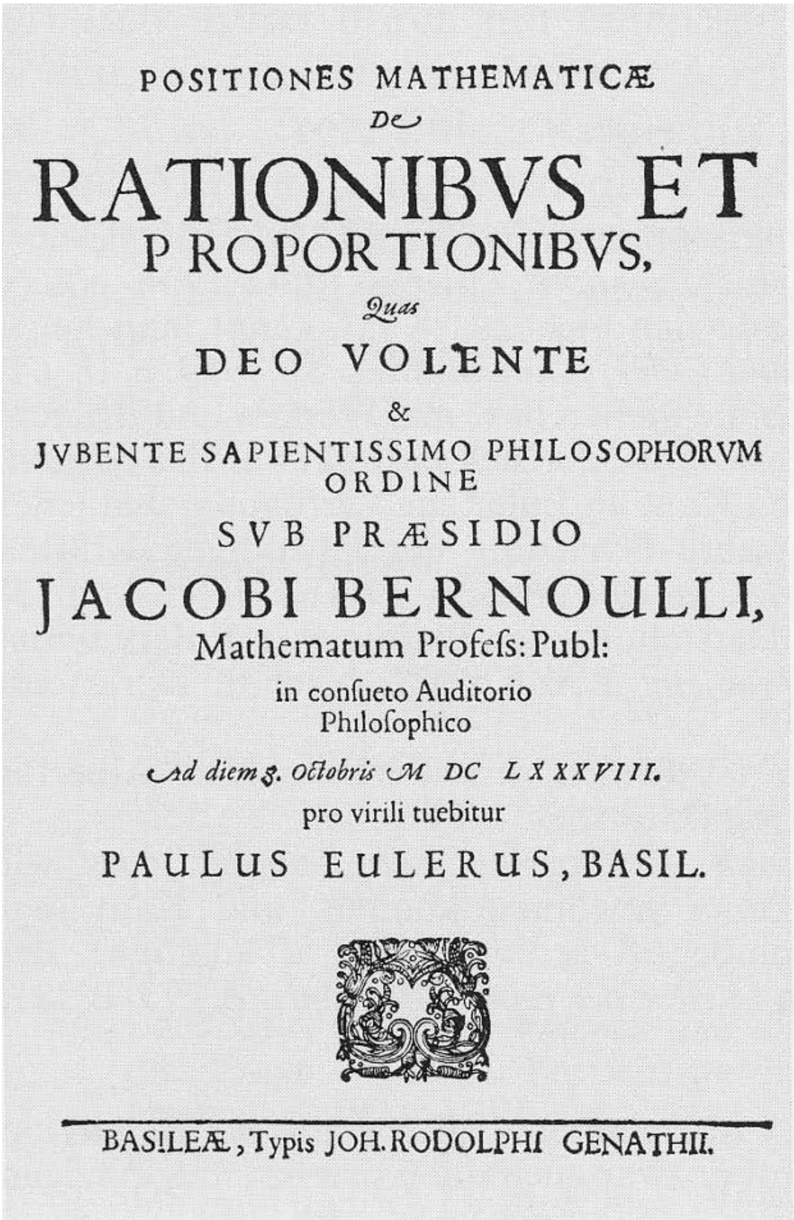


Fig. 5.7. St. Martin, indoor photograph. The sight is rather similar to that in Euler's day (extension of the church in 1694) despite the renovation in 1943. The oldest parts of St. Martin date to 950. Courtesy M. Raith, Riehen.



POSITIONES MATHEMATICÆ
De
RATIONIBVS ET
PROPORTIONIBVS,
Quas
DEO VOLENTE
&
JVBENTE SAPIENTISSIMO PHILOSOPHORVM
ORDINE
SVB PRÆSIDIO
JACOBI BERNOULLI,
Mathematum Profess: Publ:
in consueto Auditorio
Philosophico
Ad diem 8. Octobris M DC LXXXVIII.
pro virili tuebitur
PAULUS EULERUS, BASIL.

BASILEÆ, Typis JOH. RODOLPHI GENATHII.

Fig. 5.8. *De rationibus et proportionibus* (On ratios and proportions) by Paulus Euler. Under the supervision of James Bernoulli, Paulus Euler (Baccalaureus in 1687) defended 50 propositions on October 8, 1688. After that examination he started the study of theology, which was finished in 1693. It is remarkable that the disputatio was printed; however it is not sure or not whether P. Euler is its author (M. Raith). Courtesy Universitätsbibliothek Basel.

studied mathematics together with John II. He also became acquainted with Bernoulli's other sons Daniel (1700–1782) and Nicolas (1695–1726). The closer acquaintance with the three sons brought Euler into the house of Bernoulli. In his words:

> I soon found an opportunity to be introduced to a famous professor named John Bernoulli ... True, he was very busy and so refused outright to give me private lessons.[13]

However, Bernoulli supervised Euler's study by posing problems to Euler and by recommending mathematical reading. This was done on Saturday and Euler spent the rest of the week solving the problems and trying to trouble his teacher with as few questions as possible. Later on, in 1767, Leonhard Euler remembered:

> I was given permission to visit him freely every Saturday afternoon and he kindly explained to me everything I could not understand ... and this, undoubtedly, is the best method to succeed in mathematical subjects.[14]

In 1726 we first hear of Euler's own mathematical researches. In his paper "Constructio linearum isochronum ..." (Construction of isochrones in a resistant medium; E 1; EO II, 6),[15] Euler took up a special version of the brachistochrone problem posed by John Bernoulli in 1696.[16] Euler's paper was published in the journal *Acta eruditorum* in Leipzig in 1726. In the same year he wrote his paper "Meditationes super problemate nautico" (Essay on navigation; E 4; EO II, 20) dealing with the masting of ships — a prize problem of the Paris Academy of Sciences. Although Euler's paper failed to win the prize he received an honourable mention, an accessit. Later he won academy prizes twelve times, not to mention the eight prizes of his sons (substantially due to the father). Among Euler's prizes were five on navigation (an important topic at this time), and he earned 30,000 livres, a huge amount of money.

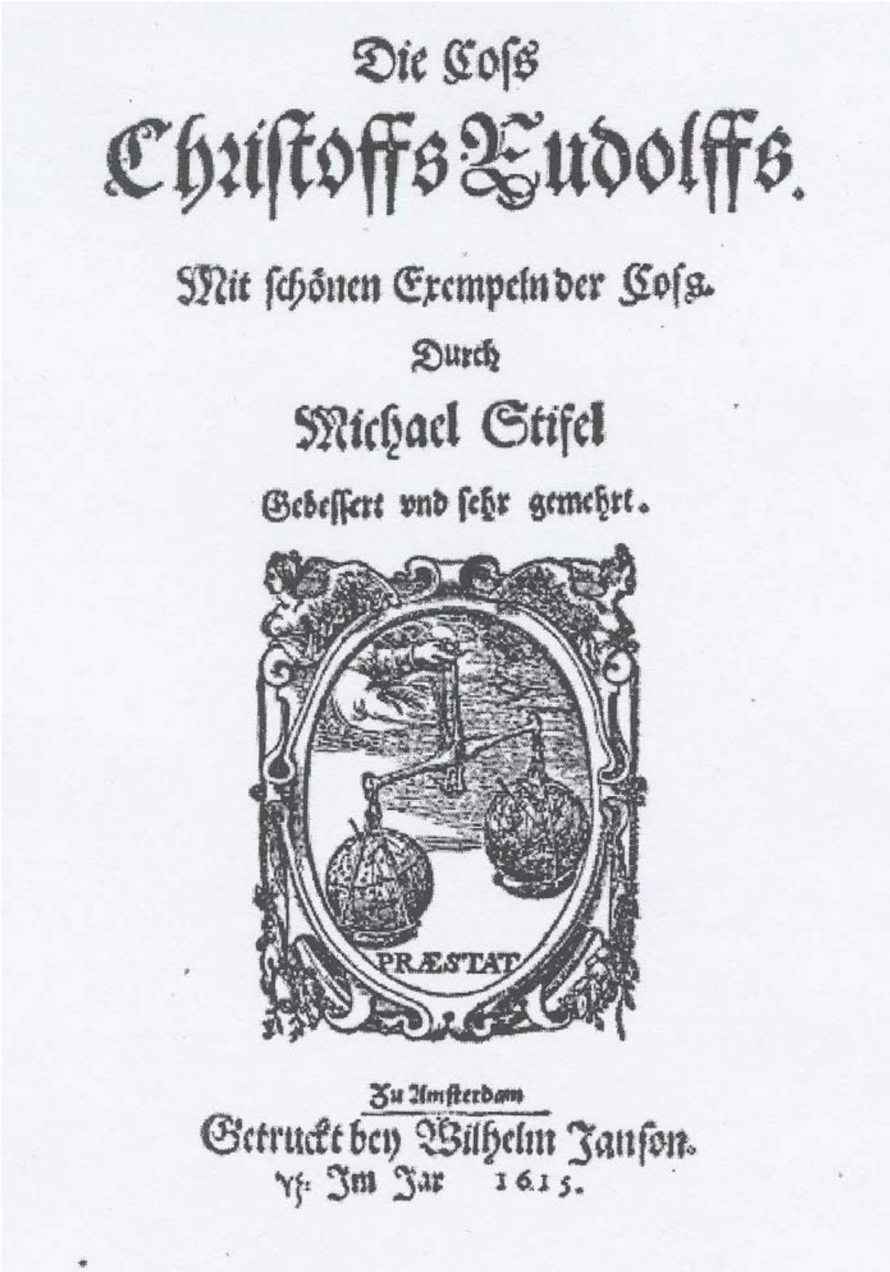
That Euler failed is not surprising when you recall the jokes about the nonexistent Swiss navy. Euler might have seen some boats on the river Rhine, but he had never seen a ship. But it comes as a surprise to see how Euler worked. He has been criticized for letting mathematics run away with his sense

[13]Euler 1767/1995, 11-13. " ... wo ich bald Gelegenheit fand mit dem berühmten Professori Johanni Bernoulli bekannt zu werden ... Privat Lectionen schlug er mir zwar wegen seiner Geschäfte gänzlich ab ... ".

[14]Euler 1767/1995, 11-13. " ... gab er mir alle Sonnabend Nachmittag einen freyen Zutritt bey sich, und hatte die Güte mir die gesammelte Schwierigkeiten zu erläutern, ... welches gewiß die beste Methode ist, um in den mathematischen Wissenschaften glückliche Progressen zu machen."

[15]E refers to the number of the Eneström register (Eneström 1910–1913); EO heads the series and the volume of the *Opera omnia Euleri* (Euler 1911–) in Roman and Hindu-Arabic figures respectively.

[16]Bernoulli 1696; see also Thiele 2002, and 2004.



Die Coss
Christoffs Rudolffs.
Mit schönen Exempeln der Coss.
Durch
Michael Stifel
Gebessert vnd sehr gemehrt.

Zu Amsterdam
Getruckt bey Wilhelm Janson.
Im Jar 1615.

Fig. 5.9. Title page of the *Coss* (Strasbourg 1525) by Ch. Rudolff in M. Stifel's edition (in a reprint of 1615, first edition in 1553). Courtesy Universitätsbibliothek Leipzig.

Fig. 5.10. John Bernoulli (left) and his son Daniel (right). Paintings by R. Huber (cut-out). Courtesy Universitätsbibliothek Basel.

of reality. The physical world was an occasion to apply mathematics, and if it failed to fit his analysis, it was the physical world, not the mathematics, which was in error. A well-known example is his investigation of a falling body given in his paper on sound ("Dissertatio physica de sono ..."; E 2; EO III, 1) in 1727. Euler considered a body falling through a tunnel from one pole of the earth to the other. His surprising and incorrect solution was that the body falls only as far as the centre of the earth and then returns. Leonhard Euler's 1726 prize paper on navigation ends with the following words:

> I did not consider it necessary to confirm my theory by any experiments. For this theory is derived from the most certain principles of mechanics. Hence there can be no doubt, whether it be true or have a place in praxi.[17]

[17]Final draft of the "Meditationes", Archive of the Russian Academy of Sciences in St. Petersburg, 136.1.213; quoted by G.K. Michailov in Sammelband 1959, p. 259. "Haud opus esse existimavi istam meam theoriam experiment confirmare, cum

However, in a draft of this paper found some years ago in the Archives in St. Petersburg we have another wording:

> I set to work to confirm my theory by experiments although this theory is deduced correctly from most certain and unquestionable principles, as far as I could see. Hence, if anybody wants to investigate the certainty of the principles or the correctness of the deduction of my theory, he will ipso facto [i.e., by his very investigation] see the truth of this theory confirmed. Consequently, no room will be left for doubt or hesitation.[18]

Concerning his theory of music Euler wrote to Daniel Bernoulli in 1740 that "experience will decide which theory is in better accordance with the truth."[19] In his late Petersburg years he still appreciated models of bridges made by the gifted Russian inventor Ivan Petrovich Kulibin (1735–1818) and went on to develop a mechanics of models (similarity mechanics).[20]

Euler, aware he was a born mathematician, applied for a professorship in Basel — and he failed. There were few opportunities for mathematicians in Switzerland. Therefore, in 1727 Daniel and Nicolas Bernoulli left Basel and became members of the Russian Academy. They promised Euler to find a position for him there. Euler received his call to Sankt Peterburg (Russian)[21] or Saint Petersburg (English) in 1727, officially as an associate of the medical section of the Academy. On April 5, 1727, a few days after Newton's death, he left Basel and he never returned to his home town, although he retained citizenship all his life. It is worth noting that Euler enrolled in the department of medicine at the University of Basel on April 2, exactly three days before his departure for Russia.

St. Petersburg

Euler started his way to St. Petersburg on the river Rhein until Mainz, continued by land until the seaport Lübeck, and finished by sea. On his journey

integra et ex certissimis et irrepugnabilibus principiis mechnicis deducta, atque adeo de illa dubitari, an vera sit ac an in praxi locum habere queat, minime possit."

[18] First draft of the "Meditationes", describing experiments; cf. footnote 17. "Sed istam meam theoriam, quamvis sit ex certissimis et de quibus neutiquam dubitar potest principiis recte quantum perspicere potui deducta, tamen eam experimentia quoque confirmare aggressus sum, quo, si quis de certitudine principiorum aut legitima deductione meae theoriae ambigere velit, veritatem ipso facto videre possit confirmatum iri, et ita, ulli dubitationi aut haesitationi nullus locus relinquatur."

[19] Eneström 1906. Letter to Daniel Bernoulli from September 15, 1740, 152. "Durch die Experientz kann man also leicht determiniren, welche Theorie mit der Wahrheit übereinkommt."

[20] Idey Eylera 1988, see N.M. Raskin, Eyler i Kulibin (in Russian), pp. 304-320.

[21] Later Petrograd (1914–1917), from 1917 Leningrad, now St. Petersburg.

Fig. 5.11. Europe in 1748. Based on a map of Sudhoff-Institut, Universität Leipzig.

Euler stopped in Marburg to meet the philosopher Christian Wolff (1679–1754). Wolff, who was a pupil of Gottfried Wilhelm Leibniz, is best known as the German spokesman of the Enlightenment. Wolff was banished from Halle in 1723 by the Prussian king Frederick William I (Friedrich Wilhelm I) (1688–1740) under penalty of hanging as a result of religious and philosophical disputes, and he fled to Marburg. In Russia, however, due to Leibniz's efforts he was made a science advisor (in absentia) to tsar Peter the Great (1672–1725, tsar from 1682, alone from 1696). He helped to establish the Russian Academy of Sciences which was founded after Peter's unexpected death in St. Petersburg in 1725 by Catherine (1684–1727, tsarina from 1725), the widow of Peter, in the same year. Wolff wrote to Euler that he would "travel into the paradise of the scholars."[22] Euler himself confirmed this statement in a letter written in Berlin in 1749:

> I and all others who had the good fortune to be some time with the Russian Imperial Academy cannot but acknowledge that we owe everything we are and possess to the favourable conditions we had there.[23]

At first, however, Euler's expectations were quickly dashed. The day he reached Russia Catherine had died and was succeeded by the boy tsar Peter II (1715–1730, tsar from 1727), a grandson of Peter the Great. Despite the regular succession, the question of succession was reopened, and Peter's reforms — Europeanization — were called in question. Peter's modernization was seen as a serious break with the past (which it was), and as a betrayal of Russia's culture. The actual power passed into the hands of a brutal orthodox faction, the Bironovshchina,[24] which looked upon the Academy as a dispensable luxury and tried to send all its foreign members home.

The Academy somehow managed to survive. In the confusion, Euler slipped into the mathematical section instead of the medical one. Political affairs slowly improved, and Euler was able to get down to solid work. He lectured to students, if there were any, and for pupils at the Gymnasium, and he gave about 10 lectures a year in the Academy. But scientific relations remained tense among the 17 members of the Academy (13 Germans, 3 Swiss, and one Frenchman).[25] The philosophical conflicts, that is to say, the controversies between Newtonians, Cartesians, and Leibnizo-Wolffians, reached St. Petersburg and came to a critical point there. Daniel Bernoulli, living with Euler, complained:

> I deplore my misfortune as much as ever I can.

---

[22]Letter from 20. 4. 1727. "Sie reisen jetzt in das Paradiess der Gelehrten." Archive of Russian Academy of Sciences in St. Petersburg. Quoted in Sammelband 1959, 276.

[23]Akademien im Briefwechsel 1961, II, 182. Letter to J. D. Schumacher from 7. (18.) 11. 1749. Johann Daniel (also Ivan Danilovich) Schumacher (1690–1761), librarian and councillor of the St. Petersburg Academy. "... auch ich und alle übrige, welche das Glück gehabt, einige Zeit bey der russischen Kaiserlichen Academie zu stehen, müssen gestehen, daß wir alles, was wir sind, den vortheilhaften Umständen, worin wir uns daselbst befunden, schuldig sind."

[24]Named after Ernst Johann Reichsgraf von [Imperial count of] Biron, from 1737 duke of Courland. German adventurer who became the influential chief adviser of the Russian empress Anna. After her death in 1740 he was exiled to Siberia; in 1762 he was finally granted an amnesty.

[25]Among them Jakob Hermann (1678–1733), Georg Bilfinger (also spelled Bülfinger) (1693–1750), Christian Goldbach (1690–1764), and Gerhard Müller (1705–1783).

Fig. 5.12. Oldest known portrait of Leonhard Euler and the only one that shows Euler with two healthy eyes. Mezzotint print by B. Sokolov (1737) after a lost painting of J. Brucker. Courtesy Sudhoff-Institut, Universität Leipzig.

Consequently, when his contract ended Daniel Bernoulli returned to Basel in 1733, as did other academicians. This was the first major change in Euler's scientific career. He succeeded Bernoulli and became professor of mathematics. Euler, at 26, stepped into the leading position in the Academy. He himself was conscious of his mathematical abilities. We read in a letter to the President of the Academy Laurentius Blumentrost (1692–1755) that

> in Europe there are only a few so advanced in mathematics and physics as I.[26]

Euler decided to settle down in Russia. Meanwhile his younger brother Johann Heinrich (1719–1750) worked as a painter in Petersburg. At the end of 1733 (old style) Euler married Katharina Gsell (1707–1773), the daughter of a Swiss painter, and he bought a wooden house. Euler was very fond of children, and the couple were to have 13, all but five of whom died very young. Johann Albrecht, Euler's first son, was born in 1734, followed by three daughters who died in infancy, and Karl in 1740.[27] Euler was able to work while children played around him.

Under Anna's reign (1730–1740), that is, under her German favourite Biron (1690–1772) (in German Bühren), Duke of Courland, Russia suffered one of its bloodiest periods. However, Euler led a quiet life interrupted only by one disaster. Since childhood he had suffered from the disease scrofula (kings' evil), unknown today, which led to terrible consequences during a fever in 1738, destroying the sight in his right eye and weakening the other eye. The common legends that this loss was caused by his geographical work or by an exhausting astronomical calculation are not true.

In the 1730s the dominant Wolffian philosophical position received a serious setback in the Academy because of the departure of some Wolffian members. However, Wolff continued to enjoy a position of high esteem in Petersburg and he maintained an active correspondence with the Academy. Euler strongly opposed Leibnizo-Wolffian rationalism, and there is no question that Wolff was aware of Euler's opposition. This caused conflicts later in Berlin.

After Anna's death, a boy tsar again succeeded the Empress in November 1740, and the political situation worsened again.

> Things looked rather dubious,[28]

as Euler himself said. At this time the new Prussian king Frederic II (Friedrich II) the Great (1712–1786, King from 1740), as young as Euler (28 and 33 years

[26] Letter from 7. (18.). 9. 1730. Archive of the Russian Academy of Sciences in St. Petersburg, AAN, f. 121, op. 2, no. 164.

[27] Johann Albrecht (1734–1800), Karl Johann (1740–1790), Katharina Helene [m. von Bell] (1741–1781), Christoph (1743–1808), Charlotte [m. van Delen] (1744–1780).

[28] Euler 1767/1995, 11-13. " . . . es bey der darauffolgenden Regentschafft ziemlich misslich auszusehen anfieng."

respectively), decided to reorganise the Berlin Academy, and Euler was invited to work in this Academy. At that time Berlin was no place of scholarship. Nevertheless, Euler accepted and finally departed with his family on 19 June 1741. During the years in Berlin, Euler's scientific work was closely connected with the Petersburg Academy to which he was appointed an honorary member (he received a pension). In Petersburg Euler wrote between 80 and 90 papers on number theory, analysis, and mechanics, of which 55 were published during his first Petersburg period, among them the following:

1736 *Mechanica sive motus scientia analytice exposita*, 2 vols. (*Analytical Mechanics*; E 15-16; EO II, 1-2),
1738 *Einleitung zur Rechen-Kunst zum Gebrauch des Gymnasii bey der Kaiserlichen Akademie in St. Petersburg* (*Introduction to the Art of Arithmetic for Use in the Grammar School Affiliated to the Imperial Academy in St. Petersburg*, 2 vols. 1738-1740; E 17, 35; EO III, 2)
1739 *Tentamen novae theoriae musicae* (*Theory of music*; E 33; EO III, 1).
1749 *Scientia navalis*, 2 vols. (*Naval Science* [fluid mechanics] prepared in 1738; E 110-111; EO II, 18-19)

Euler's international reputation rapidly increased as he regularly won Paris Academy prizes; in the period from 1737 to 1746 he won the Prix Paris for the problems posed in the years 1737, 1738, 1739, 1742, 1743, 1746 and 1747;[29] in total he received the prize a dozen times, mostly on navigation and naval science, and another one under the name of his son Johann Albrecht.

It was John Bernoulli who started Euler on his researches in several branches of mathematics. However, since the time of Euler's marriage new fields such as number theory and geography developed rapidly. Number theory had attracted many mathematicians before Euler, above all Pierre de Fermat (1601–1665). In a nutshell, while Pierre de Fermat had formulated and conjectured, Leonhard Euler proved and refuted. Euler laid the foundation of number theory as a science.

In 1725 Christian Goldbach (1690–1764), interested in number theory, became the Secretary of Conferences at the Petersburg Academy, but already in 1728 he left Petersburg in order to educate the young tsar Peter (1715–1730, tsar from 1727) at the court in Moscow. In 1729 Euler opened a very interesting correspondence[30] with Goldbach. In his first response Goldbach dealt with Fermat's conjecture that all numbers of the form

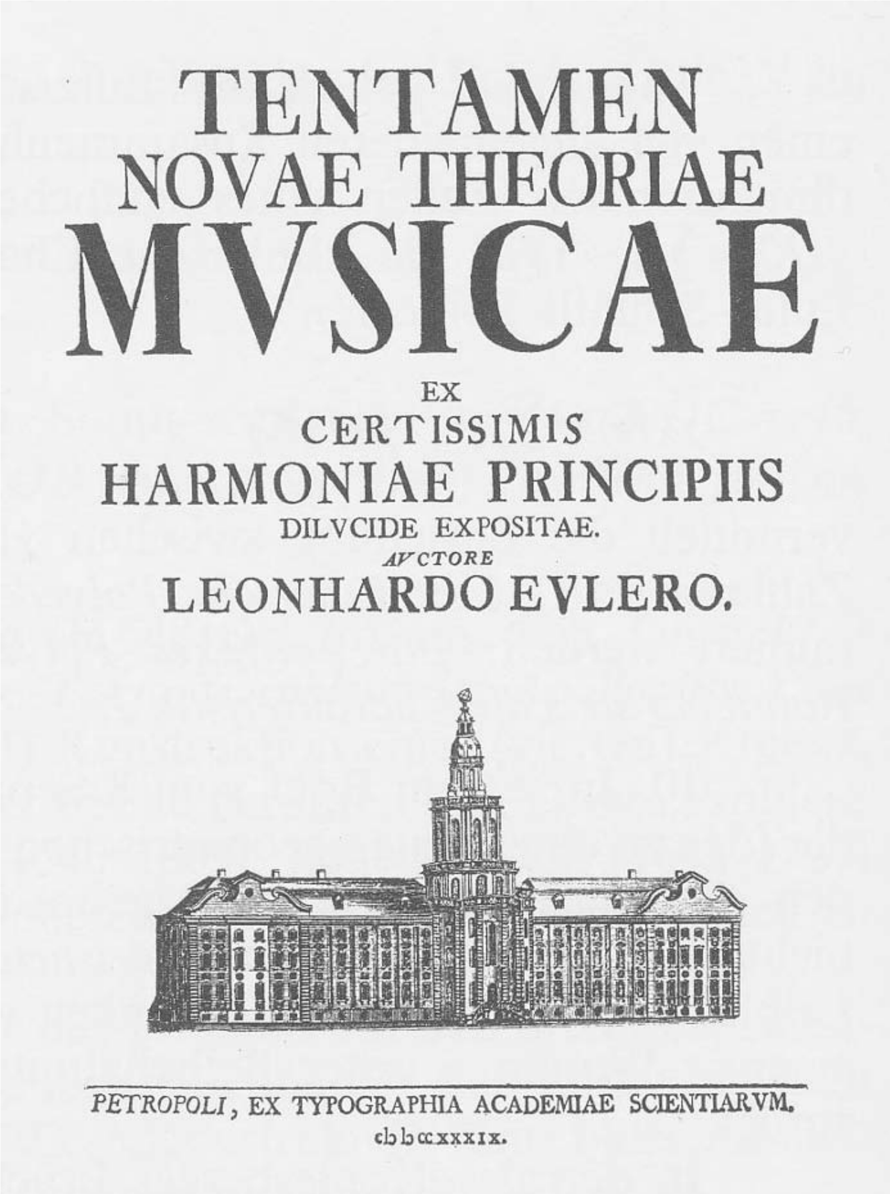
$$F_n = 2^{2^n} + 1, \qquad n = 1, 2, 3, \ldots$$

are prime. In 1732 Euler showed that F_5 is composite:

$$F_5 = 2^{32} = 4,294,967,297 = 641 \times 6,700,417$$

[29]E 34, E 78, E 57, E 108, E 109, E 150, and E 120.

[30]Euler and Goldbach 1965.



TENTAMEN
NOVAE THEORIAE
MVSICAE
EX
CERTISSIMIS
HARMONIAE PRINCIPIIS
DILVCIDE EXPOSITAE.
AVCTORE
LEONHARDO EVLERO.

PETROPOLI, EX TYPOGRAPHIA ACADEMIAE SCIENTIARVM.
cIↄIↄccxxxix.

Fig. 5.13. Euler's *Tentamen novae theoriae musicae* (*Essay on a New Theory of Music*; 1739, E 33). The title page shows the Petersburg Academy. Courtesy Universitätsbibliothek Basel.

(published in "Observationes de theoremate," 1738; E 26; EO I, 2). Soon he extended the question to numbers of the form

$$a^2 + 1,$$

and, moreover, he looked for prime factors of

$$a^2 + b^2, a^2 - b^2 \qquad \text{where } (a, b) = 1 \text{ (} a \text{ and } b \text{ are relatively prime)}$$

or even

$$ma^2 \pm nb^2 \qquad \text{where } m, n \text{ are integers.}$$

This investigation led to the basis of a general arithmetic theory of binary quadratic forms, developed by Joseph Louis Lagrange (1736–1813) and above all by Carl Friedrich Gauss (1777–1855). Euler's most remarkable discovery was the law of quadratic reciprocity, conjectured first in 1744 and completely stated in a paper "Observationes" in 1772 (published 1783; E 552; EO I, 3). But he was unable to prove it, so his conjecture was forgotten (cf. E 598 from 1775; EO I, 4).[31]

It is very impressive to see how Euler's genius developed from the minor impulse given by Goldbach to a general theory. And again it is impressive to see how a given development in mathematics is inevitable and repeats itself. We recognize the same procedure in Gauss who passed through the same development up to the law of reciprocity (his theorema aureum, the golden theorem), but more quickly and successfully, at 24 in 1801.

Incidentally, dealing with the mentioned problems Euler considered the sums

$$\begin{aligned}
0^2 &+ 41,\\
1^2 &+ 40,\\
2^2 &+ 39,\\
\vdots\ &\ \ \vdots\\
x^2 &+ (41 - x)\\
(x+1)^2 &+ (41 - x - 1)\\
\vdots\ &\ \ \vdots\\
39^2 &+ 2,\\
40^2 &+ 1,\\
41^2 &+ 0;
\end{aligned}$$

as a by-product he gave this nice example:

$$P(x) = x^2 + (41 - x) = x^2 - x + 41$$

supplies primes for $x = 0, 1, 2, \ldots, 40$, but $P(41)$ is composite (since obviously 41 is a divisor of $P(41)$).

[31] Weil 1984, §10.

Furthermore, we note the famous Königsberg[32] bridge problem originally concerning the branched river Pregel and its seven bridges. Euler generalized:

> Whatever the arrangement and division of a river into branches may be, and however many bridges there be, can one find out whether or not it is possible to cross each bridge on one tour exactly once?[33]

Euler gave his solution in 1736, founding the new branch 100 years later named topology (rather than the former analysis situs) by Johann Benedikt Listing (1808–1882), and popularized by his book *Vorstudien zur Topologie* (1847).

Berlin

Frederick the Great (Friedrich der Große), or Frederick II, transformed Prussia into an efficient and prosperous state. When his father, king Frederick William I (Friedrich Wilhelm I), died in 1740, he left an army of about 83,000 soldiers from a population of about 2,000,000. Frederick II was a brilliant military campaigner; he made Prussia the foremost military power on the Continent, enlarging Prussia's territory and undermining Austria's reign. Indeed, Frederick the Great was a great military leader but guided by a strategy of power without regard for losses. However, Henry (Heinrich) Prince of Prussia (1726–1802), a younger brother of Frederick II, preferred a more modern warfare using tactical military manoeuvres which minimized losses. These differences led to a sharp discord with the king.[34] In scientific affairs the king was later on strained terms with Euler.

On the other hand, Frederick II is often regarded as a philosopher on the throne who introduced new traits of absolutist reign. Indeed, the francophile Frederick was also interested in the sciences, although more in art and philosophy. Among his first activities was an ambitious scientific aim: to transform the unimportant Society of Science founded by his father in Berlin in 1700 into a modern academy to rival that in Paris. So, he invited top scholars to become members in his academy. When the invited Leonhard Euler finally arrived in Berlin on 25 July 1741, the Prussian king was at war and conquered Silesia (Silesian Wars 1740–1742, 1744–1745, the Seven Year's War,

---

[32] Former eastern outpost of Prussia, now Russia. Seaport. Birthplace of I. Kant and D. Hilbert.

[33] Euler 1736, p 129. "Quaecunque sit fluuii figura et distributio in ramos, atque quicunque fuerit numerus pontium, invenire, utrum per singulos pontes semel tantum transiri queat, an vero secus?"

[34] Frederick William von Steuben (1730–1794), who served in the Prussian army under Prince Henry, knew these quarrels. When Steuben served the cause of U.S. independence he supported endeavours to install a constitution like the English one rather than a republican one and, moreover, he tried to bring Henry to a leading position as a King or a Governor of the U.S.A. (letter to Henry from November 2, 1786) — but in vain (Henry's response from April 1787). Cf. von Krockow 1996.

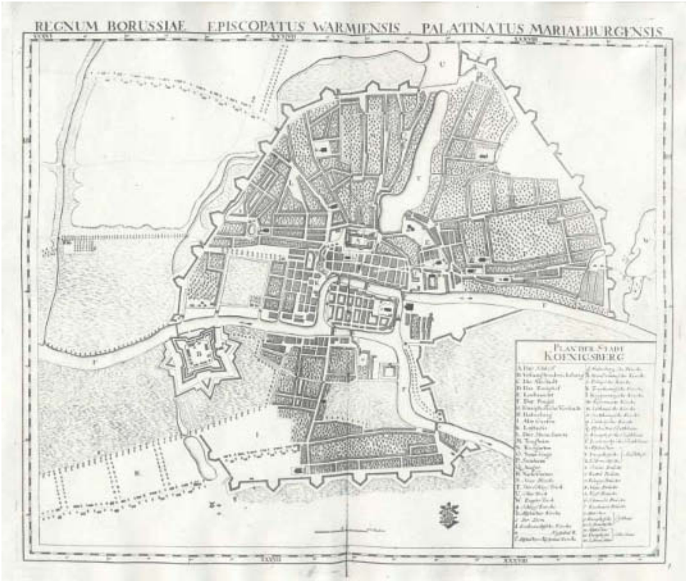


Fig. 5.14. Map of Königsberg, East Prussia, about 1760. Owing to the walks of the citizens of Königsberg over the river Pregel, the well-known Bridge Problem arose. Courtesy by Landes- und Universitätsbibliothek Halle.

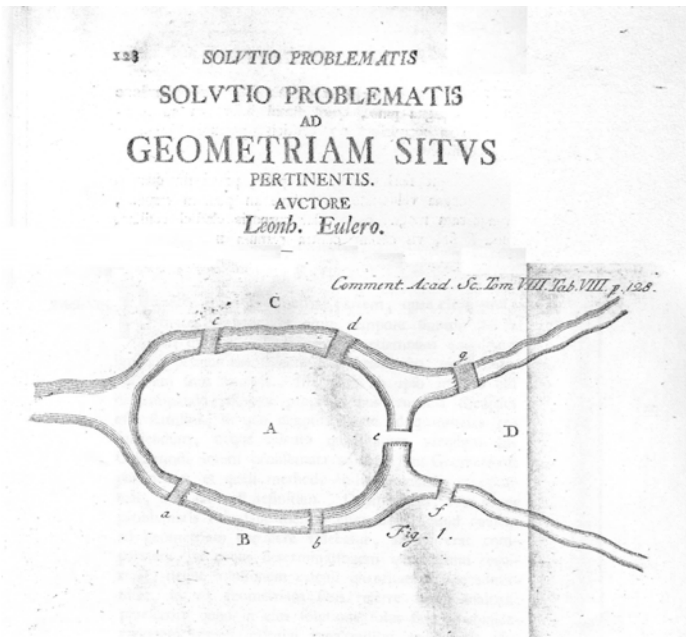


Fig. 5.15. Euler's schematic figure in his paper "Solutio problematis ad geometriam situs pertinentes" (Solution of a problem concerning geometrical position; 1736, published in 1741, E 53). Courtesy by the library of the Academy Leopoldina, Halle.

1756–1763). Because of the war the Prussian Academy was not reorganized until 1744 and formally opened in 1745.

In the 18th century universities were not centres of mathematical research, the leadership taken instead by a few royal academies. In Euler's case the Academies of Berlin and St. Petersburg (founded in 1700 and 1725 respectively), which both owed their existence to the restless ambition of Gottfried Wilhelm Leibniz, gave Euler the chance to be the mathematician we know. Euler's friend Daniel Bernoulli congratulated him from the bottom of his heart for the wonderful appointment in Berlin.[35] Incidentally, in addition to the academies in Paris and London it was above all these academies which made possible a full century of mathematical progress.

Euler became director of the mathematical class of the Academie Royale des Sciences et Belles Lettres in Berlin and even substituted for the president Pierre-Louis Moreau de Maupertuis (1698–1759) when the latter was absent. At the beginning of his Berlin period Euler was content. From the field the king addressed a letter to Euler: "A mon professeur Euler" (To my professor Euler). Euler felt flattered and wrote enthusiastically: "I can do just what I wish. [...] The king calls me his professor, and I think I am the happiest man in the world."[36] Euler bought a house in the Bärenstraße 21 [now Behrenstraße] which still exists. His mother arrived in Berlin in 1750 and stayed there until her death in 1761.

Euler's energy was inexhaustible. He supervised

- the library,
- the observatory,
- the botanical garden,
- the publication of scientific papers, and
- various financial matters including the publication of various calendars and geographical maps (the sale of which was a source of income to the Academy).

Submitting his papers to the Prussian and to the Russian Academies, Euler worked in both academies as a mathematician. Nevertheless, his endless stream of manuscripts overtaxed the publication capabilities of both academies. His busy pen left many manuscripts after his death, which the Petersburg Academy published during the next half century.

All in all, Euler led a peaceful life in Berlin for many years. However, Euler took part in several sharp philosophico-theological debates, the most famous of which was the controversial dispute on Maupertuis's principle of least action. This principle was published by Maupertuis in a paper entitled "Accord de différentes Loix de la Nature qui avoient jusq'ici paru incompatibles" (Harmony between different laws of nature which have, up to now,

[35] Fuss 1843, 2. Letter to Euler, 28. 1. 1741, 466: "Zu der herrlichen Berliner Vocation gratulire ich von Herzen."

[36] Euler and Goldbach 1965, cf. the letter to Goldbach, 2./13. 3. 1742.

Fig. 5.16. French Church in Berlin, Gendarmenmarkt. For the French refugiés (Huguenots) built by L. Cayard from 1701 to 1705, modeled after a Huguenot church in Chareton; in 1905 reconstruction changed the interior. Photo R. Thiele.

Fig. 5.17. From 1780 to 1885 the marvellous dome tower was added to the church by K. von Gontard. Owing to the French name of the tower, dôme, which sounds like the German word Dom (= Cathedral), in German the tower is named Französischer Dom (French Cathedral) even though the tower is no church. The symmetrically designed Gendarmenmarkt is one of the most beautiful squares in Germany; the counterpart of the French Church is the German Church, and both complexes have a tower and are related to the Theater (Schauspielhaus) as center. Courtesy Archive of Berlin-Brandenburgische Akademie, Berlin.

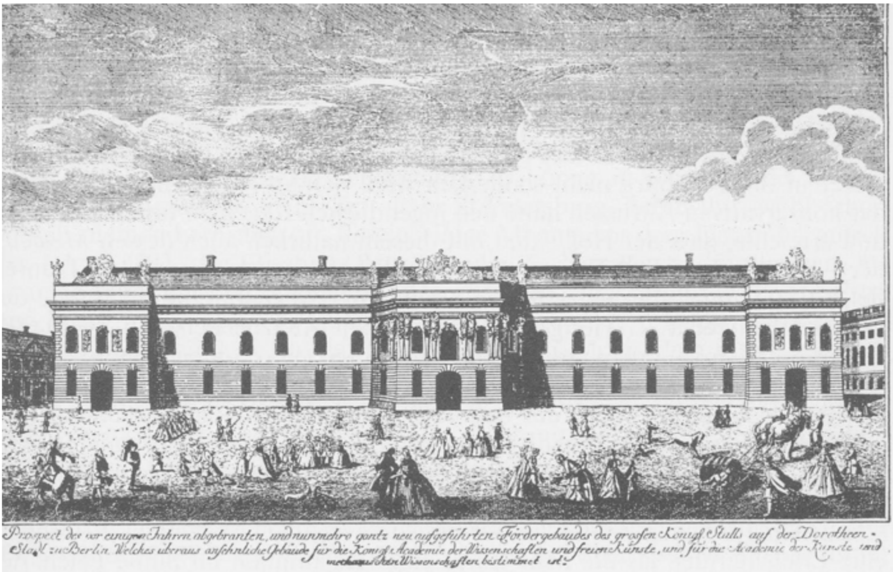


Fig. 5.18. The Prussian Academy of Sciences in Berlin (1752). Courtesy Archive of Berlin-Brandeburgischen Akademie, Berlin.

appeared incompatible) in 1744.[37] Two years later he went on to state the "principe general" in a paper "Les Loix du Mouvement et du Repos déduites d'un Principe Metaphysique" (On the Laws of Motion and of Rest Deduced by a Metaphysical Principle) in this way: "to produce some changes in nature the necessary quantity of action is the smallest that is possible."[38] By it Maupertuis hoped to unify the laws of physics and he even regarded the principle as a proof of God's existence (the supreme being). From a mathematical viewpoint, soon some scholars criticized the general principle (Gottsched circle in Leipzig 1748, Patrick Comte d'Arcy (1725–1779) 1749, 1752), and in 1751 Samuel Koenig (1712–1757) even accused Maupertuis of plagiarizing Leibniz's work.[39]

In the 1740s and 1750s this debate changed into a conflict and grew critical; Voltaire and even Frederick II were involved. Euler supported Maupertuis, but Maupertuis was finally ruined. Euler's attitude was ambiguous: because he interpreted the principle as a theological one, he was compelled to defend religion against the hated ideology of free thought. On the other hand, he

[37] Maupertuis 1744; cf. also Thiele 1996, 373-390, and 1999, 437-504.

[38] Maupertuis 1746. "Lorsqu'il arrive quelque changement dans la Nature, la Quantité d'Action, nécessaire pour ce changement, est la plus petite qu'il soit possible."

[39] Anonymous review of Maupertuis's 1744 paper, in: *Neuer Büchersaal*, vol. 7, Leipzig 1748, pp. 99-117; d'Arcy 1749 (published in 1753) and 1752; Koenig 1751.

Fig. 5.19. Frederick II (the Great), king of Prussia (1740-1786) (left); P.L.M. de Maupertuis, President of the Berlin Academy (1746-1759). Courtesy Sudhoff-Institut, Universität Leipzig.

correctly formulated the principle for some cases, and he believed that nature operates in this way. From his *Methodus inveniendi* (1744):

> For since the fabric of the Universe is most perfect and the work of a most wise Creator, nothing at all takes place in the Universe in which some rule of maximum or minimum does not appear.[40]

Euler was unpopular at Frederick's court. "Nous avons ici un gros cyclope de géomètre", Frederick II tastelessly wrote to Voltaire.[41] Euler's and the king's personalities were too different; these two important figures of the Enlightenment never became close friends. Frederick viewed science as the servant of the state — his state. King Frederick II exclusively judged science in view of its utilitarian aspects (not unlike the present arguments for cutting budgets). Inasmuch as Frederick appreciated Euler's scientific talents he engaged him in a variety of practical problems. Examples are the projects to correct the level of the Finow Canal and to build up a hydraulic system of

[40] Euler 1744, appendix *De curvis elasticis* (also in: EO I, 24, p. 231). "Cum enim Mundi universi fabrica sit perfectissima atque a Creatore sapientissimo absoluta, nihil omnio in mundo contingit, in quo non maximi minimive ratio quaepiam eluceat."

[41] Frederick II, 1849, 11. Letter to Voltaire from November 19, 1748, p. 128. "Here we have a great cyclops of mathematics [in Greek mythology cyclops are giants with one eye; Euler was nearly blind in one eye]."

pumps and pipes at the royal summer residence. Unfortunately, the fountains never worked satisfactorily.[42] The King interpreted this technical failure as a failure of science itself and commented maliciously: "Vanités des vanités! vanité de la Géometrie!"[43] Frederick preferred French culture and its representatives, such as Voltaire or Maupertuis.

When Maupertuis died in 1759 (the year of the battle of Québec), Euler continued to run the Academy, but Frederick never made him President. Euler and the King differed more and more sharply, especially in financial affairs and personnel matters concerning the Academy. Calendars were a source of revenue for the Berlin Academy. On Euler's mathematically unconvincing accounts with the king dealing with the sale of calendars Frederick responded arrogantly: "I who do not know how to calculate curves do know that sixteen thousand écus [French coin] of receipts are preferable to thirteen thousand."[44] For this and other reasons Euler began to think of leaving Berlin in 1762,[45] and during the Seven Years War he contacted the enemy Russians where Catherine the Great (tsarina from 1762) had come to power. The war ended in 1763, the same year Canada was ceded to the British by the Treaty of Paris. In 1766, February 2, Euler pleaded for royal permission to leave Prussia, but the King, now becoming aware of the immense loss, declined the request. Euler insisted. Finally, on May 2, the King agreed with the following humiliating words, showing no gratitude for Euler's incomparable work:

[42] Recently M. Eckert (Eckert 2002) pointed out that Euler is not to blame for the failure of Frederick's ambitious fountain project inspired by the French model at Versailles. Eckert quoted a letter of Euler to the king dated October 17, 1749 (EO VIA, pp. 320-330, citation p. 322, VIII): "Car sur le pied qu'elles se trouvent actuellement, il le bien certain, qu'on n'éleveroit jamais une goutte d'eau jusqu'au reservoir, et tout la force ne seroit employée qu'à la destruction de la machine et des tuyaux" (For in the situation in which they [the pumps] are at present, it is quite certain that one would never raise one drop of water as far up as the reservoir, and the entire force would be employed only for the destruction of the machine and the tubes; Eckert's translation, p. 457). But in 1778 Frederick II angrily wrote to Voltaire on the presumed mathematical execution ("exécuté géométriquement"): "Il [moulin] n'pu élever une goutte d'eau à cinquante pas du bassin" (The pumps did not elevate one drop of water fifty steps over the basin). January 25, 1778. In: Frederick II, 1853, 23, 421.

[43] Frederick II, 1853, 23. Letter to Voltaire from January 25, 1778, p. 421.

[44] Letter to Euler from June 16, 1765. In: EO VI A6, p. 390, and Frederick II, 1852, 20, 209. "Moi, qui ne sais point calculer des courbes, je sais pourtant que seize mille écus de recette en valent mieux que treize mille." As early as in 1743 Frederick II wrote to Euler: "Vous avez péché contre les règles ordinaires du calcul" (You contravened elementary rules of calculation), letter from January 21, 1743, p. 303 and 200 resp.

[45] Cf. K.-R. Biermann 1985, 91-99.

Fig. 5.20. Map showing the Northern hemisphere of the *Geographischer Atlas bestehend in 44 Land=Charten, worauf alle Theile des Erd=Creyses vorgestellt werden* (*Geographical Atlas Consisting of 44 Maps Showing All Parts of the World*). The map was drafted by the Count of Redern in 1754. The whole atlas was published by the Berlin Academy by order of L. Euler in 1762. Already in St. Petersburg Euler was involved in the edition of a map of the whole Russian Empire which appeared in 1745. Courtesy the library of the Academy Leopoldina, Halle.

> Je vous permets, sur votre lettre du 30 d'avril dernier, de quitter pour aller en Russie.[46]

Euler, at 59, left Berlin on 9 June 1766. Joseph-Louis de Lagrange succeeded him in the same year and remained in Berlin until the death of Frederick II in 1786. At the same time James Cook (1728–1779) explored the seaways and coasts of Canada (1763–1767) and began to prepare his first circumnavigation of the world in 1768–1771.

During the Berlin period Euler prepared about 380 works of which 275 were published, including these books:

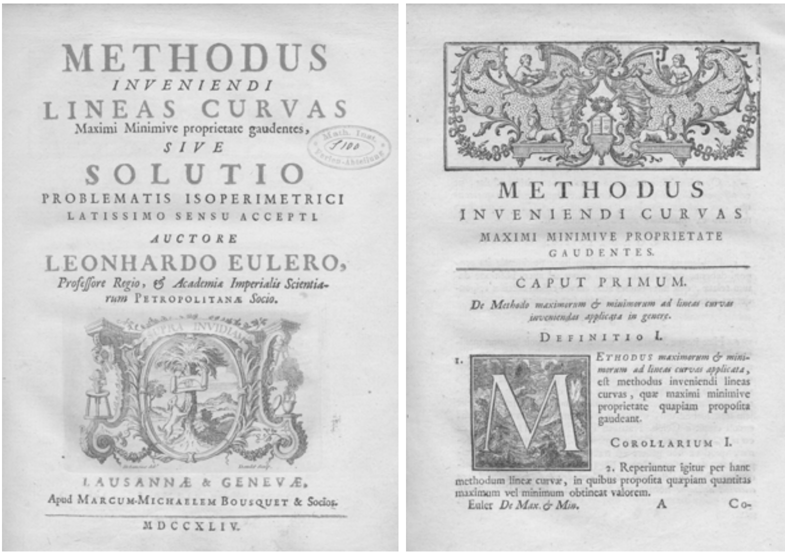
[46] Letter from Frederick II to Euler. In: EO IV A6 Basel: Birkhäuser 1986, p. 393. "Referring to your letter dated 30 of April I permit you to quit to depart for Russia."

Fig. 5.21. Euler's home in Berlin in the Behrenstraße (opposite the Komische Oper (Comic Opera), present view). Photo R. Thiele.

1744 *Methodus inveniendi*
(calculus of variations; E 65; EO I, 24)
1744 *Theoria motuum planetarum et cometarum*
(calculation of orbits; E 66; EO II, 28)
1745 *Neue Grundsätze der Artillerie*
(translated and enlarged edition of *New Principles of Gunnery* by B. Robins; E 77; EO II, 14)
(not before) 1745 *Anleitung zur Naturlehre*
(*Introduction to Natural Sciences*, published postum in *Opera posthuma* II, 1862; E 842; EO III, 1)
1746 *Gedancken von den Elementen der Körper*
(objections to monadology which were based on physical and theological nature; E 81; EO III, 2)
1747 *Rettung der göttlichen Offenbarung gegen die Einwürfe der Freygeister*
[anonymous] (*The Rescue of Divine Revelation from the Objections of the Freethinkers*; E 92; EO III, 12)
1748 *Introductio in analysin infinitorum*
(*Elements of Analysis*; E 101–102; EO I, 8-9)
1748 *Réflexions sur l'Espace et le Tems* [temps]
(*Reflections on space and time*; E 149; EO III, 2)
1749 *Scientia navalis*
(on shipbuilding and navigation, prepared in St. Petersburg; E 110–111; EO II, 18–19)
1753 *Theoria motus lunae*
(first lunar theory; E 187; EO II, 23)
1755 *Institutiones calculi differentialis*
(*Introduction to the Differential Calculus*, prepared about 1748; E 212; EO I, 10; his integral calculus written in 1763 appeared in 1768–1770; E 342, 366, 385, 660; EO I, 11–13)
1765 *Theoria motus corporum*
(second mechanics, mechanics of solids; E 289; EO II, 3–4)
[1765] *Théorie général de la dioptrique*
(*General Theory of Lenses*, prepared in Berlin, published in 1862; E 844, EO III, 9)
[1768] *Lettres à une Princesse d'Allemagne*
(philosophical letters, popular science, and philosophy, prepared in Berlin, but published in St. Petersburg; unusual success: at least 12 French, 9 English, 7 German, 4 Russian editions; E 343, 344, 417; EO III, 11–12)

In Berlin Euler's research was at the summit of his creative powers, mainly devoted to analysis and mathematical physics with a practical orientation. In his *Introductio* (E 101) Euler gave the so-called Moivre formula

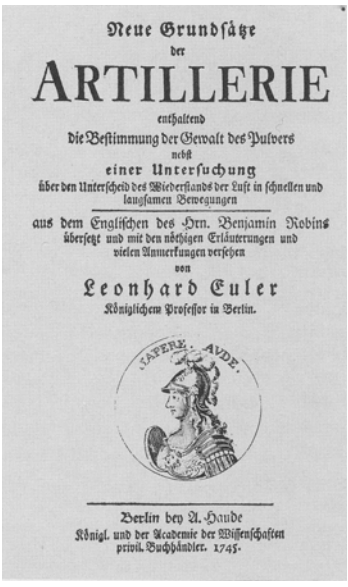
$$e^{i\varphi} = \cos\varphi + i\sin\varphi,$$



METHODUS
INVENIENDI
LINEAS CURVAS
Maximi Minimive proprietate gaudentes,
SIVE
SOLUTIO
PROBLEMATIS ISOPERIMETRICI
LATISSIMO SENSU ACCEPTI
AUCTORE
LEONHARDO EULERO,
Professore Regio, & Academiæ Imperialis Scientiarum PETROPOLITANÆ Socio.
LAUSANNÆ & GENEVÆ,
Apud MARCUM-MICHAELEM BOUSQUET & Socios.
MDCCXLIV.

METHODUS
INVENIENDI CURVAS
MAXIMI MINIMIVE PROPRIETATE GAUDENTES.
CAPUT PRIMUM.
De Methodo maximorum & minimorum ad lineas curvas inveniendas applicata in genere.
DEFINITIO I.
1. METHODUS maximorum & minimorum ad lineas curvas applicata, est methodus inveniendi lineas curvas, quæ maximi minimive proprietate quapiam proposita gaudeant.
COROLLARIUM I.
2. Reperiuntur igitur per hanc methodum lineæ curvæ, in quibus proposita quæpiam quantitas maximum vel minimum obtineat valorem.
Euler De Max. & Min. A Co-

Fig. 5.22. Title page of the *Methodus inveniendi* (calculus of variations; 1744, E 65). "One of the most beautiful mathematical works ever written" (C. Carathéodory). The poster on the tree shows a cycloid. Courtesy the library of Mathematische Institut of Universität Leipzig.



Neue Grundsätze
der
ARTILLERIE
enthaltend
die Bestimmung der Gewalt des Pulvers
nebst
einer Untersuchung
über den Unterscheid des Wiederstands der Luft in schnellen und langsamen Bewegungen
aus dem Englischen des Hrn. Benjamin Robins
übersetzt und mit den nöthigen Erläuterungen und vielen Anmerkungen versehen
von
Leonhard Euler
Königlichem Professor in Berlin.
SAPERE AUDE.
Berlin bey A. Haude
Königl. und der Academie der Wissenschaften privil. Buchhändler. 1745.

Fig. 5.23. Title page of the *Neue Grundsätze der Artillerie* (*New Principles of Gunnery*; 1745, E 77). Essential extended translation of Benjamin Robins book with the same title (1742), in a French translation (1751) read by Napoléon I Bonaparte. In the publisher's signet the slogan "Sapere aude (dare to be wise)" of the Enlightenment appears. Courtesy Universitätsbibliothek Leipzig.

which in a slightly different form had been given by Roger Cotes (1682–1716) in 1714.[47] A special case appears in the paper "De la controverse entre Mrs. Leibnitz et Bernoulli sur les logarithmes des nombres négatifs et imaginaires" (On the controversy between Messers. Leibniz and John Bernoulli on the logarithms of negative and imaginary numbers) (1749, 1751 printed; E 168; EO I, 17, pp. 195–232): $\log(-1) = i$ (discovered already in 1727), or in another form $e^{i\pi} + 1 = 0$, linking five important magnitudes in one equation.

St. Petersburg

Euler and his family (18 persons including 4 servants) arrived in St. Petersburg on 28 July 1766. He did not find any friends from his old St. Petersburg days, but he was received overwhelmingly by Catherine the Great, born a German princess. Again he lived a quiet life, and above all he lived for mathematical sciences. Clifford A. Truesdell remarked: "There is no evidence that governmental 'despotism' had any influence on Euler, who, like most scientists, accepted and rejected positions on the basis of the concrete conditions of work and pay they involved."[48]

However, he suffered some setbacks. Soon after his return he became almost completely blind. An operation only temporarily restored his sight in 1771. Owing to his phenomenal memory[49] and his enormous powers of mental calculation he was equal to the challenge and continued to work, dictating the results to his assistants, his sons Johann Albrecht (1734–1800) and Christoph (1743–1808) as well as Anders Johan Lexell (1740–1784), Wolfgang Ludwig Krafft (1743–1814), Mikhail Evseyevich Golovin (1756–1790) and Nicolas (Nicolaus) Fuss. Fuss prepared about 250 of Euler's papers and Golovin 70 which were written during Euler's second time in St. Petersburg.

In 1771 Euler's house on the embankment of the Great Neva was destroyed in a great fire, but with the help of a Swiss servant blind Euler escaped and the most of his manuscripts were saved. The house was rebuilt and is preserved, but with some changes. Nevertheless Euler, then in his sixties and almost blind, "was the principal light of Catherine II's Academy of Science" (C. A. Truesdell).[50] His papers were generally short and devoted to a particular topic; however, he completed voluminous books such as the *Theoria motum lunae* (1772) with 775 pages and the three volume *Dioptrica* (1769–1771). Almost half of his papers were written in St. Petersburg, among them:

[47]Cf. Cotes 1714.

[48]Truesdell 1984, 213.

[49]Euler was able to recite Virgil's *Aeneid* from beginning to end (almost 10,000 lines) in Latin by heart.

[50]Cf. Truesdell 1984, 338.

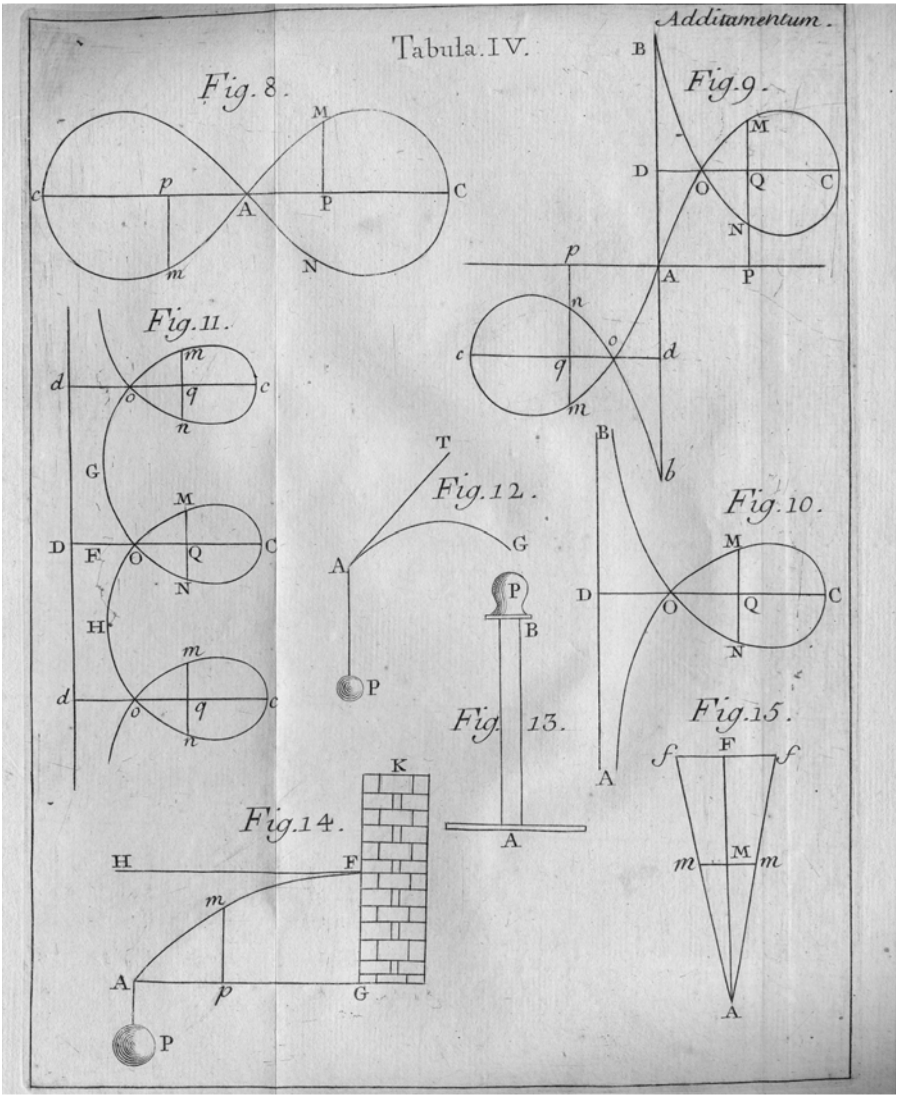


Fig. 5.24. Euler's investigations of elastic lines (of elastic materials as strips, beams, etc.) by means of the calculus of variations in "De curvis elasticis" (On elastic curves) led to nine species of which four are shown in table IV (Additamentum I in the *Methodus inveniendi*; 1744, E 65). Euler regarded an ideal elastic line to be one whose positions of stable equilibrium are characterized by the minima of the potential energy (i.e., by John Bernoulli's principal of virtual work). Cf. fig. 30. Courtesy the library of the Mathematisches Institut of Universität Leipzig.

1768 *Lettres à une Princesse d'Allemagne*
(*Letters to a German Princess* [non-existent title], 1768–1772, 3 vols.; E 343, 344, 417; EO III, 11-12; prepared in Berlin)
1768 *Institutiones calculi integralis*
(*Introduction to the Integral Calculus*, 3 vols., 1 vol. postum, 1768–1770, 1794; E 342, 366, 385, 660; EO I, 11–12)
1768 *Dioptrica*
(*Optics*, 1768–1770, 3 vols.; E 367, 386, 404; EO III, 3–4)
1770 *Vollständige Anleitung zur Algebra*
(*Complete Introduction to Algebra*, 2 vols.; E 387–388; EO I, 1)
1772 *Theoria motuum lunaea*
(second lunar theory; E 418; EO II, 22)
1773 *Théorie Complette de la Construction et de la Manoeuvre des Vaissaux*
(second naval theory; E 426; EO II, 21)

In 1773 Euler's wife, who had managed the whole household, died. Euler, at 69 and blind, intended to marry again. The children opposed their father's marriage. At first Euler gave in, but eventually he insisted. Euler announced his wedding to his children on July 20, 1776. On July 28, 1776 the wedding ceremony took place in his house and he married Salome Abigail Gsell (1723-1793), a half-sister of his late wife.[51]

Leonhard Euler died on 18 September 1783 at the age of 76 years, 5 months and 3 days, active to the end. On that very day he gave instruction to his grandson in the morning, did some mathematical calculations on the motion of balloons (recall Montgolfière), dined and discussed the orbit of Uranus (discovered in 1781 by Frederick William [Friedrich Wilhelm] Herschel, 1738–1822) with his assistants Nicolas Fuss and Andreas Johann Lexell (1740–1784). Then Euler outlined his calculations, played with his grandson, whom he had been teaching that morning, and smoked his pipe. About five o'clock Euler's pipe dropped from his hand and fell to the floor. Euler stooped, but straightened up again without the pipe and said: "Ich sterbe" (I am dying). He suffered a stroke and died the same night.

> Euler ceased to calculate and to live,[52]

said the Marquis Marie-Jean-Antoine-Nicolas de Condorcet (1743–1794), a French philosopher of the Enlightenment and Perpetual Secretary of the Paris Academy in his *Éloge de M. Euler* at the Paris Académie des Sciences. Twenty-two years earlier Euler had written in his *Lettres à une Princesse d'Allemagne* on death:

> Since death is a dissolution of the bond which links body and soul during lifetime, [...] we can derive an idea of the status of the soul

[51] Cf. the contribution "Der Zwist um die zweite Ehe Eulers (The quarrel concerning Euler's second marriage)" by G.K. Michajlow in Fellmann 1995, 112-116.
[52] Condorcet 1783. "Il cessa de calculer et de vivre."

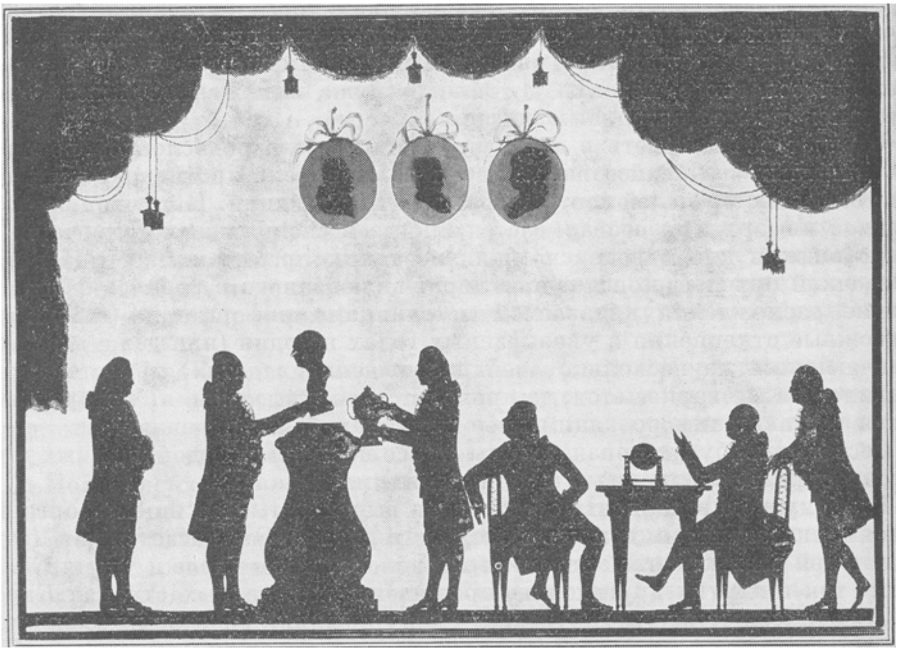


Fig. 5.25. Memorial meeting of the Petersburg Academy in honour of Euler in 1783. Silhouette by F. Anting (1784). Euler's son Johann Albrecht puts a bust of Leonhard Euler on a base. Academicians of the Mathematico-Physical Department, from left to right: A.J. Lexell (1740–1784), J.A. Euler (1734–1800) (with the bust), N. Fuss (with the amphora), I.I. Lepechin (1740–1802), P.S. Pallas (1741–1811), W.L. Krafft (1743–1814); the middle oval shadow figure on top shows the tszarina Catherine II (1729–1796, empress from 1762). Courtesy Sudhoff-Instituts of Universität Leipzig.

> after death. While the soul during life received all its information from the senses, it will no longer receive any information. [...] Sleep provides an obvious picture and at the same time an experience of this state. [...] Hence after death we shall be in a state of most perfect dreaming [...] And, I think, this is almost all that we can say [on death] in a distinct way.[53]

[53] Euler 1768-1772, Part II, Letter no. 93. Also in: EO III, 11. "La mort n'est donc autre chose que la destruction de cette liaison: ... On peut se former quelque idée de l'etat de l'ame après la mort. Comme l'ame pendant la vie tire toutes ses connoissances par le moyen des sens, étant dépoulliée par la mort de ce rapport des sens, elle n'apprend plus rien de ce qui se passe dans le monde matériel. ... Le Sommeil nous fournit aussi un bel échantillon de cet état. ... Ainsi après la mort nous nous trouverons dans un état des songes les plus parfaites, que rien ne sera plus capable de troubler. ... Et c'est à mon avis à peu près tout ce que nous saurions en dire de positif."

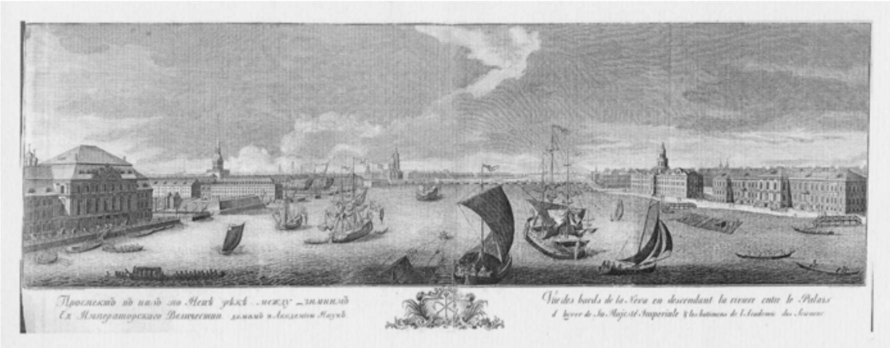


Fig. 5.26. St. Petersburg, about 1760. The view along the river Neva is downstream. The Imperial Russian Academy on the right, the Imperial Winter Palais on the left. Not far from the right side of the engraving on the embankment of the river Neva (Vasiliostrov, Vassili Island) is Euler's house in which he lived from 1766 to 1783; its present view differs essentially from the original. Copper engraving by Makhayev. Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, gr. 2°H. Russ. 434.

Fig. 5.27. Conference Room in the Imperial Academy in St. Petersburg (reconstruction). Courtesy the Archive of Berlin-Brandenburgische Akademie, Berlin.

The Concept of Function

In any lecture on Euler, who is known as analysis incarnate (l'analyse incarnée),[54] it is indispensable to say something on his analysis. Incidentally, the status of analysis at that time is described by Voltaire (1694–1778) in this way:

> The art of numbering and measuring exactly a thing whose existence cannot be conceived,

and Hermann Weyl (1885–1955) spoke of "Euler's era of happy-go-lucky analysis"[55] in order to point out the carelessness which was rooted in good and reliable intuition.

It was Euler who established the concept of function at the heart of the new branch of mathematics called analysis. Furthermore, analysis was not only an application of algebra to geometry. It became a subject in its own right, above all by making a formal theory of functions that had no need of geometrical conceptions. Therefore no figures are to be found in Euler's textbooks on analysis as for example had been the case in Isaac Barrow's *Lectiones geometricae* (1670).

But, what is a function?

"Nobody can explain that," wrote H. Weyl in *Philosophy of Mathematics and Natural Sciences*, first published in 1928 and again in 1949.[56] I will respond to this question as a historian. To understand better why Euler created functions I, therefore, provide a very brief survey of functional concepts.[57] Since the time of the Greeks magnitude has played the central role in geometry, characterised by its ability to increase and to decrease. By constructions, in general by ruler and compass, we get new magnitudes. This procedure might be called "geometrical concept of a function" used in the time of Gottfried Wilhelm Leibniz (1646–1716) and Isaac Newton.

However, new ideas arose. In Cartesian geometry (1637) in general ordinate depends on abscissa; such a relation between magnitudes is graphically represented by means of curves as well as arithmetically by means of proportions. Instead of ordinate René Descartes (1596–1650) said "la ligne est appliquée par ordre" ("the line is applied/drawn in given order"; in modern terms this means by a "given function").[58] Also the mathematical description of the real world needed concepts of relationship between observed quantities.

---

[54] "Euler, qu'on aurait pu appeler presque sans métaphore et certainement sans hyperbole, l'analyse incarnée." (Euler, whom almost without metaphor and surely without exaggeration, we can call analysis incarnate.) Arago 1854, p. 443.

[55] Weyl 1968, 124.

[56] Weyl 1928. "Niemand kann erklären, was eine Funktion ist," p. 8. In the American edition (1949) he even added: "...but this is what really matters in mathematics." However, in the index there is no entry "function."

[57] Cf. Thiele 2000, 128–181.

[58] Descartes 1637, 66.

The roots of these concepts are to be found in Galileo Galilei (1564–1642) or Isaac Newton. The latter considered all magnitudes as time-dependent. Likewise the emergence of an analytical concept arose in the calculus of variations, the typical problem of which is the determination of magnitudes, especially curves (geometrical view) or functions (analytical view). Craig Fraser (born 1951) gave an evaluation in greater detail:

> Although the theme of analysis was well established at that time [about 1730] there was in his [Euler's] work something new, the beginning of an explicit awareness of the distinction between analytical and geometrical methods and an emphasis on the desirability of the former in proving theorems of the calculus.
> Although Euler in 1744 clearly recognised the essential analytical character of the variational calculus his insight was not fully developed in his treatise. ... There was an increasing emphasis on analysis.[59]

In 1697 using a geometrical concept of function James Bernoulli posed the famous isoperimetrical problems[60] in response to his brother's brachistochrone problem. (Incidentally, the year 1997 was its tercentennial.) Dealing with his brother's new challenge John extended — or let me say transformed — this geometrical concept into an analytical one, now representing the geometrical curves analytically. In a paper published in 1706 but communicated to Leibniz already in 1698 in a letter he said:

> ... to find the curve the ordinates of which are of given power [of a line], or generally, which is the function of its ordinates expressed by other ordinates.[61]

He repeated this definition in a paper in 1718:

> Here, I call a *function* of one variable a quantity which is composed in some way of this variable quantity and of constants.[62]

The geometrical relationship, the "geometrical concept of function," disappeared. From this time forward it was Euler who developed the arithmetical concept.

There is an unpublished Latin manuscript "Calculus differentialis" by Euler written about 1727. In 30 pages he had briefly outlined the ideas which

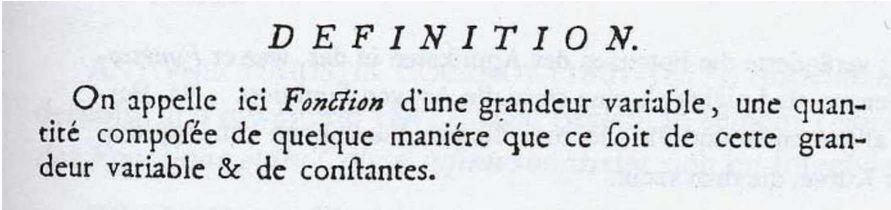
---

[59] Fraser 1997, 63.

[60] Bernoulli, Jakob 1697, 211–217. Also in: Bernoulli 1991, 271–282.

[61] Bernoulli, Johann 1706, 235: " ... trouver la Courbe telle que ses appliquées élevées à une puissance donné, ou généralement telle que les fonctions quelconques de ces appliquées, exprimées par d'autres appliquées"; already 1698 in a letter to Leibniz. See also Bernoulli 1991, 515–526, quotation p. 515.

[62] Bernoulli, Johann 1718. "On appelle ici fonction d'une grandeur variable, une quantité composée de quelque manière que ce soit de cette grandeur variable et de constantes." See also Bernoulli 1991, 527–568, quotation p. 534.



DEFINITION.

On appelle ici *Fonction* d'une grandeur variable, une quantité compoſée de quelque maniére que ce ſoit de cette grandeur variable & de conſtantes.

Fig. 5.28. John Bernoulli's concept of a function of one variable that arose by solving the isoperimetric problems, contained in the paper "Remarques" (Op. CIII) which was published in the *Mémoires de l'Academie Royal des Sciences*, Paris 1718, pp. 100–138. "We call here a quantity composed somehow by this variable quantity and by constants a function of one variable magnitude." This is the starting point of Bernoulli's disciple Euler. The first line of Euler's unpublished manuscript "Calculus Differentialis" (about 1730) reads: "Quantitas quomodocunque ex una vel pluribus quantitatibus composita appelatur ejus unius vel plurium functio" (A quantity which is composed in some way from one or more quantities is called its or their function respectively). The nature of the composition was incessantly extended by Euler. Courtesy the library of Academy Leopoldina, Halle.

he gave in the *Introductio in analysin infinitorum* (1748, E 101) about 20 years later. Euler began the manuscript by introducing the concept of function:

> One quantity composed somehow from one or a greater number of quantities is called its or their function.[63]

Then he pointed out that functions may be composed by means of algebraic operations like addition, multiplication, etc., by extraction of roots, by inversion, by taking logarithms, or by combinations of these operations. In his *Introductio* he stressed the idea that analysis is a science of functions and presented this statement:

> §4. A *function* of one variable quantity is the analytic expression (analytic formula) composed in any way from this variable quantity and numbers or constant quantities.[64]

A function is an analytic [i.e., a calculating] expression. In the course of time gradually more operations were permitted. Here, analytic expressions or formulas included transcendental functions or implicit functions or functions arrived at in the integral calculus. Euler classified functions; for example, he distinguished algebraic, transcendental, odd, even, single or multivalued functions, etc. This simple characterisation of functions by Euler makes clear

[63] Juschkevitsch 1983, 161. "Quantitas quomodocunque ex una vel pluribus quantitatibus composita appelatur ejus unius vel plurium functio."

[64] Lausanne: Bousquet 1748. Also in: EO I, 8–9. Vol. I, chap. 1, §4: "Functio quantitatis variabilis est expressio analytica quomodocunque composita ex illa quantitate variabili et numeris seu quantitabus constantibus."

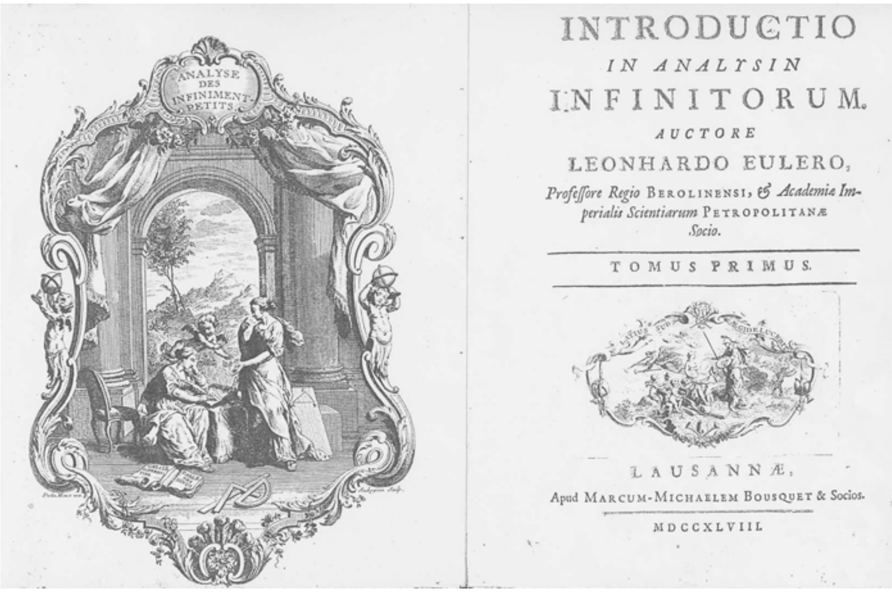


Fig. 5.29. Title page and frontispiece of the first edition of Euler's *Introductio in analysin infinitorum* (1748, E 101), vol. 1. Courtesy the library of Mathematische Institut of Universität Leipzig.

that functions were now not only a tool for mathematics but had become an object itself in mathematics. However, we should not overinterpret Euler's classification in the modern understanding as function spaces, i.e. as sets of functions with an algebraic and topological structure (because such concepts were not developed in Euler's time).

In analytic formulas variable quantities represent numbers. Therefore, the idea of variable quantity assumed — at least implicitly — a concept of any domain of numbers. When Euler explained the difference between constant [a] and variable quantities [x], he took into account also such variables as infinitesimal quantities [dx] having the strange algebraic property that

$$a + dx = a; \tag{5.1}$$

in the words of his teacher John Bernoulli in 1691/92:

> A quantity, which is increased or decreased only by an infinitely small quantity, may be considered as remaining the same.[65]

[65]Bernoulli 1922, 3: "Quantitas diminuta vel aucta quantitate infinitis minore neque diminuitur neque augetur."

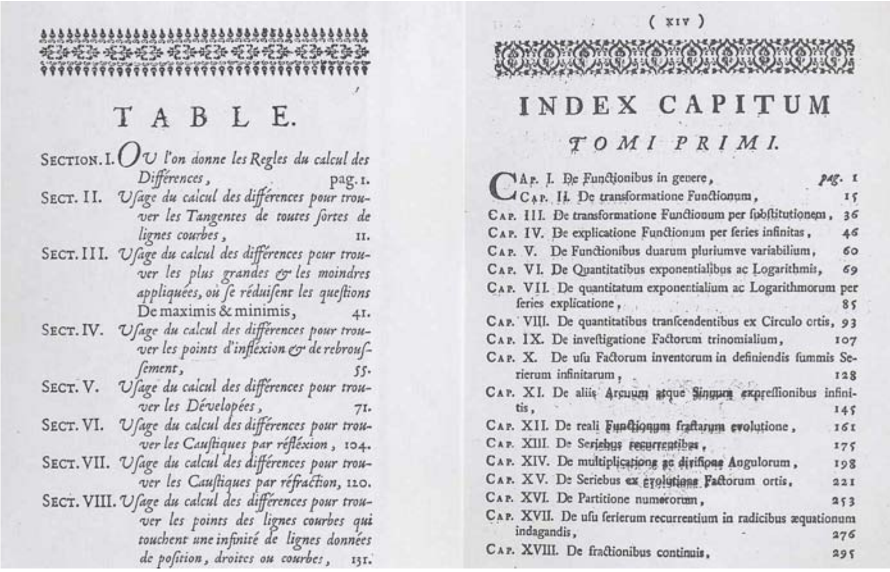


TABLE.

SECTION. I. Ou l'on donne les Regles du calcul des Différences, pag. 1.
SECT. II. Usage du calcul des différences pour trouver les Tangentes de toutes sortes de lignes courbes, 11.
SECT. III. Usage du calcul des différences pour trouver les plus grandes & les moindres appliquées, où se réduisent les questions De maximis & minimis, 41.
SECT. IV. Usage du calcul des différences pour trouver les points d'inflexion & de rebroussement, 55.
SECT. V. Usage du calcul des différences pour trouver les Dévelopées, 71.
SECT. VI. Usage du calcul des différences pour trouver les Caustiques par réflexion, 104.
SECT. VII. Usage du calcul des différences pour trouver les Caustiques par réfraction, 120.
SECT. VIII. Usage du calcul des différences pour trouver les points des lignes courbes qui touchent une infinité de lignes données de position, droites ou courbes, 131.

(xiv)

INDEX CAPITUM TOMI PRIMI.

CAP. I. De Functionibus in genere, pag. 1
CAP. II. De transformatione Functionum, 15
CAP. III. De transformatione Functionum per substitutionem, 36
CAP. IV. De explicatione Functionum per series infinitas, 46
CAP. V. De Functionibus duarum pluriumve variabilium, 60
CAP. VI. De Quantitatibus exponentialibus ac Logarithmis, 69
CAP. VII. De quantitatum exponentialium ac Logarithmorum per series explicatione, 85
CAP. VIII. De quantitatibus transcendentibus ex Circulo ortis, 93
CAP. IX. De investigatione Factorum trinomialium, 107
CAP. X. De usu Factorum inventorum in definiendis summis Serierum infinitarum, 128
CAP. XI. De aliis Arcuum atque Sinuum expressionibus infinitis, 145
CAP. XII. De reali Functionum fractarum evolutione, 161
CAP. XIII. De Seriebus recurrentibus, 175
CAP. XIV. De multiplicatione ac divisione Angulorum, 198
CAP. XV. De Seriebus ex evolutione Factorum ortis, 221
CAP. XVI. De Partitione numerorum, 253
CAP. XVII. De usu serierum recurrentium in radicibus æquationum indagandis, 276
CAP. XVIII. De fractionibus continuis, 295

Fig. 5.30. The beginning of the contents of G.F.A. de l'Hospital's *Analyse des infiniment petit* (*Analysis of the Infinitely Small*), Paris 1696 (on the left), and L. Euler's *Introductio*, vol. 1, Lausanne 1748 (on the right). Obviously, the concept of a function is totally missed in l'Hospital. Courtesy Landes- und Universitätsbibliothekbibliothek Halle and the library of Mathematische Institut of Universität Leipzig.

Consequently, Euler even wrote "dx revera $= 0$" (indeed equal zero)![66] In the differential calculus *Institutiones calculi differentialis* (published in 1755 but prepared about 1748; E 212; EO I, 10) he elaborated the calculus of zeros, in which he distinguished between arithmetical and geometrical proportions of zeros. There are various interpretations of this concept. Three examples: the "classical" understanding of infinitesimals as variables tending to zero, the view of non-standard analysis regarding infinitesimals as elements of a non-Archimedean number field, and the "physical" intuition. Euler himself wrote:

> If a [non-negative] quantity is so small that it is smaller than any given one, then it certainly could not be anything but zero. [. . .] To those who ask what the infinitely small quantity [differential] in mathematics is, we answer that it is actually zero. Hence there are not so many mysteries hidden in this concept as there are usually believed to be. The supposed mysteries have rendered the calculus of the infinitely small quite suspect to many people. Those doubts that remain we

[66]Euler 1755, Caput III, p. 78. Also in: EO I, 10. Leipzig, B.G. Teubner, Leipzig 1913, cap. III, §83, p. 69.

> shall thoroughly remove in the following pages, where we shall explain this calculus [which is only a special case of the calculus of finite differences].[67]

What before in the calculus of finite differences was assumed arbitrary is now taken to be infinitely small in the differential calculus. Taking into account the specific properties of infinitesimal quantities, Euler transformed the established rules of the calculus of finite differences into that of the differential calculus. A general principle is that in arithmetical comparison with finite quantities infinitesimal quantities can be neglected (= zero), i.e., (1) is indeed valid.[68] On the other hand, geometrical proportions of infinitesimal quantities led to the rules for differentiating functions. Euler's pragmatism here is most remarkable. Already in the early Latin manuscript he pointed out:

> It is evident that the differential calculus is a special case of the calculus which I have enunciated above [that is, the calculus of finite differences].[69]

Furthermore, in the beginning Euler believed that any analytic expression could be expressed as a power series. This idea vaguely anticipates the theorem of Weierstrass: the polynomials are dense in the continuous functions on a closed and bounded interval. His robust pragmatism may be illustrated by his remark:

> If anyone doubts that every function can be so expressed the doubt will be set aside by actually expanding functions.[70]

Obviously, this concept describes analytic functions. However, Euler extended the concepts of analytical functions and regarded other power series including so-called Puiseux and Laurent series:[71]

---

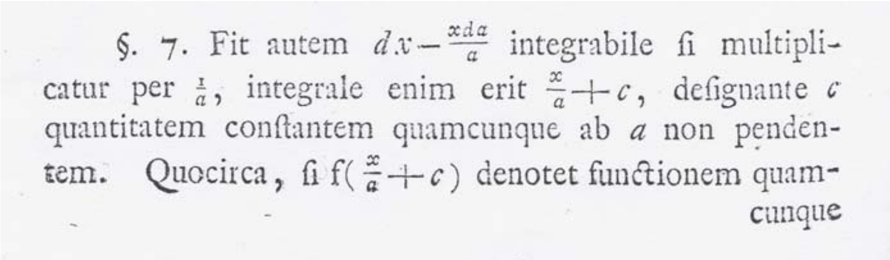
[67] Euler 1755, Caput III, §83, p. 78 (and EO I, 10). "Si enim quantitas tam fuerit parua, ut omni quantitate assignabili sit minor, ea certe non poterit non ess nulla; namque nisi esset = 0. [...] Quaerenti ergo, quid sit quantitas infinite parva in mathesi, respondemus eam ess revera = 0; neque ergo in hac idea tanta mysteria latent, quanta volgo putantur et quae pluribus calculum infinite parvorum admodum suspectum reddiderunt. Interim tamen dubia, si quae supererunt, in sequentibus, ubi hunc calculum sumus tradituri, funditus tollentur."

[68] Euler 1755, Caput III, §87, p. 80. "Hinc sequitur canon ille maxime receptus, quod infinite parva prae finitis evanescant, atque adeo horum respectu reiici queant."

[69] Cf. Juskewisch 1983, 164. "Perspicuum est Calculum differentialem, ejus, quem ante exposui, calculi esse casum specialem."

[70] Euler 1748, cap. 4, §59 (and in: EO I, 8). "Si quis dubitet, hoc dubium per ipsam evolutionem cujusque Functionis tolletur."

[71] Victor Alexandre Puiseux (1820–1883), expansion in "power" series with a fixed rational exponent (1850); Pierre Alphonse Laurent (1813–1854), (complex) power series with integral exponents (1843). Such expansions of a function are useful in the neighborhood of a pole or a singularity.



§. 7. Fit autem $dx - \frac{xda}{a}$ integrabile ſi multiplicatur per $\frac{1}{a}$, integrale enim erit $\frac{x}{a} + c$, deſignante c quantitatem conſtantem quamcunque ab a non pendentem. Quocirca, ſi f($\frac{x}{a} + c$) denotet functionem quamcunque

Fig. 5.31. The use of the universally adopted sign $f(x)$ to stand for a function of one given variable x is found firstly in an additamentum to Euler's memoir "De infinitis curvis (On infinitesimal curves)" of 1734, published in 1740 (E 44). Courtesy the library of Academy Leopoldina, Halle.

$$Az^{\alpha} + Bz^{\beta} + Cz^{\gamma} + Dz^{\delta} + \text{etc.}$$

(where α, β, γ, δ, etc. are arbitrary rational and integer numbers for Puiseux and Laurent series respectively; Cap. 4, §59). In 1905 in a paper on functions which can be represented analytically, Lebesgue showed the range of Euler's conception: when infinite expressions such as infinite series and products and continued fractions are allowed, then the class of functions is equal to that of measurable Borel functions, which for $\mathbb{R}^n$ coincide with the class of all Baire functions. In later years Euler laid less stress on the need for any particular kind of the analytical form. He had noticed that the known functions (analytic expressions) were insufficient for the requirements of analysis, especially in the debate on the vibrating string (1747). The boundary conditions of the problem lead to nonanalytic functions. A most natural shape is one with a non-differentiable point: the plucked string. If the initial state of the string is represented by an arbitrary hand-drawn curve, then we cannot expect the solution of the differential equation of the vibrating string to be an analytic expression. The debate on the new class of functions continued for another 20 years and involved prominent mathematicians, among them Euler, d'Alembert, and Daniel Bernoulli. In fact, in the discussion all the participants advanced incorrect arguments, and everybody attacked everybody. It led to ugly polemics; for instance, Jean d'Alembert (1717–1783) criticised Euler who for his part wrote to Lagrange in 1759:

> I doubt whether he [d'Alembert] is serious, unless perhaps he is thoroughly blinded by self-love.[72]

Incidentally, these were not at all the true colours of d'Alembert that the annoyed Euler saw.

To summarise the complicated debate, what is the general solution of the problem of the vibrating string? I mention the central points of disagreement.

---

[72]EO IVA5. Basel: Birkhäuser 1980, p. 420. "Je doute qu'il joue se rôle sérieusement, à moins qu'il ne soit profondément aveuglé par l'armour-propre."

Euler and d'Alembert disagreed on which kinds of functions were admitted. Euler gave up the idea of one law (formula) describing the whole function and allowed what he called discontinuous functions (that is, piecewise analytic expressions, whereas discontinuous in the sense of Euler refers to the several different analytic expressions that represent the law).[73] Daniel Bernoulli suggested trigonometric series to solve the wave equation. Euler rejected this possibility because the character of the trigonometric functions imposes certain restrictions on the form of solution that therefore cannot be fully general. This is an important question playing a crucial role in further development: can an arbitrary function be given by a particular type of series representation? Incidentally, Euler gave cursory attention to trigonometric series and on this occasion, curiously, he unconsciously did reasonable work to determine the coefficients in a trigonometric expansion.

In his *Institutiones calculi differentialis* (1755) one important result of the controversy appeared. Euler gave a general definition of a function as a quantity whose values somehow change with the changes of the independent variables:

> If, therefore, x denotes a variable quantity, all quantities which depend in some way on x or are determined by it, are called functions of this variable.[74]

The crucial words of an interpretation are "depend in some way" and "are determined." In the actual forming of such functions Euler cannot help but use the known kinds of determining; i.e., he must use the standard algebraic and transcendental operations. In other words, in his *Introductio* (1748) he had already dealt with the same concept; more or less his definitions given in 1748 and 1755 are to be regarded as being equivalent. Functionality was for Euler a matter of formal representation by calculable expressions and not so much as a description of relations by concepts. Therefore any attempt to interpret the 1755 definition in modern terms does not at all meet Euler's sight. Moreover, in a paper written in 1765 Euler remarked that the known calculus "can only be applied to curves whose nature can be contained in one analytic equation." In 1829 it was Dirichlet who showed Fourier series can represent a broad class of "arbitrary" functions, among them the functions of classical physics. In a lecture in Berlin in 1899, Herman Amandus Schwarz rightly spoke of "empirical functions" which with the help of Fourier series can be made computable.

Such modern interpretation is also often falsely attributed to the definitions of Jean Baptiste Fourier (1768–1830), Nikolai Ivanovich Lobachevsky

[73] Cf. the above definition in the *Introductio* (Euler 1748, §4), in which Euler demands one analytic expression for the law.

[74] Euler 1755, preface, p. VI. "Si igitur x denotet quantitatem variabilem, omnes quantitates, quae utcunque ab x pendent, seu per eam determinantur, eius functiones vocantur."

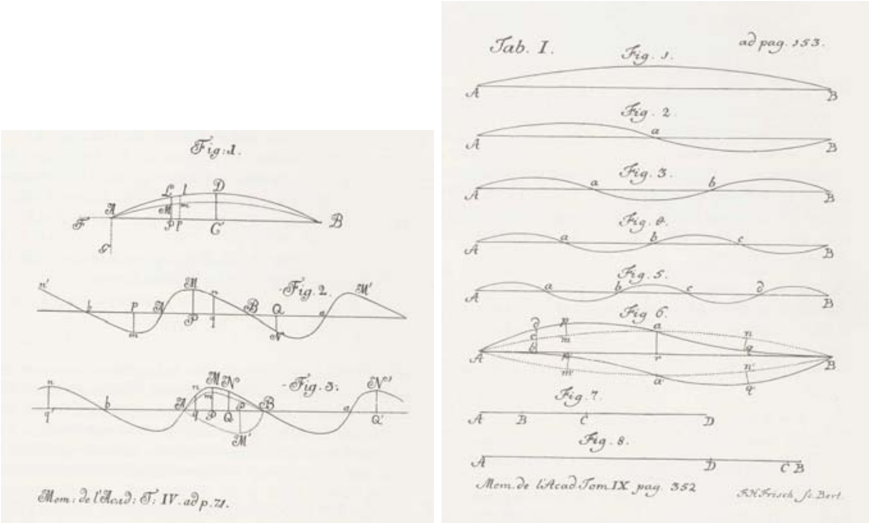


Fig. 5.32. On vibrating strings and the extension of the concept of functions. (left) Euler, "Sur la vibration des cordes" (1749, E 140). There is a Latin forerunner (E 140). Like J. le Rond d'Alembert, Euler had the general solution $y = y(x,t) = f(x+ct) + g(x-ct)$. D'Alembert regarded the arbitrary functions g and f as represented by analytic expressions, whereas Euler, extending the concept, did not demand such restricting representations.
(right) D. Bernoulli, "Réflexions et éclairissement sur les nouvelles vibrations des cordes" (1753). Daniel Bernoulli used the eigenfunctions $y_n = a_n \sin \pi x \cos n\pi t$ and got the general solution by superposition; Euler did not regard this composition as the general solution. Courtesy the library of Academy Leopoldina, Halle.

complectitur. Si igitur x *denotet quantitatem variabilem, omnes quantitates, quae vtcunque ab* x *pendent, seu per eam determinantur, eius functiones vocantur; cuius-*

Fig. 5.33. Euler's general definition of a function in the *Institutiones* (1755, E 212). Courtesy the library of Mathematische Institut of Universität Leipzig.

(1792–1856) and Johann Peter Gustav Dirichlet (1805–1859) given in 1834 and 1837, respectively. However, like Euler in his analytic expression, both regarded above all continuous changes of functions only, and therefore dealt not with modern functionality, but were guided by "mechanical" motions. Discontinuities appear in singular points only. Dirichlet's famous everywhere discontinuous function (in the modern sense) served in the first place as an example of a non-integrable function and not as an extension of the function concept. A one-to-one correspondence between arbitrary sets (= systems in Dedekind) appears 1887 in the famous paper *Was sind und was sollen die*

Zahlen? (*What are Numbers and what is their Meaning?*) by Richard Dedekind (1831–1916).[75] Whereas for numbers a topology is given by intuition, it was to be developed for arbitrary sets in order to define a continuous function on such sets.[76]

Furthermore, Euler stated that there was no need for the relation between the quantities to be given by the same law throughout an interval, nor was it necessary that the relation be given by mathematical formulas (and such functions he called discontinuous). He therefore regarded curves freely drawn by hand (cum libero manus ductu) as functions — these are the so-called mechanical (or transcendental) curves. Moreover, Euler even discussed the graph of $y(x) = (-1)^x$.

Euler did all things as easily as he could. Therefore, he fitted his concepts to the problems (not the other way round), and so he did in the case of functions too. When he built up analysis, the analytic expression was permanently extended. It is openness that characterizes Euler's concept of function. Finally, Euler classified functions as continuous and discontinuous. Continuous functions in the sense of Euler are identical with the functions he used in the *Introductio* (1748) and the *Institutiones* (1755) whereas their nature can be contained in one analytic expression (i.e., Euler's notion differs from the modern concept). Discontinuous functions, on the other hand, cannot be expressed by such a single analytic expression, but they can be piecewise composed by finitely many continuous functions and they can even be represented by hand-drawn curves.[77]

Four years after Euler's death, in 1787, the Petersburg Academy proposed the question on the nature of arbitrary functions for a prize competition. The paper "Sur la nature des fonctions arbitraire (On the nature of arbitrary functions)" by Antoine Arbogast (1759–1803) was crowned in 1790. In this prize paper Arbogast summarized Euler's view of an extended concept of functions in this way:

> Euler had the daring idea not to subject these curves to any laws [i.e., to regard arbitrary curves], and it was he who said for the first time that curves may be any line, that is, irregular and discontinuous, or composed of different parts of curves [functiones mixtas] and drawn by hand in a free movement [cum libero manus ductu], which goes with no spatial restrictions.[78]

---

[75] Dedekind 1887, §3.

[76] Cf. Thiele 2000, 128–181, here especially 170–179.

[77] Euler 1765.

[78] Arbogast 1790, 4. "M. Euler [...] eut l'idée hardie de n'assujettir ces courbes à aucune loi, & il a dit le premier, qu'elles pouvoient être quelconques, irrégulières & discontinues, c'est-à-dire, ou formées de l'assemblage de plusieurs portions de courbes différentes, ou tracées par le mouvement libre de la main qui se meut sans loi dans l'espace."

NOVA ACTA
ACADEMIAE SCIENTIARVM
IMPERIALIS
PETROPOLITANAE
TOMVS V.

PRAECEDIT HISTORIA EIVSDEM ACADEMIAE
AD ANNVM MDCCLXXXVII.

PETROPOLI
TYPIS ACADEMIAE SCIENTIARVM MDCCLXXXIX.

HISTOIRE. 5

conduit à des équations différentielles à trois & plusieurs variables: ce qui arrive même bien souvent, non seulement lorsqu'on traite des sujets de la méchanique sublime, mais surtout dans la Théorie des mouvemens des fluides: de sorte qu'on ne sauroit soûtenir rigoureusement qu'un pareil Probléme ait été résolu, avant qu'on n'ait fixé exactement la nature des fonctions arbitraires. L'Académie invite donc tous les Géomètres de décider:

Si les fonctions arbitrairés, auxquelles on parvient par l'intégration des équations à trois ou plusieurs variables, représentent des courbes ou surfaces quelconques, soit algébriques ou transcendantes, soit méchaniques, discontinues, ou produites par un mouvement volontaire de la main; ou si ces fonctions renferment seulement des courbes continues représentées par une équation algébrique ou transcendante?

Le terme du concours fut fixé jusqu'au 1 Juin 1789, & ensuite prolongé jusqu'au 1 Septembre de la même année.

Madame la Princesse de Dalchkaw prit dès le commencement de cette année un congé de six mois pour aller passer une partie de l'été sur ses terres près de Moscou: elle partit le 5 Février & revint le 31 Août. Les affaires académiques furent en attendant administrées selon les instructions que Madame la Princesse avoit données avant son départ au Secrétaire de Conférences ainsi qu'à la Chancellerie, dont elle reçut chaque semaine des rapports qui la mirent au fait de tout ce qui s'étoit passé pendant son absence dans les divers départemens de l'Académie.

Sa Majesté l'Impératrice ordonna de faire équiper quatre frégattes pour un voyage de long cours; le Collége de l'Amirauté ayant, pour cet effet, besoin de gens habiles qui sçussent déterminer, avec une exactitude suffisante, les lon-
a 3 gi-

Fig. 5.34. *Nova Acta Petropolitanae*, vol. 5 (1787), with the prize problem on the nature of functions. Courtesy Deutsche Akademie der Naturforscher, Leopoldina, Halle.

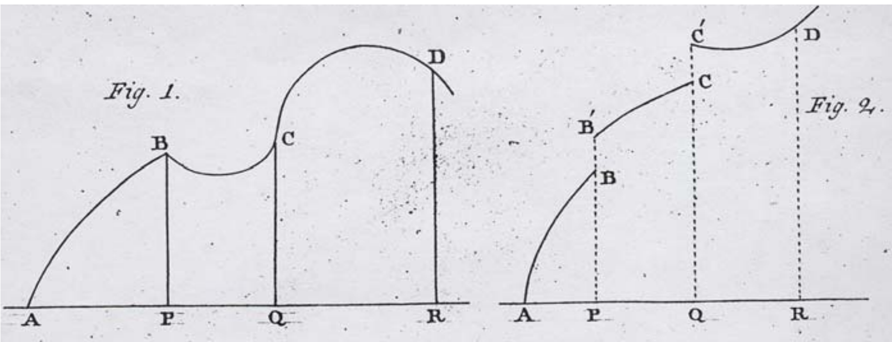


Fig. 5.35. Figures from L.F. Arbogast's paper *Sur le nature des fonctions arbitraire* (*On the Nature of Arbitrary Functions*), crowned prize paper of the Petersburg Academy 1791, in which composed functions were regarded. Both examples are discontinuous in Euler's sense; in the modern understanding only the example on the right shows a discontinuous curve (function). Courtesy Landes- und Universitätsbibliothek Halle.

The tendency to render mathematics in arithmetical terms has continued since the days of Euler. I need only remind you of Lagrange's book *Théorie des fonctions analytiques* (*Theory of analytic functions*, 1797) or of the book *Vorlesungen über reelle Funktionen* (*Lectures on real functions*, 1918) by Constantin Carathéodory (1873–1950). Felix Klein (1849–1925) spoke of the arithmetization; however, he was not pleased by these abstract arithmetizing tendencies. It was the isoperimetric problems of the calculus of variations posed in 1697 and their solutions in which the analytic expression arose that shifted the setting of analysis from geometry to arithmetic. This tendency, however, was not straightforward. In 1878 Karl Weierstrass (1815–1897) delivered the lectures "Einleitung in die Theorie der analytischen Functionen" (Introduction to the theory of analytic functions). In the lecture notes which were taken up by Adolf Hurwitz (1859-1919) we read:

> John Bernoulli first gave another and seemingly very general definition of function: if among two variable quantities there is a relation which determines along with the values of one quantity a certain number of definite values of the other one, then these quantities are called functions of each other.[79]

And some lines later:

> First of all, this definition is only valid for real numbers. But it is completely untenable and completely infertile.[80]

Why did Weierstrass come to this strange view? For him power series were the heart of analysis. Weierstrass's ultimate aim was the representation of a function.[81] To this definition he objected above all that one cannot deduce some general properties such as differentiability. Incidentally, this definition was not Bernoulli's. He neither spoke of "values of quantities" nor of "multivalued functions." The latter functions we find in Euler before 1730, but we do not find such a concept in the famous controversy on the logarithms of negative numbers between John Bernoulli and Leibniz in 1712. Using a one-infinite relation, Euler clarified the meaning of such functions in his paper "De la controverse entre Mrs. Leibnitz et Bernoulli sur les logarithmes des nombres négatifs et imaginaires" (Controversy between Mr. Leibniz and Mr.

[79] Weierstraß 1878/1988, 48.

[80] Weierstraß 1878/1988, 48. "Eine andere und scheinbar sehr allgemeine Definition einer Function gab zuerst J[ohann]. Bernoulli: Wenn zwei veränderliche Größen so miteinander zusammenhängen, daßjedem Werth der einen eine gewisse Anzahl bestimmter Werte der anderen entsprechen, so nennt man jede der Größen eine Function der anderen. ... Dieselbe gilt jedoch zunächst nur für reelle Zahlen. Sie ist aber überhaupt vollkommen unhaltbar und unfruchtbar."

[81] "Das letzte Ziel bleibt doch immer die Darstellung einer Function." Lecture notes (Mitschrift) of his cours Ausgewählte Kapitel der Functionenlehre, winter term 1885/86. Institute Mittag-Leffler, Djursholm, p. 262.

Bernoulli on the logarithms of negative and imaginary numbers)[82] in 1749. Moreover, a numerical interpretation ("value of quantity") is probably first stated by Augustin-Louis Cauchy (1789–1857) in 1821 (*Cours d'Analyse*). In comparison with Weierstrass we see the advantage of Euler's pragmatic attitude.

Only a few mathematicians have invented more than two or three symbols which are universally accepted in modern mathematics. Euler is among them (because of his influential writings). Indeed he was a great notation builder and establisher. Some examples in analysis (with the year of print):

1734	$f(x)$
1734	π
1736	e
1748	$\sin x$, $\cos x$, $\log x$
1755	Σ, Δ, Δ^2
1794	$i = \sqrt{-1}$

Let me end with a few words on Leonhard Euler himself. One of his most admirable qualities was a willingness to explain how he did mathematics, how he made discoveries. His extraordinary memory enabled him to make detailed calculations in his head (like the Austrian composer Wolfgang Amadeus Mozart (1756–1791), a younger contemporary, who had every composition in his head before he started writing). He never wanted to have the last word; on the contrary, in his papers he let the readers have many things to complete in order to encourage and to involve them. Dirk Struik (1894–2000) once said he would not like to have a cup of coffee together with the quarrelsome and envious mathematicians of the 18th century. No doubt, Euler is to be excluded. On a memorial tablet in Swiss Riehen we find a concise characterisation of nine words by Otto Spiess (1878–1966): "He was a great scholar and a kind person."[83]

Euler's enormous productivity (from 1725 to 1783 with an average about 800 pages a year) is accompanied by quality and depth of the discoveries. His prodigious output has been collected in the Leonhardi Euleri *Opera omnia*, with more than 70 volumes up to now. We have about 900 items, including 20 books, and in the correspondence there are more than 3,000 letters and 300 addresses. An entire volume is necessary to present the list of Euler's publications.[84] An evaluation of this volume supplies us with some statistics about Euler's work:[85]

[82] Euler 1749.

[83] "Er war ein großer Gelehrter und ein gütiger Mensch." Spiess's inscription on a memorial plaque at Riehen, Kirchstrasse 8, erected in 1960. Spiess was a Swiss historian of mathematics and a biographer of Euler.

[84] Eneström 1910–1913.

[85] Yushkevitch 1972, III, 37.

Years	publications	their percentage in Euler's lifetime (approximate)
1725-34	35	5%
1735-44	50	7%
1745-54	150	20%
1755-64	110	14%
1765-74	145	19%
1775-83	270	35%

Opera omnia
(Series I-III)

algebra, number theory, analysis	40%
mechanics, physics	28%
geometry, including trigonometry	18%
astronomy	11%
architecture, ballistics, philosophy, theory of music, theology, etc.	3%

Mathematics
(Series I)

integral calculus	20%
geometry, including differential geometry	17%
differential equations	13%
series	13%
number theory	13%
algebra, theory of probability	10%
foundation of analysis	7%
calculus of variations	7%

Statistics concerning Euler's publications and fields of interest (due to Eneström)

The output during his working life averaged about 800 pages a year or $15\frac{1}{2}$ pages a week. Adolf Pavlovich Youschkevich (1906–1993) noted that the 19 papers written in 1751 contain about 1,000 pages, as do the 56 papers of 1776 a quarter of a century later.[86]

Each volume of Euler's *Opera omnia* provides important texts. Thirty entries in the index of a Japanese *Encyclopedic Dictionary of Mathematics* and even 53 articles of the German *Mathematisches Wörterbuch* (*Mathematical Dictionary*) confirm Euler's influence over the more than two centuries since his death in 1783 until modern times.[87] Indeed, the esteem in which Leon-

[86]Jouschkevich 1971, 2, 741.

[87]*Encyclopedic Dictionary*, K. Ito, ed. Tokyo: Iwanami Shoten 1954, revised edition 1960, English translation Cambridge: MIT Press 1987, 2 vols.; *Mathematisches Wörterbuch*, 2 vols. J. Naas and H. Schmid, eds. Berlin: Akademie-Verlag 1961.

hard Euler was held from the beginning through our days has not diminished. Johann Heinrich Lambert (1728–1777), Euler's colleague during the Berlin period, regarded Euler and d'Alembert as the first mathematicians among his contemporaries.[88] Euler was "the most prolific mathematician in history" and the "major figure in the development of analysis in the eighteenth century" (Victor Katz, born 1942).[89] One and a half centuries before the *princeps mathematicorum*, Carl Friedrich Gauss (1777–1855) said of Euler's mathematical lifework:

> The study of Euler's work will remain the best school for the different fields of mathematics and nothing else can replace it.[90]

In his *Disquisitiones arithmeticae* (1801) Gauss spoke of "summus Euler" using an epithet he attributed nowhere else to a scholar with the only exception of Isaac Newton.[91]

It has been impossible to summarize all of Euler's important contributions to mathematics, and I did not try to do so. At best one can merely present his work in a qualified sense. In this spirit, let me quote in conclusion from Euler's younger contemporary Jonathan Swift (1667–1745):

> Elephants are always drawn smaller than life.

[88] Lambert's esteem is reported in the memories of D. Thiébault *Mes souvenirs de vingt ans de séjour à Berlin ou Frédéric le Grand* (*My memoirs of 20 years of a sojourn in Berlin or Frederick the Great*) (Paris: Buission 1804), vol. 5, p. 31f. "Le premier géomètre vivant, me répondit-il, c'est M. Euler et M. d'Alembert, ou M. d'Alembert et M. Euler: je les place au même rang." (The first living mathematician, he answered [to Thiébault's question for the most famous living mathematicians], are Mr. Euler and Mr. d'Alembert or Mr. d'Alembert and Mr. Euler. I put them on the same rank.) Incidentally, Lambert added proudly: "Le troisième c'est moi." (The third, that is me.)

[89] Katz 1993, 495.

[90] Letter to P.H. Fuss from September, 11, 1849, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Gauss-Nachlass. "Das Studium der Werke Eulers bleibt die beste Schule in den verschiedenen Gebieten der Mathematik und knn durch nichts ersetzt werden." Cf. our motto by Laplace: *Read Euler, he is the master of us all.*

[91] §151 "vir summus", §56 "vir sagacissimus". Furthermore, in his *Disquisitione arithmeticae* Gauss quoted not less than 28 papers of Euler (14 of Lagrange). Inserted in Gauss's own copy of Euler's *Methodus inveniendi* (Niedersächsische Staats- und Universitätsbibliothek Göttingen, Gauss collection) there is a traced portrait of Euler made by Gauss (communication by Prof. Dr. K. Reich, Hamburg).

References

Sources and Bibliographies

Akademien im Briefwechsel, 1961. *Die Berliner und die Petersburger Akademie der Wissenschaften im Briefwechsel Leonhard Eulers* (Juschkewitsch, A. and E. Winter, Eds.). Akademie-Verlag, Berlin vol. 2.

Arbogast, A., 1790. *Sur la nature des fonctions arbitraire.* Academy, St. Petersburg.

Bernoulli, Johann, 1691/1922. Johannis Bernoulli Lectiones de calculo differentialium (Schafheitlin, P., Ed.). *Verhandlungen der Naturforschenden Gesellschaft in Basel* 34 (1922) 1–32.

Bernoulli, Johann, 1696. Problema novum. *Acta eruditorum*, Junii 1696, p. 269. Also in: Struik 1968, 391, and in: Bernoulli 1991, 212, 258–262.

Bernoulli, Johann 1706. Solutio Du Problême proposé par M. Jacques Bernoulli. *Mémoires de l'Academie Royale des Sciences*, Paris.

Bernoulli, Johann 1718. Remarques. *Mémoires de l'Academie Royale des Sciences*, Paris, pp. 100–138.

Bernoulli, Jakob 1697. Solutio Problematum Fraternorum. *Acta eruditorum*, Maji 1697, pp. 211–217. Also in: Bernoulli 1991, 271–282.

Bernoulli, Johann and Jacob, 1991. *Die Streitschriften von Jacob und Johann Bernoulli* (Goldstine, H.H., Ed.). Birkhäuser, Basel.

Condorcet, A. de, 1783. Eloge de Euler. *Histoire de l'Académie royale des sciences pour l'année 1783 (1786).* Also in: *Opera omnia Euleri* III, 12, Zürich 1960, and Oeuvres de Condorcet, vol. 3. Paris, Didot 1847, pp. 1–42, p. 41.

Euler, L., 1839. *Œuvres complètes en français de L. Euler* (Dubois et al., eds.), 5 vols. Etablissement Géographique, Brussels (E 786).

Euler, L., 1849. *Opera minora collecta.* Commentationes arithmeticae collectae (P.H. and N. Fuss, Eds.), 2 vols. Academy, St. Petersburg 1849 (E 791).

Euler, L. 1911-, (EO I-IV). *Leonhardi Euleri Opera omnia* (Collected works), divided in four series: I Mathematics (29 vols.), II Mechanics and Astronomy (31 vols., 29 publ.), III Physics (12 vols., 11 publ.), IV A+B Correspondence and Manuscripts (8 + 7 vols., 4 publ.), edited from the Euler Committee of the Swiss Academy of Sciences. Basel: Birkhäuser, formerly Leipzig: B. G. Teubner, and Zürich (Turici): Orell-Füssli, since 1911.
[All references to Euler's collected works are given by EO followed by series and volume number in Roman and Hindu-Arabic numerals respectively: "EO ser., vol."]

Euler, L., 1737. Solutio problematis ad geometriam situs pertinentis (E 53). *Commentarii Academia Scientiarum Petropolitanae* 8 (1736), 128–140, actually published in 1741. Also in: EO I, 7, and a complete English translation in Struik 1968, 183–187.

Euler, L., 1744. *Methodus inveniendi* (E 65). Bousquet, Lausanne. Also in: EO I, 24. Zürich: Orell-Füssli 1952.

Euler, L., 1748. *Introductio in analysin infinitorum*, vol. 1. Bousquet, Lausanne 1748. Also in: EO I, 8-9. B.G. Teubner, Leipzig 1922. English translation by John Blanton, Euler 1988.

Euler, L. 1749. De la controverse entre Mrs. Leibnitz et Bernoulli sur les logarithmes des nombres négatifs et imaginaires. *Mémoires de l'Académie royale des sciences et de belles lettres* (1749). Also in: EO I, 17. B. G. Teubner, Leipzig 1914. German translation in: *Zur Theorie komplexer Funktionen*, in: Ostwalds Klassiker, No. 261 (P. Juschkewitsch, Ed.). Leipzig: Geest & Portig 1983.

Euler, L., 1755. *Institutiones calculi differentialis* (E 212). Academy, St. Petersburg. Also in: EO I, 10. Leipzig, B.G. Teubner 1913. English translation by John Blanton, Euler 2000.

Euler, L., 1765. De usu functionum (E 322). *Nova Commentarii Academia Scientiarum Petropolitanae* 11, 67–102. Also in: EO II, 25.

Euler, L., 1768-1772. *Lettres à une Princesse d'Allemagne.* St. Petersburg. Also in: EO III, 11-12; English translation *Letter of Euler to a German Princess*, 2 vols., Thoemmes Press 1977.

Euler, L., 1983. *Zur Theorie komplexer Funktionen* (A.P. Juschkewitsch, Ed.). Ostwalds Klassiker der exakten Wissenschaften, No. 261. Akademische Verlagsgesellschaft Geest & Portig, Leipzig. [Comments to and German translations of E 101 (partly), E 168, E 390, E 490, E 656, E 675, E 694.]

Euler, J.A., 1767/1995. Meines Vaters Lebens-Lauf, dictated to his son Johann Albrecht by L. Euler. Archive of the Russian Academy of Sciences in St. Petersburg. Also in: Fellmann, *Euler*. Rowohlt, Hamburg 1995, pp. 11–13.

Euler, K., 1955. Das Geschlecht Euler-Schölpi. Geschichte einer alten Familie [Genealogy]. W. Schmitz, Gießen.

Eneström, G., 1910–1913. Verzeichnis der Schriften Leonhard Eulers. Jahresberichte der Deutschen Mathematiker-Vereinigung, Ergänzungsband 4. B. G. Teubner, Leipzig. [The bibliography is organised in three parts in view of the date of publication, of the date of composition, and of the subjects. Each paper has a so-called Eneström number (denoted by E no.). Every volume of EO IV has a helpful table in which Euler's paper can be located in Series I to III by the Eneström number.]

Eneström, G., 1904–1905. Der Briefwechsel zwischen Leonhard Euler und Johann I Bernoulli. *Bibliotheca mathematica* (3) 4, 344–388, (3) 5, 248–291; (3) 6, 16-87; incomplete.

Eneström, G., 1906. Der Briefwechsel zwischen Leonhard Euler und Daniel Bernoulli. *Bibliotheca mathematica* (3) 7 126–156; incomplete.

Euler, L. and C. Goldbach, 1965. *Briefwechsel 1729–1764.* (A.P. Juskevic and E. Winter, Eds.). Akademie-Verlag, Berlin.

Frederick II, 1849, 1852, 1853. *Oeuvres de Frédéric le Grand* (Preuss, J.D.E., Ed.), t. 11; t. 20; t. 23. Decker, Berlin.

Fuss, P.H., 1843. *Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIème siècle*, 2 vols. St.-Pétersburg (Reprint: New York 1968).

Knobloch, W., 1984 (Ed.). *Leonhard Eulers Wirken an der Berliner Akademie der Wissenschaften.* 1741-1766. Spezialinventar. Regesten der Euler-Dokumente aus dem Zentralen Archiv der Akademie der Wissenschaften. Akademie-Verlag, Berlin.

Koenig, S., 1751. De universalio principio. *Acta eruditorum*, March 1751.

Maupertuis, M. de, 1744. Accord de différentes Loix. *Mémoires de l'Academie Royale des Sciences*, Paris (shortened Reprint Amsterdam 1751). Also in: *Opera omnia Euleri* ser. II, 5. Orell Füssli, Zürich 1957 (includes also some of Maupertuis' writings).

Maupertuis, P. de, 1746. Les loix du mouvement et du repos déduites d'un principe métaphysique. *Histoire de l'Académie Royale des Sciences et Belles-lettres de Berlin*, pp. 267–294. Also in: *Opera omnia Euleri* ser. II, vol. 5.

Weierstraß, K., 1878/1988. *Einleitung in die Theorie der analytischen Funktionen.* Lecture note by A. Hurwitz in 1878, edited by P. Ullrich. Vieweg, Braunschweig 1988.

Winter, E., 1957 (Ed.), *Die Registres der Berliner Akademie der Wissenschaften. 1746–1766.* Dokumente für das Wirken Leonhard Eulers in Berlin. Akademie-Verlag, Berlin.

Youshkevitch, A. and V.I. Smirnov, 1967. *Eyler. Perepiska, annoturovany ukazatel (Annoted register of Euler's correspondence).* In Russian. Nauka, Leningrad.

English Translations

Euler, L., 1988-1990. *Introduction to Analysis of the Infinite*, 2 vols. (E 101-102) Transl. and Ed. J. Blanton. Springer, New York.

Euler, L., 1754. On the general and fundamental principle of all mechanics whereon all other principles relative to the motion of solids or fluids should be established. *Gentleman's Magazine* 24, 6–7 (E 177). Also in: Fauvel, J., and J. Gray (Eds.), *The History of Mathematics.* Macmillan, London 1988, 460-462 (reprints).

Euler, L., 2000. *Foundations of Differential Calculus* (E 212) (J. Blanton, Ed.). Springer, New York.

Euler, L., 1828. *Elements of Algebra* (E 387) (J. Hewlett, Ed.). Longman, London; Reprint of the 5th edition with an introduction by C.A. Truesdell. Springer, New York 1984.

Euler. L., 1833. *Letters of Euler on Different Subjects in Natural Philosophy* (E 343, 344, 417). Trans. David Brewster with additional notes of John Griscom. Harper, New York 1833; Arno Press, New York 1975.

Euler, L., 1737. Solutio problematis ad geometriam situs pertinentis (E 53). *Commentarii Academia Scientiarum Petropolitanae* 8 (1736), 128–140, Trans. in Struik 1968, 183–187.

Proceedings of Conferences Held in Honour of Leonhard Euler

Festschrift 1907. *Festschrift zur Feier des 200. Geburtstages von Leonhard Euler.* Herausgegeben von der Berliner Mathematischen Gesellschaft (P. Schafheitlin, Ed.). B.G. Teubner; Leipzig 1907.

Sbornik 1958. *Leonard Eyler. Sbornik statey v tshest 250-letiya so dnya roshdeniya, predstablennych Akademy Nauk SSR.* In Russian. Moscow. [The 1957 conferences held in Berlin and St. Petersburg were coordinated; therefore, in the proceedings we find Russian and German abstracts respectively.]

Sammelband 1959. *Sammelband der zu Ehren des 250. Geburtstages Leonhard Eulers der Deutschen Akademie der Wissenschaften zu Berlin vorgelegten Abhandlungen* (K. Schröder, Ed.). Akademie-Verlag, Berlin.

Beiträge 1983. *Leonhard Euler. Beiträge zu Leben und Werk. Gedenkband des Kantons Basel-Stadt* (J.J. Burckhardt et al., Eds.). Birkhäuser, Basel [Very extensive bibliography by J.J. Burckhardt, pp. 511–552].

A tribute to Leonhard Euler. Special issue of *Mathematics Magazine* 56, 5 (1983).

Euler-Kolloquium 1983. *Zum Werk Leonhard Eulers.* Vorträge des Euler-Kolloquiums 1983 in Berlin. (E. Knobloch et al., eds.). Birkhäuser, Basel.

Wissenschaftliche Konferenz 1985. *Festakt und Wissenschaftliche Konferenz aus Anlaß des 200. Todestages von Leonhard Euler.* (W. Engel, ed.). Akademie-Verlag, Berlin.

Idey Eylera 1988. *Rasvitiye idey Leonardo Eylera i sovremennaya nauka.* (N.N. Bogolyubov et al., eds.) In Russian. Nauka, Moscow.

Biographies

Fuß, N., 1783. Éloge de M. Léonard Euler. Lu 23 Octobre 1783. Avec une liste [in]complette de ouvrages de M. Euler (pp. 74–124). St. Péterbourg: Academy. German translation by Fuß in: *Opera omnia Euleri* (= EO) ser. I, 1. B.G. Teubner, Leipzig 1911.

de Condorcet, Marquis, 1783. Éloge de M. Euler. *Histoire de l'Academie Royale des Sciences Paris* 1783 (printed 1786), 37–68. Also in: EO III, 11.

Wolf, R., 1862. "Euler." In: *Biographien zur Kulturgeschichte der Schweiz.* 4. Cyclus. Orell-Füssli, Zürich, pp. 87–134.

du Pasquier, L.-G., 1927. *Léonard Euler et ses amis.* Hermann, Paris.

Spieß, O., 1929. *Leonhard Euler.* Huber, Frauenfeld-Leipzig.

Fueter, R., 1948. *Leonhard Euler.* Elemente der Mathematik, Beiheft 3. Birkhäuser, Basel (Reprints[3] 1978).

Youschkevitch, A.P., 1971. "Leonhard Euler." In: *Dictionary of Scientific Biography*, vol. 4, pp. 467–484. Scribner's & Sons, New York. Reprinted in: *Biographical Dictionary of Mathematics*, vol. 2, pp. 736–753. New York: Scribner's & Sons, 1991.

Thiele, R., 1982. *Leonhard Euler.* B.G. Teubner, Leipzig. Bulgarian edition Nauka i Iskustvo, Sofia 1985; Ukrainian edition Visha shkola, Kiev 1983; Italian edition Pitagora Editrice, Bologna 2000.

Fellmann, E.A., 1995. *Leonhard Euler.* Rowohlt, Hamburg.

Selected References

Arago, F., 1854. *Oeuvres complètes.* J.A. Barral, ed. Tom 2: Notices biographiques. Paris: Gide.

Bernoulli, R., 1983. Leonhard Eulers Augenkrankheiten. In: *Leonhard Euler.* Gedenkband des Kantons Basel-Stadt. (J.J. Burckhardt et al., Eds.). Birkhäuser, Basel, pp. 471–488.

Biermann, K.-R., 1985. Wurde Leonhard Euler durch J.H. Lambert aus Berlin vertrieben? In: *Festakt und Wissenschaftliche Konferenz 1983 in Berlin* (Engel, W., Ed.). Akademie-Verlag, Berlin, pp. 91–99.

Breidert, W., 1983. Leonhard Euler und die Philosophie. In: *Leonhard Euler.* Gedenkband des Kantons Basel-Stadt. (J.J. Burckhardt et al., Eds.). Birkhäuser, Basel, pp. 447–458.

Calinger, R., 1975–1976. Euler's letters to a princess of Germany as an expression of his mature scientific outlook. *Archives of the History of the Exact Sciences* 15, 211–233.

Calinger, R., 1996. Leonhard Euler. The first St. Peterburg years (1727-1741). *Historia Mathematica* 23, 121–166.

Cotes, R., 1714. Logometria. *Philosophical Transactions* 1714 (1717); reprinted in: *Harmonia mensurarum.* London 1722.

d'Arcy, P., 1749, 1752. Réflexions sur le principe de la moindre action, and Réplique à une mémoire de M. de Maupertuis, both in *Mèmoires de l'Academie Royale des Sciences*, 1749 (published in 1753) and 1752.

Dedekind, R., 1887. *Was sind und was sollen die Zahlen?* Vieweg, Braunschweig. Reprints. English translation in: *Essays on the theory of numbers* (Transl. W.W. Beman). Dover, New York 1963.

Descartes, R., 1637/1954. *La Géométrie.* La Haye, 1637. Cf. D. E. Smith's bilingual edition of Dover, New York 1954.

Dunham, W., 1999. *Euler. The Master of Us All.* Dolciani Mathematical Expositions No, 22. MAA, Washington (DC).

Eckert, M., 2002. Euler and the fountains of Sanssouci. *Archive for History of Exact Science* 56, 451–468.

Engelsman, S.B., 1990. What you should know about Euler's *Opera omnia. Nieuw Archief voor Wiskunde* 8,1, 67–79.

Fraser, C., 1991. Mathematical technique and physical conception in Euler's investigation of the elastica. *Centaurus* 34, 211–246.

Fraser, C., 1994. The origin of Euler's variational calculus. *Archives of the History of the Exact Sciences* 47, 103–141.

Fraser, C., 1997. The background to and early emergence of Euler's analysis. In: *Analysis and Synthesis in Mathematics* (M. Otte and. M. Panza, eds.). Kluwer, Dordrecht.

Gauss, C.F., 1801/1973. *Disquisitiones arithmeticae.* Göttingen 1801. Also in: *Werke*, vol. 1. Reprint Hildesheim, Olms 1973.

Hakfoort, C., 1995. *Optics in the age of Euler.* University Press, Cambridge.

Hankel, H., 1882. Untersuchungen über die unendlich oft oszillierenden und unstetigen Fuktionen. *Mathematische Annalen* 20, 63–112.

Hartweg, F.G., 1979. Leonhard Eulers Tätigkeit in der französisch-reformierten Kirche von Berlin. *Die Hugenottenkirche* 32, 4, 14–15; 32, 5, 17–18.

Hult, J., 1985. Eulers Briefe an eine deutsche Prinzessin. Populärwissenschaft höchster Vollendung. In: *Festakt und Wissenschaftliche Konferenz aus Anlaß des 200. Todestages von Leonhard Euler* (W. Engel, Ed.). Akademie-Verlag, Berlin, pp. 83–90.

Juschkewitsch, A.P., 1959. Euler und Lagrange über die Grundlagen der Analysis. In: *Sammelband* (K. Schröder, ed.). Akademie-Verlag, Berlin, pp. 224–244.

Juschkewitsch, A.P., 1983. Euler's unpublished manuscript Calculus Differentialis. In: *Leonhard Euler.* Gedenkband des Kantons Basel-Stadt (J. J. Burckhardt et al., Eds.). Birkhäuser, Basel, pp. 161–170.

Katharina II, 1986. *Memoiren*, 2 vols. Insel, Leipzig (German translations of the French and Russian memoirs).

Katz, V., 1993. *A History of Mathematics.* Harper Collins, New York.

Kline, M., 1983. Euler and infinite series. *Mathematics Magazine* 56, 5, 307–315.

von Krockow, C. Graf, 1996. *Die preußischen Brüder.* Dtv, München.

Laugwitz, D., 1983. Die Nichtstandard-Analysis: Eine Wiederaufnahme der Ideen und Methoden von Leibniz und Euler. In: *Leonhard Euler.* Gedenkband des Kantons Basel-Stadt (J. J. Burckhardt et al., Eds.). Birkhäuser, Basel, pp. 185–198.

Lichtenberg, G.C., 1972. Vermischte Gedanken über die aërostatischen Maschinen. *Schriften und Briefe* (Promis, W., Ed.), vol. 3, pp. 62–63. C. Hanser, München.

Lützen, J., 1983. Euler's vision of a general partial differential calculus for a generalized kind of function. *Mathematics Magazine* 56, 5, 299–306.

Maltese, G., 2000. On the relativity of motion in Leonhard Euler's science. *Archive for History of Exact Sciences* 54, 319–348.

Outram, D., 1996. *The Enlightenment.* Cambridge 1996 (second edition).

Pulte, H., 1989. *Das Prinzip der kleinsten Wirkung und die Kraftkonzeptionen der rationalen Mechanik.* Sonderheft 19 der Studia Leibnitiana. Steiner, Stuttgart.

Soreth, M., 2000. Die Eulerkreise in Eulers Briefen an eine deutsche Prinzessin. In: *Mathesis* (R. Thiele, Ed.). GNT-Verlag, Berlin, pp. 55–81.

Raith, M., 1983. Der Vater Paulus Euler. Beiträge zum Verständnis der geistigen Herkunft Leonhard Eulers. In: *Leonhard Euler.* Gedenkband des Kantons Basel-Stadt. (J. J. Burckhardt et al., Eds.). Birkhäuser, Basel, pp. 459–470.

Snow, C.P., 1959. *The Two Cultures and the Scientific Revolution.* University Press, Cambridge (Eng.).

Stieda, W., 1931. Die Übersiedlung Leonhard Eulers von Berlin nach St. Petersburg. *Berichte über die Verhandlungen der Sächsischen Akademie der Wissenschaften in Leipzig.* Bd. 83, Heft 3. Hirzel, Leipzig.

Struik, D.J., 1968, *A Source Book in Mathematics.* University Press, Princeton.

Szabó, A., 1987. *Geschichte der mechanischen Prinzipien* (3rd ed.). Wissenschaft und Kultur, Band 32. Birkhäuser, Basel.

Thièbault, D.: Mes souvenirs de vingt ans de séjour à Berlin. Paris: Buisson 1803-04.

Thiele, R., 1996. Euler und Maupertuis vor dem Horizont des teleologischen Denkens. Über die Begründung des Prinzips der kleinsten Aktion. In: *Schweizer im Berlin des 18. Jahrhunderts* (Fontius, M., and H. Holzhey, Eds.). Akademie-Verlag, Berlin, pp. 373–390.

Thiele, R., 1997. Das Zerwürfnis Johann Bernoullis mit seinem Bruder Jakob. *Acta historica Leopoldina* 27, 257–276.

Thiele, R., 1999. Ist die Natur sparsam? Betrachtungen zum Prinzip von Maupertuis aus mathematikhistorischer Sicht. In: *Pierre Louis Moreau de Maupertuis. Eine Bilanz nach 300 Jahren.* Berlin-Verlag/Nomos Verlagsgesellschaft, Berlin/Baden-Baden, pp. 437–503.

Thiele, R., 2000. Frühe Variationsrechnung und Funktionsbegriff. In: *Mathesis.* Festschrift zum siebzigsten Geburtstag von Matthias Schramm (R. Thiele, Ed.). GNT-Verlag, Berlin, pp. 128–181.

Thiele, R., 2002. 300 Jahre Brachistochronenproblem. In: *Medium Mathematik* (Löffladt, G., and M. Toepell, Eds.) Franzbecker, Hildesheim, pp. 76–99.

Thiele, R., 2005a. Leonhard Euler. In: *Mathematics and the Divine* (L. Bergmans and T. Koetsier, Eds.). Elsevier, Amsterdam.

Thiele, R., 2005b. *Von der Bernoullischen Brachistochrone zum Kalibrator-Konzept.* Brepols, Turnhout.

Thiele, R., in print. Leonhard Euler. In: [Ueberwegs] *Grundriss der Geschichte der Philosophie des 18. Jahrhunderts.* (H. Holzhey et al., Eds.) Schwabe, Basel, forthcoming.

Truesdell, C.A., 1968. *Essays in the History of Mechanics.* Springer; Berlin.

Truesdell, C.A. 1984. *An Idiot's Fugitive Essays on Science.* Springer, New York.

Weil, A., 1984. *Number Theory. An Approach through History.* Birkhäuser, Basel.

Weyl, H. 1928/1949. *Philosophie der Mathematik und Naturwissenschaft.* Oldenbourg, München 1928; *Philosophy of Mathematics and Natural Science.* Enlarged American edition. University Press, Princeton 1949.

Weyl, H., 1968. Obituary: David Hilbert, 1862-1943, no. 131. In: *Gesammelte Abhandlungen.* J. Springer, Berlin.

Yushkevitch, A.P. 1972. *Istoriya matematika*, vol. 3. Nauka, Moscow.

Acknowledgments

I take the opportunity of expressing my gratitude to those whose support has assisted me. In 1997 I spent some weeks at the Institute for the History and Philosophy of Science and Technology, Victoria College, University of Toronto as guest of Craig Fraser. It is a pleasure to thank him for the kind invitation and to acknowledge his excellent hospitality. Especially I would like to express my thanks to Craig Fraser for many discussions and for his help in improving this text as also did the editors Glen Van Brummelen and Michael Kinyon for this revised reprint. Furthermore, I would like to extend my thanks to my colleague Gerhard Betsch, Tübingen, Germany, and for her technical assistance to Mrs. Ulrike Rau, Leipzig.

I acknowledge the involved libraries and archives for their readiness to support my work and to give permission to reprint illustrations. In detail I thank

Universitätsbibliothek Basel, Handschriftenabteilung und Portraitsammlung,
Berlin-Brandenburgische Akademie, Archiv,
Landes- und Universitätsbibliothek Halle,
Deutsche Akademie der Naturforscher, Leopoldina, Halle, Bibliothek,
Universitätsbibliothek Leipzig,

Bibliothek des Sudhoff-Instituts der Universität Leipzig,
Michael Raith, Riehen.

The final word belongs to the organiser Tom Drucker of the 23rd Annual Meeting of the SCHPM/CSHPM. It would be remiss of me if I did not acknowledge to them the honour to be an invited speaker.

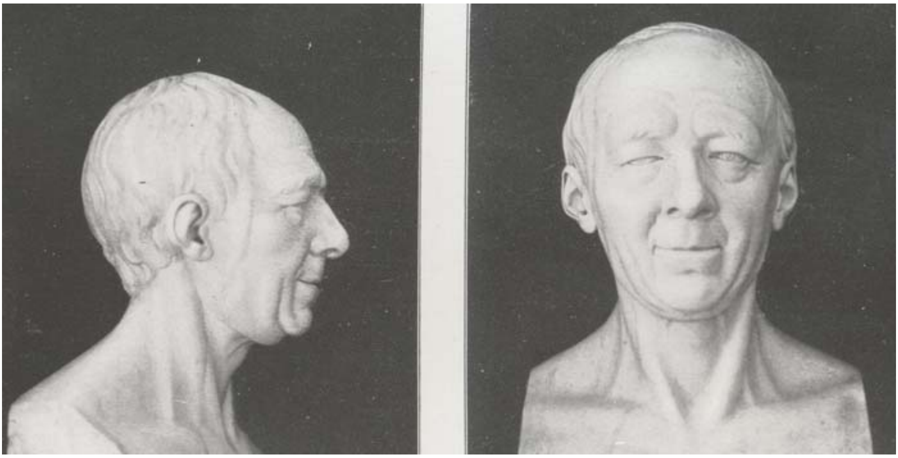


Fig. 5.36. Leonhard Euler. Bust in form of a hermes statue, about the end of 18th century. Courtesy Sudhoff-Institut, Universität Leipzig.

6

Mathematics in Canada before 1945: A Preliminary Survey*

Thomas Archibald[1] and Louis Charbonneau[2]

[1] Department of Mathematics and Statistics, Acadia University
[2] Département de Mathématiques, Université de Québec a Montréal

In 1932, Professor J. C. Fields of the University of Toronto summarized the achievements of Canadian mathematics up to that point:

> progress in mathematics in Canada up to the present has not been all that might have been hoped for, things look more promising for the future. There is a small but increasing group of the younger men who are interested in mathematical research, and some of the later appointments have been encouraging.[3]

For once, this rather bleak assessment is not the product of Canadian understatement. Unlike the situation in the U.S., Fields noted, the Canadian mathematical community was so tiny that virtually all those conducting research were Fellows of the Royal Society of Canada. Since section III of that society was devoted to Mathematics, Chemistry and the Physical Sciences, and since the total membership in 1932 of that section was about 100, we see that we are looking at quite a small group indeed. This situation was to alter in a few years, with the founding of the Canadian Mathematical Congress and the *Canadian Journal of Mathematics*, both called for by Fields in this 1932 retrospective view.

Mathematics has an old history in Canada, however, and like the country itself it represents two cultures based on and evolving from distinct national traditions. French-Canadian mathematics is much older, stemming from the earliest years of the colony of Nouvelle-France. The breaking of contact with metropolitan France following the conquest curtailed French development after a healthy start, however. From that time onward, the parallels in mathematical development within the two cultural contexts are quite striking. In order to have a mathematical culture, the first requirement is a population or regime which recognizes that a portion of the population must have basic

*First published in Peter Fillmore, ed., *Mathematics in Canada*, vol. I, Ottawa, ON: Canadian Mathematical Society, pp. 1–90.

[3]Fields (1932), p. 112

or advanced mathematical skills. These skills are then promoted via elementary or higher-level schools, by teachers in these schools, and by the materials these teachers employ. Hence an essential part of the history of mathematics in Canada is closely bound up with the development of the educational infrastructure. Exactly who is doing the teaching is also of interest, since the single individual can have a very strong influence in a situation where few individuals are involved and when the system is in a nascent state. The content of curriculum at all levels is of course important, as are the origins and the goals of that curriculum. An additional aspect of the local setting is the book trade: publishers and printers played an important role in the diffusion of basic mathematical knowledge.

As with most aspects of Canadian culture, and indeed Canadian nationhood itself, mathematics in Canada manifested itself rather gradually, emerging from French, British and U.S. antecedents between the late eighteenth and the early twentieth centuries. The phases of this emergence into a mathematical maturity are quite similar to what we see elsewhere in the Americas, chronology apart: an initial phase when all necessary mathematical skill was imported, or the property of first-generation immigrants educated elsewhere; a recognition of a requirement for some mathematical training in an elite of bankers, accountants, merchants, surveyors, navigators, etc., and the corresponding development of an infrastructure for providing such training using materials and teachers from abroad; the broadening of this infrastructure to include a wider portion of the population, the development of secondary and post-secondary education, the local training of teachers, and the creation of texts which suited local conditions and educational objectives; and finally the gradual development of a fully articulated mathematical community, engaged in teaching and research at all levels, publishing, and integrated into an international mathematical community. These phases can and do overlap in the Canadian context, as in most places.

In this paper we will attempt to survey briefly these developments. The authors apologize for the fact that many individual stories of interest could not be told in a short space. In particular, materials on the more recent history, while abundant, were in most cases difficult to access, and the result is a rather unbalanced account which will as usual favour central Canada. A little less usually, we will spend more space on the earlier portion of the history.

Little has been written about these events. Apart from Charbonneau's article in the *Canadian Encyclopedia*, and Yves Gingras' *Les origines de la recherche scientifique au Canada: le cas des physiciens*, which has some points of contact with our story, most of the information presented here comes from archival research, biographical compilations, memorial volumes, and institutional histories of Departments of Education and universities. We are fortunate to have Karpinski's *Bibliography of Mathematics in the Americas through 1850*, which, though incomplete, gives a helpful picture of the situation in Canada in the first half of the century. In addition, R.S. Harris's *History of Higher Education in Canada, 1663–1960* provides a good deal of important background information. In general, French Québec has been better served by historians than English Canada, and the following account reflects this.

Nouvelle-France (1635–1760)

Under the French regime the economy of the French colony in America depended to a great extent on navigation of the Saint Lawrence river, and on exploration of new lands. The population of the colony grew slowly in the seventeenth century: around 11,000 in 1685, 18,500 in 1713, and 55,000 in 1754. In 1754 Quebec City had only about 8000 people.[4] In this setting, mathematical requirements were restricted to a few practical applications, notably in surveying, cartography, and navigation. However, the increase in the population, especially in Quebec City, led the Jesuits to begin offering a complete classical course at the Collège de Québec beginning in 1659. The old system had required five years, to which they now added two more, devoted to "philosophy."[5] In keeping with the curriculum in French Jesuit Colleges of the period (which not so long before had produced Descartes) mathematics teaching was concentrated in these last two years.

The royal chair of mathematics and hydrography (1660-1760)[6] Problems involving property and the fixing of land boundaries are common in any sedentary society, and Nouvelle-France was no exception. Champlain (1567-1635), on his return to Québec in 1632, declared himself "engineer in chief of the colony," and concerned himself with such issues. At Champlain's death, Jean Bourbon, an engineer in the Compagnie des Cents-Associés, took on the job and served as judge of land boundary questions until his own death in 1668. In 1674, the Sovereign Council of Nouvelle-France required that all surveying instruments be approved by Martin Boutet, Sieur de Saint-Martin (1616-1683), who was at the time the professor of mathematics at the Collège de Québec.

Jean Talon, Talon, Jean the *intendant* or chief administrator of the colony from 1665 to 1681, grasped the fact that the economic future of the colony would require a better knowledge of the country's geography, as well as depending on pilots that were competent to navigate on the Saint Lawrence. Since 1661, Boutet had given mathematics courses at the Collège which were oriented towards surveying and navigation. Soon after taking up his position, Talon requested that these courses be extended to include the training of pilots. The need was urgent: the census of the same year, 1666, shows only 22 "marins" (sailors) in the entire colony. Besides the chronic shortage of navigators, there was also a need for accurate maps. Once again the natural choice to provide such training was Boutet. In 1671, Talon named him Professor of

[4] Kerr (1966), p. 24.

[5] Audet (1971), t. 1, p. 174.

[6] The information in this section comes from Audet (1971), t. 1, pp. 192–202, and from Chartrand, Duchesne, Gingras (1987), pp. 20–34.

Hydrography. Despite repeated efforts by the colonial administration, it was only in 1678 that he was awarded the official title — and the corresponding salary of royal engineer, with a mandate to "teach hydrography, piloting, and other parts of mathematics." This appointment occurred in a flood of creation of such royal chairs of mathematics and hydrography in France in the 1670s.

Even though Boutet's teaching was undertaken at the express request of Talon, no remuneration accompanied it until 1678. This, it turns out, was symptomatic of future events. Until the end of the seventeenth century, the colony had difficulty ensuring continuity in the teaching of practical mathematics applied to surveying and hydrography. Jean-Batiste Louis Franquelin (1652-1718) received the title of "Hydrographe du roi à Québec" in 1687, three years after the death of Martin Boutet. Prior to this appointment, he had given hydrography courses privately and had prepared a map of New France. After getting the appointment, he spent a number of years in France, leaving the hydrography courses without a teacher. In 1697, Louis Jolliet (1565-1700) took up the post. Jolliet, famous for his explorations, had trained with the Jesuits in Québec, and thus became the first Canadian-born Royal Hydrographer. Unfortunately he died only three years after taking up the post. A renewed search led to the appointment of Jean Deshayes (d. 1706), a French astronomer and cartographer, in 1703. Deshayes had visited Québec in 1685, during which time he had observed an eclipse of the moon which allowed a determination of the longitude of Québec by the French astronomer Jean-Dominique Cassini.[7] At his death, Deshayes left what was probably the first scientific library in New France: it contained about fifteen volumes, including the Marquis de l'Hôpital's 1696 *Analyse des infiniment petits pour l'intelligence des lignes courbes*.

These recurring difficulties in staffing the post of hydrographer led the authorities to petition the crown to hand the chair over to the Jesuits in perpetuity. The request was a reasonable one: since 1700, the Jesuits had given a hydrography course at the Collège de Québec; they may have done so as well at Montréal. The request was successful, and from 1708 until 1759 the king granted a Chair in Hydrography to the Jesuits at Québec. The priests who held the chair had an average tenure of five years in the period up to 1741. During the winter the students resided in Québec, where they took courses in geometry, trigonometry, and physics as well as courses in naval theory. In the summers, they apprenticed as pilots under the direction of the second captain of the port. The best-known of the occupants of the chair is Father Joseph-Pierre de Bonnécamps (1707-1790) , who taught hydrography from 1741 until the fall of Québec in 1759. In addition to his teaching duties, Bonnécamps took part in expeditions as cartographer, keeping up a correspondence with astronomers and men of science in the mother country. Thus he published a memoir on the aurora borealis in the *Mémoires de Trévoux*, a scientific periodical published by the Jesuits. Another memoir, presented to

[7] Chartrand, Duchesne, Gingras (1987), pp. 26–27.

the Paris Academy of Sciences by the astronomer Jean-Nicolas Delisle (1688-1768), reported calculations of the longitude of Québec made by Bonnécamps and de Lotbinière. The cumulative precision of observations resulted in an improvement over the earlier calculations of Deshayes and Cassini.[8]

In the final years of the French regime, a distinguished scientist lived in Québec as an officer under Montcalm, namely Louis-Antoine de Bougainville (1729-1811). When he arrived in Canada in 1756, he had just completed the second part of his *Traité du calcul intégral pour servir de suite a l'analyse des infiniments-petits de M. le marquis de l'Hôpital*, the first part of which had appeared in 1754. His relations with Father Bonnécamps seem to have been good, since he wrote a letter of recommendation for the priest on the latter's return to France in 1757.[9]

The Collège de Québec[10] Founded in 1635 by the Jesuits, the Collège de Québec was the main seat of intellectual training in the colony for the entire French regime, its importance increasing with the growth of the colony. In 1651, the entire responsibility for teaching a course of study with a normal duration of five years rested on two of the ten priests, together with six brothers residing at the college who assisted. One of the two priests taught mathematics, doubtless mostly elementary and commercial mathematics. At that time there were only a dozen students, including the "petite école" where the preparatory teaching was done. As mentioned earlier, the classes of philosophy (two more years) were added in 1659, and mathematics was placed in one of these years, as was the custom in France. Both Talon and the first bishop of Québec, Mgr François de Montmorency Laval, concerned themselves with the curriculum.[11] During the second half of the eighteenth century, in French colleges, the mathematics program was the following:[12]

[8]The longitude of Québec is $73°33'$, while Deshayes and Cassim obtained $72°13'$ and Bonnécamps and de Lotbinière, $72°30'$. Chartrand, Duchesne, Gingras (1987), p. 34.

[9]For more about Bougainvile and his time in America, see Struik, Dirk, J. (1956). Among the French engaged in mathematical activity in Nouvelle-France, we may also mention Joseph Bernard Chabert who, for geodesic purposes, made astronomical observations along the Atlantic coast near Louisbourg in 1750-1751. See Struik, Dirk J. (1976), p. 102.

[10]Unless otherwise mentioned, the information in this section is from Audet.

[11]In most of the French colleges, the philosophy course was spread out over three years. Audet (1971), t. 1, pp. 172–189. Note that the 1775 program of the philosophy class of the Séminaire de Québec corresponds to the French program except that "mathématiques mixtes" was reduced to practical geometry with some commercial mathematics.

[12]Dainville, Francois de, (1964), p. 52.

Pure mathematics:
- arithmetic, algebra, geometry, plane trigonometry

"Mixed" mathematics:
- practical geometry: measurement of length, area and volume
- mechanics: the science of forces and the actions of bodies
- hydrostatics
- spherical astronomy
- use of the gnomon (*i.e.* solution of triangles in practical circumstances)
- optics: perspective, mirrors and lenses
- fortification
- pyrotechnics (sometimes)

This is an impressive-sounding program. However, it should be remembered that mathematics at this time was taught in connection with only one of the four parts of philosophy (logic, metaphysics, ethics and physics), namely physics, and that physics was done first.[13] The instruction had virtually nothing in common with how mathematics is taught now: there were no exercises, and practical applications were limited to those things which were dictated in the notes. In France, students entering the philosophy class had often received no mathematical training at all. The Abbé Sauri, author of an introductory textbook called *Institutions mathématiques*, noted as much in his introduction:[14]

> I would advise messieurs the philosophy teachers to teach my institutions at the beginning of the course, or at least, to teach arithmetic, the first four rules of algebra, and the notions of geometry contained in No. 1, p. 141 to No. 7, p. 147 inclusive. This will place their students in a better position to understand logic and metaphysics. In logic itself one often talks of triangles, circles, etc. How can we expect that young people who have no acquaintance with these figures will understand any of the professor's explanations?

There is no reason to think the situation would be otherwise in Québec.

Still, we should distinguish between the Collège courses and those given in connection with the Chair of mathematics and hydrography. Although the history of the Chair is closely associated with the college, it seems that the professors of hydrography did not simply teach the same material as was given in the college philosophy courses.[15] It should also be noted that, on this side of the Atlantic anyway, the classical course given by the college led to only one end: the priesthood. The liberal professions were, so to speak,

[13] Chartrand, Duchesne, Gingras (1987), pp. 34–35.

[14] Sauri (1786), p. xvii. Even if the book appeared after 1760, teaching in the French colleges did not alter significantly between 1750 and 1780.

[15] So it seems from the short list of mathematics professors at the college given by Audet. Audet (1971), t. 1, p. 185.

closed to Canadians. There were no lawyers in the colony, and notaries needed only a minimal training. Physicians came almost uniquely from France. This could not help but push the teaching in a direction that was not particularly favourable to independent teaching in the sciences.

By the time of the defeat of Montcalm, Bougainville, Bonnécamps, and their associates had left the colony, and the college closed.

English Canada from 1760 to the Union: Creating an Infrastructure

Reflecting their later start as colonists, English mathematical production and education start later than in the French colony. The only 18th century English work mentioned by Karpinski is a ready reckoner printed at Quebec in 1790 by Major Williams. It ran to a second edition, is described in the *Quebec Herald* of 1789-91, and appears lost. It deals with rates of exchange, an important subject since local currencies were in pounds, shillings and pence, and exchange rates varied.

A second early paramathematical book deserves mention as showing the state of affairs in the early nineteenth century. This is the 1822 work by Arthur Fessenden, titled *Tables, showing the interest at six per cent of any sum from 1 pound to 1000 pounds, from one day to one hundred days, and from one month to twelve months*. Fessenden was an accountant for the Bank of Canada, then at Montreal. His tables were published in Montreal by Nahum Mower, a newspaper printer/publisher, and later editions appeared in Montreal in 1830 and Halifax in 1832. Subsequent editions of 1837, 1841, and 1847 were extended to 365 days. Of course several things appear odd to us about this. Only one rate is tabulated, though the existence of reprints suggest that this was "the going rate" for the entire 25-year period. The tabulation of many principal amounts shows that while addition was a staple of the counting-house, multiplication was not. The work bears a recommendation from the President and Cashier of the Bank of Canada, and from the President of the Montreal Bank.

In the early years of the nineteenth century the main emphasis was on elementary education (whether for children or not), and the first years of mathematical study were devoted almost exclusively to arithmetic. The earliest English Canadian arithmetic book we have identified is the 1809 treatise of John Strachan, later Bishop of York, which he wrote in order to have an appropriate book for his classes at the Cornwall Grammar School. The work was titled *Concise introduction to practical arithmetic: for the use of schools*, and published in Montreal. Strachan was educated at King's College, Aberdeen and at St. Andrews; he had a strong interest in science, and the Scotland of his day offered in general a better scientific education than was available in England. Strachan was instrumental in the founding of both McGill and King's College (the earliest component of the University of Toronto), and was an ardent advocate of proper grammar schools (i.e. secondary schools). He was also the first Inspector of Schools in Upper Canada, and an ardent proponent of the standardized education that would require the creation of such a post. Strachan was also important in setting the stage for a domestic production of textbooks through his ardent opposition to U.S. "democratic" influences, an opposition widely shared in loyalist British North America.

Both before and after Strachan, British and American texts were often used by Canadian teachers, a pattern which continued until the 1850s. An example of a widely approved English book is Francis Walkingame's *The tutor's assistant, being a compendium of arithmetic and complete question-book.* A Canadian edition of the work appeared in Montreal in 1818, though the English original appeared first in 1751, and passed through an enormous number of editions over the next century. In fact, the *Dictionary of National Biography* states the following:

> A so-called 71st edition appeared in 1831 ... Except for the section dealing with the rule of three, which needed improvement, the work remained little altered down to 1854.[16]

Little wonder then that Karpinski lists nine Canadian editions prior to 1850, printed at places such as Toronto, Picton, and St. John. A quick look at the book, however, suggests why it would be of limited use for Strachan and his contemporaries. It is essentially a book directed at adults, covering a great deal of ground (e.g. cube root, single-entry bookkeeping, and basic algebra) in a short span. As the title suggests, it appears to have been intended as an aid for teachers, structuring the curriculum and providing worked examples. Answer keys were published soon after its original appearance, and may well have existed for the Canadian editions as well.

An American competitor for such British arithmetics was provided by Daniel Adams (1773-1864), a New England physician and educator, who published *The Scholar's Arithmetic* in 1801 expressly to provide a suitable school text.[17] The work ran to many editions, and was revised in 1827 as *Adams' New Arithmetic*, under which title it appeared in a Canadian version in 1833. The title, given the American origin of the book, is a bit misleading: *Adams' new arithmetic, suited to Halifax currency, in which the principles of operating by numbers are analytically explained and synthetically applied; thus combining the advantages of the inductive and synthetic mode of instructing. The whole made familiar by a great variety of useful and interesting examples, calculated at once to engage the pupil in the study, and to give him a full knowledge of figures in their application to all the practical purposes of life. Designed for*

[16]DNB, v. 20, p. 548. It is described as the "most popular arithmetic both in England and America down to the time of Colenso" that is, to the late nineteenth century.

[17]Adams also published works on grammar, oratory, and geography. See the *Dictionary of American Biography*, 1, pp. 54–55.

use in the schools and academies in the British provinces. The first Canadian edition was printed at Stanstead, in the Eastern Townships, with or without the approval of the author. A second edition was published at Sherbrooke in 1849. We may note the methods of analysis and synthesis being touted here as pedagogical strategies. It seems likely that this reflects some awareness of contemporary British debates on analysis versus synthesis as appropriate means of procedure in learning higher mathematics, but we have no other evidence of this at present.

In 1832, hence shortly before the publication of Adams' book, appeared a second work written specifically for Canadian schools, with William Phillips as author. Phillips, apparently based in York, is described as a teacher "in Ladies' Schools", as well as a private tutor. In abbreviated form, the title is: *A new and concise system of arithmetic, calculated to facilitate the improvement of youth in Upper Canada.* The work was published by subscription under the patronage of Sir John Colborne, then Lt. Governor of the Province of Upper Canada. Other listed subscribers included Strachan, by then the Archdeacon of York, and Dr. Harris, Principal of the newly-founded Upper Canada College.

This list of patrons sheds light on some historical issues that require a bit of background to be completely understood, and which bring us back to John Strachan. Strachan was a very political creature, one who saw his teaching as a means of extending his personal influence; in 1817 he had declared with satisfaction that "all my pupils [are] now the leading characters in many parts of the province."[18] Strachan had long argued for the establishment of a University in Canada on the "Scottish or German" model. He viewed this as a Christian institution, though Methodists and other dissenting sects should be excluded from governance of the institution, which should be firmly in the hands of the established church (i.e. the Church of England). This stance fit in well with his generally Tory outlook, one which had enabled him to successfully cultivate Lord Maitland, Colborne's predecessor. With Colborne's accession to the Lieutenant-Governor's post, Strachan's influence waned, as did his efforts to establish the newly chartered King's College. Colborne felt that Grammar schools of high quality were essential, so that the children of the appropriate classes in the colonies could return to England for university education when desirable. Upper Canada College was founded on Colborne's initiative, and somewhat against Strachan's wishes, as such a preparatory school.[19]

---

[18] Craig (1986) p. 755.

[19] Similar doctrinal battles formed the background for the establishment of institutions of higher education in Nova Scotia where King's College (Windsor) was founded in 1789 and was restricted to members of the Church of England, then about 20% of the population. This eventually led to the founding of Pictou Academy by Thomas McCulloch. The strength of mathematics and science there was one of McCulloch's key arguments in attempting to obtain provincial funding. Eventually McCulloch became the first president of Dalhousie (1838).

The Methodists, disinclined to be excluded by Strachan from control of higher education, had meanwhile been moved to action, establishing Upper Canada Academy at Cobourg, Ontario. Though this school originally experienced financial difficulty, Egerton Ryerson, a Canadian-born Methodist, worked to obtain a royal charter for this school, which eventually became Victoria College in Toronto. Ryerson was the College's first principal, inducted in 1841. Ryerson became Superintendent of Schools for Canada West in 1844, an appointment which had eventual repercussions for the world of elementary mathematics in Canada. He was also the major architect of the system of Normal Schools (or education colleges) in Upper Canada.

Despite his "democratic" leanings, Ryerson was a staunch believer in the importance of a firm adherence to the British Empire and of a close association between the colonies. This contrasts somewhat with the position of Strachan, who in his later years became an advocate of Upper Canadian colonial autonomy. The imperial focus is reflected in a contemporary arithmetic by G. and J. Gouinlock. *A complete system of practical arithmetic, for the use of schools in British America, to which are added, a set of book-keeping by single entry, and a practical illustration of mental arithmetic, federal money, receipts, bills of exchange, inland and foreign, explanation of commercial terms, etc. The whole adapted to the business of real life, to the circumstances of the Country, and to the present improved state of commerce* was printed in Hamilton in 1842. The Gouinlocks describe themselves on the title page as "formerly British teachers of long experience and extensive practice". In addition to standard operations with whole numbers, fractions and decimals, it includes proportion, simple and compound interest, and the following advanced topics apart from those mentioned in the title: British exchange of moneys, with a number of countries (incl. Europe, W. Indies, E. Indies and Canton in China); Alligation, Involution, Evolution, Square root and Cube root, duodecimal multiplication, tonnage of ships, and permutation.[20]

French Canada as an English Colony, from 1760 to the Union

The first arithmetic books. The period from the conquest to the Union of Upper and Lower Canada, in 1840, is characterized by a desire on the part of the authorities to establish a system of primary public education. This idea of training a large number of people to a minimal level entailed certain changes in both curriculum and pedagogy. As in English Canada, arithmetic was emphasized.

The publication of the first arithmetic books in Quebec should be seen against this shifting backdrop. The first such work, written by Jean-Antoine Bouthillier (1782-1835), appeared in Quebec in 1809 and was titled *Traité*

[20] Karpinski, p. 438

d'arithmétique pour l'usage des écoles. Bouthillier had studied at the Collège Saint-Raphaël in Montreal before apprenticing as a surveyor.[21] He never taught, though he worked at a variety of professions, among them journalist, translator, inspector of highways, and justice of the peace.[22] His book dealt with the elementary operations on whole numbers and fractions, as well as rules useful for merchants (rules of three, rule of false position, simple and compound interest, and exchange). Despite being intended for schools, its success was limited at first. This was perhaps in part because it was oriented toward the memorization of rules, rather than comprehension. One might also ask whether, at the time of its appearance, the number of schools with instruction going beyond the four elementary operations justified the scope of the work. However, after 1830 the use of this book became more widespread, and it appeared in numerous editions until 1864.

Another student from the same school in Montreal was Michel Bibaud, who published a manual of arithmetic in 1816. This work was titled: *L'arithmétique en quatre parties, savoir: l'arithmétique vulgaire, l'arithmétique marchande, l'arithmétique scientifique, l'arithmétique curieuse, suivie d'un précis sur la tenue des livres de comptes, principalement pour ceux qui veulent apprendre l'Arithmétique d'eux-mêmes et sans Maître, ou s'y perfectionner*. As the title indicates, this was intended not for the schools but for an audience of autodidacts. Nevertheless it was used in many schools. The "four parts" of the title are: common arithmetic (the four elementary operations on numbers and fractions, calculation of areas and volumes); commercial arithmetic (rule of three, exchange); scientific arithmetic (decimals, powers, roots, proportions and logarithms) and recreational arithmetic (riddles, games, puzzles, etc.). Apart from the recreational portion, Bibaud's content corresponds to that in Bouthillier's book. Bibaud's work is in fact a compilation, and he indicates his sources. The common and scientific arithmetic were inspired by the 1786 text of Abbé Sauri, which Bibaud had doubtless used as a student at Collège St. Raphaël.[23] The commercial arithmetic comes from Walkingame, a popular English text of the time. Finally, the recreational material was based on a book by a M. Despiau called *Choix d'amusements physiques mathématiques*.[24] Bibaud was to rework the contents of his book, republishing it in 1832 under the title *L'Arithmétique a l'usage des écoles élémentaires du Bas-Canada* without the recreational portion.[25]

---

[21] This became the Collège de Montréal in 1806.

[22] See Lavoie, Paul (1994), chap. 5, section 5.5.

[23] Sauri (1786). One finds in the Archives du Collège de Montréal a *Compendium des institutions mathématiques de l'Abbé Sauri* copied at Québec the 7 of August 1785. A copy of the fourth edition (1786) of the Sauri book belonged to the Collège de Montréal. It is now located at the Bibliothèque Nationale in Montréal.

[24] This book seems to have been published in London in 1800, with a translation in 1801 also at London. See Lavoie, Paul (1994), p. 276, note 1.

[25] Lavoie, Paul (1994), pp. 274–282.

A third book deserves mention, written by Casimir Ladreyt (1797-1877) and published in 1836. As with many books of the period, its title tells us a great deal: *Nouvelle arithmétique raisonnée ou cours complet de calcul théorique et pratique, a l'usage des collèges et des maisons d'éducation de l'un et de l'autre sexe, des personnes qui veulent apprendre cette science en peu de temps et sans le secours d'un maître, et de celles qui veulent se livrer au commerce; suivi de quelques leçons sur la plannimétrie et la stéréométrie (arpentage et cubage), ou Toisé des surfaces et des volumes.* We know little about the author beyond what is on the title page: a "former French trader, now a teacher". Though Ladreyt covers much the same ground as Bibaud and Bouthillier, he distinguishes himself by his pedagogical care. His page layout is more adapted to the logical hierarchy of the ideas, and reasoning is important because of its importance in training the faculty of judgement. For this reason he does not cite the rule of three as such, preferring that the reader reason out the problem rather than blindly applying a rule. Despite these qualities, or perhaps because of them, Ladreyt's book had a rather limited success.[26]

Finally we note a book which had a limited distribution in the region near Québec, Joseph Laurin's *Traité d'arithmétique: contenant une claire et familière explication de ses principes: et suivi d'un traité d'Algèbre*, which corresponded closely in its content to the books of Bouthillier and Bibaud.

The popularity of the books of Bouthillier and Bibaud around 1840 and even over the next two decades indicates that much mathematics teaching was individually based. This state of affairs was to evolve following the arrival in 1837 of the Frères des Écoles chrétiennes (FÉC), to which we return below.

Secondary teaching: the classical Collèges. In 1757, the state of war between France and England had led to the closing of the Collège de Québec, which subsequently remained closed.[27] England did not permit the recruitment of Jesuits any longer, and besides many returned to France. Those who stayed tried to take up teaching again, but with little success. In 1765, on the order of Mgr. Jean-Olivier Briand, the Séminaire de Québec took up the task.[28] The Séminaire organized its teaching directly in the tradition of the Collège de Québec; from 1765 to 1770, since there were no priests, the course was limited to the classes of Letters and Humanities concentrating on Latin, French, English and Greek.[29] In 1770, the philosophy classes began again, though there was no mathematics at all for another three years. In Montréal, the Collège

[26] Lavoie, Paul (1994), pp. 282–292.

[27] Galarneau, Claude (1978), p. 16.

[28] Founded in 1668, the Séminaire de Québec had never provided secondary education in a continuous way prior to the arrival of the English. The seminarians attended courses given by the Jesuits at the Collège de Québec. At times when they were dissatisfied with the Jesuit courses (as in 1732) they organized their own philosophy courses, but this was never done on a regular basis. Audet, Louis-Philippe (1971), t. 1, p. 373.

[29] The Letters classes, often called Humanities, consist of the first six years of the course.

Saint-Raphaël, already mentioned in connection with its illustrious students Bouthillier and Bibaud, began to offer courses in the heart of the city in 1773. This began by offering only humanities (the introductory years of the course), with those interested in completing the philosophy course being obliged to go to Québec. Not until 1790, following pressure from parishioners, did the Bishop of Québec appoint a philosophy professor. In fact, the parishioners also argued for a professor who could teach arithmetic and mathematics as well as a course in writing, complaining that such instruction was available only in the Protestant schools.[30] Such courses began in 1791.[31] We may assume that they revolved around commercial arithmetic, in accordance with the changing economic nature of Montréal at the time.

The content of the mathematics courses in the philosophy classes of the Séminaire de Québec are known to us in part because of two presentations of "theses" in mathematics, one in 1775 by the "physics students of Mr. Thomas Bédard Diacre", the other between 1786 and 1790 by the "students of Mr. Edmund Burke, priest".[32] These are thesis defences in the medieval sense: public academic exercises during which physics students of the philosophy class discussed a few mathematical propositions. On these occasions, the Séminaire printed a pamphlet which summarized the propositions debated. The pamphlet of 1775 has nine pages, and we find there material from elementary arithmetic, algebra, and the calculation of proportions. This is followed by the solution of equations in one to four unknowns, theorems on arithmetic and geometric progressions, and quadratic equations. There are also propositions of elementary geometry, practical geometry, and trigonometry. The 1790 pamphlet consists of ten pages of statements of theorems, and goes well beyond the earlier work in including conic sections, spherical trigonometry, eleven propositions on differential and integral calculus, and a large number on various aspects of mechanics. The content of the 1790 pamphlet corresponds much more to the program in French colleges of the time than did that of 1775. This evolution may well have begun with the arrival in Québec in 1775 of M. J.-B. Lahaille, a French Jesuit from Bordeaux, who succeeded Thomas Bédard for one year at that time. It appears to have continued in the hands of Charles Chauveaux, who taught physics and mathematics from 1776 to 1786, and was one of the students who participated in the "defence" of 1775.[33] His course notes contain many of the subjects treated in the 1790 work, though several of the mechanical topics, as well as conic sections, are missing.

Subsequently the purely mathematical content of the philosophy classes ceased to evolve. Nevertheless, there was a change in the attitude to science.

[30] Charbonneau, Louis, 1984, p. 43.

[31] Galarneau, Claude, 1978, p. 18.

[32] Bédard, Thomas (1775), Burke (n.d.).

[33] These notebooks are held at the Archives of the Collège de Montréal, of the Séminaire de Saint-Hyacinthe, and of the Séminaire de Québec. Galarneau, Claude (1977), pp. 86–87.

This change may be seen in the lesson plans of the philosophy class and even in the names given to each of the two years.[34] In 1790, the first year was called "Logic", and logic, metaphysics and morals were covered. The second year was called "Physics", and covered physics and mathematics. This division corresponds faithfully to that in the French Collèges and doubtless to that in the Collège de Québec. In 1816, the earlier counsel of the Abbé Sauri was at last heeded and we see in the course of study the addition of "a part of mathematics" to the program in the Logic year. Finally, from 1838, the name of the first year is no longer Logic, but Mathematics. The program of study consisted of: algebra, geometry, differential and integral calculus, and conic sections. The second year was still called physics, and covered physics and chemistry. Despite appearances, logic, metaphysics and ethics were still taught, but clearly mathematics and science were considered of greater importance than at the beginning of the century. It is worth noting that in England at this time differential and integral calculus were part of the curriculum at Cambridge, but for the most part not elsewhere, and even Cambridge had provided them only for about twenty years. Hence it is not surprising that we do not find courses of corresponding sophistication in English Canada at this time.

Prior to 1835, it was also the case that the philosophy courses alternated: that is, there was one professor, who taught the two courses of Logic and Physics in alternate years. Thus the students did not necessarily follow the course in order. From 1835, because of the growth of the student population, three teachers divided the task by discipline. Abbé Jérôme Demers taught "intellectual" philosophy (logic, metaphysics and morals), while mathematics was taught by Abbé Normandin, and physics by Abbé J. L. Casault. At about the same time, we see mathematics enter the lower school, under the impulsion of Demers and his colleague Abbé John Holmes, a transplanted American convert. Thus 8^e and 7^e (the preparatory years) contained arithmetic; fractions and decimals in 6^e and 5^e; bookkeeping and the metric system in 4^e; and then algebra, and elementary geometry in 3^e, 2^e, and 1^{re}. The other colleges and seminaries of Lower Canada followed suit, though there were delays in implementing the program depending on the college.[35]

What provoked this exodus of mathematics from the philosophy classes? Certainly we may mention the pressing needs for men outfitted with commercial mathematics. We already mentioned pressures of this kind originating from the parishioners of Notre-Dame de Montréal at the end of the eighteenth century. Such demands were to be repeated throughout the nineteenth century, the more so because the majority of students left the classical Collèges without even beginning the two years of philosophy. There is more to it than this, however: the attitude of the Church had been ambivalent until about 1840. Under the influence of Abbé Demers, however, the Séminaire de Québec

[34] Charbonneau, Louis, 1984, p. 43.
[35] Lamonde, Yvan (1980) p. 76.

and other Collèges adopted a position that is related to that enunciated by the Jesuits in their *Ratio Studiorum* of 1832:

> The times require us to give more importance than formerly to physical sciences and mathematics .... For the fact that these sciences have been abused to oppose our holy religion is not a reason to abandon them, but on the contrary a reason that our people should devote themselves to them with all the more ardour, in order to seize the weapons of the enemy and to employ in defence of the truth the means which they abused to attack it.[36]

Given these developments it seems likely that after the mid-1830s students entering the philosophy classes were better prepared to undertake study in intermediate-level mathematics. One may wonder, however, if in fact the teaching was any different from that which had gone before. Was the training of mathematics teachers who taught in the philosophy classes of a kind to improve this teaching? Many of these teachers had traditionally been quite young. We have already mentioned Thomas Bédard and his student Charles Chauveaux, the latter having begun teaching the year after completing his philosophy course. It was traditional in the Collèges for brilliant young seminarians to give courses immediately after completing their own studies. Such a tradition was hardly propitious for an improvement of teaching, since the young professor would simply repeat courses which he had heard himself a few months before.[37] Nonetheless, there were notable exceptions. Burke, for example, who succeeded Chauveaux, had received a solid education in Paris.[38] Unfortunately, his career at the Séminaire ended in 1790. Abbé Houdet, a Sulpician father who immigrated to Canada to escape the anticlerical laws of the French revolution, was responsible for science and philosophy at the Collège de Montréal between 1798 and 1826.[39] The most remarkable of the philosophy professors of the first half of the nineteenth century is undoubtedly Jérôme Demers, who taught both philosophy classes at Québec from 1800 to 1835, then restricted himself to intellectual philosophy from 1835 to 1849. His influence was felt in all the Collèges of the province. It is interesting to note that Abbé Demers had worked for some time as a surveyor between the end of his classical studies and entering the seminary in 1795. His course notes on mathematics show no special originality, but those in physics indicate that he was up to date on the latest discoveries, especially in electricity and magnetism.[40] Demers also encouraged higher study elsewhere, and was involved in

---

[36] Cited in Simard, G., *Tradition et Évolution dans l'enseignement classique*, Ottawa, 1923, p. 10.

[37] Martineau, Armand (1967), pp. 215–216.

[38] Galarneau, Claude (1977), p. 87.

[39] Galarneau, Claude (1977), p. 90–93.

[40] Many student notebooks are held at the Archives of the Collège de Montréal, the majority dating from 1811.

sending the Abbés Isaac and François Desaulniers to Georgetown University, a Jesuit institution in Washington, where they obtained M.A.'s in science. On their return, they too became philosophy professors, Isaac at St.-Hyacinthe and Francois at Nicolet.[41] Despite these examples, young, inexperienced teachers without additional training (such as Michel Racine at Québec) continued to teach mathematics even after 1835, and the overall quality of these courses showed little improvement.[42]

English Canada from the Act of Union to Confederation

Following the Act of Union, the educational system in the newly-established Canadas was substantially reorganized. Indeed, following a general movement in Britain, between 1842 and about 1853 all the British North American colonies restructured education to establish common schools available to all boys, to encourage teacher training, and to develop uniform textbooks. In the Canadas, the standardization of texts took the form of a list of approved books which schools could select, with subsidies being withheld if approved books were not chosen. This led to a substantial market for publishers to tap; there were 2500 elementary schools in Canada West in 1844.

The most successful of these publishers was John Lovell (1810–1893) of Montreal. Originally a printer and newspaper publisher, Lovell had turned to literary efforts in the 1850s; from the late 1850s he produced a highly successful school series of which arithmetic was an important part. His principal arithmetic author was John Herbert Sangster (1831–1904), a London-born teacher who had immigrated to Canada at an early age, and who had studied at Upper Canada College and at Victoria University, Cobourg (M.A. 1861, and M.D. 1864). Shortly after obtaining the M.D. he became headmaster of the Normal School at Toronto which Ryerson had founded. In addition to his arithmetic works, Sangster wrote a treatise on natural philosophy, one on chemistry, and an introduction to algebra.[43]

Lovell's virtual monopoly began to be broken after Confederation, presumably in part because of expansion of the market: in 1871 the Schools Act in Ontario eliminated tuition for elementary schools and instituted a preliminary form of compulsory attendance. The main publishers to come onto the scene

[41] Science teaching had clearly become important to the directors of the Séminaire de Nicolet a decade before the Desaulniers had been sent to the U. S. For example, in 1824 and 1829, they had published a work called *Nouveau Traité abrégé de la Sphere d'aprés le systéme de copernic, par demandes et par réponses* for the students of the seminary. It seems likely that this was a reprint of a French textbook. See Lessard, C. (1980), p.265. Icon Lortie (1955, p. 39) attributes this 24 page work to Isaac Desaulniers which seems implausible given that he was only 13 years old at the time of the first printing.

[42] Provost, Honorius (1959), p. 667.

[43] MacMillan's *Dictionary of Canadian Biography* (1963), p. 668.

at this time in Toronto were W. Gage, Copp Clark, and Rose (later Hunter Rose). English models (and English authors) continued to be of importance after Confederation, however, though the phrase "for Canadian schools" was usually tacked on to an adaptation of a British work. Key among the English authors were Barnard Smith (1810-1876), a fellow of Peterhouse (Cambridge) and later rector of Glaston and Rutland; and his cousin J. Hamblin Smith (1829–1901). These two men authored a large number of school textbooks at a variety of levels. Barnard Smith's books were adapted in many editions by Archibald MacMurchy for Copp Clark; James Hamblin Smith's arithmetic was likewise adapted, by Thomas Kirkland and William Scott, and appeared in a number of editions published by Adam Miller and later by his successor William Gage.

MacMurchy (1832-1912) came to Canada in 1840, and after early education at Rockwood Academy graduated B.A. (1861) and M.A. (1868) from the University of Toronto. From 1858 he was the mathematical master at the Toronto Grammar School (later Jarvis Collegiate), becoming rector in 1872 until his retirement in 1900.[44] In addition to his role as an adaptor, he was editor of the *Canadian Educational Monthly* for many years; we shall have occasion to refer to some of his reviews below.

A further textbook author of the 1870s and after is James Alexander McLellan (1832-1907). Besides his contributions as an author of books on arithmetic and algebra, McLellan is significant in displaying some of the attributes now associated with professional educators. In particular, he wrote on educational psychology from 1889 onwards, collaborating with the U.S. philosopher and psychologist Thomas Dewey on *The Psychology of Number* (1903).[45]

The Beginnings of Higher Mathematics

Sangster, MacMurchy, and others mark a point around the time of Confederation by which Canadian-educated writers were producing materials for elementary education in the Canadian context. These were required to meet Provincial standards, and included extensive answer keys and examination practice materials. By the same period, Canadian-authored upper school texts began to appear, not only in "advanced arithmetic" which included problems in compound interest, book-keeping, and so on, but also in Euclidean Geometry and in Algebra.[46]

[44]Wallace (1963), p. 481.

[45]Wallace (1963), p. 475. McLellan taught at a number of schools, including Upper Canada College, and became Director of Normal Schools for Ontario in 1875. He was later Principal of the Ontario Normal College at Hamilton (from 1885).

[46]In fact, there is an earlier algebra text by Sangster (1853) though we have not seen it. Its title, *Algebraic formulae: showing the method of deducing the most*

During the period in question it is difficult to distinguish works that are produced for upper secondary and for post-secondary teaching. In some cases the same book was used for both, the choice depending rather precisely on the audience and on local conditions. For example, an incoming class of engineering students from a variety of schools might very well require an algebra course more or less identical to that given in the better high schools, though the pace might be faster. We will therefore not try to distinguish them too carefully here, the more so because the same authors wrote for both groups.

This overlap between secondary and post-secondary education extends far beyond the mathematics courses, however. Virtually all Canadian universities and colleges of this period began their history as high schools, a gradual differentiation occurring only when an adequate supply of students became available to sustain post-secondary education.[47] To illustrate this in a mathematical context, let us briefly consider the courses of mathematical study at some Canadian institutions at various periods.

The oldest one known is from King's College, Windsor, N.S. in 1814 (the College had been chartered in 1789, but began granting degrees only in 1807). In this case, Oxford was the model; hence there was a very heavy emphasis on the classics. There were two professors, one of whom taught "Euclid and Wood's algebra"[48] in the third year. In Fredericton in 1824 all instruction was provided by the principal; the amount of mathematics appears to have been similar.

By 1860, an increased emphasis on mathematics is generally evident. This is in part due to the Scottish model, urged by Strachan (hence influential in Ontario, English Quebec, and, via a Royal Commission of 1854, in New Brunswick). Thomas McCulloch in Nova Scotia also urged its adoption. Details of the mathematical requirements for two B.A. degrees at this time are provided by Harris:

Queens:
- Year One: Euclid 1-6, algebra, plane trigonometry, logarithms
- Year Two: Euclid 11, part of 12; plane and spherical trigonometry, conics, calculus
- Year Three: Principia 1-3; hydrostatics

Trinity:
- Previous Examination: Euclid 1-4, 6; algebra to the binomial theorem
- B.A. Exam. Algebra to the end of the binomial theorem, trigonometry and solution of triangles, mechanics, hydrostatics

At the same time, Toronto required mathematics in each of the four years, and McGill in the first three of four. At Toronto mathematics could be avoided

important rules of arithmetic and mensuration suggests something other than a simple introduction to algebra.

[47] For details on the development of Canadian Universities, see Harris (1976).

[48] Harris (1976), p. 30. The Wood in question is James Wood of late eighteenth-century Cambridge.

after first year by obtaining first class honours in the first year. Other institutions had similar requirements. It is interesting to note that these were required of all students; as a corollary it follows that the standards in mathematics were doubtless not very high. Honours students at Toronto or McGill, who wrote specialized examinations in the chosen subject, doubtless had to meet a higher standard than did general students. Materials for these courses were originally imported, though as mentioned earlier Canadian texts at a higher level started to appear in the 1860s. These developments in mathematics education reflected an increased interest in science and engineering in the universities. Besides an expansion of these programs, this development is marked by the founding of two scientific societies: the Canadian Institution (later the Royal Canadian Institute), founded at Toronto in 1849; and the Royal Society of Canada, established in 1882. Both of these provided a forum for members to discuss scientific and technological issues of general interest, and for the presentation and publication of research.

In Ontario the connection between school and university textbook writers was particularly close in this period, in part owing to the membership of many of these men in the Royal Canadian Institute. Thomas Kirkland, William Scott, John H. Sangster and Alexander MacMurchy all were members, as were several university professors, among them James Loudon and J. B. Cherriman of the University of Toronto, and Alexander MacKay of MacMaster. The elementary and secondary teachers among these men were at the top of their profession by the mid-1880s and early 1890s; all were either Principals of leading schools, or of one of the Ontario Normal Schools, and many had taught at Upper Canada College at some point.

Of the authors of the 1870s and 1880s, three stand out because of their involvement in university teaching, as well as in production of texts for people at the university level. These are: John Bradford Cherriman (1823-1908), professor of mathematics at Toronto from 1850 to 1875; his successor James Loudon (1841-1916), professor of mathematics and physics at Toronto; and Nathan Fellowes Dupuis (1836-1917), professor at Queen's.

Cherriman was born in England, and had graduated from St. John's College, Cambridge, as sixth wrangler in 1845, the year William Thompson was second. He came to Toronto as an Assistant Professor in 1850, and was promoted to Professor in 1853. Cherriman published a dozen papers in the *Canadian Journal*, the publication of the Canadian Institute, and three in the first volume of the Transactions of the Royal Society of Canada after the founding of that Society in 1882. However, as Gingras has pointed out, these are very much either recreational or teaching-related. They do not build on earlier work, and are at best new proofs of old results.[49] Two titles will suffice to

[49]Gingras (1991), 17-5 1. In this regard they rather resemble the weaker papers in the *Cambridge and Dublin Mathematical Journal*, with which Cherriman was no doubt well-acquainted.

convey their flavour: *Note on the composition of parallel rotations*, and *Note on the bishop's move in chess.*

Loudon was a product of the Upper Canadian system to which we have been devoting our attention. He studied at the Toronto Grammar School, Upper Canada College, and the University of Toronto, receiving a B.A. in 1862 and an M.A. in 1864. He then became a tutor in classics, but moved to mathematics, eventually becoming the University's professor of mathematics and physics in 1875 (and the first Canadian-born professor). In 1887 he became professor of physics only, and became president of the University in 1892. This set the stage for the first mathematical research at Toronto. Like Cherriman, Loudon published a number of articles in the Proceedings of the Royal Canadian Institute; and like Cherriman, these mostly arose from teaching concerns.

Dupuis taught school from 1857 to 1863, presumably following Normal School training.[50] At that point he was able to enter the University of Queen's College (Frontenac County was his home), where he worked as an observer in the Kingston observatory and as librarian. Obtaining a B.A. (with eight others) in 1866, he went on to obtain an M.A. from the same institution in 1868. At that point he succeeded Robert Bell as Professor of Chemistry and Natural History, and his writing career began in the same year with a textbook on geometrical optics. During his time at Queen's Dupuis taught physics, geology, mineralogy, biology, mathematics, and various engineering courses. He was an important institution builder, instrumental in the establishment of engineering and a medical school on a firm footing.

For reasons that are unclear, Dupuis began to teach mathematics in 1880. He was at once concerned to provide his students with an up-to-date course of study, and the two books that he produced in the 1880s attest to this concern. The first of these, *Junior Algebra* (1882) eventually became *The Principles of Elementary Algebra*, published by MacMillan in 1892. The term "junior" in the title may refer to the third year, as in U.S. nomenclature, or it may refer to a preparatory work, presumably for first year students. It was characterized by Dupuis as an "intermediate algebra", and owes a good deal to Chrystal's *Algebra.* The content includes a certain amount of formal algebra, so that terms such as "commutative" are introduced, but the work is very much concerned with applications, as Dupuis noted in the Preface:

> Probably the most distinctive feature of the work is the importance attached to the interpretation of algebraic expressions and results... the results arrived at have little interest and no special meaning until they are interpreted. This interpretation is either Arithmetical, that is, into ideas involving numbers and the operations performed upon numbers; or Geometrical, that is, into ideas concerning magnitudes and their relations.[51]

[50] Most information about Dupuis is from Varkaris (1980).

[51] Dupuis (1892), p. iv.

At the higher end, the book includes such subjects as simultaneous quadratics, the remainder theorem, approximation of roots, annuities, continued fractions, series and determinants.

Probably prepared in outline at around the same time as the algebra, but not in print until 1889, is *Geometry of the Point, Line and Circle in the Plane*. This was probably Dupuis' most successful book, appearing in at least five editions to 1914. This is at about the same level (again described as "junior") as the algebraic work, and is definitely intended as preliminary to a study of analytic geometry and the calculus. The work shows a good deal of originality in not simply reorganizing Euclid. Instead, lines and curves are treated as plane loci, so that triangles are distinguished from the regions that they bound. One of the inspirations cited is Sylvester:

> The principle of motion in the transformation of geometric figures, as recommend by Dr. Sylvester, and as a consequence the principle of continuity are freely employed, and an attempt is made to generalize all theorems which admit of generalization.[52]

There is a certain naiveté in this last statement, one which is probably quite genuine. Although Dupuis had worked hard to put together good preliminary courses for his students, there is little evidence that he had a grasp of higher geometry as it was practiced in his day. However the book does introduce such topics as inversion in the circle, pole and polar, homographies and involutions.

These books were well-received in the community for which they were intended, as their success in the U.S. and Canadian markets attests. Favourable reviews aided in the process. "It is safe to say", noted the *Canada Educational Journal*, "that a student will learn more of the science [of Geometry] from this book in one year than he can learn from the old-fashioned translations of a certain ancient Greek treatise in two years". The same review urged every mathematical master to study the book "in order to learn the logical method of presenting the subject to beginners."[53] As for the algebra, it was described in *The Schoolmaster* as "one of the most able expositions of algebraic principles that we have yet met with ... emphatically a book for teachers."[54]

These books are symptomatic of yet another generational change in Canadian mathematics, one in which students at the universities begin to see rigorous mathematics beyond the elementary Euclidean level, and in which a variety of subjects are treated in an incipiently rigorous way. In addition, the courses written by Dupuis are clearly prefatory to acquiring a higher level of expertise, and the debt to more advanced work is acknowledged for the student to see. These are not simply "everything you need to know" about a certain subject for the purposes of application. Students prepared under

[52] Dupuis (1914), p. vi.
[53] Review reprinted at the end of Dupuis (1914).
[54] Reprinted at the end of Dupuis (1914).

this regime would clearly be more susceptible of recognizing the existence of more advanced mathematics, and of beginning to value the subject outside of its applications (though the latter were clearly very important). Perhaps significantly, Dupuis' title for a time was Professor of Pure Mathematics.

By the 1890s, most of the established institutions had one professor of mathematics, though in some cases duties were still split between mathematics and physics or engineering. At some schools, notably Toronto and McGill, there was also an assistant or lecturer. Thus in the course of the nineteenth century we see that mathematics in English Canada has transformed. An institutional base for elementary mathematics education was firmly in place by the time of Confederation, and was closely linked to developing institutions at the secondary and post-secondary levels. University mathematics was established as a teaching subject in most institutions by the 1890s, though research was still in a nascent state; and it was to remain in that state at least until the time of Fields' comments cited at the beginning of the paper, despite a general expansion of mathematical education at the universities in the intervening decades.

Reasons for the lack of research late in the nineteenth century are not hard to find. While in fields such as physics and chemistry it was still possible at that time for an relatively inexperienced student to undertake experimental work of a meaningful sort, participation in mathematical research at an international level required an extensive exposure to the literature and, ideally, to working research mathematicians. Such literature was difficult to come by in Canada. Furthermore, there were few rewards for engaging in research beyond personal satisfaction. The tiny Canadian mathematics community of the late nineteenth century was instead fully involved with teaching a variety of largely introductory courses in institutions which were widely separated geographically.

More than this, the Canadian university was a combination of the English and Scottish models. Thus liberal and practical education vied for position in the curriculum, and professors were either British-trained or trained in Canada according to this model. Research had long taken a secondary position in the British schools, and where it was undertaken an emphasis on practical results was emphasized. The education-related, physically oriented papers of Loudon and Cherriman reflect these values, and several decades were to pass before the views of Fields, that pure mathematical research was desirable in the universities, came to find widespread favour in Canada.

Québec from the Union to the Foundation of the Université de Montréal (1840–1920)

The union of Upper and Lower Canada had followed political disturbances that were felt in each of the two colonies. In Québec some small changes for mathematics are detectable, though these changes were mostly felt at the lower levels (due to a new trend in pedagogy) and in connection with the practical applications of mathematics. In the classical colleges, mathematics teaching in the philosophy courses stagnated. We also see in this period the beginning of the publication of local works which go beyond the elementary level. Among these only one displays any originality. Later, however, the founding of the École Polytechnique de Montréal in 1873 marks the beginning of mathematics teaching beyond the elementary level outside the classical colleges.

Primary teaching: "l'enseignement mutuel"[55] The arrival in Canada of the Frères des Écoles chrétiennes (Brothers of the Christian Schools) marks a turning point in the history of education in the Province. This French community, which had been founded in 1684 by Jean-Baptiste de la Salle (1651-1719), had a long tradition of teaching students in groups. When the first four brothers arrived in Montreal in November of 1837, they brought with them this tradition, which permitted them to address the problems of mass education head on. Their school took in 200 students in the first year, and by 1840 they had 860 students.[56] In mathematics, they were innovative, and their approach was rather more dynamic than that of their predecessors. This is indicated by the title of their first textbook, published only a year after their arrival: *Nouveau traité d'arithmétique: contenant toutes les operations ordinaires du calcul, les fractions et les différentes reductions de fractions, les règles de trois, d'intérêt, de société, d'alliage, l'extraction des racines, les principes pour mesurer les surfaces et la solidité des corps; enrichi de 400 problèmes a résoudre, pour servir d'exercice aux élèves: a l'usage des écoles chrétiennes des frères.* The book was a reedition of a recent (1833) French book, hastily modified to take into account the peculiarities of the British colony (notably currency and the system of weights and measures).[57] Unlike its predecessors, it contains a large number of problems. Even more indicative of a new approach, most of the problems were not presented with solutions. This indicates the importance which the brothers gave to exercises to be completed by the students themselves, and it also presupposes the competence of the teacher, at a certain level at least. Thus the rule-example-rule presentation of earlier manuals was replaced by a long, careful presentation of each topic, calling on examples, and concluding with a concise statement of the rule. This was then followed by exercises and problems. The brothers even attempted to give their book a deductive structure adapted to the level of the students.

[55]Lavoie, Paul (1994), pp. 384–410. Here there are analogies with two U. S. books: Dilworth, Thomas, *Schoolmaster's Assistant: Being a Compendium of Arithmetic Both Practical and Theoretical*, 1773, and Adams, Daniel, *Arithmetic, in which the Principles of Operating by Numbers are Analytically Explained and Synthetically applied; thus Combining the Advantages to be Derived both from the Inductive and Synthetic Mode of Instruction*, 1801. The latter had many Canadian editions.

[56]Audet, Louis-Philippe (1971), t. 1, p.370.

[57]F.É.C (1833).

Another textbook, more properly Québecois than that of the brothers, appeared around 1843. While emphasizing British currency and weights and measures, it harks back to its predecessors in presenting few explanations. The following year, a book of solutions to the exercises was published.[58]

These books had a large number of editions, and became the paradigmatic texts for secondary teaching. To follow the developments which follow at this level would take us too far from our main subject. However, we note that these works were often used in the Humanities classes of the classical colleges.[59]

The classical colleges. In 1840, mathematics had solidified its position in the colleges, particularly in the Humanities classes and in the two years of philosophy. Léon Lortie has argued that 1840-1850 was the golden age of mathematics teaching in Québec, a statement which needs some added nuance.[60] Certainly mathematics instruction progressed greatly from 1800 to 1840, but this progress is linked to the construction of a system of elementary education in the province. The classical colleges participated in these developments by opening the door to mathematics in the Humanities class. At the upper level, the progress seems slight, if indeed there was any. The notes for the course given by Abbé Alexis Pelletier in 1862–1863 in the philosophy class at the Séminaire de Québec are revealing in this regard.[61] It is divided into three main parts, the first on algebra, the second on geometry, and the third on plane trigonometry. The first contains, among other things, sections on ratio and proportion, the rule of three, and questions with a commercial flavour, for example on interest. This content is exactly what was taught at the end of the previous century at Québec. The fact that such a notebook exists also shows that the taking of notes remained an important part of the student's activities. This is confirmed by reading the memoirs of another student at the Séminaire, where we find remarks on the course given by Abbé Théophile-Étienne Hamel, who replaced Pelletier in 1866–1867.[62] Hamel had returned from Paris ten years previously, with a license in science obtained at the École des Carmes[63], and was thus the third college professor to have studied physics or mathematics abroad. His student Gosselin reminisced as follows:

> He didn't have time to rework the course which he had taken in Paris, and to adapt it to another milieu, to reduce its proportions, to make it accessible to students of whom many knew almost nothing, even in

[58] F.É.C. (1842) and F.É.C. (1843). The date is uncertain. See Lavoie, Paul (1994), p. 404.

[59] For a complete and detailed overview of the teaching of arithmetic in Québec in the nineteenth century, see Lavoie, Paul (1994).

[60] I (L.C.) endorsed this view in (Charbonneau, Louis (1984)).

[61] These notes were taken by Jean-Alfred Charlebois and may be found in the Archives de l'Université Laval, box P 211.

[62] Gosselin, D. (1908), pp. 174–175.

[63] Gingras, Yves, (1991), p. 34.

> arithmetic. The hour of the course was spent copying—at full steam—what he dictated to us. Under the heading "theory of limits", incomprehensible to the vast majority, I scribbled four hundred lines of tiny, cramped writing ... he did not often compel us to grind the mill of problems, contenting himself to deliver the formulas without which it could not be moved .... Furthermore, arithmetic took the lion's share of the time, almost six months out of ten. Only four months remained for algebra, geometry, and trigonometry. Of this last we got only a bird's-eye glimpse.

As we can see, the students often arrived in philosophy with minimal mathematics, so it is hardly surprising that arithmetic got the lion's share. What could the students possibly take away from a mathematics course that consisted of dictation? While the professors of mathematics were no longer just out of the philosophy course, the teaching itself did not evolve.

Mathematics had come into the Humanities course of the colleges in the context of the general reform of the 1830s, but for several reasons this did not extend to the philosophy classes.[64] For one thing, the reform movement was halted by the events following the rebellion of 1837. As Jarrell has argued, there appears to have been a detouring of intellectual energy away from the sciences and toward political concerns.[65] As a result, the small number of French-Canadians who had participated in scientific and literary societies was not sufficient to orient the colleges towards science teaching after 1837. Gradually an image developed of the French-Canadian who was naturally drawn toward the "moral and political sciences, history, literature and the arts", in contrast with the English, who had an affinity for the "mathematical, physical and natural sciences". The sciences became tributary to these values, to the extent that they could be, and lay society no longer came knocking at the door of the colleges to insist that the science and mathematics taught there met the needs of the industrial society developing in Québec. In addition, the Catholic Church in whose hands this education lay was concerned primarily at this time with developing its weapons against the new social philosophies; its priorities did not include the teaching of the sciences, on which a number of these philosophies claimed to be based. As Abbé J. S. Raymond of the Séminaire of Saint-Hyacinthe noted in 1872, "Is it not to the deeper knowledge of the sciences that the materialist movement in which our century so prides itself is due?" No more was needed to assure that Latin and Greek would be seen as the most important tools for developing the spirit, to the detriment of mathematics. The formative qualities of mathematical training were not contested, but they had the major flaw of abetting the "materialist movement". As a result, when we compare the program of 1863, as evidenced by the notes of Abbé Pelletier, with the program of 1921 at Laval for graduates

[64]The citations in this paragraph are taken from Charbonneau, Louis (1984), pp. 29–31.

[65]Jarrell, R.A. (1977).

of the philosophy class, the similarity of the contents is striking. Whatever may have happened in the lower schools since 1840, in the upper school the attitude toward mathematics was, by 1900, one of mere toleration.

The first intermediate level mathematics books in French. In a context where French-Canadians developed a bit of a complex with regard to the sciences and mathematics, it isn't too surprising to find that between 1840 and 1920, apart from school texts, only four mathematical books were published. What is perhaps surprising is that the "mathematically-inclined" English Canadians did little better.

The *Traité élémentaire de calcul différentiel et de calcul intégral*, attributed to Jean Langevin, (1821-1892) was possibly intended to help in teaching differential and integral calculus in the colleges.[66] However, the level goes well beyond what could have been taught there. Langevin was professor of mathematics at the Séminaire de Québec in 1838, even before his ordination, and is best-known as the first principal of the École Normale Laval, and later as the first bishop of Rimouski.

The book is divided into three parts, followed by three notes. The first two parts, differential and integral calculus, rest on the idea of the differential with the Leibnizian notation. Each of the two parts is divided into two chapters, the first of which gives rules for the differentiation and integration of the principal functions and the second of which gives "applications" such as maxima and minima or volumes of solids of revolution. The third part, "Method of Limits", uses the difference quotient idea to get at series development. The notes deal with further series expansions using Newton's binomial series, and with the method of undetermined coefficients. The few "problems" in the book are generally accompanied by a solution, except in the last note where only statements are given.

This treatise could well have been written at the beginning of the nineteenth century. Allusion is made to the fact that there are several methods which "all lead to the same result" and that "the difference between them is more metaphysical than mathematical." The methods to which the author refers are those of Newton and Leibniz, but also those of Landen, d'Alembert and Lagrange. There is no mention of Cauchy. The treatise concludes with a reference (for additional information) to Lacroix, Hind, and Boucharlat, thus revealing its mixed ancestry, both English and French.[67] While it is not completely original, it is clearly written, and permits the reader to develop the basic methods of the calculus.

The *Premier livre des elements de géometrie d'Euclide, a l'usage des étudiants au Collège nautique du Canada* (1853) is considerably less original

[66] Lortie, Léon (1955, p. 40) gives the origins of this attribution.

[67] John Hind (1796-1866) and Jean Louis Boucharlat (1775-1848). The work of Hind referred to by Langevin remains unidentified. As for Boucharlat, the book is certainly *An elementary treatise on the differential and integral calculus*. Translated from the French by R. Bladelock, Cambridge, 1828.

than the work we have just been speaking of. The author's name is not given, but the work is in fact a translation of a brief portion of the well- known book of Robert Simson, *Euclid, Elements, First six books, with the 11th and 12th and Euclid's Data.*[68] An examination of the statements of the propositions suggests that the anonymous translator was inspired by the French translation of Euclid's works made by Peyrard at the beginning of the century.[69]

The third of these home-grown works is arguably the most interesting, not only because of its length and its originality, but also because of its author, Charles Baillargé (1826-1906). The book is *Nouveau traité de géométrie et de trigonométrie rectiligne et sphérique, suivi du toisé des surfaces et des volumes et accompagné de tables de logarithmes des nombres et sinus, etc. naturels et logarithmiques et d'autres tables utiles. Ouvrage théorique et pratique illustré de plus de 600 vignettes, avec un grand nombre d'exemples et de problemes a l'usage des Arpenteurs, Architectes, Ingenieurs, Professeurs et eleves, Etc.* Weighing in at 900 pages, it appeared in 1866. Baillargé came from a family of architects and engineers which had come to Canada in 1741, and he continued the family tradition, working for example as the on-site architect of the Parliament Buildings in Ottawa from 1863 to 1865.[70] Here he was able to put his mathematical skills to good use, defending himself unsuccessfully from charges of overspending on construction by calculating the volume of materials employed in the irregularly-shaped buildings. His book aimed at introducing higher elementary mathematics to a wide audience, and Baillargé did not hesitate to redo classical treatments to achieve this end, cutting down the number of propositions in Euclid's first six books by half.[71] His inspirations included Legendre and Davies, while his trigonometry seems to have come from Playfair and from the old *Institutions mathématiques* of Sauri.[72] He was proudest of his chapter on the measurement of areas and volumes, noting the originality of some of his results, notably his proposition 1521.[73] This dis-

[68] Simson's book was first published in 1756. We have consulted the 25th edition, dating from 1841.

[69] Peyrard, F. (1819).

[70] Cameron, Christina (1989).

[71] Playfair, John, *Elements of geometry: containing the first six books of Euclid, with a supplement on the quadrature of the circle, and the geometry of solids: to which are added elements of plane and sphericale trigonometry*, 8th ed., Edinburgh: s.n., 1831. First edition, 1795. There was a tenth edition in 1846.

[72] Legendre, Adrien-Marie, *Elements de géométrie*, First published in 1794 with many editions throughout the nineteenth century. Legendre was also translated into English: Legendre, Adrien-Marie, *Elements of geometry and trigonometry; with notes*, Edinburgh: Olivier & Boyd, 1824. Davies, Charles, *Elements of geometry and trigonometry from works of A.M. Legendre: adapted to the course of mathematical instruction in the United States*, New York: A.S. Barnes & Co., 1862 (reprinted 1871).

[73] The statement of the theorem is as follows (p. 662 of the *Nouveau traité. . .*): "Of every prism or right cylinder or oblique – of every pyramid, regular or irregular, or every cone, whether right or oblique – of every truncated pyramid or cone between

covery was the theoretical basis for his most famous production, the *Tableau stéréométrique*, a wooden box on a base five feet by three feet, about five inches deep which contained 200 wooden models of geometric shapes.[74] While the object could be used as an introduction to solids for classes, its main use was to assist engineers and architects in the measurement of volumes. Baillargé promoted the work energetically, and succeeded beyond his wildest expectations: between 1872 and 1876, the tableau earned him 13 medals and 17 diplomas from eight different countries.[75] In Russia, for example, its adoption was urged not only in the primary schools but also in Polytechnics.[76] In Québec, the device was investigated, following Baillargé's request for official recognition, by Jean Langevin, who passed it on to Thomas-Étienne Hamel, now director at the Séminaire de Québec. Together with the Abbé Mainguy of the Séminaire he declared it satisfactory, and Mainguy even published a tract on the device. Following this expert examination, the Council of Public Instruction recommended its use in the schools of the Province.[77]

In 1882, at the founding of the Royal Society of Canada, Baillargé was a founding member of Section III on mathematics, physics and chemistry. He regularly presented papers to the annual meetings of the society, among which there are three on mathematics. However, as with English-language writers of the day, his main preoccupations were toward pedagogy and applications, as his 1882 "Utility of Geometry as applied to the Arts and Sciences" and his "Hints to Geometers for a new Edition of Euclid" suggest.[78]

A fourth treatise, the *Théorie élémentaire des nombres d'après Buler, Legendre, Gauss et Cauchy* (1870) is a work which falls outside the patterns we have seen so far. Of course, the appearance of Buler for Euler in the title doesn't exactly inspire confidence in the anonymous author. It is a 22 page pamphlet listing relatively elementary results in number theory, including material on Gaussian residues. Without further information, there are only questions about this work: could such information be found in Québec at the time? Who could the author have imagined to be the public for such a work?[79]

---

parallel bases – of the sphere ... [many such figures omitted]: the volume is equal to the sum of the surface of the base, if there is only one, or of its parallel bases, if there are two, and four times the surface of a section midway between the bases, between the base and the vertex, or between opposite vertices, multiplied by one sixth of the height of the solid."

[74] A complete bibliography of Baillargé's works is in Cameron, Christina (1989), pp. 161–166. On the *Nouveau Traité...*, see Chapter 11, pp. 131–138.

[75] Prospectus du Tableau Stéréométrique Baillargé, [Quebec]: n.p. [1871].

[76] Cameron, Christina (1989), p. 136.

[77] Cameron, Christina (1989), pp. 132–133.

[78] The complete list of papers given before the Royal Society is given in Cameron, Christina (1989), p. 191, note 62.

[79] The author mentions that the proofs of one of the propositions was given to him by a Prof. Wantzella. Anonymous (1870), p. 12–13.

The École Polytechnique de Montréal. In 1852, the Université Laval was founded in Québec, with four faculties: theology, law, medicine and arts.[80] The beginnings, as with most institutions of the period, were difficult, with shortages of both professors and students except possibly in medicine (where the Faculty was just a restructuring of an existing school).[81] The main aim of the Arts faculty was to control the quality of teaching in the classical colleges.[82] However, the shortage of professors led to the mission of T.-É Hamel to Paris, from which he returned with his licence in mathematics as we mentioned above.[83] The lukewarm attitude toward science and particularly applied science at, Laval at the time is well-illustrated by the story of the founding of the École polytechnique de Montréal.[84] In 1870, the Chauvau government had offered Laval a grant to create a school of applied sciences. After much hesitation, this was refused: the university did not want to see the government mixed up in its affairs. Such a school remained a political priority, and in 1873 the new government of Gédéon Ouimet began negotiation with the Commission of Catholic Schools of Montréal, and in particular with the director of the Catholic Commercial Academy of Montreal, Urgel-Eugène Archambault, which led to the introduction of a "Cours scientifique et industriel" at that school. In 1876, the Academy became the École polytechnique de Montréal. Its beginnings were modest, with 114 diplomas awarded during the years from 1877 to 1904. Of these graduates, only 54 came from classical colleges, reflecting the lowly status given to the engineering profession by the colleges. As Abbé Hamel, himself a science graduate, said in 1876, "We don't hesitate to encourage such [engineering] studies for those of our young people who are not destined for the priesthood."[85]

It is hard to assess the level of instruction at the École in its early years.[86] The two professors charged with mathematics instruction during the first decades of its existence, Frédéric André and Emile Balète, had no university training. André taught for the most part the introductory course from 1875 to

[80] Chartrand, Duchesne, Gingras (1987), pp. 222–227.

[81] Chartrand, Duchesne, Gingras (1987), pp. 216–220.

[82] Following Léon Lortie, some authors have argued that, in its efforts toward uniformization, the Faculty of Arts brought about a decline in the quality of mathematics and science teaching at the Séminaire de Québec. We instead support the view of Chartrand, Duchesne, Gingras (1987), p. 220 who attribute this point of view to an error.

[83] The kind of licence obtained be Abbé Hamel in Paris in unclear. Chartrand, Duchesne, Gingras (1987), p. 217, suggest a qualification in mathematics while Gingras, Yves (1991), p. 34, speaks of a *licence en sciences*, for other references indicating a *licence en mathématiques*, see Charbonneau, Louis (1984), p. 33, note 29.

[84] Chartrand, Duchesne, Gingras (1987), pp. 227–230 and Gagnon, Robert (1991), pp. 39–44.

[85] Gagnon, Robert (1991), p. 70.

[86] Gagnon, Robert (1991), pp. 64–68.

1911.[87] Balète was in charge of the rest of the mathematical program, teaching from 1875 to 1908 and serving as director of the École from 1882 to 1909. He had been trained at the French military Collège at Saint-Cyr, and immigrated to Canada in 1872 following an apparently disappointing military career. Despite a lack of advanced mathematical training, Balète also became professor of mathematics at the Arts faculty of the Montreal Branch of the Université Laval in 1900-1901.[88] A look at the annual programs gives us some idea of the content of courses at this time, though the testimony of the programs should be accepted with caution. It appears that differential and integral calculus became somewhat more important during the period in question.[89] After 1910, more profound changes occurred, with two new professors, each more distinguished mathematically than their predecessors. Victor Elzéar Beaupré and Conrad Manseau were both graduates of the school, and both engaged in scientific activity extramurally. Beaupré, later a professor of mathematics at the newly-founded Faculty of Science at the Université de Montréal, was an actuary, indeed the first French-Canadian to become a member of the Society of Actuaries of America.[90] Marseau, on the other hand, was an astronomer, having obtained a *licence* from the Sorbonne in 1914.[91]

Expanding Horizons

By the turn of the century, signs of a new agenda for mathematics in Canadian universities begin to be seen. This agenda both tended away from applications—though this varied from one school to another—and showed an awareness of the importance of research and of the diversity of the mathematics then being practiced internationally. This awareness, as we shall see, seems to have been concentrated in a few individuals. It manifests itself by the beginning of doctoral programs, by an effort to increase exposure of students to a broader range of mathematical ideas, and, where doctoral programs did not exist, by preparing students for advanced study in other institutions. We will take each of these developments in turn.

[87] Even in retirement he continued to teach, until just before his death in August 1923. Archives de l'École polytechnique, dossier Frédéric André, dossier 320-300-22.

[88] According to the *Annuaire of the Université Laval* for 1900–1901.

[89] In the *Bulletin annuel* of the École polytechnique for 1878–1879, the "first principles of the theory of derivatives" form part of the algebra course in first year. The "complete theory of derivatives" and series expansions constitute the first part of the second year algebra course. The booklet *Programmes des travaux techniques et questionnaires des examens généraux* of 1896-1897 shows clearly that there was a calculus course in the third year on differential and integral calculus.

[90] *La Voix Nationale*, Dec. 1935, from a clipping in the dossier Beaupré (dossier 329-300-22), Archives de l'École polytechnique.

[91] Gagnon, Robert (1991), p. 132.

Building mathematics in the universities. At around the turn of the century, as was the case in the U.S., young Canadian-trained mathematicians began to study abroad. This led to the first Canadian Ph.D's in mathematics, people with an exposure to a European research tradition who had not only undertaken research on their own, but who had seen the necessary conditions in which such research could flourish. The most important of these were those who returned to Canada to academic positions. Best-known of these is J. C. Fields (1863-1932), a University of Toronto gold medallist as an undergraduate in 1884, who took his Ph.D. at Hopkins in 1887. After a short teaching stint in the U.S., Fields went to Europe, where he spent approximately ten years in the mathematical centres of Berlin and Paris. Fields was appointed to the Faculty at Toronto in 1902. At this time, another young Toronto-educated mathematician was also in the Department. A. T. DeLury had undertaken graduate work at Clark and in Paris. He had joined the Toronto faculty in 1892. The department at the time thus consisted of Fields, DeLury, Loudon, Baker, and M. A. Mackenzie, an actuary. Of these, Fields was perhaps the most important in developing research at Toronto; his student, Samuel Beatty, was the first mathematics Ph.D. at Toronto, obtaining his degree in 1915. In the next twenty years, Toronto was to produce eight doctorates in mathematics, two of them women. This production of Ph.D.'s was however supervised mostly by imported mathematical talent: W. J. Webber, from Cambridge, and J. L. Synge from Dublin.[92]

At other institutions, mathematics was either in a nascent state (as were the institutions themselves in many cases), or else–as at McGill for example–it remained firmly subsidiary to other programs, particularly Engineering. McGill is an interesting case in point. Its faculty included J. Harkness, a Cambridge mathematician with international connections who could have fostered research interests, though not active in research himself. It is interesting to look at the situation there and elsewhere through the eyes of Henry Marshall Tory (1864-1947), a McGill alumnus of 1891 who taught in the department with Harkness, and was an important figure in the development of scientific research in Canada.

Tory was from a Nova Scotia Methodist family, and had the good luck to be "discovered" by a reasonably well-educated teacher who gave him books on algebra, geometry and trigonometry. His own reminiscences stress the paucity of books available, then and later. Determining to go to university, he worked as a clerk to save enough to spend six months completing a teacher's certificate. In 1886, having planned to go to Mt. Allison, he heard of McGill, where the chances of summer employment would be better. Despite his weaknesses in the classics he graduated at the top of his class in 1890 and was immediately engaged as a lecturer for the fall of 1891, thus "to be associated as a colleague with the men by whom I had been taught", as he put it.[93] At the time the

[92]See Robinson (1979) for details of the situation at Toronto.

[93]Corbett (1954) p. 41.

mathematics faculty included himself, Harkness, George Chandler, who was the head of the Mathematical Department in the Engineering Faculty, John Cox, "a Tripos man" who was adept at the style of mathematical physics then taught at Cambridge, and Alexander Johnson, trained in Dublin, then Dean of Arts. A student later noted of Tory: "he used to joke about the relative capacities of himself and his associate Dr. Harkness–a Cambridge Scholar of very high attainments. Dr. Tory admired and liked him but admitted that he himself was a better choice for 1st yr work."[94] The student quoted Tory, "The poor youngsters don't know what Harkness is talking about."

The McGill emphasis on engineering and British-style applied mathematics in those years was all the more influential because of McGill's activities in franchising its program. In order to be viable, a good supply of students was necessary, especially at McGill, which unlike the Ontario and Atlantic schools received no government subsidy. For this reason, McGill's admnistration supported efforts to develop affiliated institutions elsewhere in the country which would provide instruction equivalent to the first two years at McGill. Students with satisfactory performance could then travel to Montreal for their final year or years, completing a degree at a well-equipped institution with a good name. Such an arrangement had begun in Vancouver and Victoria in 1899 and 1902 respectively, where the leading high schools began to offer McGill first year, then second year. These arrangements would later lead to the establishment of the University of British Columbia. Tory himself was instrumental in these arrangements.

It is clear, then, that what impetus existed to produce research mathematicians in this period came from the German and French models. England was in the process of changing, but those with English-style training at this time were inclined to the Cambridge view. While research in other sciences, notably engineering and physics, was getting under way, mathematics remained as a service department at most schools.

The role of the First World War in this story is problematic. It must certainly be true that the general mobilization required in Canada cut into any efforts to expand graduate studies during the war years, something which was not the case in the U. S. for example. On the other hand, the war pointed out the pressing importance of individuals trained to do research, both to the British and to those in the Dominions, as Canada then was. The war had led to a general call throughout the Empire for highly qualified researchers. It was then discovered that there were without exaggeration more trained scientists in a few of the large German industries than could be found in the whole Empire.[95] Following urging from Britain, the National Research Council was established in 1916, though the results for mathematics were again slight at first. The continued emphasis on applications did however begin to produce some results as far as the use of mathematics in the physical sciences and

[94] Corbett (1954), quoting Susan Cameron Vaughan, p. 43.

[95] Corbett (1954) p. 154.

engineering, as well as in actuarial work and other applied areas. This is shown nicely by the International Congress of 1924.

The 1924 International Congress of Mathematicians. In 1920, the International Mathematical Union determined to hold its next quadrennial Congress in New York. By 1922 New York backed out, and the Union determined to retain an American location by selecting Toronto, apparently at the urging of J. C. Fields. Fields thus took on the job of chairing the Organizing Committee. We can speculate on his motives: by assembling a strong group of the world's leading research mathematicians, he would illustrate the diversity of mathematics and the importance accorded by other nations to research in mathematics. The meeting would also provide a forum for Canadian researchers. Fields was successful in attracting a rather surprising amount of outside funding: the federal and provincial governments supplied $27,000 each, with $6,500 from the Carnegie Corporation and $2000 from the University of Toronto. There were also contributions from the private sector: Eaton's gave $500, for example, as did Imperial Oil and several private citizens.

This sponsored a congress, by all accounts successful, with 444 attendees, among them 107 Canadians and 191 Americans. Following the dictates of the war settlement, Germany was still excluded. Those who have been involved in organizing conferences recently will doubtless appreciate the complexity of the task when many participants were arriving by ship, and where the ancillary travel options included a rail trip to the west coast. Fifteen of the papers or abstracts were given by Canadians, and as this should give a fairly complete picture of research mathematics in the country at the time, we list the mathematicians giving them: J. C. Fields (Toronto–algebra); F. H. Murray (Dalhousie–partial differential equations); N. B. MacLean (Manitoba–geometry); C. T. Sullivan (McGill–geometry); J. L. Synge (Toronto–geometry); Daniel Buchanan (UBC–mechanics); L. V. King (McGill–numerical analysis); H. B. Dwight (Westinghouse, Hamilton–electrical engineering); Alan Ferrier (RCAF, Ottawa–aeronautics); T. R. Rosebrugh (Toronto–electrical engineering); T. R. Wilkins (Brandon–ballistics); R. W. Angus (Toronto–hydraulics); R. H. Coats (Dominion Statistician, Ottawa–descriptive statistics); H. H. Wolfenden (Consulting Actuary, Grimsby–actuarial science). The breadth of topics is particularly interesting, comprising what would now be called the mathematical sciences broadly conceived. Only one third of the papers are in pure mathematics as we think of it today, and these are really not at the leading edge of research. For the most part these consist of elaborations on the authors' theses, Synge's paper constituting an exception.

The congress undoubtedly brought together a large portion of the Canadian mathematical community, establishing new ties and reaffirming old ones. Nevertheless, Fields' assessment, quoted at the beginning of this paper, is quite accurate: there was very little mathematical research being done in the country in the period immediately before or after the Congress. This had to await further developments in the university sector and internationally.

After 1920: The University Period in Québec

With the exception of the École polytechnique, as of 1920 higher level mathematics in Québec had not evolved since the mid-nineteenth century. Mathematics was thus out of step with Québec society as a whole in this regard. Particularly in Montréal, industrialization made evident the absence of French-Canadians in scientific and engineering areas. Many observers suggested that it was important to create institutions where a scientific education of high quality would be offered. Those institutions which did exist had a primarily practical orientation. As with the École polytechnique, mathematics was seen as a tool in these schools. For a research community to develop, it was first necessary that those interested in science and mathematics begin to act in a concerted fashion. After 1920, a mathematical community gradually formed; this was the necessary prelude to the more extensive transformation which followed the second world war.[96]

The universities: mathematics in the service of the other sciences. On February 14, 1920, the Montreal branch of the Université Laval received its charter and began to function independently as the Université de Montréal. In the original plan, the faculty of science would be created at the École polytechnique. Mgr Georges Gauthier, the university's first rector, saw to it that the science faculty got rolling quickly, not least because a substantial grant from the Rockefeller Foundation required solid improvements in the level of scientific training provided to medical students. Thus one of the principal raisons-d'être of the new Faculty of Science was to ensure that students admitted to the medical faculty had acquired a good level of training in basic science. Under these circumstances, the university decided to create an independent Faculty of Science outside the École polytechnique, one which would share laboratory facilities with the Faculty of Medicine.[97] This was probably not the best thing for mathematics, which would likely have found an affiliation with the École polytechnique more stimulating.

Courses and teaching in the Faculty were organized on the French model. Students enrolled in a one-year "certificate" course in a given discipline. The accumulation of three certificates entitled the student to a *licence*.[98] By far

[96]For an overview of this movement, see Chartrand, Duchesne, Gingras (1987), chap. 8, pp. 239–272.

[97]The history of this decision is sketched by Léon Lortie in a note in the Archives de l'Université de Montréal, box 3523 (18-8-5-1), probably written in 1970. See also Gagnon, Robert (1991), p. 182.

[98]The decision to imitate the French system may have been taken, among other reasons, from the fact that the Baccalaureate in Sciences was given to those who

the most popular certificate was the "PCN", that is, physics, chemistry and natural sciences, which was the certificate required for entrance to the medical faculty. From 1920 to 1945, this program accounted for between 40% and 80% of the students in the faculty. Between 1920 and 1944, only eight *licences* in mathematical sciences were awarded.[99] The number of mathematics professors was very small. At the time of creation of the faculty, Arthur Léveillé, who had an honours B.A. in mathematics from London, was named professor of mathematics, permitting him to leave his previous post as clerk in a bookstore. Victor-Elzéar Beaupré of the École polytechnique also became professor in the Faculty of Science. It was not until 1936 that a graduate continued to higher study, when Abel Gauthier went to Columbia to undertake a Master's. He obtained this in 1939 with a thesis called *Theory of Group Representation by Matrices*. In the same year he was hired by the Faculty of Science at the Université de Montréal. He published several articles in the period from 1936 and 1941, and continued his education with courses at Chicago, Columbia and Brown in the early forties. This set a new tone, and soon Maurice L'Abbé, Francois Mumer and Jacques Saint-Pierre went abroad to complete doctorates. On their return, they brought a new mathematical culture with them. In 1947, after the death of Arthur Léveillé, Abel Gauthier became director of the Department of Mathematics; during his tenure as head, until 1957, the Department positioned itself to become the research centre it was to be in the following decade.[100]

As for Laval, it founded its École supérieure de chimie in 1921. The first mathematics courses were given there by Athéod Tremblay, a surveyor and geometer from Québec. More important for mathematics in Québec was the appointment the following year of Adrien Pouliot, who was to become the guiding spirit of Québec mathematics.[101] Pouliot had graduated from the École polytechnique de Montréal, and in 1928 obtained a licence in mathematics from the Sorbonne. From 1929 to 1939 he spent his summers in Chicago in order to improve his mathematics. He was the only professor of mathematics at Laval until 1936, when the Abbé Alexandre LaRue joined him.[102] In 1923, Pouliot founded the Société mathématique de Québec; in 1929 he became known to a wider public by sparking a lively debate on the quality of science teaching in the classical Collèges and in secondary schools generally. University science programs had continual difficulties in finding students who

---

failed the rhetoric examination for the Baccalaureate in Arts. Rhetoric was the final year of the classical course prior to the philosophy classes. See the text of Leon Lortie mentioned in n. 97.

[99] Charbonneau, Louis (1988), pp. 8–9.

[100] Formerly, following the French model, the group of mathematics professors was termed the *Institut de mathématiques*. With the progressive abandonment of the French system after 1945, the institutes became departments.

[101] For a detailed biography of Adrien Pouliot, see Ouellette, Danielle (1986).

[102] Althéod Tremblay continued teaching for many years. Richard, Guy W. (1982), p. 18.

were sufficiently qualified to undertake university-level studies in science. The controversy initiated by Pouliot's criticisms started the long process of inserting science properly into secondary education in the province, an end which wasn't really achieved until the 1960s.[103] In 1939, two years after the founding of a Faculty of Science at Laval, Pouliot organized a Department of Pure and Applied Mathematics. This had six members in 1945, including Basil White, educated abroad, and Paul Lorrain, a physicist whose later career was spent principally at the Université de Montréal. The first five mathematics degrees were awarded in 1951.

During the period from 1920 to 1945, then, we see little that resembles mathematical research in Québec. This was also the case at the École polytechnique, though that institution remained central to mathematics in the province under the direction of Augustin Frigon. But even though research was still not taking place, the formation of a mathematical community was occurring.

Mathematics in scientific societies, 1923-1945.[104] In addition to the Société mathématique de Québec, founded at Quebec by Pouliot in 1923, the Société de Mathématiques et d'astronomie du Canada was established at Montreal in April of the same year.[105] Although there are traces of activities from that date until 1942, this group only met regularly in the period from 1925 to 1932, during which it organized four to six lectures a year which were intended for a general audience, notably teachers.[106] At the peak of its activity it had 34 members, of whom the most active came from the École polytechnique. These included André-V. Wendling, Lorenzo Brunotto, and Beaupré, as well as Léveillé from the Université de Montréal. The lectures don't seem to have provoked any real scientific interaction or discussion, with the possible exception of one by Léveilleé given in November 1927 which provoked responses from Beaupré and Jules Poivert (another polytechnicien) at the December meeting. After a decade of inactivity, Brunotto tried in 1942 to reactivate the society, which was eventually reorganized under the name Société mathématique de Montréal in 1944 with 15 members. This society, as well as Pouliot's SMQ,

[103] Chartrand, Duchesne, Gingras (1987) pp. 257–260 and Galarneau, Claude (1978) p. 221-228.

[104] For an overview of the activities of mathematical associations in Québec, see Richard, Guy W. (1982).

[105] The name *Société de Mathématiques et d'Astronomie du Canada* might lead to confusion. The founders hoped that mathematicians and astronomers from elsewhere in Canada would eventually join the society. In reality, the society stayed essentially Montréal-based, and even francophone.

[106] Minutes of the *Société de Mathématiques et d'Astronomie*, in the Archives de l'École polytechnique, box 999-303-87 (24). The minutes cover the first 32 meetings, the last of which took place on the 16 of August 1931. Activities continue after that date, as Guy W. Richard (1982) has stressed. Unfortunately he does not mention his sources.

both became members of ACFAS, the umbrella group, founded in 1923.[107] After a decade of little activity, ACFAS began organizing large numbers of lectures in the 1930s, and from 1933 has had an annual meeting. Between 1933 and 1945, there were 32 communications in mathematics at these meetings, the mathematical level of which was higher than that given at the meetings of the Montreal or Quebec societies. These were nevertheless intended as expository, and hence did not reach a research level. Pouliot gave 12 of these papers, and his colleague Althéod Tremblay gave eight, with participation by others from the universities and the École polytechnique.

Conclusion

In the years 1935–1945 there are distinct signs of research mathematics beginning to come to Canada. Synge returned to Toronto in 1930 as the head of a new Department of Applied Mathematics, which later included Alexander Weinstein and Leopold Infeld. In addition, the Nuremberg Laws brought the first refugee mathematician of what would later be a large and productive group: Richard Brauer came to Toronto in 1935. Brauer's appointment was apparently made at the suggestion of Emmy Noether, as Robinson reports.[108] However, Robinson also reports that "Our chairman was anxious to build up the department, and the suggestion was immediately accepted", while Morawetz notes "It is hard to imagine today the struggle to make that appointment".[109] This is just one example of history that this brief article has not been able to unravel. In addition to the refugee influx–in part unwilling, as enemy aliens arrested in Britain and transported–we see a general growth in interest in both pure and applied mathematics across the country after the war. Young mathematicians began to leave the country for mathematical study. Their return to teaching posts, their research activity, and the founding of the Canadian Mathematical Society/Société mathématique du Canada brought new perspectives to mathematics in Canada. Mathematics began to develop in an independent fashion, with new contacts with the world mathematics community.

Bibliography

Adams, Daniel. *Adams' new arithmetic, suited to Halifax currency, in which the principles of operating by numbers are analytically explained and synthetically applied; thus combining the advantages of the inductive and synthetic mode of instructing. The whole made familiar by a great variety of useful and*

[107] For a history of ACFAS, see Gingras, Yves (1994).
[108] Robinson (1979), p. 41.
[109] Morawetz (1993). p. 14.

interesting examples, calculated at once to engage the pupil in the study, and to give him a full knowledge of figures in their application to all the practical purposes of life. Designed for use in the schools and academies in the British provinces. Stanstead, L.C. Walton and Gaylord (1883). A second edition was published at Sherbrooke in 1849.

Anand, Kailash K. "Canadian women Mathematicians from the Early Nineteenth Century to 1960–A More Comprehensive Study". *CMS Notes*, 21 no. 5, 1989, 31-42.

Anonymous, *Nouveau Traité abrégé de la Sphère d'après le système de Copernic, par demandes et par reponses, Nouvelle édition a l'Usage du Séminaire de Nicolet*, Trois-Rivières: Ludger Duvernay Imprimeur, 1824. The 1829 edition is a l'usage des Ecoles de cette Province)

Anonymous (attributed to Jean Langevin), *Traité élémentaire de calcul différentiel et de calcul intégral*, Québec: Imprimerie d'Aug. Côté et Cie, 1848.

Anonymous, *Théorie élémentaire des nombres d'après Buler, Legendre, Gauss et Cauchy, 1er fascicule*, Montréal: Eugene Senécal, 1870.

Audet, Louis-Philippe, *Histoire de l'enseignement au Québec*, 2 vol., Montréal, 1971.

Baillairgé, Charles, *Nouveau traité de géométrie et de trigonometrie rectiligne et sphérique. Suivi du toisé des surfaces et des volumes et accompagné de tables de logarithmes des nombres et sinus, etc. naturels et logarithmiques et d'autres tables utiles. Ouvrage théorique et pratique illustré de plus de 600 vignettes, avec un grand nombre d'exemples et de problèmes a l'usage des Arpenteurs, Architectes, Ingénieurs, Professeurs et Élèves, Etc.*, Québec : C. Darveau, 1866.

Bibaud, Michel, *L'arithmétique en quatre parties, savoir : l'arithmétique vulgaire, l'arithmétique marchande, l'arithmétique scientifique, l'arithmétique curieuse, suivie d'un précis sur la tenue des livres de comptes, principalement pour ceux qui veulent apprendre l'Arithmétique d'eux-même et sans Maître, ou s'y perfectionner*, Montréal: Nahum Mower, 1816.

Bibaud, Michel, *L'arithmétique a l'usage des écoles élémentaires du Bas-Canada*, Montréal Workman & Bowman, 1832. Nouvelle impression en 1847.

Burke, Edmund, *Thèses de mathématique et de physique qui seront soutenues au Séminaire de Québec* sans lieu d'édition, sans date, mais entre 1786 et 1790.

Bouthillier, Jean Antoine, *Traité d'arithmétique pour l'usage des écoles, Québec*: John Neil-son, 1809. Une seconde édition en 1829 et une neuvième et dernière édition en 1864. (Voir Lavoie, Paul (1994), la Bibliographie)

Cameron, Christina, *Charles Baillairgé, Architect & Engineer*, Montréal, Kingston, 1989.

Charbonneau, Louis, "L'abbé Jérôme Demers (1774-1853), 'L'homme qui lit dans les astres'," *Bulletin AMQ (Association mathématiques du Québec)*, décembre 1983, pp. 4–6.

Charbonneau, Louis, "L'enseignement des mathématiques dans les collèges classiques du Québec au XIX^e siécle", *Bulletin AMQ (Association mathématiques du Québec)*, mai 1984, pp. 41–44 et octobre 1984, pp. 29–34.

Charbonneau, Louis, "Les mathématiques a Montréal, 1920-1960. *Bulletin AMQ* May 1988, pp. 8–13.

Chartrand, Luc, Duchesne Raymond, Gingras, Yves, *Histoire des sciences au Québec*, Montréal, 1987.

Corbett, E. A. *Henry Marshall Tory, Beloved Canadian.* Ryerson Press, 1954.

Craig, G. M. "John Strachan" in *Dictionary of Canadian Biography*, v. IX (1861 to 1870), pp. 751–766. Toronto, University of Toronto Press,

Dainville, Francois de, "L'enseignement scientifique dans les collèges des Jésuites", in Taton, R., *Enseignement et diffusion des sciences en France au XVIII^e siècle*, Paris, 1964, pp. 27–65.

F.É.C., *Nouveau traité d'Arithmétique décimale, contenant toutes les opé rations ordinaires du calcul, les fractions, la racine carrée, etc. enrichie de 1316 problémes a résoudre, pour servir d'exercice aux élèves*, Paris, Lille, 1833.

F.É.C., *Nouveau traité d'arithmétique: contenant toutes les opérations ordinaires du calcul, les fractions et les différentes reductions de fractions, les règles de trois, d'intérêt, de société, d'alliage, l'extraction des racines, les principes pour mesurer les surfaces et la solidité des corps; enrichi de 400 problèmes a résoudre, pour servir d'exercice aux éléves: a l'usage des écoles chrétiennes des frères*, Montréal,: C.P. Leprehon, 1838.

F.É.C., *Traite d'arithmétique contenant toutes les opérations ordinaires du calcul, les fractions, l'extraction des racines, les principes pour mesurer les surfaces et la solidité des corps, enrichi d'un grand nombre de problèmes a résoudre, pour servir d'exercice aux éléves: a l'usage des écoles chrétiennes*, Montréal, 1842.

F.É.C., *Solutions des problèmes avec leurs réponses du Traité d'arithmétique a l'usage des écoles chrétiennes*, Montréal, 1843.

Fessenden, Arthur. *Tables, showing the interest at six per cent, of any sum from 1 pound to 1000 pounds, from one day to one hundred days and from one month to twelve months.* Montreal, Nahum Mower, 1822. Editions: 1830, Montreal, printed by Workman and Bowen for E.C. Tuttle; 1832, Halifax, C.S. Belcher, 1837, extended to 365 days; 1841.

Fields, J.C. "The Royal Society of Canada and Canadian Mathematics", in *Fifty Years Retrospect*, Royal Society of Canada, 1932.

Gagnon, Robert. *Histoire de l'École polytechnique de Montréal, 1873-1990* La Montée des lngémieurs, Montreal, 1991.

Galarneau, Claude, *Les Collèges classiques au Canada français*, Montréal, 1978.

Galarneau, Claude, "L'enseignement des sciences au Québec et Jérôme Demers (1765-1835)", *Revue de l'Université d'Ottawa*, vol. 47, no. 1-2, 1977, pp. 84–94.

Gingras, Yves. *Les Origines de la recherche scientifique au Canada: le cas des physiciens.* Montréal, Borèal, 1991.

Gingras, Yves. *Pour l'avancement des sciences: histoire de l'ACFAS, 1923-1993*, Montréal, 1994.

Gosselin, D., *Les étapes d'une classe au Petit Séminaire de Québec, 1859-1868*, Québec, 1908.

Gouinlock, G. and J. *A complete system of practical arithmetic, for the use of schools in British America, to which are added, a set of book-keeping by single entry, and a practical illustration of mental arithmetic, federal money, receipts, bills of exchange, inland and foreign, explanation of commerical terms, etc. The whole adapted to the business of real life, to the circumstances of the Country, and to the present improved state of commerce.* Hamilton, Ontario. J. Ruthven, 1842, 215pp.

Harris, Robin S. *A History of Higher Education in Canada, 1663-1960.* Toronto, University of Toronto Press, 1976.

Jarrell, R.A., "The Rise and Decline of Science at Quebec, 1824-1844", *Social History / Histoire sociale*, vol. 10, no. 1, 1977, pp. 77–91.

Karpinski, Louis. *Bibliography of Mathematics in America through 1850*, Ann Arbor, University of Michigan Press and London, Humphrey Milford, Oxford University Press, 1940.

Kerr, D.G.G., *A Historical Atlas of Canada*, Don Mills, 1966.

Ladreyt, Casimir, *Nouvelle arithmétique raisonnée ou cours complet de calcul théorique et pratique, a l'usage des collèges et des maisons d'éducation de l'un et de l'autre sexe, des personnes qui veulent apprendre cette science en peu de temps et sans le secours d'un maître, et de celles qui veulent se livrer au commerce; suivi de quelques leçons sur la plannimétrie et la stéréométrie (arpentage et cubage), ou toisé des surfaces et des volumes*, Montréal: s.d., 1836.

Lamonde, Yvan, *La philosophie et son enseignement au Québec (1665-1920)*, Montréal: HMH, 1980.

Lavoie, Paul, *Contribution a une histoire des mathématiques scolaires au Québec: L'arithmétique dans les écoles primaires (1800-1920)*, Thèse, Université Laval, Québec, 1994.

Lessard, C., *Le Séminaire de Nicolet, 1803-1969*, Trois-Rivières: Editions du Bien Public, 1980.

Lortie, Léon, "Les mathématiques de nos ancêtres", *Mémoire de la Société royale du Canada*, t. XLIX, troisiéme série, Section 1, juin 1955.

Martineau, Armand, "Programme des études au Canada durant la période 1760-1790", *Revue de l'Université d'Ottawa*, vol. 37, no. 2, 1967, pp. 206–230.

Morawetz, Cathleen. "The early history of applied mathematics in Canada". *CMS Notes*, 25, 12–15, 1993.

Ouellet, Damelle. *Adrien Pouliot, un homme en avance sur son temps*, Montréal, 1986.

Peyrard, F., *Les oeuvres d'Euclide, traduites littéralement d'après un manuscrit grec très-ancien, resté inconnu jusqu à nos jours*, Paris: Chez C.-F. Patris, 1819.

Phillips, William. *A new and concise system of arithmetic, calculated to facilitate the improvement of youth in upper Canada. Published by subscription under the patronage of his excellency Sir John Colborne, K.C.B., Lt. Governor of the Province of Upper Canada, Inc. Under the partronage of the honourable and venerable the Archdeacon of York, the reverend Dr. Harris, Principal of the Collège, & c. York*, [1832] Eastwood and Skinner.

Provost, Honorius, "Documents pour une histoire du Séminaire de Québec (Suite), CXXXII–Renseignement sur le Séminaire, 28 décembre 1836," *Revue de l'Université Laval*, vol. XIII, n 7, mars 1959, pp. 665–669.

Robinson, Gilbert de B. *The Mathematics Department in the University of Toronto*, Toronto, University of Toronto Press, 1979.

Sauri, *Institutions mathématiques, servant d'introduction a un cours de philosophie, a l'usage des universités de France*, Paris, quatrième edition, 1786.

Simson, Robert, *Euclid, Elements, First 6 books, with the 11th and 12th, Euclid's Data*, 25th edition, London, 1841.

Strachan, John. *Concise introduction to practical arithmetic: for the use of schools.* Montreal, 1809.

Struik, Dirk J., "Mathematicians at Ticonderoga," *The Scientific Monthly*, vol. 82, no. 5. May, 1956, pp. 236–240.

Struik, Dirk J., "Mathematics in Colonial and Early Republican America", in *Men and Institutions in American Mathematics*, Graduate Studies, Texas Tech. university, No. 13, 1976.

The University of Toronto and Its Collèges, Toronto, The University Library, 1906.

Varkaris, Jane and Costas, *Nathan Fellowes Dupuis, Professor and Clockmaker of Queen's University, and his Family.* Toronto, Ontario Genealogical Society

Walkingame, Francis, *The Tutor's Assistant: Being a Compendium of Arithmetic, and Complete Question-book ... to Which is Added A compendium of Book-keeping*, 51st edition, Montréal: Nahum Mower, 1818.

Wallace, William Stewart. *The MacMillan Dictionary of Canadian Biography*, Toronto, MacMillan, 1963.

Wallace, William Stewart, et al. *The Royal Canadian Institute Centenary Volume, 1849-1949.* Toronto, Royal Canadian Institute, 1949.

Acknowledgments

L.C. would like to thank the personnel of the Archives de l'Université Laval, the Archives de l'École polytechnique and the Archives de l'Université de Montréal for their speed and assistance in searching their respective collections.

T.A. gratefully acknowledges the support of SSHRC during the conduct of the research which led to this paper, and thanks the National Museum of American History, Smithsonian Institution, for kind hospitality. He also wishes to thank Peter Fillmore for provoking this work, and Louis Charbonneau for agreeing to make available his expertise and the results of many years of research. Thanks are also due to Jan Marontate for encouragement and assistance, and to the editors of this volume [8] for the opportunity to republish this still-provisional work.

7

The Emergence of the American Mathematical Research Community

Karen Hunger Parshall *

Departments of History and Mathematics, University of Virginia

Various factors affected the emergence of a community of mathematical researchers in the United States in the closing quarter of the nineteenth century. In our book on this subject, David Rowe and I focus principally on the roles of three men—James Joseph Sylvester, Felix Klein, and Eliakim Hastings Moore—the institutions in which they worked—the Johns Hopkins University, Göttingen University, and the University of Chicago—and their productions—both mathematical and organizational—in the formation of this community. Largely through their work and efforts, we argue, American mathematicians united in common cause to create and disseminate research-level mathematics.

James Joseph Sylvester, a sixty-two-year-old British algebraist seemingly well past his prime in 1876; Felix Klein, a rising German geometer recovering from nervous exhaustion in the mid-1880s; and Eliakim Hastings Moore, a young and unproven American mathematician in the early 1890s—what could these three men, separated by generation, mathematical training, and cultural background, possibly have to do with the emergence of a mathematical research community in the United States between 1876 and 1900?[1]

---

*This text was originally given as a talk in June of 1994 and drew extensively from the manuscript then in press of Karen Hunger Parshall and David E. Rowe, *The Emergence of the American Mathematical Research Community, 1876–1900: J. J. Sylvester, Felix Klein, and E. H. Moore*, HMATH, vol. 8 (Providence: American Mathematical Society and London: London Mathematical Society, 1994). It presents an overview of the argument of that book.

The bibliography given here is necessarily abbreviated. For the complete list of the sources upon which the book was based, see Parshall and Rowe, pp. 455-485. Much additional work on the American mathematical scene in the nineteenth and early twentieth centuries has been done in the decade since 1994; it is not reflected in the bibliography presented here.

I thank the American Mathematical Society for permission to publish the present text.

[1]The substance of this and the next eight paragraphs closely follows the argument in the preface of *ibid.*, pp. ix-xv. Here, and in what follows, for the full range of

Sylvester, after all, worked as an actuary from 1844 to 1855 and taught at the Royal Military Academy in Woolwich from 1855 until his forced retirement in 1870. In neither of these posts did he have an opportunity to train his own countrymen, much less aspiring Americans, in research-level mathematics.[2] Indeed, training at the research level did not even form part of the university mission in nineteenth-century Britain. Felix Klein taught and conducted his mathematical research in Germany: first at Erlangen, next at Munich, then at Leipzig, and finally at Göttingen. These institutions and their respective mathematical traditions were not only geographically remote from late nineteenth-century America but also intellectually far-removed from a country where institutions of higher education functioned primarily at a collegiate—as opposed to a university—level and where basic mathematical research received little encouragement.[3] Finally, Eliakim Hastings Moore, thirty years old when the University of Chicago opened its doors in 1892 with him as acting head of its Department of Mathematics, had received reasonably solid training at Yale, had studied abroad, and had even done a bit of original—if unexciting—research. His own development into a major researcher was, however, by no means assured. He had never taught students at the graduate level, and his youth, inexperience, and Midwestern vantage point all seemed to militate against his chances of becoming a major voice in a community of mathematicians which, insofar as it was discernible at all, had begun to coalesce around the fledgling East Coast undertaking—the New York Mathematical Society—founded with six members in 1888.[4]

On the surface, then, these three men would appear to be unlikely protagonists in the story of the emergence of an American mathematical research community. Moreover, the period in which their impact on American mathematics was most immediate and decisive, the last quarter of the nineteenth century, would seem too early for the detection of significant contributions to higher mathematics from a country known more for its "Yankee ingenuity" than for the cultivation of abstract ideas. Nevertheless, a confluence of historical trends and events made this disparate trio *the* formative figures in the creation of a community of mathematical researchers on American shores in the years from 1876 to 1900. Moreover, the analysis of these trends and events fills a conspicuously large gap in the literature on the history of American science.

In 1986 the book, *Historical Writing on American Science: Perspectives and Prospects*, appeared, owing to an initiative taken by the History of Science Society. Representing the collective effort of over a dozen specialists, this work aimed both to survey the various areas of the history of American science

---

sources upon which the argument and material draw, see the notes on the cited pages of *ibid.*

[2] On Sylvester's early career, see *ibid.*, pp. 59-75.

[3] On Klein's career trajectory, see *ibid.*, pp. 167-187.

[4] For an account of Moore's biography, see *ibid.*, pp. 279-285.

and to suggest fruitful avenues for further research.[5] In the volume's preface, editors Sally Gregory Kohlstedt and Margaret Rossiter explained that by the mid-1980s "[s]everal Americanists were ready to assess the current state of various specialties and to indicate what 'needs and opportunities' remained after more than a decade of significant activity."[6] In their collective assessment, institutional history along with science in medicine, religion, and the federal government constituted the four so-called "classical themes"; the "newer areas" of native American scientific knowledge in addition to science and technology, war, and public policy received special attention; and the history of the scientific specialties of geology, astronomy, chemistry, biology, physics, and the social sciences were singled out for analysis.[7] Notably absent from the specialties treated? Mathematics.

To be sure, the history of mathematics in general and the history of American mathematics in particular have been relatively neglected in the last several decades by the American community of historians of science. The "glory days" of the 1930s and 1940s when George Sarton declared the primacy of mathematics within the history of science from his lofty positivist heights have long since passed. "[T]he history of mathematics should really be the kernel of the history of culture," he wrote in 1937. "Take the mathematical developments out of the history of science, and you suppress the skeleton which supported and kept together all the rest. Mathematics gives to science its innermost unity and cohesion, which can never be entirely replaced with props and buttresses or with roundabout connections, no matter how many of these may be introduced."[8] Yet despite such pronouncements, even during Sarton's era, the history of *American* mathematics and, in fact, the history of American science failed to satisfy the prescripts of a generation of internalist historians of science which largely adhered to a "great name" approach to the discipline.

Since the 1950s and in response initially to the diverse points of view reflected in the work of Alexandre Koyré and Thomas Kuhn, among others, historians of science have increasingly broadened their purview to embrace issues like the impact of philosophical and religious ideas on science, the role of external, social factors in the development of scientific thought, and the interrelations between science and society at large.[9] This changed historio-

---

[5] Sally Gregory Kohlstedt and Margaret W. Rossiter, *Historical Writing on American Science: Perspectives and Prospects* (Baltimore: The Johns Hopkins University Press, 1985).

[6] *Ibid.*, p. 7.

[7] *Ibid.*, pp. 9-15.

[8] George Sarton, *The Study of the History of Mathematics* (Cambridge, MA: Harvard University Press, 1937; reprint ed., New York: Dover Publications, Inc., 1957), p. 4.

[9] See, for example, Alexandre Koyré, *Études galiléennes*, 3 parts, 1935–1939; reprinted in one volume (Paris: Hermann, 1939); and *From the Closed World to the Infinite Universe* (Baltimore: The Johns Hopkins University Press, 1957); as well as Thomas Kuhn, *The Structure of Scientific Revolutions* (Chicago: The University

graphical climate has encouraged a new interest in American science, and especially in those sciences perceived as impinging most upon society. Medicine, biology, and physics (particularly in the twentieth century with relativity theory and the atomic bomb)—sciences like these lend themselves naturally to this sort of analysis. Mathematics, however, with its abstruse language and arcane symbolism, with its seemingly insulated practitioners, and with its unapparent impact on public policy or daily life, seems uncongenial—and so uninteresting—within such an historiographical framework.[10] While this explanation may shed some light on the relative indifference toward the history of American mathematics from the 1970s through the 1980s, a period characterized by Kohlstedt and Rossiter as one of "significant activity" in the history of American science, it fails to illuminate the deeper aspects of the relation of the subdiscipline of the history of mathematics to the history of science as a whole.

In a 1990 position paper published in *History of Science*, Ivor Grattan-Guinness stated the case bluntly: the history of mathematics has been largely ignored, making it one of the least developed subdisciplines within the history of science.[11] In his view, historians of science have failed to treat the historical development both of mathematics *per se* and of mathematics as related to other sciences due to a fundamental fear of the subject, their mostly empty discussion of notions like "mathematization" aside. As he put it, "Historians of Science, like most of the population, do not like mathematics, or at least find nothing particularly interesting or appealing in it."[12] As for the other constituency that might have advanced the subdiscipline, the mathematicians, Grattan-Guinness contended that "[e]ven those ... who become somewhat interested in history usually assert its importance only for trivial reasons of anecdote and general heuristic without consideration of basic questions of historiography. Further and more importantly," he continued, "they usually view history as the record of a 'royal road to me'—that is, an account of how a particular modern theory arose out of older theories instead of an account of those older theories in their own right."[13] While, as he fully admitted, counterexamples to this general assessment of the recent historiography of the

of Chicago Press, 1962). For a concise and cogent discussion of these developments within the history of science, see Allen G. Debus, *Science and History: A Chemist's Appraisal* (Coimbra: Serviço de Documentação e Publicações da Universidade de Coimbra, 1984), pp. 17-33.

[10]This refers to the naïve perception of pure mathematics. It goes without saying that applied probability and statistics have affected public policy in key and obvious ways. Perhaps due to their relatively short histories in the United States, not even these areas of the history of American mathematics have received much attention from historians of science.

[11]Ivor Grattan-Guinness, "Does the History of Science Treat of the History of Science? The Case of Mathematics," *History of Science* 28 (1990):149-173.

[12]*Ibid.*, p. 155.

[13]*Ibid.*, p. 157.

history of mathematics certainly exist, they are indeed rare. As he claimed, the neglect of the historians of science coupled with the mostly ahistorical approach of the mathematicians have rendered the history of mathematics "a classical example of a ghetto subject: too mathematical for historians and too historical for mathematicians."[14]

Grattan-Guinness's analysis of the state of the history of mathematics within the history of science in particular, and within history in general, goes far to explain the absence of a discussion of the history of mathematics in Kohlstedt and Rossiter's book. In light of the book's prefatory remarks, the omission of mathematics from this otherwise competent study of the historiography of American science thus raises several obvious questions. Was there no one "ready to assess the current state" of the history of American mathematics? Were there no "needs and opportunities" for historical research in this scientific discipline? Was there simply no history of American mathematics to survey in the mid-1980s? Or, perhaps, were the historians of American science merely ignoring the case of mathematics?

Like the history of American astronomy, which did receive treatment in *Historical Writing on American Science*, the history of American mathematics—although just as clearly "underdeveloped"—was not nonexistent in the mid-1980s.[15] In the massive, three-volume collection of original and reprinted essays published in 1988 and 1989 to commemorate the centenary of the American Mathematical Society, Uta Merzbach provided much of Kohlstedt and Rossiter's "missing chapter" on mathematics.[16] Although she treated the period from 1969 to the late 1980s cursorily, Merzbach discussed in some detail the work done from 1890, when Florian Cajori published his book, *The Teaching and History of Mathematics in the United States*, through the 1930s, when David Eugene Smith and Jekuthiel Ginsburg produced *A History of Mathematics in America before 1900*, and up to the 1960s, when Clifford Truesdell and Kenneth O. May conceived journals to promote and encourage high-level research in the history of mathematics irrespective of geographical and political boundaries.[17] While much of the research Merzbach surveyed did not deal exclusively with the history of *American* mathematics, her article made clear that there was a small but extant body of research which, if surveyed,

---

[14] *Ibid.*, p. 158.

[15] Marc Rothenberg, "History of Astronomy," pp. 117-131 on p. 131 in Kohlstedt and Rossiter.

[16] Uta C. Merzbach, "The Study of the History of Mathematics in America: A Centennial Sketch," in *A Century of Mathematics in America–Part III*, ed. Peter Duren *et al.* (Providence: American Mathematical Society, 1989), pp. 639-666.

[17] Florian Cajori, *The Teaching and History of Mathematics in the United States* (Washington, D. C.: Government Printing Office, 1890); and David Eugene Smith and Jekuthiel Ginsburg, *A History of Mathematics in America before 1900* (Chicago: Mathematical Association of America, 1934; reprint ed., New York: Arno Press Inc., 1980). Truesdell founded the *Archive for History of Exact Sciences* while May began *Historia Mathematica*.

would have highlighted precisely those questions and areas begging for further study and analysis. Moreover, a closer look at the works she cited—such as the volume by Smith and Ginsburg—underscored the fact that the existing historical literature on American mathematics suffered from a paucity of archival sources, a near total absence of substantive discussions of the mathematics actually produced by the Americans, and a failure to situate this research within the broader context of the history either of mathematics or of American science. The book, *The Emergence of the American Mathematical Research Community, 1876–1900*, thus aims to lay a solid foundation for further research by defining and documenting one crucial process and one key period in the history of American mathematics, namely, the emergence of a mathematical research community in the United States between 1876 and 1900.

The notion of periodization inherent here is central to the argument. To focus on the *period* from 1876 to 1900 explicitly draws the boundaries of two other periods in the historical development of American mathematics. In the first period, the century from 1776 to 1876, mathematics evolved not as a separate discipline but rather within the context of the general structure-building of American—as opposed to colonial—science.[18] The colleges formed a primary locus of scientific activity, but, by and large, they did little to encourage the pursuit of research for the advancement of science. At the same time, however, the concept of research in American science—as in other academic disciplines—emerged as scientists looked toward Europe as their model and measured themselves against the yardstick of European scientific achievement. Indeed, the Americans could point to the work of Nathaniel Bowditch in celestial mechanics; the scientific accomplishments at the United States Coast Survey of Superintendents Ferdinand Hassler, Alexander Dallas Bache, and Benjamin Peirce; the influence of engineering mathematics at West Point with its curriculum modeled on that of the *École polytechnique*; and the astronomical research of Simon Newcomb and George William Hill. However, few of the achievements associated with these institutions and individuals had any lasting effect on the generation that followed in the last quarter of the nineteenth century.

Quite simply, prior to 1876, nothing even remotely resembling a *mathematical research* community existed in the United States, nor did the time appear ripe for its imminent emergence. Rather, the century from 1776 to 1876 witnessed the formation of an American *scientific* community, which, loosely characterized, earned its living primarily through undergraduate teaching (although to some extent also through federal-governmentally-supported jobs) but which defined itself in terms of its extracurricular research.[19] General sci-

[18] For a fuller discussion of this period, see Parshall and Rowe, pp. 1-51.

[19] For more on the issues discussed in this and the next paragraph, see Parshall and Rowe, pp. 33-49. See also John C. Greene, *American Science in the Age of Jefferson* (Ames: The Iowa State University Press, 1984); George H. Daniels, *American Science in the Age of Jackson* (New York: Columbia University Press, 1968); and

entific societies and their publications, like the American Association for the Advancement of Science and the National Academy of Sciences, provided the communications outlets for the scientific community, since critical numbers of practitioners of the individual sciences did not yet exist to sustain specialized societies or journals.[20] On the educational front, colleges broke from the confines of the colonial era by expanding their faculties with scientists and their curricula with the sciences.[21] Concomitantly, the traditional mathematics curriculum, which had largely been restricted to Euclid's *Elements*, incorporated pedagogical innovations issuing mostly from France and began to include the calculus, among other topics. These changes within higher education, however, did not imply the existence of institutional support for or an encouragement of basic scientific research. In fact, the lack of support for research within the institutions of higher education fundamentally distinguishes the periods before and immediately after 1876.

Before that time, when it was fostered at all, research was promoted primarily within the federal government—in agencies like the Coast Survey and the Nautical Almanac Office—but only exceptionally within the colleges.[22] As a result, the research done had, by and large, an applied flavor. The accomplishments of Hill and others notwithstanding, American mathematics, as it unfolded after 1876, had little in common with the research "tradition" of the previous era. The next generation—associated with institutions of higher education in which departmental structures discouraged the kind of cooperation generally needed for applied research—focused its attention almost exclusively on the pure side of the mathematical spectrum, rather than pursuing areas like celestial mechanics as had Bowditch, Benjamin Peirce, and Hill. Moreover, its leading figures reinforced their mathematical predilections by forging a viable community during this period, which successfully incorporated research-level mathematics into the intellectual fabric of the country.

As noted, three men and the institutions within which they worked largely shaped this second period: the Englishman James Joseph Sylvester, at The Johns Hopkins University; the German Felix Klein, first from Leipzig but more crucially from Göttingen; and the American Eliakim Hastings Moore, at the University of Chicago.[23] Indeed, their respective periods of involve-

---

Robert V. Bruce, *The Launching of Modern American Science: 1846–1876* (New York: Alfred A. Knopf, 1987).

[20]See, for example, Sally Gregory Kohlstedt, *The Formation of the American Scientific Community: The American Association for the Advancement of Science 1848–60* (Urbana: University of Illinois Press, 1976).

[21]Stanley Guralnick made this argument in "The American Scientist in Higher Education: 1820–1910," pp. 99-141 in *The Sciences in the American Context: New Perspectives*, ed. Nathan Reingold (Washington: Smithsonian Institution Press, 1979).

[22]On the government's role in the support of science, see A. Hunter Dupree, *Science in the Federal Government: A History of Policies and Activities* (Baltimore: The Johns Hopkins University Press, 1986).

[23]This paragraph follows Parshall and Rowe, pp. xiv-xv.

ment in American mathematics clearly define three distinct phases centered around three separate programs. An analysis of these programs and of the nature of the influence of their progenitors reflects not only the progressive deepening of research standards and output in the United States throughout the twenty-five-year period but also the generational differences separating Sylvester, Klein, and Moore. It also highlights, by focusing on the students and research issuing from these institutions, the process of maturation of an American mathematical research community which had fully emerged by 1900.

A British import, Sylvester assumed the first professorship in mathematics in 1876 at The Johns Hopkins University, an institution pivotal in the history of higher education in America and, by extension, in the history of American mathematics.[24] Unlike other, older colleges and universities in the United States, Hopkins emphasized graduate education, although it did not neglect the important function of the undergraduate college as a sort of feeder into its graduate programs. Having as two of its primary goals the training of future researchers and the maintenance of high levels of research productivity among its faculty, this institution departed radically from the traditional model of the undergraduate teaching college.

At Hopkins, Sylvester was able to continue his own lines of inquiry unfettered by the heavy teaching loads and burdensome duties that encumbered most of his colleagues around the country.[25] Motivated—and often talented—students accepted graduate fellowships and came to Baltimore determined to do original work under his guidance. In mathematics, that work tended to reflect Sylvester's interests in the theories of invariants, partitions, and algebras, but geometry and mathematical logic also found their respective proponents in William Story, whom Sylvester had stolen from Harvard to be his Teaching Associate, and in the mathematician-logician-philosopher Charles Peirce. Likewise, the work tended to reflect Sylvester's own spontaneous approach to mathematical research. In conducting his so-called Mathematical Seminarium, Sylvester challenged his students to fill in the details—big and small—in his presentations and then to present their findings before the group in polished form. Not all of the research produced in this environment was exciting; not all of it was important; but most of it was solid; and some of it was genuinely remarkable. Of real importance, however, is not so much the *quality* of the research as the evidence of the conviction, shared by the members of the Department of Mathematics (and by the University as a whole for that matter), of the importance and primacy of the production of new knowledge. Thanks

[24]On The Johns Hopkins University and its importance for the development of research-level mathematics in the United States, see Parshall and Rowe, pp. 53-58 and 75-94. See also Hugh Hawkins, *Pioneer: A History of the Johns Hopkins University, 1874–1889* (Ithaca: Cornell University Press, 1960).

[25]On the research done at The Johns Hopkins during Sylvester's tenure there, see Karen Hunger Parshall, "America's First School of Mathematical Research: James Joseph Sylvester at The Johns Hopkins University, 1876–1883," *Archive for History of Exact Sciences* 38 (1988):153-196, and Parshall and Rowe, pp. 99-146.

to a singular intellectual environment and a common drive towards an ideal of research, the Department of Mathematics at The Johns Hopkins defined America's first school of research mathematics.

After Sylvester returned to Britain in 1883 to take up the Savilian Chair of Mathematics at Oxford's New College, prospective American mathematicians suffered from the absence, on American shores, of mathematical instruction comparable to that available on the Continent.[26] While a dozen or so schools—among them, Harvard, Yale, Princeton, Cornell, and the Universities of Virginia, North Carolina, Texas, Michigan, and Wisconsin—offered programs ostensibly at the graduate level, none of these institutions yet had a staff of researchers teaching or working at the frontiers of mathematics.[27] The Hopkins-trained students emerged from their mathematically charged environment only to find themselves unable to maintain their research momentum in other home institutions. Sylvester himself recognized and commented on this deplorable state of affairs in his farewell address to the Hopkins community. Rhetorically questioning his audience, he asked:

> What happens to them? They are absorbed by inferior though valuable colleges and institutions, and their work droops. They write to me or to their friends, "We miss the stimulus of the Johns Hopkins." What a great thing it would be if means were found for providing traveling scholarships or Fellowships for a year or two, that they might prolong their studies, and come in contact with scientific men and science in England and on the Continent of Europe.[28]

Of course, Sylvester offered only a stopgap solution to the problem. His students would have returned from their trips abroad only to face the same mathematical isolation. For example, one of his students, William Durfee, took a professorship of mathematics at teaching-intensive Hobart College in Geneva, New York in 1884, became Dean in 1888, and dropped from the research ranks. A similar fate befell Sylvester's two number-theoretically-oriented students. Oscar Mitchell accepted a position at another small college, Marietta College in Ohio, and George Ely became an examiner at the United States Patent Office. As late as 1888, Sylvester's last student, Ellery William Davis, *was* the Mathematics Department at the University of South Carolina and

[26]This and the next paragraph follow Parshall and Rowe, pp. 144-145.

[27]See Cajori for descriptions of these various programs, and Karen Hunger Parshall, "A Century-Old Snapshot of American Mathematics," *The Mathematical Intelligencer* 12 (3) (1990):7-11.

[28]Remarks of Prof. Sylvester at a Farewell Reception Tendered to him by the Johns Hopkins University, Dec. 20, 1833 (Reported by Arthur S. Hathaway)," 24 typescript pages, Daniel Coit Gilman Papers Ms. 1, Special Collections Division, The Milton S. Eisenhower Library, The Johns Hopkins University. For the quote, see p. 12.

tried, but failed, to institute a graduate program there before moving on to somewhat greener pastures at the University of Nebraska in 1893.[29]

To get a sense of what it was like for these students, consider the case of Sylvester's student Arthur Hathaway at Cornell. By the standards of the day, Cornell supported a huge mathematics faculty at seven strong, counting all ranks from professor to instructor.[30] The leader of this group, James Edward Oliver, had graduated in 1849 from a midcentury Harvard dominated mathematically by Benjamin Peirce and had returned in the mid-1850s to pursue Peirce's advanced course in mathematics through the Lawrence Scientific School.[31] Earning his living in Cambridge at the federally supported Nautical Almanac Office, Oliver entered the academic ranks in 1871 when he accepted the assistant professorship of mathematics at Cornell. Only two years later, he assumed the mathematical chair and directed Cornell mathematics from this vantage point until his death in 1895. A modest and unassuming man, Oliver pursued mathematics more for his own pleasure than for the reputation publication might have brought, yet he was fully attuned to the growing importance and desirability of publication within the emergent mathematical community. In Oliver's words,

> [w]e [at Cornell] are not unmindful of the fact that by publishing more, we could help to strengthen the university, and that we ought to do so if it were possible. Indeed, every one of us five is now preparing work for publication or expects to be doing so this summer, but such work progresses very slowly because the more immediate duties of each day leave us so little of that freshness without which good theoretical work can not be done.[32]

The duty that sapped their energies most completely was teaching.

During the 1886–1887 academic year, Oliver, together with his four colleagues, Associate Professor Lucien Augustus Wait, Assistant Professor George William Jones, and Instructors James McMahon and Arthur Hathaway, taught an average of seventeen to twenty hours each week. The next year two more Instructors, Duane Studley and George Egbert Fisher, joined the staff, but this 40% increase in the teaching faculty hardly lessened the burden. According to Oliver, "our department's whole teaching force, composed of only about one-eleventh of all resident professors, has to do about one-ninth of all the teaching in the University."[33]

In spite of this load, which Oliver clearly viewed as inequitable, he and his colleagues managed to make a fair showing publication-wise in 1887–1888. Oliver seemed quite proud to report that

---

[29] Parshall, "A Century-Old Snapshot of American Mathematics," pp. 9-10.

[30] On the situation at Cornell, see *ibid.*, and Parshall and Rowe, pp. 269-271.

[31] For an idea of the advanced nature of Peirce's curriculum, see Cajori, pp. 137-138, reproduced as Table 1.1 in Parshall and Rowe, p. 50.

[32] Cajori, p. 180.

[33] *Ibid.*, p. 186.

> Professor Oliver has sent two or three short articles to the [*Annals of Mathematics*], and has read, at the National Academy [of Sciences'] meeting in Washington, a preliminary paper on the Sun's rotation, which will appear in the *Astronomical Journal.* Professor Jones and Mr. Hathaway have lithographed a little *Treatise on Projective Geometry.* Mr. McMahon has sent to the [*Annals*] a note on the circular points at infinity, and has also sent to the *Educational Times*, London, solutions (with extensions) of various problems. Other work by members of the department is likely to appear during the summer, including a new edition of the *Treatise on Trigonometry.*[34]

The latter work comprised part of the popular series of textbooks by Oliver, Wait, and Jones designed primarily for use in the college classroom. Thus, the Cornell faculty, although perhaps more active in textbook writing than in original research, was nonetheless alive mathematically. In Oliver's view, only a sufficiently high level of vitality would successfully attract that increasingly desirable entity—the graduate student—to the department. Apparently, he and his colleagues attained the necessary level, for their program attracted eleven graduate students in the 1887–1888 academic year.

Relatively speaking, then, the situation was positively rosy at Cornell, but Cornell was the exception rather than the rule. In the 1880s, the United States simply did not yet support the critical mass of mathematicians necessary for the sustenance of a specialized research community, and with Sylvester's departure to England, it no longer had the means to train such a community's membership effectively. With unencouraging educational prospects at home, Americans turned to Europe, and particularly to the lecture halls of Felix Klein, for their mathematical training. For roughly a decade following Sylvester's departure, Klein actively served as the mathematical standard-bearer for the United States. Why was Klein the main conduit for the sudden transfusion of abstract mathematics in the German style that so decisively enlivened the fledgling community of American mathematicians? An understanding of this requires penetrating beyond domestic factors and external causes to the man himself and the unusually rich sources that defined and shaped his career.[35]

Klein's mathematics embodied many of the ideals characteristic of German scholarship in the nineteenth century.[36] Even from his youth, he sought to attain a unified conception of mathematical knowledge that embraced the

[34] *Ibid.*, p. 181. See James Oliver, Lucien Wait, and George Jones, *A Treatise on Trigonometry*, 4th ed. (Ithaca: G. W. Jones, 1890).

[35] On Klein and his influence in training American mathematicians, see Parshall and Rowe, pp. 147-259.

[36] This and the next paragraph follow Parshall and Rowe, pp. 147-148. See also David E. Rowe, "The Early Geometrical Works of Sophus Lie and Felix Klein," in *The History of Modern Mathematics*, ed. David E. Rowe and John McCleary, 2 vols. (Boston: Academic Press, 1989), 1:209-273.

achievements of his predecessors. To this end, he strove for and attained an extraordinary breadth of knowledge, much of which he acquired in discussions with colleagues and friends. A master of give-and-take, he cultivated scientific relations with many of the leading mathematicians of his day and then imparted the ideas so gained to the students in his lecture courses. Furthermore, he freely shared the hard-won insights which allowed him to capture the essence of a mathematical theory. Klein's approach clearly met with success, for by the age of thirty, he had already begun to attract talented students from outside of Germany, many of whom—like Maxime Bôcher and William Fogg Osgood at Harvard—went on to prominent positions as scholars and leaders within their respective scientific communities.

In contrast to Sylvester, whose mathematical style was dominated by complex computations and directed towards fairly restricted problems within specialized branches of algebra, Klein tended to soar above the terrain that occupied ordinary workaday mathematicians to take in vast expanses of mathematical knowledge. His one glaring weakness, as Richard Courant once put it, was that he often found it difficult to land his plane.[37] Klein had no patience for thorny problems that required abstruse technical arguments. For him, what counted was the "big picture," and he drew it largely from the work of his forerunners—Gauss, Abel, Galois, Riemann, and Weierstrass—as well as from leading contemporary figures, including Lie, Schwarz, and Dedekind. Throughout the course of his career, Klein made important contributions to geometry, group theory, Riemannian function theory, Galois theory, rigid-body mechanics, and even general relativity theory. In his own mind at least, all of this seemingly disparate work was of a piece, and, what is more, he ultimately viewed it as largely embedded within the mathematical tradition associated with Göttingen University. Thus, to learn mathematics from Felix Klein meant gaining an overview of a substantial part of the deepest parts of nineteenth-century mathematics as put forth by some of its foremost proponents. Furthermore, it meant participating in the challenging and highly competitive seminar setting.

Klein generally organized his seminars by choosing a subject closely related to the content of his lecture courses.[38] In sharp contrast to Sylvester, Klein tended to orchestrate every detail of his seminar in advance and to unveil his overall plan during the seminar's initial session. At that time, too, he discussed his goals and assigned the specific topics that he wanted his individual students to research and present in a series of formal lectures. Before the student could step to the podium, however, he or she had to meet with Klein in order to discuss the content of the lecture. Although himself a born pedagogue, Klein believed that good teaching was an acquired characteristic that could be inculcated in properly receptive minds. Motivated by this belief as well

[37] Richard Courant, "Felix Klein," *Die Naturwissenschaften* 37 (1925):765-772 on p. 772.

[38] This paragraph follows Parshall and Rowe, p. 191.

as by his deeply ingrained sympathy for holistic solutions, he conducted his seminars in accordance with the principle that the activities of teaching and learning must be viewed as integrally related, and his students, both American and otherwise, imbibed this philosophy.

Unlike Sylvester, however, who trained American students in the United States, Klein taught in Germany.[39] His influence on the emergent mathematical research community in the United States was nevertheless far more pervasive than that of Sylvester. No fewer than six of Klein's students went on to become President of the American Mathematical Society, and thirteen served as Vice President.[40] None of these Americans, though, perpetuated the distinctly Kleinian approach to mathematics. In later years, in fact, Klein expressed some bitterness that the research he and his contemporaries had cultivated, and which had captivated the attention of his American students in Göttingen, had fallen out of fashion. This had partly to do with the intensely personal nature of Klein's mathematical vision. As one commentator remarked, "few of Klein's ... contemporaries were willing to assimilate his singularly personal methods. ... Klein's mathematics demanded too much knowledge of too many things for mastery in a reasonable time, and in addition it frequently presupposed a facility in spacial [sic] linguistics beyond the capacities of most mathematicians."[41]

In sum, Klein's influence on his American students had less to do with his specific research program than with his general ability to inspire them and to train them to do mathematical research. Indeed, to many Americans, Klein represented an emissary of mathematical culture at large, and it was that culture that they very much wanted to transplant to the United States. In 1893, E. H. Moore and his Chicago colleagues helped secure Klein as the keynote speaker at the Mathematical Congress associated with the World's Columbian Exposition in order to bring that avatar of mathematical culture directly to American shores. Klein's appearance at this event, however, foreshadowed not only his desire to step back from his involvement in the American scene but also the appearance of E. H. Moore as the third formative figure in research mathematics in the United States.[42]

Together with his colleagues, Oskar Bolza and Heinrich Maschke, Moore established an environment at Chicago conducive to the training of research-level mathematicians.[43] Like Sylvester at Hopkins, Moore worked within an

[39]This and the next paragraph follow Parshall and Rowe, p. 253-254.

[40]See Table 5.3 in Parshall and Rowe, p. 259.

[41]Eric Temple Bell, *The Development of Mathematics* (New York: McGraw-Hill, 1945), pp. 511-512.

[42]On the Mathematical Congress at the World's Columbian Exposition, see Parshall and Rowe, pp. 295-330. See also Karen Hunger Parshall and David E. Rowe, "Embedded in the Culture: Mathematics at the World's Columbian Exposition of 1893," *The Mathematical Intelligencer* 15 (2) (1993):40-45.

[43]On Moore and the Chicago program in mathematics, see Karen Hunger Parshall, "Eliakim Hastings Moore and the Founding of a Mathematical Community in

institutional setting which greatly facilitated this achievement. At the undergraduate level, he helped fashion a program which would not only expose the student body to mathematics as part of general culture but also prepare the mathematically inclined for more advanced work at the graduate level. For those already at this higher stage, he worked to put together a broad range of courses designed to bridge the gap between studying mathematics and doing mathematical research. As researchers, the Chicago mathematicians embraced many of the leading areas of late nineteenth- and early twentieth-century mathematics.[44] Moore, who had started his career working in algebraic geometry, had switched into group theory by the 1890s, and had moved into axiomatics and the foundations of analysis by the turn of the century. His colleague, Bolza, initially focused on the theories of hyperelliptic and elliptic integrals and functions, spending the 1890s in such pursuits before shifting into the calculus of variations after 1901. Finally, Maschke had first devoted his energies to the theory of finite linear groups before taking up the invariant theory of differential forms from 1900 until his death in 1908. All three mathematicians shared their work and their new ideas with their advanced students and guided them to open problems. Yet, it was E. H. Moore who, sniffing the changing mathematical winds, uncannily and successfully shifted his mathematical course and brought generations of students—including Leonard E. Dickson, Oswald Veblen, Robert L. Moore, and George D. Birkhoff—along in his wake. Finally, at the post-doctoral level, he and his colleagues strove to heighten research activity both within the department, through the regular meetings of their Mathematical Club (an institution reminiscent of Klein's seminar), and nationally, through their involvement particularly in the American Mathematical Society.[45]

Sylvester, Klein, and Moore, looking at these men within their respective institutional settings allows for an examination of the environmental factors that encouraged them to train students at the research level as well as to pursue their own research objectives. Furthermore, an analysis of the training of this first research-oriented generation of American mathematicians provides an illuminating look at the different styles these men employed in training young men and women to do mathematical research as well as at the research so produced. Viewing the students' research output against the backdrop of late-nineteenth-century mathematics highlights the great strides made in mathematical research in the United States between 1876 and 1900. In a key shift that partially delineates the first and second periods in the history of American mathematics, academic institutions in the United States, led by

America, 1892–1902," *Annals of Science* 41 (1984):313-333, and Parshall and Rowe, pp. 275-294 and 363-372.

[44]On the research that issued from the Chicago department in its early years, see Parshall and Rowe, pp. 372-401.

[45]On the broader community activiism of the Chicago department, see Parshall, "Eliakim Hastings Moore and the Founding of a Mathematical Community in America," and Parshall and Rowe, pp. 401-419.

The Hopkins, slowly adopted and adapted a research ethic that had become firmly entrenched in German higher education early in the nineteenth century. This translated into an increasing emphasis on—and an increasing production of—research as an officially sanctioned and supported endeavor in the emergent university setting and, by intimate association, in the emergent American mathematical community.

Although the training of mathematicians at the research level represented a critical ingredient in the emergence of this community, the formation of a *community*—an interacting group of people linked by common interests—required more than just advanced training in mathematics.[46] Various organizational activities proved crucial in forging the requisite communications links. In 1878, Sylvester founded, under the auspices of The Johns Hopkins University, the *American Journal of Mathematics*, the oldest research-level mathematics journal in the United States; and in 1899, E. H. Moore established the *Transactions of the American Mathematical Society*. Neither of these journals suffered for lack of high-quality material for publication, in contradistinction to the numerous failed attempts at research-level mathematical periodicals before 1876.[47] As already mentioned, the American Mathematical Society was founded as the New York Mathematical Society in 1888 and grew from a membership of six to over two hundred in three years. The first major mathematics meeting was organized by E. H. Moore and his Chicago colleagues in 1893 as the Mathematical Congress of the Chicago World's Columbian Exposition. On that occasion, Felix Klein served not only as the keynote speaker but also as an official cultural emissary of the Prussian government at the fair. He brought with him contributed papers from some of the most influential mathematicians in Germany, and the meeting attracted an audience totaling some forty-five Americans. Following the Congress, Klein proceeded to Northwestern University, where the same Chicago contingent had organized the so-called Evanston Colloquium, the first research-level colloquium on American shores, again with Klein as the featured speaker.[48] There, he gave a two-week-long series of ten lectures attended by some two dozen specialists or specialists-to-be, in which he surveyed the mathematical landscape of the late nineteenth century and propagandized for his own unique vision of and approach to the subject. As a result of the late-nineteenth-century conjunction of these and other innovations and their innovators, of changed attitudes as to the value and desirability of research at both an individual and institutional level, and of the existence of a critical mass of practitioners, an American mathematical research community had emerged by 1900.

[46]For this definition, see Parshall and Rowe, p. xvi.

[47]On these journals, see Parshall and Rowe, pp. 88-94 and 411-415, respectively. A list of America's failed journals appears in Table 1.2 in *ibid.*, p. 51.

[48]For more on the Evanston Colloquium and Klein's role in it, see Parshall and Rowe, pp. 331-361.

As this last statement suggests, the community of research-level mathematicians clearly extended beyond the immediate spheres defined by Sylvester, Klein, and Moore.[49] However, using their careers as a point of departure, it is possible to examine meaningfully a broad spectrum—although by no means the complete roster—of participants in the field at the turn of the twentieth century. An identification of their students and, to some extent, their students' students and colleagues penetrates beyond the *crème de la crème* to underscore the existence of an extended population of mathematicians—some talented and some not, some well-known and some obscure—based at colleges and universities throughout the country who actively shared an interest in mathematics at the research level.

In summary, three factors that had been altogether absent before 1876 serve to define the emergent period of research-level mathematics in America.[50] The first—the founding of research-oriented universities beginning with The Johns Hopkins in 1876, followed by Clark University in 1889, and culminating with the University of Chicago in 1892—reflected a fundamental change in the American academic climate. Within these new institutions, research in pure mathematics as well as in other fields attained a degree of credibility and support previously lacking. Moreover, the existence and example of these schools prompted the older institutions—Harvard, Columbia, Yale, and Princeton—to develop viable graduate programs during this era or soon thereafter. It had quickly become clear to leading American educators, many of them inspired by the German university model, that graduate training and research went hand in hand, and those institutions that pursued this policy soon established a competitive edge within a fast-changing academic environment.

As a measure of the rapidity of the transformation that ensued, consider the following statistics on the number of doctoral degrees in mathematics earned by Americans between 1875 and 1900.[51] Before 1875, American universities had conferred a total of only six degrees in the field. During the next fifteen years, thirty-nine Americans took doctorates in the United States, and another fifteen earned their degrees abroad. These figures were dwarfed again by those of the final decade of the century, which witnessed a total of 107 new Ph.D.s in mathematics, eighty-four of them earned at home.

A second factor crucial in the emergence of the American mathematical research community during the last quarter of the nineteenth century was the founding of the New York Mathematical Society in 1888. In the wake of the Chicago Congress and the prototypic Evanston Colloquium, the organization assumed national dimensions and changed its name to the American

[49]Compare Parshall and Rowe, pp. xvi-xvii.

[50]This and the next four paragraphs follow Parshall and Rowe, pp. 429-431.

[51]R. G. D. Richardson, "The Ph.D. Degree and Mathematical Research," *American Mathematical Monthly* 43 (1936):199-215; reprinted in *A Century of Mathematics in America–Part II*, ed. Peter Duren *et al.* (Providence: American Mathematical Society, 1989), pp. 361-378 on p. 366.

Mathematical Society (AMS) in 1894. Spurred by E. H. Moore and Henry White, a group of Midwestern mathematicians established the Chicago Section of the AMS in 1897. Following the Chicago lead, twenty West Coast Society members—including Sylvester's student W. Irving Stringham and the Klein-trained Mellon W. Haskell—met and founded a San Francisco Section of the Society in May of 1902. By December of 1906, the Göttingen Ph.D. Earle Raymond Hedrick, the Russian-born Alexander Chessin, and another Sylvester student, Ellery Davis, among others, had met in Columbia, Missouri and had formed a Southwestern Section initially numbering some thirty-five strong. Thus, from the mid-1890s onward, the AMS served as the principal organizational vehicle for meetings, colloquia, and other activities of interest to the budding community of research mathematicians across the country. By the turn of the century, it had also initiated its *Transactions* as a complement both to its *Bulletin* (begun in 1891) and to the two older, research-oriented periodicals, the *American Journal of Mathematics* and the *Annals of Mathematics*, the latter founded by astronomer Ormond Stone at the University of Virginia in 1884. Thus, by the end of this pivotal period, the United States had already developed sufficient research capacity to support four major mathematical journals.

While it would be difficult to overestimate the role the American Mathematical Society played in promoting mathematical research in the United States, its early success depended on a third factor: a generation of Americans not only interested in mathematics but also possessing the requisite knowledge to educate their successors. The older members of this group necessarily turned to European scholars, notably J. J. Sylvester and Felix Klein, for their training and inspiration. Whether they studied in Baltimore, Göttingen, Paris, or Leipzig, this generation went on to establish a new standard for mathematics instruction at colleges and universities in every region of the United States. Led by E. H. Moore at Chicago, a few select institutions—most notably Harvard and Princeton—also developed solid doctoral programs, the graduates of which shaped and directed American mathematics in the opening decades of the twentieth century.[52]

These three factors interacted to transform what amounted to a few scattered pockets of mathematical expertise into a cohesive and extensive mathematical community. Quantitatively speaking, from 1891 to 1906 or during the first fifteen years of publication of the American Mathematical Society's *Bulletin*, over 1,000 geographically dispersed participants at various levels of interest and activity formed a pyramidal community broadly based upon more than 500 interested—if not active—participants and tapering to an apex of some sixty highly active and productive researchers.[53] Within this community, those most active in research spanned the landscape of pure mathematics

[52] For a sketch of these developments, see Parshall and Rowe, pp. 432-453.

[53] See Della Dumbaugh Fenster and Karen Hunger Parshall, "A Profile of the American Mathematical Research Community, 1891–1906," pp. 228-261 in *The His-*

from algebra to analysis to geometry, while tending to favor the subdisciplines of group theory, the theory of automorphic functions, the calculus of variations, real and complex function theory, classical projective, algebraic, and differential geometry, and foundational studies. (Applied areas, although represented, held less attraction for turn-of-the-century American research mathematicians, due, in part, to the fact that they drew their inspiration from pure mathematicians, like Sylvester, Klein, and E. H. Moore.) This broad support of pure mathematics at the research level that solidified from 1876 to 1900 undoubtedly contributed to the sustained activity in American mathematics that characterized the ensuing period of consolidation and growth from 1900 to 1933. The final quarter of the nineteenth century thus marks a true watershed in the history of American mathematics. The major institutional structures and research traditions of American mathematics stood firmly in place by the end of the first decade of the twentieth century. No dramatic new qualitative changes affected the community's overall contours until the influx of European refugees began in the mid-1930s. This would mark the beginning of a fourth period in the history of American mathematics—from 1933 to roughly 1960—characterized by this infusion of Europeans as well as by the large-scale governmental funding of basic research during and after the Second World War.

Just as internal institutional and external international influences affected the shift in the historical development of American mathematics from its third period of quiet consolidation and expansion to the tumultuous era that followed, so similar forces had shaped the crucial second period during which a self-sustaining American mathematical research community emerged.[54] Individuals both at home and abroad, educational institutions both domestic and foreign, general developments in science and its social and cultural status, broader philosophies of education, political rivalries, and the encroachment of modernity in its several guises, these were among the components of the matrix from which research-level mathematics evolved in the United States during the last quarter of the nineteenth century.

Selected References

Bell, Eric Temple. *The Development of Mathematics.* New York: McGraw-Hill, 1945.

Bruce, Robert V. *The Launching of Modern American Science: 1846–1876.* New York: Alfred A. Knopf, 1987.

Cajori, Florian. *The Teaching and History of Mathematics in the United States.* Washington, D. C.: Government Printing Office, 1890.

---

tory of Modern Mathematics, vol. 3, ed. Eberhard Knobloch and David E. Rowe (Boston: Academic Press, Inc., 1994).

[54]This paragraph follows Parshall and Rowe, p. 453.

Courant, Richard. "Felix Klein." *Die Naturwissenschaften* 37 (1925):765-772.

Daniels, George H. *American Science in the Age of Jackson.* New York: Columbia University Press, 1968.

Debus, Allen G. *Science and History: A Chemist's Appraisal.* Coimbra: Serviço de Documentação e Publicações da Universidade de Coimbra, 1984.

Dupree, A. Hunter. *Science in the Federal Government: A History of Policies and Activities.* Baltimore: The Johns Hopkins University Press, 1986.

Fenster, Della Dumbaugh and Parshall, Karen Hunger. "A Profile of the American Mathematical Research Community, 1891–1906." Pp. 228-261. In *The History of Modern Mathematics.* Vol. 3. Ed. Eberhard Knobloch and David E. Rowe. Boston: Academic Press, Inc., 1994.

Grattan-Guinness, Ivor. "Does the History of Science Treat of the History of Science? The Case of Mathematics." *History of Science* 28 (1990):149-173.

Green, John C. *American Science in the Age of Jefferson.* Ames: The Iowa State University Press, 1984.

Guralnick, Stanley M. "The American Scientist in Higher Education: 1820–1910." Pp. 99-141. In *The Sciences in the American Context: New Perspectives.* Ed. Nathan Reingold. Washington, D. C.: Smithsonian Institution Press, 1979.

Hawkins, Hugh. *Pioneer: A History of the Johns Hopkins University, 1874–1889.* Ithaca: Cornell University Press, 1960.

Kohlstedt, Sally Gregory. *The Formation of the American Scientific Community: The American Association for the Advancement of Science 1848-60.* Urbana: University of Illinois Press, 1976.

Kohlstedt, Sally Gregory and Rossiter, Margaret W. *Historical Writing on American Science: Perspectives and Prospects.* Baltimore: The Johns Hopkins University Press, 1985.

Koyré, Alexandre. *Études galiléennes.* 3 Parts, 1935–1939; Reprinted in one volume. Paris: Hermann, 1939.

Koyré, Alexandre. *From the Closed World to the Infinite Universe.* Baltimore: The Johns Hopkins University Press, 1957.

Kuhn, Thomas. *The Structure of Scientific Revolutions.* Chicago: The University of Chicago Press, 1962.

Merzbach, Uta C. "The Study of the History of Mathematics in America: A Centennial Sketch." Pp. 639-666. In *A Century of Mathematics in America–Part III.* Ed. Peter Duren *et al.* Providence: American Mathematical Society, 1989.

Oliver, James; Wait, Lucien; and Jones, George. *A Treatise on Trigonometry.* 4th Ed. Ithaca: G. W. Jones, 1890.

Parshall, Karen Hunger. "America's First School of Mathematical Research: James Joseph Sylvester at The Johns Hopkins University 1876–1883." *Archive for History of Exact Sciences* 38 (1988):153-196.

Parshall, Karen Hunger. "A Century-Old Snapshot of American Mathematics." *The Mathematical Intelligencer* 12 (3) (1990):7-11.

Parshall, Karen Hunger, "Eliakim Hastings Moore and the Founding of a Mathematical Community in America, 1892–1902." *Annals of Science* 41 (1984):313-333.

Parshall, Karen Hunger and Rowe, David E. *The Emergence of the American Mathematical Research Community, 1876–1900: J. J. Sylvester, Felix Klein, and E. H. Moore.* HMATH. Vol. 8. Providence: American Mathematical Society and London: London Mathematical Society, 1994.

Parshall, Karen Hunger and Rowe, David E. "Embedded in the Culture: Mathematics at the World's Columbian Exposition of 1893." *The Mathematical Intelligencer* 15 (2) (1993):40-45.

Richardson, R. G. D. "The Ph.D. Degree and Mathematical Research." *American Mathematical Monthly* 43 (1936):199-215; Reprinted in *A Century of Mathematics in America–Part II.* Ed. Peter Duren *et al.* Providence: American Mathematical Society, 1989. Pp. 361-378.

Rothenberg, Marc. "History of Astronomy," Pp. 117-131. In *Historical Writing on American Science: Perspectives and Prospects.* Ed. Sally Gregory Kohlstedt and Margaret Rossiter. Baltimore: The Johns Hopkins University Press, 1985.

Rowe, David E. "The Early Geometrical Works of Sophus Lie and Felix Klein." In *The History of Modern Mathematics.* Ed. David E. Rowe and John McCleary. 2 Vols. Boston: Academic Press, Inc., 1989, 1:209-273.

Sarton, George. *The Study of the History of Mathematics.* Cambridge, MA: Harvard University Press, 1937; Reprint ed., New York: Dover Publications, Inc., 1957.

Smith, David Eugene and Ginsburg, Jekuthiel. *A History of Mathematics in America before 1900.* Chicago: The Mathematical Association of America, 1934; Reprint ed, New York: Arno Press Inc., 1980.

Sylvester, James Joseph. "Remarks of Prof. Sylvester at a Farewell Reception Tendered to him by the Johns Hopkins University, Dec. 20, 1833 (Reported by Arthur S. Hathaway)." 24 typescript pages. Daniel Coit Gilman Papers Ms. 1. Special Collections Division. The Milton S. Eisenhower Library. The Johns Hopkins University.

8

19th Century Logic Between Philosophy and Mathematics*

Volker Peckhaus

Universität Paderborn, Fakultät für Kulturwissenschaften – Philosophie

Summary. The history of modern logic is usually written as the history of mathematical or, more general, symbolic logic. As such it was created by mathematicians. Not regarding its anticipations in Scholastic logic and in the rationalistic era, its continuous development began with George Boole's *The Mathematical Analysis of Logic* of 1847, and it became a mathematical subdiscipline in the early 20th century. This style of presentation cuts off one eminent line of development, the philosophical development of logic, although logic is evidently one of the basic disciplines of philosophy. One needs only to recall some of the standard 19th century definitions of logic as, e.g., the art and science of reasoning (Whateley) or as giving the normative rules of correct reasoning (Herbart). In the paper the relationship between the philosophical and the mathematical development of logic will be discussed. Answers to the following questions will be provided:

1. What were the reasons for the philosophers' lack of interest in formal logic?
2. What were the reasons for the mathematicians' interest in logic?
3. What did "logic reform" mean in the 19th century? Were the systems of mathematical logic initially regarded as contributions to a reform of logic?
4. Was mathematical logic regarded as art, as science or as both?

8.1 Introduction

Most 19th century scholars would have agreed to the opinion that philosophers are responsible for research on logic. On the other hand, the history of late 19th century logic indicates clearly a very dynamic development instigated not by philosophers, but by mathematicians. The central feature of this development was the emergence of what has been called the "new logic", "mathematical logic", "symbolic logic", or, since 1904, "logistics". This new logic came from Great Britain, and was created by mathematicians in the second half of the 19th century, finally becoming a mathematical subdiscipline in the early 20th century.

*First published in *Bulletin of Symbolic Logic* vol. 5 (1999), 433-450.

Charles L. Dodgson, better known under his pen name Lewis Carroll (1832–1898), published two well-known books on logic, *The Game of Logic* ([13]) of 1887 and *Symbolic Logic* of 1896 ([14]) of which a fourth edition appeared already in 1897. These books were written “to be of real service to the young, and to be taken up, in High Schools and in private families, as a valuable addition of their stock of healthful mental recreations” ([14, p. xiv]). They were meant “to *popularize* this fascinating subject,” as Carroll wrote in the preface of the fourth edition of *Symbolic Logic* ([14, p. xiv]). But, astonishingly enough, in both books there is no definition of the term “logic”. Given the broad scope of these books the title “Symbolic Logic” of the second book should at least have been explained.

Maybe the idea of symbolic logic was so widely spread at the end of the 19th century in Great Britain that Carroll regarded a definition as simply unnecessary. Some further observations support this thesis. They concern a remarkable interest by the general public in symbolic logic, after the death of the creator of the algebra of logic, George Boole, in 1864.

Recalling some standard 19th century definitions of logic as, e.g., the art and science of reasoning (Whately) or the doctrine giving the normative rules of correct reasoning (Herbart), it should not be forgotten that mathematical or symbolic logic was not set up from nothing. It arose from the old *philosophical* collective discipline logic. The standard presentations of the history of logic ignore the relationship between the philosophical and mathematical side of its development; they sometimes even deny that there has been any development of philosophical logic at all. Take for example William and Martha Kneale’s programme in their eminent *The Development of Logic*. They wrote ([32, p. iii]): “But our primary purpose has been to record the first appearances of these ideas which seem to us most important in the logic of our own day,” and these are the ideas leading to mathematical logic.

Another example is J. M. Bocheński’s assessment of “modern classical logic” which he dated between the 16th and the 19th century. It was for him a noncreative period in logic which can therefore justly be ignored in a problem history of logic ([7, p. 14]). According to Bocheński classical logic was only a decadent form of this science, a dead period in its development ([7, p. 20]).

Such assessments show that the authors adhered to the predominant views on logic of our time, i.e., actual systems of mathematical or symbolic logic. As a consequence, they have not been able to give reasons for the final divorce between philosophical and mathematical logic, because they have ignored the seed from which mathematical logic has emerged. Following Bocheński’s view Carl B. Boyer presented a consistent periodization of the development of logic ([11, p. 633]): “The history of logic may be divided, with some slight degree of oversimplification, into three stages: (1) Greek logic, (2) Scholastic logic, and (3) mathematical logic.” Note Boyer’s “slight degree of oversimplification” which enabled him to skip 400 years of logical development and ignore the fact that Kant’s transcendental logic, Hegel’s metaphysics and Mill’s inductive logic were called “logic”, too.

In discussing the relationship between the philosophical and the mathematical development of logic, at least the following questions will be answered:

1. What were the reasons for the philosophers' lack of interest in formal logic?
2. What were the reasons for the mathematicians' interest in logic?
3. What did "logic reform" mean in the 19th century? Were the systems of mathematical logic initially regarded as contributions to a reform of logic?
4. Was mathematical logic regarded as art, as science or as both?

This paper focuses not only on the situation in Great Britain, but also on the development in Germany. British logicians of that time regarded Germany as the logical paragon. John Venn can be regarded as a chief witness. He deplored, in the second edition of his *Symbolic Logic* of 1894, the lack of a tradition in logic in Great Britain which caused problems in creating the collection of books on logic for the Cambridge University Library ([67, p. 533]):

> At the time when I commenced the serious study of Symbolic Logic many of the most important works which bore on the subject were not to be found in any of those great libraries in this country to which one naturally refers in the first place, and could therefore only be obtained by purchase from abroad. [...] I suppose that the almost entire abandonment of Logic as a serious academic study, for so many years in this country at least, had prevented the formation of those private professorial libraries, the frequent appearance of which in the market has kept the secondhand booksellers' shops in Germany so well supplied with works on this subject.

It should be stressed, however, that when speaking of German logic Venn wasn't referring to contemporary German logical sytems, but to the great 18th century rationalistic precursors of the British algebra of logic beginning with Gottfried Wilhelm Leibniz and ending with the Swiss, Johann Heinrich Lambert.

In the following sections surveys are given of the philosophical and mathematical contexts in which the new logic emerged in Great Britain and Germany. The strange collaboration of mathematics and philosophy in promoting the new systems of logic will be discussed, and finally answers to the four questions already posed will be given.

8.2 Contexts

8.2.1 The philosophical context in Great Britain

The development of the new logic started in 1847, completely independent of earlier anticipations, e.g., by the German rationalist Gottfried Wilhelm Leibniz (1646–1716) and his followers (cf. [45]; [48, ch. 5]). In that year the British mathematician George Boole (1815–1864) published his pamphlet *The Mathematical Analysis of Logic* ([9]). Boole mentioned that it was the struggle for priority concerning the quantification of the predicate between the Edinburgh philosopher William Hamilton (1788–1856) and the London mathematician Augustus De Morgan (1806–1871) which encouraged this study. Hence, he referred to a startling philosophical discussion which indicated a vivid interest in formal logic in Great Britain. This interest was, however, a new interest, not even 20 years old. One can even say that neglect of formal logic could be regarded as a characteristic feature of British philosophy up to 1826 when Richard Whately (1787–1863) published his *Elements of Logic*.[1] In his preface Whately added an extensive report on the languishing research and education in formal logic in England. He complained ([69, p. xv]) that only very few students of the University of Oxford became good logicians and that

> by far the greater part pass through the University without knowing any thing of all of it; I do not mean that they have not learned by rote a string of technical terms; but that they understand absolutely nothing whatever of the principles of the Science.

Thomas Lindsay, the translator of Friedrich Ueberweg's important *System der Logik und Geschichte der logischen Lehren* ([63], translation [64]), was very critical of the scientific qualities of Whately's book, but he, nevertheless, emphasized its outstanding contribution for the renaissance of formal logic in Great Britain ([38, p. 557]):

> Before the appearance of this work, the study of the science had fallen into universal neglect. It was scarcely taught in the universities, and there was hardly a text-book of any value whatever to be put into the hands of the students.

One year after the publication of Whately's book, George Bentham's *An Outline of a New System of Logic* appeared ([6]) which was to serve as a commentary to Whately. Bentham's book was critically discussed by William Hamilton in a review article published in the *Edinburgh Review* ([20]). With the help of this review Hamilton founded his reputation as the "first logical name in Britain, it may be in the world."[2] Hamilton propagated a revival of the Aristotelian scholastic formal logic without, however, one-sidedly preferring the syllogism. His logical conception was focused on a revision of the standard forms by quantifying the predicates of judgements.[3] The controversy

[1] [53] lists nine editions up to 1848 and 28 further printings to 1908. Van Evra ([66, p. 2]) mentions 64 printings in the USA to 1913.

[2] This opinion can be found in a letter of De Morgan's to Spalding of 26th June, 1857 (quoted in [23, p. xii]) which was, however, not sent. Hamilton was for George Boole one of the "two greatest authorities in logic, modern and ancient" ([9, p. 81]). The other authority is Aristotle. This reverence to Hamilton might not be without irony because of Hamilton's disregard of mathematics.

[3] Cf. [22, vol. 4, p. 287].

about priority arose, when De Morgan, in a lecture "On the Structure of the Syllogism" ([16]) given to the Cambridge Philosophical Society on 9th November 1846, also proposed quantifying predicates. Neither had any priority, of course. Application of the diagrammatic methods of the syllogism proposed, e.g., by the 18th century mathematicians and philosophers Leonard Euler, Gottfried Ploucquet, and Johann Heinrich Lambert, presupposed quantification of the predicate. The German psychologistic logician Friedrich Eduard Beneke (1798–1854) suggested quantifying the predicate in his books on logic [4] and [5], the latter of which he sent to Hamilton. In the context of this paper it is irrelevant to solve the priority question. It is, however, important that a dispute of this extent arose at all. It indicates, there was new interest in *research* on formal logic.

This interest represented only one side of the effect released by Whately's book. Another line of research stood in the direct tradition of Humean empiricism and the philosophy of inductive sciences: the inductive logic of John Stuart Mill (1806–1873), Alexander Bain (1818–1903) and others. Boole's logic was in clear opposition to inductive logic. It was Boole's follower William Stanley Jevons (1835–1882; cf. [29]) who made this opposition explicit.

Boole referred to the controversy between Hamilton and De Morgan, but this influence should not be overemphasized. In his main work on the *Laws of Thought* ([10]) Boole went back to the logic of Aristotle by quoting from the Greek original. This can be interpreted as indicating that the influence of contemporary philosophical discussion was not as important as his own words might suggest. In writing a book on logic he was doing philosophy, and it was thus a matter of course that he related his results to the philosophical discussion of his time. This does not mean, of course, that his thoughts were really influenced by this discussion.

8.2.2 The philosophical context in Germany

It seems clear that, in regard to the 18th century dichotomy between German and British philosophy represented by the philosophies of Kant and Hume, Hamilton and Boole stood on the Kantian side. There are some analogies between the situations in Great Britain and Germany, where philosophical discussion on logic after Hegel's death was determined by the Kantian influence. In the preface to the second edition of his *Kritik der reinen Vernunft* of 1787 ([31]), Immanuel Kant (1723–1804) wrote that logic has followed the safe course of a science since earliest times. For Kant this was evident because of the fact that logic had been prohibited from taking any step backwards from the time of Aristotle. But he regarded it as curious that logic hadn't taken a step forward either ([31, B VIII]). Thus, logic seemed to be closed and complete. Formal logic, in Kant's terminology the analytical part of general logic, did not play a prominent rôle in Kant's system of transcendental philosophy. In any case, it was a negative touchstone of truth, as he stressed ([31, B 84]). Georg Wilhelm Friedrich Hegel (1770–1831) went further in denying any relevance of formal logic for philosophy ([24, I, Introduction, pp. XV-XVII]). Referring to Kant, he maintained that from the fact that logic hadn't changed since Aristotle one could infer that it needed a complete rebuilding ([24, p. XV]). Hegel created a variant of logic as the foundational science of his philosophical system, defining it as "the science of the pure idea, i.e., the idea in the *abstract element of reasoning*" ([25, p. 27]. Hegelian logic thus coincides with metaphysics ([25, p. 34]).

This was the situation when after Hegel's death philosophical discussion on formal logic in Germany started again. This discussion on logic reform stood under the label of "the logical question", a term coined by the Neo-Aristotelian Adolf Trendelenburg (1802–1872). In 1842 he published a paper entitled "Zur Geschichte von Hegel's Logik und dialektischer Methode" with the subtitle "Die logische Frage in Hegel's Systeme". But what is the logical question according to Trendelenburg? He formulated this question explicitly towards the end of his article: "Is Hegel's dialectical method of pure reasoning a scientific procedure?" ([62, p. 414]). In answering this question in the negative, he provided the occasion of rethinking the status of formal logic within a theory of human knowledge without, however, proposing a return to the old (scholastic) formal logic. In consequence the term "the logical question" was subsequently used in a less specific way. Georg Leonard Rabus, the early chronicler of the discussion on logic reform, wrote that the logical question emerged from doubts concerning the justification of formal logic ([49, p. 1]).

Although this discussion was clearly *connected* to formal logic, the so-called reform did not *concern* formal logic. The reason was provided by the Neo-Kantian Wilhelm Windelband who wrote in a brilliant survey on 19th century logic ([70, p. 164]):

> It is in the nature of things that in this enterprise [i.e., the reform of logic] the lower degree of fruitfulness and developability power was on the side of formal logic. Reflection on the rules of the correct progress of thinking, the technique of correct thinking, had indeed been brought to perfection by former philosophy, presupposing a naive world view. What Aristotle had created in a stroke of genius, was decorated with the finest filigree work in Antiquity and the Middle Ages: an art of proving and disproving which culminated in a theory of reasoning, and after this constructing the doctrines of judgements and concepts. Once one has accepted the foundations, the safely assembled building cannot be shaken: it can only be refined here and there and perhaps adapted to new scientific requirements.

Windelband was very critical of English mathematical logic. Its quantification of the predicate allows the correct presentation of extensions in judgements, but it "drops hopelessly" the vivid sense of all judgements, which tend to claim or deny a material relationship between subject or predicate. It is "a logic of the conference table", which cannot be used in the vivid life of science, a "logical sport" which has, however, its merits in exercising the final acumen ([70, pp. 166–167]).

The philosophical reform efforts concerned primarily two areas:

1. the problem of a foundation of logic which itself was approached by psychological and physiological means, leading to new discussion on the question of priority between logic and psychology, and to various forms of psychologism and anti-psychologism (cf. [50], [33]);
2. the problem of logical applications focusing interest on the methodological part of traditional logic. The reform of applied logic attempted to bring philosophy in touch with the stormy development of mathematics and sciences of the time.

Both reform procedures had a destructive effect on the shape of logic and philosophy. The struggle with psychologism led to the departure of psychology (especially in its new, experimental form) from the body of philosophy at the beginning of the 20th century. Psychology became a new, autonomous scientific discipline. The debate on methodology resulted in the creation of the philosophy of science which was separated from the body of logic. The philosopher's ignorance of the development of formal logic caused a third departure: Part of formal logic was taken from the domain of the competence of philosophy and incorporated into mathematics where it was instrumentalized for foundational tasks.

8.2.3 The mathematical context in Great Britain

As mentioned earlier, the influence of the philosophical discussion on logic in Great Britain on the emergence of the new logic should not be overemphasized. Of greater importance were mathematical influences. Most of the new logicians can be related to the so-called "Cambridge Network" ([12, pp. 29–71]), i.e., the movement which aimed at reforming British science and mathematics which started at Cambridge. One of the roots of this movement was the foundation of the Analytical Society in 1812 (cf. [17]) by Charles Babbage (1791–1871), George Peacock (1791–1858) and John Herschel (1792–1871). In regard to mathematics Joan L. Richards called this act a "convenient starting date for the nineteenth-century chapter of British mathematical development" ([51, p. 13]). One of the first achievements of the Analytical Society was a revision of the Cambridge Tripos by adopting the Leibnizian notation for the calculus and abandoning the customary Newtonian theory of fluxions: "the principles of pure D-ism in opposition to the Dot-age of the University" as Babbage wrote in his memoirs ([1, p. 29]). It may be assumed that this successful movement triggered by a change in notation might have stimulated a new or at least revived interest in operating with symbols. This new research on the calculus had parallels in innovative approaches to algebra which were motivated by the reception of Laplacian analysis. Firstly the development of symbolical algebra has to be mentioned. It was codified by George Peacock in his *Treatise on Algebra* ([42]) and further propagated in his famous report for the British Association for the Advancement of Science ([43], especially pp. 198–207). Peacock started by drawing a distinction between arithmetical and symbolical algebra, which was, however, still based on the common restrictive understanding of arithmetic as the doctrine of quantity. A generalization of Peacock's concept can be seen in Duncan F. Gregory's (1813–1844) "calculus of operations". Gregory was most interested in *operations* with symbols. He defined symbolical algebra as "the science which treats of the combination of operations defined not by their nature, that is by what they are or what they do, but by the laws of combinations to which they are subject" ([18, p. 208]). In his much praised paper "On a General Method in Analysis" ([8]) Boole made the calculus of operations the basic methodological tool for analysis. However in following Gregory, he went further, proposing more applications. He cited Gregory who wrote that a symbol is defined algebraically "when its laws of combination are given; and that a symbol represents a given operation when the laws of combination of the latter are the same as those of the former" ([19, pp. 153–154]). It is possible that a symbol for an arbitrary operation can be applied to the same operation ([19, p. 154]). It is thus necessary to distinguish between arithmetical algebra and symbolical algebra which has to take into account symbolical, but non-arithmetical fields of application. As an example Gregory mentioned the symbols a and $+a$. They are isomorphic in arithmetic, but in geometry they need to be interpreted differently. a can refer to a point marked by a line whereas the combination of the signs $+$ and a additionally expresses the direction of the line. Therefore symbolical algebra has to distinguish between the symbols a and $+a$. Gregory deplored the fact that the unequivocity of notation didn't prevail as a result of the persistence of mathematical practice. Clear notation was only advantageous, and Gregory thought that our minds would be "more free from prejudice, if we never used in the general science symbols to which definite meanings had been appropriated in the particular science" ([19, p. 158]).

Boole adopted this criticism almost word for word. In his *Mathematical Analysis of Logic* of 1847 he claimed that the reception of symbolic algebra and its principles was delayed by the fact that in most interpretations of mathematical symbols the idea of quantity was involved. He felt that these connotations of quantitative relationships were the result of the context of the emergence of mathematical symbolism, and not of a universal principle of mathematics ([9, pp. 3–4]). Boole read the principle of the permanence of equivalent forms as a principle of independence from interpretation in an "algebra of symbols". In order to obtain further affirmation, he tried to free the principle from the idea of quantity by applying the algebra of symbols to another field, the field of logic. As far as logic is concerned this implied that only the principles of a "true Calculus" should be presupposed. This calculus is characterized as a "method resting upon the employment of Symbols, whose laws of combination are known and general, and whose results admit of a consistent interpretation" ([9, p. 4]). He stressed (*ibid.*):

> It is upon the foundation of this general principle, that I purpose to establish the Calculus of Logic, and that I claim for it a place among the acknowledged forms of Mathematical Analysis, regardless that in its objects and in its instruments it must at present stand alone.

Boole expressed logical propositions in symbols whose laws of combination are based on the mental acts represented by them. Thus he attempted to establish a psychological foundation of logic, mediated, however, by language. The central mental act in Boole's early logic is the act of election used for building classes. Man is able to separate objects from an arbitrary collection which belong to given classes, in order to distinguish them from others. The symbolic representation of these mental operations follows certain laws of combination which are similar to those of symbolic algebra. Logical theorems can thus be proven like mathematical theorems. Boole's opinion has of course consequences for the place of logic in philosophy: "On the principle of a true classification, we ought no longer to associate Logic and Metaphysics, but Logic and Mathematics" ([9, p. 13]).

Although Boole's logical considerations became increasingly philosophical with time, aiming at the psychological and epistemological foundations of logic itself, his initial interest was not to reform logic but to reform mathematics. He wanted to establish an abstract view on mathematical operations without regard to the objects of these operations. When claiming "a place among the acknowledged forms of Mathematical Analysis" ([9, p. 4]) for the calculus of logic, he didn't simply want to include logic in traditional mathematics. The superordinate discipline was a *new* mathematics. This is expressed in Boole's writing: "It is not of the essence of mathematics to be conversant with the ideas of number and quantity" ([10, p. 12]).

8.2.4 The mathematical context in Germany

The results of this examination of the British situation at the time when the new logic emerged–a reform of mathematics, with initially a lack of interest in a reform of logic, by establishing an abstract view of mathematics which focused not on mathematical objects, but on symbolic operations with arbitrary objects–these results could be transferred to the situation in Germany without any problem. The most important representative of the German algebra of logic was the mathematician Ernst Schröder (1841–1902) who was regarded as having completed the Boolean period in logic (cf. [7, p. 314]). In his first pamphlet on logic, *Der Operationskreis des Logikkalkuls* ([55]), he presented a critical revision of Boole's logic of classes, stressing the idea of the duality between logical addition and logical multiplication introduced by William Stanley Jevons in 1864. In 1890 Schröder started on the large project, his monumental *Vorlesungen über die Algebra der Logik* ([56], [57], [58], [59]) which remained unfinished although it increased to three volumes with four parts, of which one appeared only posthumously. Contemporaries regarded the first two volumes alone as completing the algebra of logic (cf. [68, p. 196]).

Schröder's opinion concerning the question as to the end to which logic is studied (cf. [44], [46]) can be drawn from an autobiographical note, published in 1901 (and written in the third person), the year before his death. It contains Schröder's own survey of his scientific aims and results. Schröder divided his scientific production into three fields:

1. A number of papers dealing with some of the current problems of his science.
2. Studies concerned with creating an "absolute algebra," i.e., a general theory of connections. Schröder stressed that such studies represent his "very own object of research" of which only little was published at that time.
3. Work on the reform and development of logic.

Schröder wrote ([60]) that his aim was

> to design logic as a calculating discipline, especially to give access to the exact handling of relative concepts, and, from then on, by emancipation from the routine claims of spoken language, to withdraw any fertile soil from "cliché" in the field of philosophy as well. This should prepare the ground for a scientific universal language that, widely differing from linguistic efforts like Volapük [a universal language like Esperanto, very popular in Germany at that time], looks more like a sign language than like a sound language.

Schröder's own division of his fields of research shows that he didn't consider himself a logician: His "very own object of research" was "absolute algebra," which was similar to modern abstract or universal algebra in respect to its basic problems and fundamental assumptions. What was the connection between logic and algebra in Schröder's research? From the passages quoted one could assume that these fields belong to two separate fields of research, but this is not the case. They were intertwined in the framework of his heuristic idea of a general science. In his autobiographical note he stressed ([60]):

> The disposition for schematizing, and the aspiration to condense practice to theory advised Schröder to prepare physics by perfecting mathematics. This required deepening–as of mechanics and geometry–above all of arithmetic, and subsequently he became by the time aware of the necessity for a reform of the source of all these disciplines, logic.

Schröder's universal claim becomes obvious. His scientific efforts served to provide the requirements to found physics as the science of material nature by "deepening the foundations," to quote a famous metaphor later used by David Hilbert ([26, p. 407]) in order to illustrate the objectives of his axiomatic programme. Schröder regarded the formal part of logic that can be formed as a "calculating logic," using a symbolical notation, as a *model* of formal algebra that is called "absolute" in its last state of development.

But what is "formal algebra"? The theory of formal algebra "in the narrowest sense of the word" includes "those investigations on the laws of algebraic operations [...] that refer to nothing but general numbers in an unlimited number field without making any presuppositions concerning its nature" ([54, p. 233]). Formal algebra therefore prepares "studies on the most varied number systems and calculating operations that might be invented for particular purposes" ([54, p. 233]).

It has to be stressed that Schröder wrote his early considerations on formal algebra and logic without any knowledge of the results of his British predecessors. His sources were the textbooks of Martin Ohm, Hermann Günther Graßmann, Hermann Hankel and Robert Graßmann. These sources show that Schröder was a representative of the tradition of German combinatorial algebra and algebraic analysis (cf. [48, ch. 6]).

Like the British tradition, but independent of it, the German algebra of logic was connected to new trends in algebra. It differed from its British counterpart in its combinatorial approach. In both traditions, algebra of logic was invented within the enterprise to reform basic notions of mathematics which led to the emergence of structural abstract mathematics. The algebraists wanted to design algebra as "pan-mathematics", i.e., as a general discipline embracing all mathematical disciplines as special cases. The independent attempts in Great Britain and Germany were combined when Schröder learned about the existence of Boole's logic in late 1873, early 1874. Finally he enriched the Boolean class logic by adopting Charles S. Peirce's theory of quantification and adding a logic of relatives according to the model of Peirce and De Morgan.

The main interest of the new logicians was to utilize logic for mathematical and scientific purposes, and it was only in a second step, but nevertheless an indispensable consequence of the attempted applications, that the reform of logic came into the view. What has been said of the representatives of the algebra of logic also holds for the proponents of competing logical systems such as Gottlob Frege or Giuseppe Peano. They wanted to use logic in their quest for mathematical rigor, something questioned by the stormy development in mathematics.

8.3 Accepting the New Logic

Although created by mathematicians, the new logic was widely ignored by fellow mathematicians. In Germany Schröder was only known as the algebraist of logic, and regarded as rather exotic. George Boole was respected by British mathematicians, but his ideas concerning an algebraical representation of the laws of thought received very little published reaction. He shared this fate with Augustus De Morgan, the second major figure of symbolic logic at that time. In 1864, Samuel Neil, the early chronicler of British mid 19th century logic, expressed his thoughts about the reasons for this negligible reception: "De Morgan is esteemed crotchety, and perhaps formalizes too much. Boole demands high mathematic culture to follow and to profit from" ([41, p. 161]).

One should add that the ones who had this culture were usually not interested in logic. The situation changed after George Boole's death in 1864. In the following comments only some ideas concerning the reasons for this new interest are hinted at. In particular the rôles of William Stanley Jevons and Alexander Bain are stressed which exemplify "the strange collaboration of mathematics and philosophy in promoting the new systems of logic" mentioned in the introduction.

8.3.1 William Stanley Jevons

A broader international reception of Boole's logic began when William Stanley Jevons made it the starting point for his influential *Principles of Science* ([28]). He used his own version of the Boolean calculus introduced in his *Pure Logic* of 1864. Among his revisions were the introduction of a simple symbolical representation of negation and the definition of logical addition as inclusive "or". He also changed the philosophy of symbolism ([27, p. 5]):

> The forms of my system may, in fact, be reached by divesting his [Boole's] of a mathematical dress, which, to say the least, is not essential to it. The system being restored to its proper simplicity, it may be inferred, not that Logic is a part of Mathematics, as is almost implied in Professor Boole's writings, but that the Mathematics are rather derivatives of Logic. All the interesting analogies or samenesses of logical and mathematical reasoning which may be pointed out, are surely reversed by making Logic the dependent of Mathematics.

Jevons' interesting considerations on the relationship between mathematics and logic representing an early logicistic attitude will not be discussed. Similar ideas can be found not only in Gottlob Frege's work, but also in that of Hermann Rudolf Lotze and Ernst Schröder. In the context of this paper, it is relevant that Jevons abandoned mathematical symbolism in logic, an attitude which was later taken up by John Venn. Jevons attempted to free logic from the semblance of being a special mathematical discipline. He used the symbolic notation only as a means of expressing general truths. Logic became a tool for studying science, a new language providing symbols and structures. The change in notation brought the new logic closer to the philosophical discourse of the time. The reconciliation was supported by the fact that Jevons formulated his *Principles of Science* as a rejoinder to John Stuart Mill's *A System of Logic* of 1843, at that time the dominating work on logic and the philosophy of science in Great Britain. Although Mill called his logic *A System of Logic Ratiocinative and Inductive*, the deductive parts played only a minor rôle, used only to show that all inferences, all proofs and the discovery of truths consisted of inductions and their interpretations. Mill claimed to have shown "that all our knowledge, not intuitive, comes to us exclusively from that source" ([40, Bk. II, ch. I, §1]). Mill concluded that the question as to what induction is, is the most important question of the science of logic, "the question which includes all others." As a result the logic of induction covers by far the largest part of this work, a subject which we would today regard as belonging to the philosophy of science.

Jevons defined induction as a simple inverse application of deduction. He began a direct argument with Mill in a series of papers entitled "John Stuart Mill's Philosophy Tested" ([29]). This discourse proved that symbolic logic could be of importance not only for mathematics, but also for philosophy.

Another effect of the attention caused by Jevons was that British algebra of logic was able to cross the Channel. In 1877, Louis Liard (1846–1917), at that time professor at the Faculté de lettres at Bordeaux and a friend of Jevons, published two papers on the logical systems of Jevons and Boole ([34], [35]). In [36] he added a booklet entitled *Les logiciens anglais contemporaines* which had five editions until 1907, and was translated into German in [37]. Although Herman Ulrici had published a first German review of Boole's *Laws of Thought* as early as 1855 ([65]; cf. [47]), the knowledge of British symbolic logic was conveyed primarily by Alois Riehl, then professor at the University of Graz, in Austria. He published a widely read paper "Die englische Logik der Gegenwart" ("English contemporary logic") in 1877 ([52]) which reported mainly Jevons' logic and utilized it in a current German controversy on the possibility of scientific philosophy.

8.3.2 Alexander Bain

Finally a few words on Alexander Bain (1818–1903): This Scottish philosopher was an adherent of Mill's logic. Bain's *Logic*, first published in 1870, had two parts, the first on deduction and the second on induction. He made explicit that "Mr. Mill's view of the relation of Deduction and Induction is fully adopted" ([2, I, p. iii]). Obviously he shared the "[...] general conviction that the utility of the purely Formal Logic is but small; and that the rules of Induction should be exemplified even in the most limited course of logical discipline" ([2, p. v]). The minor rôle of deduction showed up in Bain's definition "*Deduction* is the application or extension of Induction to *new cases*" ([2, p. 40]).

Despite his reservations about deduction, Bain's *Logic* was quite important for the reception of symbolic logic because of a chapter of 30 pages entitled "Recent Additions to the Syllogism." In this chapter the contributions of William Hamilton, Augustus De Morgan and George Boole were introduced. Presumably many more people became acquainted with Boole's algebra of logic through Bain's report than through Boole's own writings. One example is Hugh MacColl (1837–1909), the pioneer of the calculus of propositions (statements) and of modal logic. He created his ideas independently of Boole, eventually realizing the existence of the Boolean calculus by means of Bain's report. Even in the early parts of his series of papers "The Calculus of Equivalent Statements" he quoted from Bain's presentation when discussing Boole's logic ([39]). In 1875 Bain's logic was translated into French, in 1878 into Polish. Tadeusz Batóg and Roman Murawski ([3]) have shown that it was Bain's presentation which motivated the first Polish algebraist of logic, Stanisław Piątkiewicz (1848–?) to begin his research on symbolic logic.

The remarkable collaboration of mathematics and philosophy can be seen in the fact that a broader reception of symbolic logic commenced only when its relevance for the philosophical discussion of the time came to the fore.

8.4 Conclusions

Finally, these are the answers to the initial questions:

1. What were the reasons for the philosophers' lack of interest in formal logic?

In Germany philosophers shared Kant's opinion that formal logic was a completed field of knowledge. They were interested primarily in the foundations and application of logic. In Great Britain there was hardly any vivid logical tradition. Philosophy was dominated by empiricist conceptions. New systems of formal logic therefore had difficulties in gaining a footing in the philosophical discussion.

2. What were the reasons for the mathematicians' interest in logic?

Foundational problems and problems in grasping new mathematical objects forced some mathematicians to look intuitively at the logical foundations of their subject. The interest in formal logic was thus a result of the dynamic development of late 19th century mathematics. One should not assume, however, that this was a general interest. Most mathematicians did not (and still do not) care about foundations.

3. How did the mathematicians' logical activities fit into the reform of logic conceptions of the time?

In Germany in the second half of the 19th century, logic reform meant overcoming the Hegelian identification of logic and metaphysics. In Great Britain it meant enlarging the scope of the syllogism or elaborating the philosophy of science. Mathematicians were initially interested in utilizing logic for mathematical means, or they used it as a language for structuring and symbolizing extra-mathematical fields. Applications were, e.g., the foundation of mathematics (Boole, Schröder, Frege), the foundation of physics (Schröder), the preservation of rigour in mathematics (Peano), the theory of probabilities (Boole, Venn), the philosophy of science (Jevons), the theory of human relationships (Alexander Macfarlane), and juridical questions. The mathematicians' preference for the organon aspect of formal logic seems to be the point of deviation between mathematicians and philosophers who were not interested in elaborating logic as a tool.

4. Was mathematical logic regarded as art or as science?

From the applicational interest it follows that it was mainly regarded as an art. The scientific aspect grew, however, with the insight into the power of logical calculi. Nevertheless, in an institutional sense the new logic was established only in the beginning of the 20th century as an academic subject, i.e., as an institutionalized domain of science.

References

1. Charles Babbage, *Passages from the life of a philosopher*, Longman, Green, Longman, Roberts, & Green, London 1864; reprinted Greeg, Westmead, 1969.
2. Alexander Bain, *Logic*, 2 volumes, pt. 1: *Deduction*, pt. 2: *Induction*, Longmans, Green, & Co., London, 1870.
3. Tadeusz Batóg and Roman Murawski, Stanisław Piątkiewicz and the beginnings of mathematical logic in Poland, *Historia Mathematica*, vol. 23 (1996), pp. 68–73.
4. Friedrich Eduard Beneke, *Syllogismorum analyticorum origines et ordinem naturalem*, Mittler, Berlin, 1839.
5. ________, *System der Logik als Kunstlehre des Denkens*, 2 volumes, F. Dümmler, Berlin, 1842.
6. George Bentham, *An outline of a new system of logic: With a critical examination of Dr. Whately's "Elements of logic"*, Hunt and Clark, London, 1827; reprinted Thoemmes, Bristol, 1990.
7. Joseph Maria Bocheński, *Formale Logik*, Alber, Freiburg/München, 1956 (= *Orbis Academicus*, vol. III, no. 2), ^{4}1978.
8. George Boole, On a general method in analysis, *Philosophical Transactions of the Royal Society of London for the Year MDCCCXLIV*, pt. 1 (1844), pp. 225–282.
9. ________, *The mathematical analysis of logic: Being an essay towards a calculus of deductive reasoning*, Macmillan, Barclay, and Macmillan, Cambridge/George Bell, London, 1847; reprinted Basil Blackwell, Oxford, 1951.
10. ________, *An investigation of the laws of thought, on which are founded the mathematical theories of logic and probabilities*, Walton & Maberly, London, 1854; reprinted Dover, New York, ca. 1951.
11. Carl B. Boyer, *A history of mathematics*, John Wiley & Sons, New York/London/Sydney, 1968.
12. Susan Faye Cannon, *Science in culture: The early Victorian period*, Dawson Scientific History Publications, New York, 1978.
13. Lewis Carroll, *The game of logic*, MacMillan, London, 1887; reprinted in [15].
14. ________, *Symbolic logic*, MacMillan, London, 1896; reprinted in [15].
15. ________, *Mathematical recreations of Lewis Carroll: Symbolic logic and The game of logic (both books bound as one)*, Dover Publications and Berkeley Enterprises, New York, 1958.
16. Augustus De Morgan, On the syllogism I: On the structure of the syllogism, and on the applications of the theory of probabilities to questions of arguments and authority, *Transactions of the Cambridge Philosophical Society*, vol. 8 (1846), pp. 379–408.
17. Philip C. Enros, The analytical society (1812–1813): Precursor of the renewal of Cambridge Mathematics, *Historia Mathematica*, vol. 10 (1983), pp. 24–47.

18. Duncan Farquharson Gregory, On the real nature of symbolical algebra, *Transactions of the Royal Society of Edinburgh*, vol. 14 (1840), pp. 208–216.
19. ________, On a diffculty in the theory of algebra, *The Cambridge Mathematical Journal*, vol. 3 (1842), pp. 153–159.
20. William Hamilton, Logic. In reference of the recent English treatises on that science, *Edinburgh Review*, vol. 66 (April 1833), pp. 194–238; again in [21], pp. 16–174.
21. ________, *Discussions on philosophy and literature, education and university reform. Chiefly from the Edinburgh Review; corrected, vindicated, enlarged, in notes and appendices*, Longman, Brown, Green and Longmans, London/Maclachlan and Stewart, Edinburgh, 1851.
22. ________, *Lectures on metaphysics and logic*, 4 volumes (H. L. Mansel and J. Veitch, editors), William Blackwood and Sons, Edinburgh/London, 1859–1866.
23. Peter Heath, Introduction, in Augustus De Morgan, *On the syllogism and other logical writings* (Peter Heath, editor), Routledge & Kegan Paul, London, 1966 (= Rare masterpieces of philosophy and science), pp. vii–xxxi.
24. Georg Wilhelm Friedrich Hegel, *Wissenschaft der Logik*, vol. 1: *Die objective Logik*, Johann Leonhard Schrag, Nürnberg, 1812/1813; critical edition: Hegel, *Wissenschaft der Logik. Erster Band. Die objektive Logik* (Friedrich Hegemann and Walter Jaeschke, editors), Felix Meiner, Hamburg, 1978 (= Hegel, *Gesammelte Werke*, vol. 11).
25. ________, *Encyclopädie der philosophischen Wissenschaften im Grundrisse. Zum Gebrauch seiner Vorlesungen. Dritte Ausgabe*, Oßwald'scher Verlag, Heidelberg, 1830. Critical edition: Hegel, *Enzyklopädie der philosophischen Wissenschaften im Grundrisse (1830)* (Wolfgang Bonsiepen and Hans-Christian Lucas, editors), Felix Meiner, Hamburg, 1992 (= Hegel, *Gesammelte Werke*, vol. 20).
26. David Hilbert, Axiomatisches Denken, *Mathematische Annalen*, vol. 78 (1918), pp. 405–415.
27. William Stanley Jevons, *Pure logic or the logic of quality apart from quantity with remarks on Boole's system and the relation of logic and mathematics*, E. Stanford, Londone, 1864; reprinted in [30], pp. 3–77.
28. ________, *The principles of science. A treatise on logic and scientific method*, 2 volumes, Macmillan and Co., London, 1874 [New York, 1875].
29. ________, John Stuart Mill's philosophy tested, *The Contemporary Review*, vol. 31 (1877/78), pp. 167–182 and 256–275; vol. 32 (1878), pp. 88–99; again in [30], pp. 137–172.
30. ________, *Pure logic and other minor works* (Robert Adamson and Harriet A. Jevons, editors), Macmilland and Co., London/NewYork, 1890; reprinted Thoemmes Press, Bristol, 1991.
31. Immanuel Kant, *Critik der reinen Vernunft*, 2nd ed., Johann Friedrich Hartknoch, Riga, 1787; again in *Kant's gesammelte Schriften*, vol. 3 (Königlich Preußische Akademie der Wissenschaften, editor), Reimer, Berlin, 1911.
32. William Kneale and Martha Kneale, *The development of logic*, Clarendon Press, Oxford, 1968; reprinted 1986.
33. Martin Kusch, *Psychologism. A case study in the sociology of philosophical knowledge*, Routledge, London/New York, 1995 (= *Psychological issues in science*).
34. Louis Liard, Un nouveau système de logique formelle. M. Stanley Jevons, *Revue philosophique de la France et de l'Étranger*, vol. 3 (1877), pp. 277–293.

35. ———, La logique algébrique de Boole, *Revue philosophique de la France et de l'Étranger*, vol. 4 (1877), pp. 285–317.
36. ———, *Les logiciens anglais contemporains*, Germer Baillière, Paris, 1878, ^{5}1907.
37. ———, *Die neuere englische Logik* (J[ohannes] Imelmann, editor), Denicke's Verlag, Berlin, 1880, ^{2}1883.
38. Thomas M. Lindsay, On recent logical speculation in England, in [64], pp. 557–590.
39. Hugh MacColl, The calculus of equivalent statements (second paper), *Proceedings of the London Mathematical Society*, vol. 9 (1877–1878), pp. 177–186.
40. John Stuart Mill, *A system of logic, ratiocinative and inductive: Being a connective view of the principles of evidence and the methods of scientific investigation*, 2 volumes, J. W. Parker, London, 1843.
41. Samuel Neil, John Stuart Mill, *The British Controversialist and Literary Magazine* n.s. (1864), pp. 161–173 and 241–256.
42. George Peacock, *A treatise on algebra*, J. & J. J. Deighton, Cambridge/G. F. & J. Rivington, London, 1830.
43. ———, Report on the recent progress and present state of certain branches of analysis, in *Report of the third meeting of the British Association for the Advancement of Science held at Cambridge in 1833*, John Murrary, London, 1834, pp. 185–352.
44. Volker Peckhaus, Ernst Schröder und die 'pasigraphischen Systeme' von Peano und Peirce, *Modern Logic*, vol. 1 (Winter 1990/1991), no. 2/3, pp. 174–205.
45. ———, Leibniz als Identifikationsfigur der britischen Logiker des 19. Jahrhunderts, in *VI. Internationaler Leibniz-Kongreß. Vorträge I. Teil, Hannover, 18.-22.7.1994*, Gottfried-Wilhelm-Leibniz-Gesellschaft, Hannover, 1994, pp. 589–596.
46. ———, Wozu Algebra der Logik? Ernst Schröders Suche nach einer universalen Theorie der Verknüpfungen, *Modern Logic*, vol. 4 (1994), pp. 357–381.
47. ———, *Hermann Ulrici (1806–1884): Der Hallesche Philosoph und die englische Algebra der Logik. Mit einer Auswahl von Texten Ulricis und einer Bibliographie seiner Schriften*, Hallescher Verlag, Halle a. S., 1995 (= *Schriftenreihe zur Geistes- und Kulturgeschichte. Texte und Dokumente*).
48. ———, *Logik, Mathesis universalis und allgemeine Wissenschaft. Leibniz und die Wiederentdeckung der formalen Logik im 19. Jahrhundert*, Akademie Verlag, Berlin, 1997 (= *Logica Nova*).
49. Georg Leonhard Rabus, *De neuesten Bestrebungen auf dem Gebiete der Logik bei den Deutschen und Die logische Frage*, Deichert, Erlangen, 1880.
50. Matthias Rath, *Der Psychologismus in der deutschen Philosophie*, Karl Alber Verlag, Freiburg/München, 1994.
51. Joan L. Richards, *Mathematical visions: The pursuit of geometry in Victorian England*, Academic Press, Boston, etc., 1988.
52. Alois Riehl, Die englische Logik der Gegenwart, *Vierteljahrsschrift für wissenschaftliche Philosophie*, vol. 1 (1877), pp. 51–80.
53. Wilhelm Risse, *Bibliographica Logica. Verzeichnis der Druckschriften zur Logik mit Angabe ihrer Fundorte*, vol. 2: 1801–1969, Olms, Hildesheim/New York, 1973 (= *Studien und Materialien zur Geschichte der Philosophie*; 1).
54. Ernst Schröder, *Lehrbuch der Arithmetik und Algebra für Lehrer und Studirende*, vol. 1 [no further volumes published]: *Die sieben algebraischen Operationen*, B. G. Teubner, Leibzig, 1873.

55. ________, *Der Operationskreis des Logikkalkuls*, Teubner, Leipzig, 1877; reprinted as "special edition", Wissenschaftliche Buchgesellschaft, Darmstadt, 1966.
56. ________, *Vorlesungen über die Algebra der Logik (exakte Logik)*, vol. 1, B. G. Teubner, Leipzig, 1890; reprinted in [61].
57. ________, *Vorlesungen über die Algebra der Logik (exakte Logik)*, vol. 2, B. G. Teubner, Leipzig, 1891; reprinted in [61].
58. ________, *Vorlesungen über die Algebra der Logik (exakte Logik)*, vol. 3, pt. 1: *Algebra und Logik der Relative*, B. G. Teubner, Leipzig, 1895; reprinted in [61].
59. ________, *Vorlesungen über die Algebra der Logik (exakte Logik)*, vol. 2, pt. 2 (Karl Eugen Müller, editor), B. G. Teubner, Leipzig, 1905; reprinted in [61].
60. ________, Grossberzoglich Badischer Hofrat Dr. phil. Ernst Schröder[,] ord. Professor für Mathematik an der Technischen Hochschule in Karlsruhe i. Baden, in *Geistiges Deutschland. Deutsche Zeitgenossen auf dem Gebiete der Literatur, Wissenschaften und Musik*, Adolf Eckstein, Berlin-Charlottenburg, 1901.
61. ________, *Vorlesungen über die Algebra der Logik (exakte Logik)*, ["second edition"], 3 volumes, Chelsea, Bronx, New York, 1966.
62. Friedrich Adolf Trendelenburg, *Zur Geschichte von Hegel's Logik und dialektischer Methode. Die logische Frage in Hegel's Systeme. Eine Auffoderung [sic!] zu ihrer wissenschaftlichen Erledigung, Neue Jenaische Allgemeine Literatur-Zeitung*, vol. 1 (23 April 1842), no. 97, pp. 405–408; (25 April 1842) no. 98, pp. 409–412; (26 April 1842) no. 99, pp. 413–414; separately published in Trendelenburg, *Die logische Frage in Hegel's System. Zwei Streitschriften*, Brockhaus, Leipzig, 1843.
63. Friedrich Ueberweg, *System der Logik und Geschichte der logischen Lehren*, Adolph Marcus, Bonn, 1857.
64. ________, *System of logic and history of logical doctrines*, translated from the German with notes and appendices by Thomas M. Lindsay, Longmans, Green, and Co., London, 1871; reprinted Thoemmes Press, Bristol, 1993.
65. Hermann Ulrici, Review of [10], *Zeitschrift für Philosophie und philosophische Kritik*, vol. 27 (1855), pp. 273–291.
66. James Van Evra, Richard Whately and the rise of modern logic, *History and Philosophy of Logic*, vol. 5 (1894), pp. 1–18.
67. John Venn, *Symbolic logic*, 2nd ed., "revised and rewritten," Macmillan & Co., London, 1894; reprinted Chelsea Publishing, Bronx, New York, 1971.
68. Alexander Wernicke, Review of [57], *Deutsche Litteraturzeitung*, vol. 12 (1891), col. 196–197.
69. Richard Whately, *Elements of logic: Comprising the substance of the article in the Encyclopaedia Metropolitana: with additions &c.*, J. Mawman, London, 1826.
70. Wilhelm Windelband, Logik, in Windelband (ed.), *Die Philosophie im Beginn des zwanzigsten Jahrhunderts. Festschrift für Kuno Fischer*, vol. 1, Carl Winter, Heidelberg, 1904, pp. 163–186.

9

The Battle for Cantorian Set Theory

Joseph W. Dauben*

Department of History, Herbert H. Lehman College, CUNY

Ph.D. Program in History, The Graduate Center, City University of New York

The substance of Georg Cantor's revolutionary mathematics of the infinite is well-known: in developing what he called the arithmetic of transfinite numbers, he gave mathematical content to the idea of actual infinity. In so doing he laid the groundwork for abstract set theory and made significant contributions to the foundations of the calculus and to the analysis of the continuum of real numbers. Cantor's most remarkable achievement was to show, in a mathematically rigorous way, that the concept of infinity is not an undifferentiated one. Not all infinite sets are the same size, and consequently, infinite sets can be compared with one another. But so shocking and counter-intuitive were Cantor's ideas at first that the eminent French mathematician, Henri Poincaré, condemned Cantor's theory of transfinite numbers as a "disease" from which he was certain mathematics would one day be cured.[1] Leopold Kronecker, one of Cantor's teachers and among the most prominent members of the German mathematics establishment, even attacked Cantor personally, calling him a "scientific charlatan," a "renegade," and a "corrupter of youth."[2]

*This paper was first presented as the invited centennial address for history of mathematics on the occasion of the 100th anniversary of the American Mathematical Society, held jointly with the Mathematical Association of America in Atlanta, Georgia, January, 1988. A revised version of that lecture, "Cantor and the Epistemology of Set Theory," was presented in Montreal on June 3, 1995, as one of the Kenneth O. May lectures of the Canadian Society for History and Philosophy of Mathematics, at its Annual Meeting. Although a videotape of the AMS centennial lecture is distributed by the American Mathematical Society, the lecture itself has not been published until now, and I am grateful to the Canadian Society for History and Philosophy of Mathematics for including it in this memorial volume of May lectures. I am especially pleased to dedicate this paper to the memory of Kenneth O. May, from whom I learned so much in the course of working with him from 1977-1979 as Managing Editor of *Historia Mathematica*, the journal he founded in 1974.

[1]Poincaré 1908, p. 182.

[2]For Kronecker's criticism of Cantor, see Schoenflies 1927, p. 2.

It is also well-known that Cantor suffered throughout his life from a series of "nervous breakdowns" which became increasingly frequent and debilitating as he got older. Some have tried to link this to his dangerous flirtations with the infinite, but in the opinion of Karl Pönitz, who treated Cantor at the Halle Nervenklinik:

> Cantor's illness was basically endogenous, and probably showed some form of manic-depression: exogenous factors, such as the difficulties of his researches and the controversies in Halle University, are likely to have played only a small part in the genesis of his attacks, little more than the clap that starts the avalanche. Thus he would have suffered his attacks if he had pursued only an ordinary mundane career.[3]

Nevertheless it was all too easy for his early biographers to present Cantor, who was trying to defend his complex theory, as the hapless victim of the infinite, due to his increasingly long periods of mental breakdown that began in the 1880s, all of which were exacerbated by the persecutions of his contemporaries.[4] But such accounts distort the truth by trivializing the genuine intellectual concerns that motivated some of the most thoughtful contemporary opposition to Cantor's theory. They also fail to credit the power and scope of the defense he offered for his ideas in the battle to win acceptance for transfinite set theory.

At first Cantor himself resisted the implications of his research–because he had always believed that the idea of the actual infinite could not be consistently formulated, and so had no place in rigorous mathematics. Nevertheless, by his own account, he soon overcame his "prejudice" regarding the transfinite numbers because he found they were indispensable for the further development of his mathematics.[5] Because of his own early doubts he was able to anticipate opposition from diverse quarters, which he attempted to meet with philosophical and theological arguments as well as mathematical ones. Moreover, when he was called upon to respond to his critics, he was able to muster

[3]The description is from Grattan-Guinness 1971, pp. 268-69. Among the doctors who were responsible for treating Cantor at the *Universitätsnervenklinik Halle* were Karl Pönitz and a Dr. Mekus. See Purkert and Ilgauds 1985, pp. 52-59, and pp. 118-119; Purkert and Ilgauds 1987, pp. 79-82, and pp. 193-95. Cantor's mental condition has been analyzed in detail by the French Lacanian psychiatrist, Nathalie Charraud. See especially her chapter, "La maladie," in Charraud 1994, pp. 193-216.

[4]This is the view, among others, of Schoenflies 1927; and Bell 1937, chapter 29. Schoenflies's account, it should be noted, was concerned exclusively with Cantor's first major breakdown in 1884, and it was not difficult for him to draw explicit lines between Cantor's illness and specific anxieties which the climate of Cantor's research had produced. But his later bouts of manic-depression seem to reflect no such concerns or connections. According to the doctor who treated his next major episode of manic-depression in 1899, it was most likely triggered by the broken engagement of one of his daughters; see Purkert and Ilgauds 1985, p. 118; and Purkert and Ilgauds 1987, p. 193.

[5]Cantor 1883, p. 175.

his ideas with considerable force. His mental illness, far from playing an entirely negative role, in its manic phases may well have contributed to the energy and single-mindedness with which he promoted and defended his theory, just as the theological dimension of Cantor's understanding of the infinite also reassured him–in fact convinced him–of its absolute truth, regardless of what opponents like Kronecker might say against the theory.

Before it is possible to appreciate the origins, scope and significance of Cantor's battle to win acceptance for his transfinite numbers–the alephs–it will be helpful to say something, briefly, about his life and the early development of set theory.

Cantor's Early History

Georg Ferdinand Ludwig Philip Cantor was born on March 3, 1845, in St. Petersburg.[6] His mother, a Roman Catholic, came from a family of notable musicians; his father, a successful tradesman, was the son of a Jewish businessman, but a devout Lutheran, having been raised in a Lutheran mission in St. Petersburg. Cantor's father passed on his own deep religious convictions to his son. According to Eric Temple Bell's widely-read book, *Men of Mathematics*, first published in 1937, Georg Cantor's insecurities in later life stemmed from a ruinous Freudian conflict with his father, but surviving letters and other evidence concerning their relationship indicate quite the contrary. Georg's father appears to have been a sensitive man who was attentive to his children and took a special but not coercive interest in the welfare and education of his eldest son.[7]

When the young Cantor was still a child the family moved from Russia to Germany, and it was there that he began to study mathematics. After receiving his doctorate from the University of Berlin in 1868 for a dissertation on the theory of numbers, two years later he accepted a position as *Privatdocent*, or instructor, at the University of Halle, a respected institution but not as prestigious for mathematics as the universities at Göttingen or Berlin. One of his colleagues at Halle was Heinrich Eduard Heine, who was then working on the theory of trigonometric series, and he encouraged Cantor to take up the difficult problem of the uniqueness of such series. In 1872, when he was twenty-seven, Cantor published a paper that included a very general solution to the problem, along with his theory of real numbers, which contained

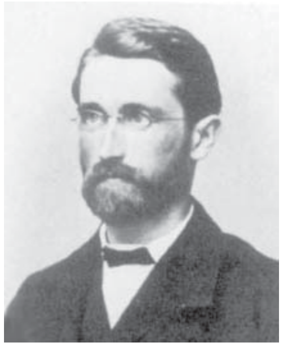
[6]The details provided here of Cantor's life and early career are drawn largely from my study, Dauben 1979, especially pp. 271-299. Other biographical studies worth consulting include Meschkowski 1967, Grattan-Guinness 1971, Purkert and Ilgauds 1985, and Purkert and Ilgauds 1987.

[7]Bell 1937, chapter 29. For alternative evaluations of Cantor's relationship with his father, see Dauben 1979, esp. pp. 272-280, and Charraud 1994.

the seeds of what would develop later into his theory of transfinite sets and numbers.[8]

Cantor in the 1870s

Eduard Heine
(1821–1881)



Richard Dedekind
(1831–1916)

Cantor's Discovery that the Real Numbers are Uncountably Infinite

Cantor was not alone in studying the properties of the continuum of real numbers in rigorous detail. In 1872, the same year that Cantor's paper on

[8]Cantor 1872. Dauben 1971 provides a detailed analysis of this paper and four others that preceded it, which led Cantor from an early version of his representation theorem in 1870 proving that if a function $f(x)$ is represented by a trigonometric series convergent for every value of x, then the series is unique, to the theorem of 1872 which established the uniqueness of the representation even for an infinite number of "exceptional" points. Remarkably, the trick Cantor needed to establish a limited version of this theorem (published in 1871), namely that for a certain finite number of values of x either the representation of the function or the convergence of the trigonometric series could be given up, was provided by Kronecker (see Dauben 1979, p. 34). Within the year Cantor realized that he could establish the uniqueness theorem even if an infinite number of exceptional points were permitted, provided that they were distributed in a particular way. This was to lead eventually to the first stages of Cantor's development of set theory in terms of his theory of point sets. To describe how the exceptional point sets were distributed over the continuum in the case of the representation theorem, he also found that he needed first to develop a rigorous theory of real numbers. Thus from a series of papers on the representation of functions by trigonometric series, Cantor was led to consider a rich concatenation of ideas that would prove especially fertile for his thinking about sets, the structure and nature of the continuum, and eventually, his theory of transfinite numbers over the next few decades.

trigonometric series appeared, the German mathematician Richard Dedekind also published an analysis of the continuum that was based on infinite sets. In his monograph on the subject of continuity and the irrational numbers, Dedekind articulated an idea that Cantor later made more precise:

> The line L is infinitely richer in point-individuals than is the domain R of rational numbers in number-individuals.[9]

Dedekind's statement, however, conceals a serious weakness. If anyone had asked Dedekind how much richer the infinite set of points in the continuum was than the infinite set of rational numbers, he could not have replied. Cantor's major contribution to this question was published in 1874, in Crelle's *Journal für die reine und angewandte Mathematik*. What Cantor showed was the nondenumerability of the real numbers, a discovery that was soon to transform much of modern mathematics. Cantor's paper was short, three pages, and bore a very strange title:

Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen.

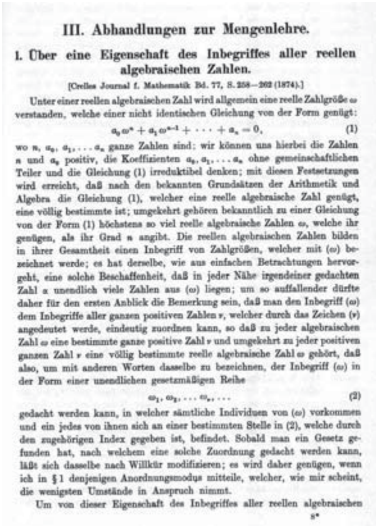
On a Property of the Collection of All Real Algebraic Numbers.[10]

No one scanning the title of this short paper would have guessed that this was the paper that disclosed Cantor's revolutionary discovery of the nondenumerability of the continuum of real numbers, which established that some infinite sets were larger than others, in particular that the set of natural numbers N was of a lower magnitude of infinity than the set of real numbers R. Instead, the article bore a deliberately misleading title suggesting that its major result was a theorem about algebraic numbers, thus failing even to hint at the more significant point that the paper actually contained. What could possibly have prompted Cantor to choose such an inappropriate title for what now, in retrospect, strikes any mathematician as one of the most important discoveries in modern mathematics?

The answer hinges on one of Cantor's teachers at Berlin, Leopold Kronecker. Having studied with Kronecker, Cantor was well-acquainted with Kronecker's work in number theory and algebra, and with his highly conservative philosophical views with respect to mathematics. By the early 1870s, Kronecker was already vocal in his opposition to any infinitary arguments, including the Bolzano-Weierstrass theorem and upper and lower limits, as well as to irrational numbers in general. Kronecker's later pronouncements against analysis and set theory, as well as his adamant and well-known insistence upon using the natural numbers to provide the only satisfactory foundation for mathematics, were simply extensions of these early views.

With this in mind, it is not unreasonable to suspect that Cantor had good reason to anticipate Kronecker's opposition to his proof of the nondenumerability of the real numbers, which proved that they comprised a set

[9] Dedekind 1872, p. 9.
[10] Cantor 1874.



III. Abhandlungen zur Mengenlehre.

1. Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen.

[Crelles Journal f. Mathematik Bd. 77, S. 258–262 (1874).]

Unter einer reellen algebraischen Zahl wird allgemein eine reelle Zahlgröße ω verstanden, welche einer nicht identischen Gleichung von der Form genügt:

$$a_0\omega^n + a_1\omega^{n-1} + \cdots + a_n = 0, \tag{1}$$

wo $n, a_0, a_1, \ldots a_n$ ganze Zahlen sind; wir können uns hierbei die Zahlen n und a_0 positiv, die Koeffizienten $a_0, a_1, \ldots a_n$ ohne gemeinschaftlichen Teiler und die Gleichung (1) irreduktibel denken; mit diesen Festsetzungen wird erreicht, daß nach den bekannten Grundsätzen der Arithmetik und Algebra die Gleichung (1), welcher eine reelle algebraische Zahl genügt, eine völlig bestimmte ist; umgekehrt gehören bekanntlich zu einer Gleichung von der Form (1) höchstens so viel reelle algebraische Zahlen ω, welche ihr genügen, als ihr Grad n angibt. Die reellen algebraischen Zahlen bilden in ihrer Gesamtheit einen Inbegriff von Zahlgrößen, welcher mit (ω) bezeichnet werde; es hat derselbe, wie aus einfachen Betrachtungen hervorgeht, eine solche Beschaffenheit, daß in jeder Nähe irgendeiner gedachten Zahl α unendlich viele Zahlen aus (ω) liegen; um so auffallender dürfte daher für den ersten Anblick die Bemerkung sein, daß man den Inbegriff (ω) dem Inbegriffe aller ganzen positiven Zahlen ν, welcher durch das Zeichen (ν) angedeutet werde, eindeutig zuordnen kann, so daß zu jeder algebraischen Zahl ω eine bestimmte ganze positive Zahl ν und umgekehrt zu jeder positiven ganzen Zahl ν eine völlig bestimmte reelle algebraische Zahl ω gehört, daß also, um mit anderen Worten dasselbe zu bezeichnen, der Inbegriff (ω) in der Form einer unendlichen gesetzmäßigen Reihe

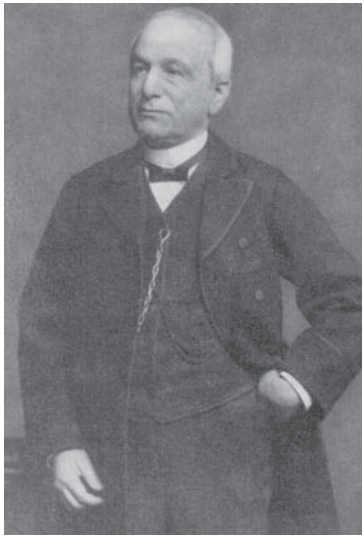
$$\omega_1, \omega_2, \ldots \omega_\nu, \ldots \tag{2}$$

gedacht werden kann, in welcher sämtliche Individuen von (ω) vorkommen und ein jedes von ihnen sich an einer bestimmten Stelle in (2), welche durch den zugehörigen Index gegeben ist, befindet. Sobald man ein Gesetz gefunden hat, nach welchem eine solche Zuordnung gedacht werden kann, läßt sich dasselbe nach Willkür modifizieren; es wird daher genügen, wenn ich in § 1 denjenigen Anordnungsmodus mitteile, welcher, wie mir scheint, die wenigsten Umstände in Anspruch nimmt.

Um von dieser Eigenschaft des Inbegriffes aller reellen algebraischen

8*

Cantor's paper in which he proves the nondenumerability of the real numbers.



Leopold Kronecker (1823–1891)

infinitely larger than the set of integers. Certainly any result (such as Cantor's) which confirmed the existence of transcendental numbers–against which Kronecker's opinions were well-known–would have been subject to criticism from Kronecker.

Worse yet, Kronecker was on the editorial board of the journal to which Cantor submitted his paper. Had Cantor been more direct with a title like "The Set of Real Numbers is Non-Denumerably Infinite," or "A New and Independent Proof of the Existence of Transcendental Numbers," he could have counted on a strongly negative reaction from Kronecker. In fact, when Ferdinand Lindemann later established the transcendence of π in 1882, meaning that it was not only irrational but also not an algebraic number, Kronecker asked what value the result could possibly have, since irrational numbers did not exist.[11] As Cantor contemplated publishing his paper on the nondenumerability of the real numbers in 1874, an innocuous title was clearly a strategic choice. Reference only to algebraic numbers would have had a much better chance of passing Kronecker's eye unnoticed, for there was nothing to excite either immediate interest or censure.

If the idea that Cantor may have harbored fears about Kronecker's opposition to his work at such an early date seems unwarranted, it is worth noting that Kronecker had already tried to dissuade Cantor's colleague at Halle, namely Heine, from publishing an article on trigonometric series in

[11]Kronecker made this remark in a lecture at the *Berliner Naturforscher-Versammlung* in 1886; see Weber 1893, p. 15; Kneser 1925, p. 221; and Pierpoint 1928, p. 39.

Crelle's Journal. Although Heine's article eventually appeared, Kronecker was at least successful in delaying its appearance, about which Heine was particularly vocal in letters to Schwarz, also a friend of Cantor's. Heine explained the circumstances of the situation in a letter to Schwarz on May 26, 1870:

> My little work "On Trigonometric Series," of which I now have the page-proofs in hand, and which at present is still under debate with Kronecker, who wanted to persuade me to retract it (the particulars below), appears in the current volume (71) of the journal p. 353, and has made me very happy; I had sent it to Borchardt in February, where Kr[onecker] saw it and to whom it was given and he kept it without my knowing it, until I came to Berlin.[12]

Heine to Schwarz, May 26, 1870.

Cantor and his wife, Vally, about 1880.

Doubtless Schwarz and Heine would both have brought Kronecker's readiness (and ability) to block ideas with which he disagreed to Cantor's attention. Indeed, several years later Kronecker also delayed publication of a paper Cantor had written on the invariance of dimension.[13] This so angered Cantor that

[12]The complete letter is transcribed in Dauben 1979, pp. 308-09.

[13]Cantor 1878; Cantor wrote a bitter letter to Dedekind about the incident and even planned to withdraw the paper from *Crelle's Journal*,a step Dedekind persuaded him not to take. See Cantor's letter to Dedekind of October 23, 1887, in Cantor/Dedekind 1937, p. 40. Cantor gave a lecture in Braunschweig in 1897 in which he recalled the incident and said that it was Weierstrass who had interceded on his behalf, thanks to which the paper was eventually published. See Fraenkel 1930, p. 10; full details are given in Dauben 1979, pp. 66-70.

he never submitted anything to *Crelle's Journal* again. A decade later he regarded Kronecker as both a private and a public menace–not only because he was condemning set theory openly, but Weierstrassian analysis as well.

There was however a positive side to Kronecker's opposition to Cantor's work, for it forced Cantor to evaluate the foundations of set theory as he was in the process of creating it. This concern prompted long philosophical passages in Cantor's major publication of the 1880s on set theory, his *Grundlagen einer allgemeinen Mannigfaltigkeitslehre* of 1883. It was here that Cantor issued one of his most famous pronouncements about mathematics, namely that the essence of mathematics is exactly its freedom.[14] This was not simply an academic or philosophical message to his colleagues, for it also carried a hidden and deeply personal subtext. It was, as he later admitted to David Hilbert, a plea for objectivity and openness among mathematicians. This, he said, was inspired directly by the oppression and authoritarian closed-mindedness that he felt Kronecker represented, and worse, had wielded in a flagrant and damaging way against those he opposed.

Thus at the very beginning of his career, even before he had begun to develop any of his more provocative ideas about transfinite set theory, Cantor had experienced his first bitter taste of Kronecker's opposition to his work. Doubtless Cantor knew that he could expect more trouble in future.

Set Theory Begins to Develop

Meanwhile, Cantor devoted himself to developing further the ideas about point sets which he had first investigated in the context of representing functions by trigonometric series in the 1870s. By the end of the decade he had married Vally Guttman, and shortly thereafter he began to publish an important series of papers on linear point sets. These eventually led to his *Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen*, which emphasized transfinite ordinal numbers introduced rather vaguely in terms of what Cantor termed "principles of generation." The *Grundlagen* also offered a detailed philosophical defense of his new ideas on the infinite.[15] It was in 1883 that Cantor first tried vigorously to establish his Continuum Hypothesis in the version that the set of real numbers was the next largest after the denumerable set of natural numbers.

Despite his vigorous efforts to prove the correctness of the Continuum Hypothesis, he was greatly frustrated by his inability to do so. Early in 1884 he thought he had found a proof, but a few days later he reversed himself completely and thought he could actually disprove the hypothesis. Finally he

[14] Cantor 1883, p. 182.

[15] Cantor 1883, pp. 131-132.

realized that he had made no progress at all, as he reported in letters to his friend and editor of *Acta Mathematica*, Gösta Mittag-Leffler in Stockholm.[16]

Cantor in the 1880s.

All the while Cantor was facing mounting opposition and threats from Kronecker, who said he was preparing an article to show that "the results of modern function theory and set theory are of no real significance."[17]

Soon thereafter, in May of 1884, Cantor suffered the first of his serious nervous breakdowns. Although his lack of progress on the Continuum Hypothesis or stress from Kronecker's attacks may have helped to trigger the breakdown, it now seems clear that such events had little to do with its underlying cause. The illness took over with startling speed and lasted somewhat longer than a month. At the time, only the manic phase of manic-depressive psychosis was recognized as a symptom. When Cantor "recovered" at the end of June, 1884, and entered the depressive phase of his illness, he complained that he lacked the energy and interest to return to rigorous mathematical thinking. He was content to take care of trifling administrative matters at the university, but felt capable of little more.

Although Cantor eventually returned to mathematics, he also became increasingly absorbed in other interests. He undertook a study of English history and literature and became engrossed in a scholarly diversion that was taken

[16]See especially letters Cantor wrote to Mittag-Leffler between August and November of 1884, in Meschkowski 1967, pp. 240-241 and p. 243; and in Schoenflies 1927, p. 12 and pp. 17-18.

[17]Quoted by Cantor in a letter to Mittag-Leffler dated January 26, 1884, in Schoenflies (1927), p. 5.

with remarkable seriousness by many people at the time: namely, the supposition that Francis Bacon was the true author of Shakespeare's plays. Cantor also tried his hand without success at teaching philosophy instead of mathematics, and he began to correspond with several theologians who had taken an interest in the philosophical implications of his theories about the infinite. This correspondence was of special significance to Cantor because he was convinced that the transfinite numbers had come to him as a message from God. But more about the significance of this, as already promised, in a moment.

Transfinite Cardinal Numbers: Cantor's Alephs

There is still one last element of Cantor's technical development of transfinite set theory that needs to be mentioned as part of his continuing efforts to mount a convincing and satisfactory mathematical defense of his ideas, namely, the nature and status of the transfinite cardinal numbers. The evolution of Cantor's thinking about the transfinite cardinals is curious, because although the alephs are probably the best-known legacy of Cantor's creation, they were the last part of his theory to be given either rigorous definition or a special symbol. Indeed, it is difficult in the clarity of hindsight to reconstruct the obscurity within which Cantor must have been groping, and up to now his work has been discussed here largely as if he had already recognized that the power of an infinite set could be understood as a cardinal number. In fact, beginning in the early 1880s, Cantor first introduced notation for his infinite (actually transfinite) sequence of derived sets P^ν, extending them well beyond the limitation he had earlier set himself to sets of the first species.[18] At this time he spoke of the indexes only as "infinite symbols" with no reference to them in any way as numbers.

By the time he wrote the *Grundlagen* in 1883, the transfinite ordinal numbers had finally achieved independent status as numbers and were given the familiar omega notation, ω being the first transfinite ordinal number following the entire sequence of finite ordinal numbers, i.e., $1, 2, 3, \ldots, \omega$. However, there was no mention whatsoever of transfinite cardinal numbers, although Cantor clearly understood that it is the power of a set that establishes its equivalence (or lack thereof) with any other set; from this he would eventually develop his concept of transfinite cardinal numbers. But in the *Grundlagen*, he carefully avoided any suggestion that the power of an infinite set could be interpreted as a number.

Soon, however, he began to equate the two concepts and in September of 1883, did so in a lecture to mathematicians at a meeting in Freiburg. Even

[18] An infinite set P was said to be of the *first* species if there were some finite number ν for which the νth derived set of limit points of P was empty, i.e., $P^\nu = \emptyset$. Infinite sets for which there was no such ν were sets of the *second* species. See Cantor 1872, and Dauben 1979, pp. 43-45.

so, no symbol was as yet provided for distinguishing one transfinite cardinal number from another. Since he had already adopted the symbol ω to designate the least transfinite ordinal number, when Cantor finally introduced a symbol for the first transfinite cardinal number, it was borrowed from the symbols already in service for the transfinite ordinals. By 1886, in correspondence, Cantor had begun to represent the first transfinite cardinal as $\overset{*}{\omega}$; the next larger he denoted $\overset{*}{\Omega}$. This notation was not very flexible, and within months Cantor realized the need for a more general notation capable of representing the entire ascending hierarchy of transfinite cardinals. Temporarily he used fraktur $\mathfrak{o}$'s, obviously derivatives from his omegas, to represent his sequence of cardinal numbers. For a time Cantor actually used superscripted stars, bars, and his fraktur $\mathfrak{o}$'s interchangeably for transfinite cardinals, without feeling any need to decide upon one or the other notation as preferable.[19]

The Paradoxes of Set Theory

However, in 1893 the Italian mathematician Giulio Vivanti was preparing a general account of set theory, and Cantor realized it was time to adopt a standard notation. Only then did he choose the Hebrew aleph ($\aleph$) for the transfinite cardinals, because he thought the familiar Greek and Roman alphabets were too common and already widely employed in mathematics for other purposes. His new numbers deserved something distinct and unique–and the Hebrew alphabet was readily available among the type fonts of German printers. The choice was particularly clever, as Cantor was pleased to admit, because the Hebrew aleph was also the symbol for the number one. Since the transfinite cardinal numbers were themselves infinite unities, the aleph could be taken to represent a new beginning for mathematics. Cantor designated the cardinal number of the first number class $\aleph_1$ in 1893, but in 1895 changed his mind; from then on, he used $\aleph_0$ to represent the first and least transfinite cardinal number, the number he had previously designated by $\overset{*}{\omega}$. From $\aleph_0$, he went on to designate the cardinal number of the second number class as $\aleph_1$, after which there followed an unending sequence of transfinite cardinal numbers.

Cantor made his last major contributions to set theory in 1895 and 1897. He had already used his famous method of diagonalization in 1891 to show that the cardinal number of any set P is always less than the cardinal number of its power-set, the set of all subsets of P. A few years later he presented a corollary to this result, namely that the cardinal number of the continuum is equal to $2^{\aleph_0}$, and he hoped this result would soon lead to a solution of the Continuum Hypothesis–which he could now express as $2^{\aleph_0} = \aleph_1$.

[19] For a detailed discussion of the evolution of Cantor's notation for the transfinite cardinal numbers, see Dauben 1979, pp. 179-183.

The arguments Cantor used in his proof about the cardinal number of the power-set of all subsets of any given set, however, led to far different conclusions. Rather than leading to a resolution of the Continuum Hypothesis, they led directly to the discovery of the paradoxes of set theory, for the fact that there could be no "largest" transfinite cardinal number immediately raised the question of the cardinality of the set of "all" transfinite cardinal numbers. Cantor resolved the problem by excluding this possibility entirely; the aggregate of "all" transfinite numbers was what he called an "inconsistent" aggregate, and therefore was simply not to be considered as a "set." Bertrand Russell, in contemplating this problem, drew far more problematic conclusions, for what he discovered was that a paradox can be derived in set theory by considering those sets that do not include themselves as members.[20]

Russell's paradox suggested that there was something fundamentally wrong with Cantor's definition of a set, and the consequences of this realization immediately became an important problem in 20th-century mathematics. Even before Bertrand Russell, however, Cantor had already come upon his own version of the paradoxes of set theory in the form of contradictions he associated with the idea of a largest ordinal or cardinal number. This was all explained in letters first to Hilbert in 1897, and then to Dedekind in 1899. As Cantor wrote to Dedekind on August 3, 1899, if one considers the collection of *all* transfinite ordinal numbers Ω, "the system Ω of all numbers is an inconsistent, absolutely infinite aggregate."[21]

But it is possible that Cantor may have been aware of the paradoxes of set theory much earlier, perhaps as early as the 1880s when his difficulties with Kronecker were weighing on his mind and as he was just beginning to experience his first serious technical problems with set theory. For example, in his *Grundlagen* of 1883, Cantor referred to collections that are too large to be comprehended as a well-defined, completed, unified entity. At the time he wrote obscurely, with references to absolute sets in explicitly theological terms, explaining that "the true infinite or Absolute, which is in God, permits no determination."[22] Was this a hint that he already understood that the collection of all transfinite ordinal numbers was inconsistent, and therefore not to be regarded as a set? Later, Cantor said that it was–that he meant this to be a veiled sign, even then, that he was aware of the paradoxical results that followed from trying to determine what transfinite ordinal number should correspond to the well-ordered set of all transfinite ordinal numbers.

---

[20] In contemplating the result of Cantor's diagonalization proof, Russell considered the implications of the fact that there could be no one-to-one correspondence between the elements of a set P and the elements of its power set. In asking himself what elements of the power set were left out of such a correspondence, Russell was led to the discovery of his paradox of sets which are not members of themselves. For details see Russell 1907; Dauben 1979, pp. 261-263; and Chapter IV of Garciadiego 1992: "Russell's discovery of the 'paradoxes'," pp. 81-130.

[21] Cantor to Dedekind, August 3, 1899, in Cantor 1932, p. 445.

[22] Cantor 1883, Note 2 to Section 4 of the *Grundlagen*, p. 205.

By the mid-1890s, Cantor could no longer be so vague about absolute entities, and was forced to be much more explicit about the paradoxes that arose from contemplating the sets of *all* transfinite ordinal or cardinal numbers. The solution Cantor chose for dealing with such mathematical paradoxes was simply to exclude them from set theory. Anything that was too large to be comprehended as a well-defined, unified, consistent set was declared inconsistent. These were "absolute" collections, and lay beyond the possibility of mathematical determination. This, in essence, is what Cantor communicated first to Hilbert in 1897, and somewhat later to Dedekind in Cantor's letters of 1899.[23]

Foundations and Philosophy of Mathematics

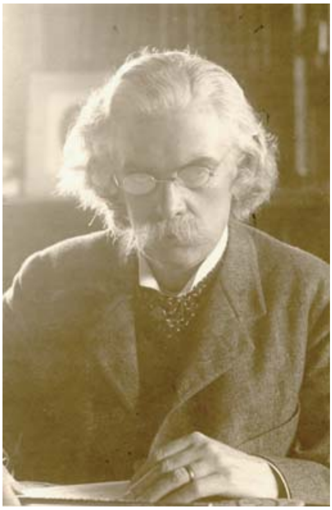
Whatever the extent of Cantor's awareness of the paradoxes may have been in the early 1880s, he was certainly sensitive to Kronecker's, growing and increasingly vocal opposition. Above all, it is clear that explicitly philosophical concerns expressed in his *Grundlagen* were in Cantor's opinion strategically crucial for a comprehensive defense of his new theory. This was unusual at the time; it still is. When Mittag-Leffler arranged to publish French translations of Cantor's papers on set theory for Mittag-Leffler's newly-founded journal, *Acta Mathematica*, he persuaded Cantor that it would be best to omit all of the philosophical portions of the *Grundlagen* as unnecessary (and possibly repugnant) to mathematicians who might find the theory of interest but the philosophy unacceptable.[24]

The philosophical arguments, however, were essential to Cantor, if not to Mittag-Leffler. They were essential because they were part of the elaborate defense he had begun to construct to subvert opposition from any quarter, but especially from Kronecker. One particularly important ploy was to advance a justification of transfinite set theory based upon the freedom of mathematics to admit any self-consistent theory. Applications might eventually determine which mathematical theories were useful, but for mathematicians, Cantor insisted that the only real question was consistency. This of course was just the interpretation he needed to challenge an established mathematician like Kronecker. Cantor clearly felt obliged, early in his career, to plead as best he could for a fair hearing of his work. So long as it was self-consistent it should

[23]Cantor's letters to Hilbert about the "absolute" character of the collection of all transfinite numbers were long thought to be lost (or nonexistent), but two of Cantor's letters on this subject to Hilbert dated September 26 and October 2, 1887, have recently been published by Walter Purkert and Hans Joachim Ilgauds; see their *Georg Cantor 1845-1918* (1987), pp. 224-227. These same two letters are also reproduced in Meschkowski 1991, pp. 388-390.

[24]Mittag-Leffler, in a letter to Cantor dated March 11, 1883, in the archives of the Institut Mittag-Leffler, Djursholm, Sweden; cited in Dauben 1979, p. 297.

be taken as mathematically legitimate, and the constructivist, finitist criticisms of Kronecker might be disregarded by most mathematicians, for whom consistency alone should be the viable touchstone.



Gösta Mittag-Leffler (1846-1927)

The Freedom of Mathematics

Cantor put his philosophy about the freedom of mathematics into action early in the 1890s, when his career had reached the point where he could do more than simply write about it. During the 1880s he had already begun to lay the strategic foundations for an independent union of mathematicians in Germany. The specific goal of such a union, as he often made clear in his correspondence, was to provide an open forum, especially for young mathematicians. The union (as Cantor envisaged it) would guarantee that anyone could expect free and open discussion of mathematical results without prejudicial censure from members of the older establishment, whose conservatism might easily ruin the career of an aspiring mathematician. This was primarily needed in cases where the ideas in question were at all new, revolutionary or controversial.

Cantor labored intensively for the creation of the Deutsche Mathematiker-Vereinigung.[25] Eventually, agreement was reached and the Union of German Mathematicians held its first meeting in conjunction with the annual meeting of the Gesellschaft Deutscher Naturforscher und Ärzte at Halle in 1891.

[25]For details, see Dauben 1979, pp, 160-163.

Cantor was elected the Union's first president, and at its inaugural meeting he presented his now famous proof that the real numbers are nondenumerable using his new method of diagonalization.[26]

The German Union was not the end of Cantor's vision. He also recognized the need to promote international forums, and thus began lobbying for international congresses shortly after formation of the DMV. These were eventually organized through the cooperative efforts of many mathematicians, and not directly as a result of Cantor's exclusive efforts by any means. The first of these was held in Zürich in 1897, the second in Paris in 1900.[27]

Promoting new avenues for discussion of mathematics was one way Cantor reacted to opposition of the sort his own research had provoked. Despite criticisms, especially from Kronecker, Cantor persevered, even in the face of his own repeated failure to resolve some of the most basic questions about set theory (notably his Continuum Hypothesis), and even though he began to suffer increasingly serious cycles of manic-depression. Ironically, like his conflicts with Kronecker, Cantor's manic-depression may have served a useful purpose. In his own mind it was closely linked to the infallible support set theory drew from his strongly-held religious convictions. Letters (and the testimony of colleagues who knew him) reveal that Cantor believed he was chosen by God to bring the truths of set theory to a wider audience. He also regarded the successive waves of manic-depression that began to plague him in the 1880s–peaks of intense activity followed by increasingly prolonged intervals of introspection–as divinely inspired. Long periods of isolation in hospital provided opportunities for uninterrupted reflection during which Cantor envisioned visits from a muse whose voice reassured him of the absolute truth of set theory, whatever others might say about it.

In promoting set theory among mathematicians, philosophers, and theologians (he even wrote to Pope Leo XIII at one point on the subject of the infinite), Cantor was convinced he would succeed in securing the recognition that transfinite set theory deserved.[28] By stressing self-consistency and the intrinsic freedom of mathematics, he also advanced an essential element of any intellectual inquiry, namely that the mind must be free to pursue the truth wherever it may lead. Inspiration should be encouraged, not confounded by arbitrary prejudice, and for Cantor this meant that theories should be judged upon standards of consistency and utility.

[26] Cantor 1891.

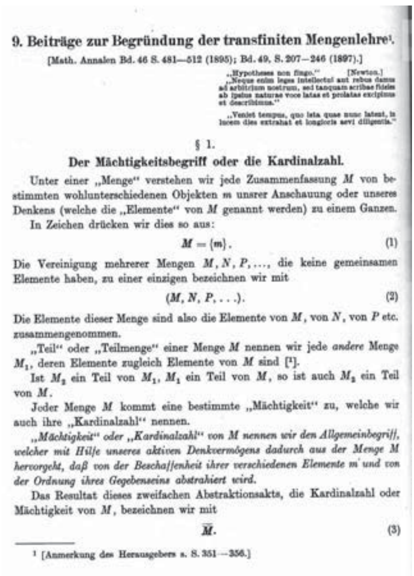
[27] For details, see Dauben 1979, pp. 163-165.

[28] Cantor wrote in Latin to Pope Leo XIII, February 13, 1896; transcribed in Purkert and Ilgauds 1987, pp. 198-199; published in Latin with German translation in Meschkowski 1991, p. 383.

Transfinite Mathematics and Cantor's Manic-Depression

To understand Cantor's tenacious promotion and defense of set theory, especially in his later years after the publication of the *Beiträge* of 1895/1897, it is important to appreciate the connection between Cantor's faith in God, his mental illness, and his mathematics. Certain documents suggest that in addition to enforcing periodic intervals of contemplation and withdrawal from daily affairs, Cantor's periods of depression were productive in other ways. In fact, he was often able to pursue his mathematical ideas in the solitude of the hospital or quietly at home. This may have supported his belief that the transfinite numbers had been communicated to him from God. In fact, as he noted in the third motto he chose to head his last publication, the *Beiträge* of 1895:

> *Veniet tempus, quo ista quae nunc latent, in lucem dies extrahat et longioris aevi diligentia.*
>
> The time will come when these things which are now hidden from you will be brought into the light.[29]



9. Beiträge zur Begründung der transfiniten Mengenlehre[1].

[Math. Annalen Bd. 46 S. 481—512 (1895); Bd. 49, S. 207—246 (1897).]

„Hypotheses non fingo." [Newton.]

„Neque enim leges intellectui aut rebus damus ad arbitrium nostrum, sed tanquam scribae fideles ab ipsius naturae voce latas et prolatas excipimus et describimus."

„Veniet tempus, quo ista quae nunc latent, in lucem dies extrahat et longioris aevi diligentia."

§ 1.

Der Mächtigkeitsbegriff oder die Kardinalzahl.

Unter einer „Menge" verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unsrer Anschauung oder unseres Denkens (welche die „Elemente" von M genannt werden) zu einem Ganzen. In Zeichen drücken wir dies so aus:

$$M = \{m\}. \tag{1}$$

Die Vereinigung mehrerer Mengen $M, N, P, \ldots$, die keine gemeinsamen Elemente haben, zu einer einzigen bezeichnen wir mit

$$(M, N, P, \ldots). \tag{2}$$

Die Elemente dieser Menge sind also die Elemente von M, von N, von P etc. zusammengenommen.

„Teil" oder „Teilmenge" einer Menge M nennen wir jede *andere* Menge M_1, deren Elemente zugleich Elemente von M sind [1].

Ist M_2 ein Teil von M_1, M_1 ein Teil von M, so ist auch M_2 ein Teil von M.

Jeder Menge M kommt eine bestimmte „Mächtigkeit" zu, welche wir auch ihre „Kardinalzahl" nennen.

„Mächtigkeit" oder „Kardinalzahl" von M nennen wir den Allgemeinbegriff, welcher mit Hilfe unseres aktiven Denkvermögens dadurch aus der Menge M hervorgeht, daß von der Beschaffenheit ihrer verschiedenen Elemente m und von der Ordnung ihres Gegebenseins abstrahiert wird.

Das Resultat dieses zweifachen Abstraktionsakts, die Kardinalzahl oder Mächtigkeit von M, bezeichnen wir mit

$$\overline{\overline{M}}. \tag{3}$$

[1] [Anmerkung des Herausgebers s. S. 351—356.]

Cantor's *Beiträge* of 1895.

This is a familiar passage from the Bible and reflects Cantor's belief that he was an intermediary serving as the means of revelation. It may also be taken

[29] The Bible, I Corinthians 4:5.

to reflect Cantor's faith that despite any prevailing resistance to his work, it would one day enjoy recognition and praise from mathematicians everywhere.

It is easy, of course, to misinterpret the religious element in Cantor's thinking, as popularizers often do. This was certainly the case in an article that appeared in 1977 in the French magazine *La Recherche*, which supplied the following caricatures to illustrate an expository article about Cantor, his religious convictions, psychological illness, and transfinite set theory. The first drawing depicts Cantor in ecstasy, as it were, receiving the divine message:

Cantor in ecstasy!

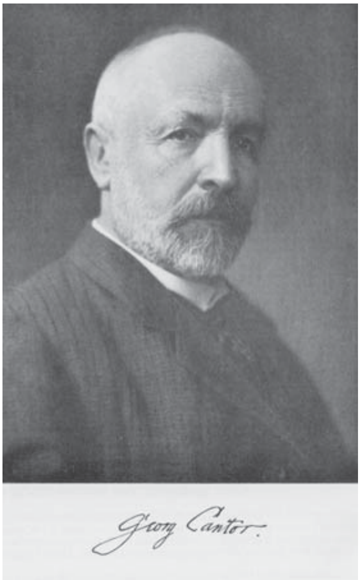
The precarious balance.

In the second illustration, the figure with the gun is meant to be Leopold Kronecker–with God helping Cantor to maintain his balance–all of which rests precariously on a transfinite aleph.[30]

But there is a very serious side to all of this and it deserves to be emphasized. For example, following a long period of hospitalization in 1908, Cantor wrote to a friend in Göttingen, the British mathematician Grace Chisholm Young. As he described it, his manic-depression took on a strikingly creative quality:

[30]The two drawings above from the article on Cantor by Pierre Thuillier [Thuillier 1977] are reproduced here by kind permission of the artist, André Barbe, and the editors of *La Recherche*.

> A peculiar fate, which thank goodness has in no way broken me, but in fact has made me stronger inwardly... has kept me far from home– I can say also far from the world... In my lengthy isolation neither mathematics nor in particular the theory of transfinite numbers has slept or lain fallow in me.[31]



Elsewhere, Cantor actually described his conviction about the truth of his theory explicitly in quasi-religious terms:

> My theory stands as firm as a rock; every arrow directed against it will return quickly to its archer. How do I know this? Because I have studied it from all sides for many years; because I have examined all objections which have ever been made against the infinite numbers; and above all, because I have followed its roots, so to speak, to the first infallible cause of all created things. [32]

Later generations might dismiss the philosophy, look askance at Cantor's abundant references to St. Thomas or to the Church Fathers, overlook his metaphysical pronouncements, and miss entirely the deeply religious roots of Cantor's later faith in the absolute truth of his theory. But all these commitments contributed to his resolve not to abandon the transfinite numbers.

[31] Cantor to Grace Chisholm Young, June 20, 1908; transcribed in Meschkowski 1971, pp. 30–34; translated in Dauben 1979, p. 290.

[32] Cantor to K. F. Heman, June 21, 1888; quoted from Dauben 1979, p. 298.

Opposition seems to have strengthened his determination. His forbearance, as much as anything else Georg Cantor might have contributed, ensured that set theory would survive the early years of doubt and denunciation to flourish eventually as a vigorous, revolutionary force in 20th-century mathematics.

Bibliography

Charraud, Nathalie, *Infini et Inconscient. Essai sur Georg Cantor* (Paris: Anthropos, 1994).

Cantor, Georg, "Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen," *Mathematische Annalen*, 5 (1872), pp. 123–132; in Cantor 1932, pp. 92–102; translated as "Extension d'un théorème de la théorie des series trigonométriques," *Acta Mathematica*, 2 (1883), pp. 336-3-48.

———, "Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen," *Journal für die reine und angewandte Mathematik*, 77 (1874), pp. 258–262; in Cantor (1932), pp. 115–118; translated as "Sur une propriété du système de tous les nombres algébriques reels," *Acta Mathematica*, 2 (1883), pp. 205–310.

———, *Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen* (Leipzig: B.G. Teubner, 1883). In Cantor (1932), pp. 165–208. Translated, in part, into French as "Fondaments d'une Théorie générale des ensembles," *Acta Mathematica*, 2 (1883), pp. 381-408; and into English as "Foundations of the Theory of Manifolds" (trans. U. Parpart), *The Campaigner* (The Theoretical Journal of the National Caucus of Labor Committees), 9 (January and February, 1976), pp. 69–96; a better translation into English is by W. B. Ewald, "Foundations of a General Theory of Manifolds: A Mathematico-Philosophical Investigation into the Theory of the Infinite," in *From Kant to Hilbert: a Source Book in the Foundations of Mathematics*, ed. W. B. Ewald (New York: Oxford University Press, 1996), vol. 2, pp. 878–920.

———, "Über eine elementare Frage der Mannigfaltigkeitslehre," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 1 (1891), pp. 75–78; in Cantor 1932, pp. 278–280.

———, *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, ed. E. Zermelo (Berlin: J. Springer; rep. Hildesheim: Olms, 1966, and Berlin: Springer, 1990).

Cantor, Georg, and Richard Dedekind, *Briefwechsel Cantor-Dedekind*, eds. E. Noether and J. Cavaillès (Paris: Hermann, 1937).

Dauben, Joseph W., "The Trigonometric Background to Georg Cantor's Theory of Sets," *Archive for History of Exact Sciences*, 7 (1971), pp. 181–216.

_______, *Georg Cantor: His Mathematics and Philosophy of the Infinite* (Cambridge, MA: Harvard University Press, 1979; rep. Princeton: Princeton University Press, 1990).

Dedekind, Richard, *Stetigkeit und irrationale Zahlen* (2nd ed. Braunschweig: Vieweg, 1892); English translation: Dedekind 1901/1963.

_______, *Essays on the Theory of Numbers, Continuity of Irrational Numbers, the Nature and Meaning of Numbers*, trans. W.W. Beman (Chicago: Open Court, 1901; repr. New York: Dover, 1963).

Fraenkel, A., "Georg Cantor," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 39 (1930), pp. 189–266.

Garciadiego, Alejandro R., *Bertrand Russell and the Origins of the Set-theoretic 'Paradoxes'* (Basel: Birkhäuser Verlag, 1992).

Grattan-Guinness, Ivor, "Towards a Biography of Georg Cantor," *Annals of Science*, 27 (1971), pp. 345–391.

Kneser, A., "Leopold Kronecker," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 33 (1925), pp. 210–228.

Meschkowski, Herbert, *Probleme des Unendlichen. Werk und Leben Georg Cantors* (Braunschweig: Vieweg, 1967).

_______, "Zwei unveröffentlichte Briefe Georg Cantors," *Der Mathematikunterricht*, 4 (1971), pp. 30-34.

_______, *Georg Cantor–Briefe*, eds. Herbert Meschkowski and Winfried Nilson (Berlin: Springer-Verlag, 1991).

Moore, Gregory H., *Zermelo's Axiom of Choice* (New York: Springer-Verlag, 1982).

Pierpont, J., "Mathematical Rigor, Past and Present," *Bulletin of the American Mathematical Society*, 34 (1928), pp. 23–53.

Poincaré, Henri, "L'avenir des mathématiques," *Atti del IV Congresso Internazionale dei Matematici*, Rome, 6-11 April, 1908 (Rome: Tipografia della R. Accademia dei Lincei, C.V. Salviucci, 1909), pp. 167–182.

Purkert, Walter, and Hans Joachim Ilgauds, *Georg Cantor (Biographien hervorragender Naturwissenschaftler, Techniker und Mediziner*, vol. 79, Leipzig: B.G. Teubner, 1985).

_______, *Georg Cantor 1845-1918* (*Vita Mathematica*, vol. 1, Basel: Birkhäuser Verlag, 1987).

Russell, Bertrand, "On Some Difficulties in the Theory of Transfinite Numbers and Order Types," *Proceedings of the London Mathematical Society*, 4 (1907), pp. 29-53.

Schoenflies, A., "Die Krisis in Cantor's mathematischem Schaffen," *Acta Mathematica* 50 (1927), pp. 1–23.

Thuillier, Pierre, "Dieu, Cantor et l'infini," *La Recherche*, 84 (December, 1977), pp. 1110–1116.

Weber, A., "Leopold Kronecker," *Mathematische Annalen*, 43 (1893), pp. 1–25.

10

Hilbert and his Twenty-Four Problems

Rüdiger Thiele

Karl-Sudhoff-Institut für Geschichte der Medizin und der Naturwissenschaften, University of Leipzig

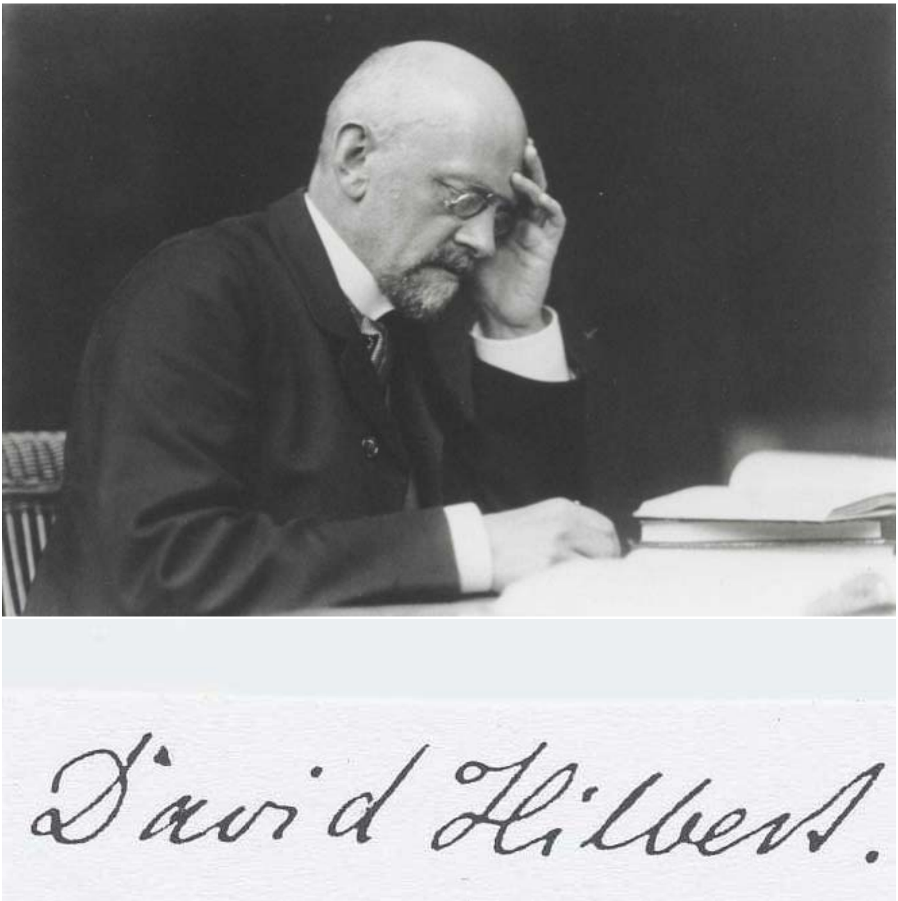
> If you can look into the seeds of time,
> And say which grain will grow, and which will not,
> Speak then to me.
> William Shakespeare, Macbeth (I, 3)

At the turn of the 19th century David Hilbert (1862-1943) was well known for fundamental results in invariant theory, for his profound *Zahlbericht* (*Report on the Theory of Numbers*), and the far-reaching *Grundlagen der Geometrie* (*Foundations of Geometry*; 14 German editions) which opened the way for the axiomatic method in mathematics. A substantial part of Hilbert's fame, however, rests on his address "Mathematical Problems", delivered at the second International Congress of Mathematicians (ICM) in Paris. Although the discussion which followed Hilbert's lecture on that summer morning of August 8, 1900 was desultory, "it became quite clear" [by the printed versions of the lecture in some languages], as Constance Reid (born 1917) remarked, "that David Hilbert had captured the imagination of the mathematical world with his list for the 20th century. His practical experience seemed to guarantee that they met the criteria which he had set up in his lecture, his judgment, that they could actually be solved in the years to come."[1]

Elie Cartan (1869–1951) emphasized the importance of this speech in a letter to Constantin Carathéodory (1873–1950) which was written shortly after Hilbert's death: "We will never hear the like of such a talk at congresses."[2] On the other hand, it was Carathéodory who pointed out that these 23 problems divided Hilbert's career into two parts. In Germany up to this speech Hilbert was a respected mathematician; after the talk his international fame grew rapidly (partly because the American Mathematical Society very quickly supplied English readers with both a report and a translation

[1] Reid, 1996, p. 84.

[2] Carathéodory 1943, p. 350. "On n'entendra plus dans les Congrès de conférence pareille."



David Hilbert
1862-1943
Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Voit Collection.

of Hilbert's speech).[3] In the second period David Hilbert gathered more and more pupils and his seed began to grow. Hilbert had become one of the most famous mathematicians of the day, his fame possibly exceeded only by that of Henri Poincaré (1854–1912). Anyone able to solve one of the problems could instantly make a reputation for himself.

Immediately mathematicians set about their work and the mathematical community watched each contribution attentively. Hermann Weyl (1885–1955) once remarked:

[3] Hilbert 1902, 2000; G.B. Halsted, 1900; Scott 1900.

Nachrichten

von der

Königl. Gesellschaft der Wissenschaften

zu Göttingen.

Mathematisch-physikalische Klasse

aus dem Jahre 1900.

Göttingen,
Commissionsverlag der Dieterich'schen Universitätsbuchhandlung
Lüder Horstmann.
1900.

Mathematische Probleme.

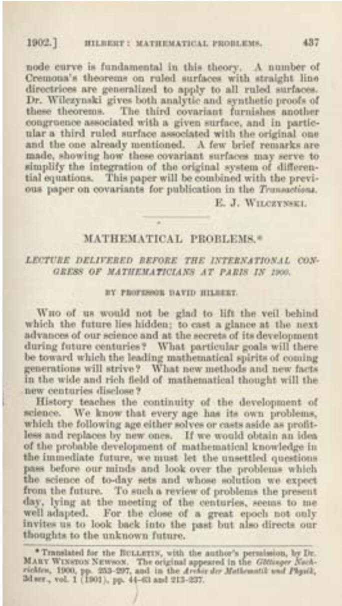
Vortrag, gehalten auf dem internationalen Mathematiker-Kongreß zu Paris 1900.

Von

D. Hilbert.

Wer von uns würde nicht gern den Schleier lüften, unter dem die Zukunft verborgen liegt, um einen Blick zu werfen auf die bevorstehenden Fortschritte unsrer Wissenschaft und in die Geheimnisse ihrer Entwickelung während der künftigen Jahrhunderte! Welche besonderen Ziele werden es sein, denen die führenden mathematischen Geister der kommenden Geschlechter nachstreben? welche neuen Methoden und neuen Thatsachen werden die neuen Jahrhunderte entdecken — auf dem weiten und reichen Felde mathematischen Denkens?

Fig. 10.1. Hilbert's Paris lecture "Mathematische Probleme (Mathematical Problems)" was first published in *Göttinger Nachrichten* 1900. Courtesy Mathematisches Institut, Universität Leipzig.



1902.] HILBERT: MATHEMATICAL PROBLEMS. 437

node curve is fundamental in this theory. A number of Cremona's theorems on ruled surfaces with straight line directrices are generalized to apply to all ruled surfaces. Dr. Wilczynski gives both analytic and synthetic proofs of these theorems. The third covariant furnishes another congruence associated with a given surface, and in particular a third ruled surface associated with the original one and the one already mentioned. A few brief remarks are made, showing how these covariant surfaces may serve to simplify the integration of the original system of differential equations. This paper will be combined with the previous paper on covariants for publication in the *Transactions*.

E. J. Wilczynski.

MATHEMATICAL PROBLEMS.*

LECTURE DELIVERED BEFORE THE INTERNATIONAL CONGRESS OF MATHEMATICIANS AT PARIS IN 1900.

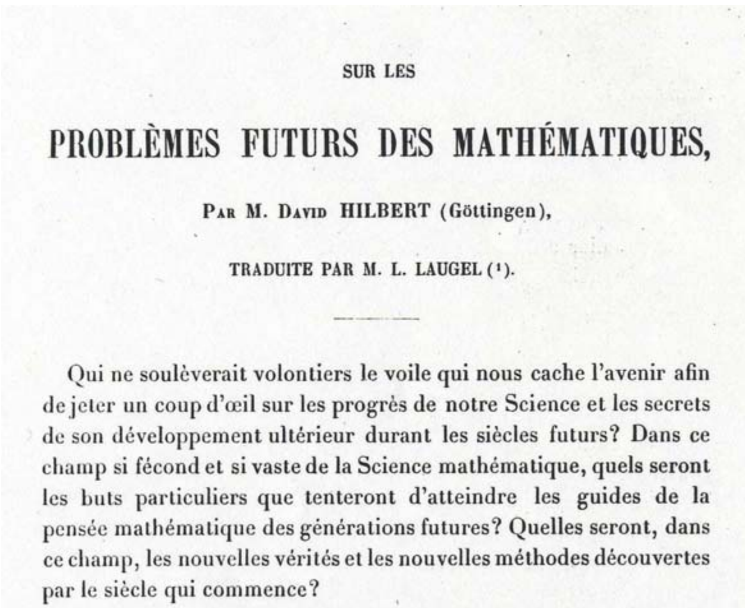
BY PROFESSOR DAVID HILBERT.

Who of us would not be glad to lift the veil behind which the future lies hidden; to cast a glance at the next advances of our science and at the secrets of its development during future centuries? What particular goals will there be toward which the leading mathematical spirits of coming generations will strive? What new methods and new facts in the wide and rich field of mathematical thought will the new centuries disclose?

History teaches the continuity of the development of science. We know that every age has its own problems, which the following age either solves or casts aside as profitless and replaces by new ones. If we would obtain an idea of the probable development of mathematical knowledge in the immediate future, we must let the unsettled questions pass before our minds and look over the problems which the science of to-day sets and whose solution we expect from the future. To such a review of problems the present day, lying at the meeting of the centuries, seems to me well adapted. For the close of a great epoch not only invites us to look back into the past but also directs our thoughts to the unknown future.

* Translated for the Bulletin, with the author's permission, by Dr. Mary Winston Newson. The original appeared in the *Göttinger Nachrichten*, 1900, pp. 253-297, and in the *Archiv der Mathematik und Physik*, 3d ser., vol. 1 (1901), pp. 44-63 and 213-237.

Fig. 10.2. English translation of Hilbert's Paris lecture by M.W. Newson, *Bulletin of the American Mathematical Society* 8 (1901/02). Courtesy Mathematisches Institut, Universität Leipzig.



SUR LES

PROBLÈMES FUTURS DES MATHÉMATIQUES,

Par M. David HILBERT (Göttingen),

TRADUITE PAR M. L. LAUGEL (1).

Qui ne soulèverait volontiers le voile qui nous cache l'avenir afin de jeter un coup d'œil sur les progrès de notre Science et les secrets de son développement ultérieur durant les siècles futurs? Dans ce champ si fécond et si vaste de la Science mathématique, quels seront les buts particuliers que tenteront d'atteindre les guides de la pensée mathématique des générations futures? Quelles seront, dans ce champ, les nouvelles vérités et les nouvelles méthodes découvertes par le siècle qui commence?

Fig. 10.3. French translation of Hilbert's Paris lecture by L. Laugel in *Compte Rendu Deuxième Congrès International des Mathématiciens*. Paris: Gauthier 1902. Courtesy Mathematisches Institut, Universität Leipzig.

> We mathematicians have often measured our progress by checking which of Hilbert's questions have been settled in the meantime.[4]

Elsewhere he added:

> The problems of mathematics are not isolated problems in a vacuum; there pulses in them the life of ideas which realize themselves in concreto [in practice] through our human endeavors in our historical existence, but forming an indissoluble whole transcend any particular science.[5]

Hilbert, best known for his axiomatic foundations of mathematics and his formalist viewpoint, knew the value of important problems. As his disciple and biographer Otto Blumenthal (1876–1944) put it: "Hilbert is the man of problems. He collects and solves existing problems; he poses new ones."[6] Indeed, it is just by the solution of concrete problems that mathematics will be developed; in the end, problem solving and theory building go hand in hand. That's why Hilbert risked offering a list of unsolved problems instead of presenting new methods or results, as was usually done at meetings. "He who seeks for methods without having a definite problem in mind seeks for the most part in vain,"[7] Hilbert told his Paris audience.

Let us examine Hilbert's career from the first stages up to the Paris lecture. From 1880 to 1884 Hilbert studied at the University of Königsberg in East Prussia, far from European scientific centers. At that time, Adolf Hurwitz (1859–1919) (three years older than Hilbert) was appointed professor in Königsberg, and Hermann Minkowski (1864–1909) (two years younger) was a brilliant student and close friend of Hilbert's at Königsberg. It was this mathematical community that had great influence on Hilbert the student. Ferdinand Lindemann (1852–1939), famous for his 1882 proof that π is not an algebraic but a transcendental number, had been professor at Königsberg since 1883. Under Lindemann's influence Hilbert became interested in the then-flourishing theory of invariants, the area of research in which he wrote his "Inauguraldissertation" (his Ph. D.) *Über die invarianten Eigenschaften specieller binärer Formen* (*On the Invariant Properties of Special Binary Forms*) in 1884. Most of Hilbert's work was devoted to the theory of algebraic invariants during his Königsberg period.

This theory of invariants appears also as an example in the canceled 24th problem, and therefore we will go into some details. In the last decades of the 19th century, besides group theory, above all invariants show how structural

---

[4] Weyl 1968, p. 466.

[5] Weyl 1944, p. 615.

[6] Blumenthal 1922, p. 67. "Hilbert ist der Mann der Probleme. Er sammelt und löst vorhandene, er weist neue."

[7] Hilbert 1900/1901, p. 32/444. "Denn wer, ohne ein bestimmtes Problem vor Augen zu haben, nach Methoden sucht, dessen Suchen ist meist vergeblich."

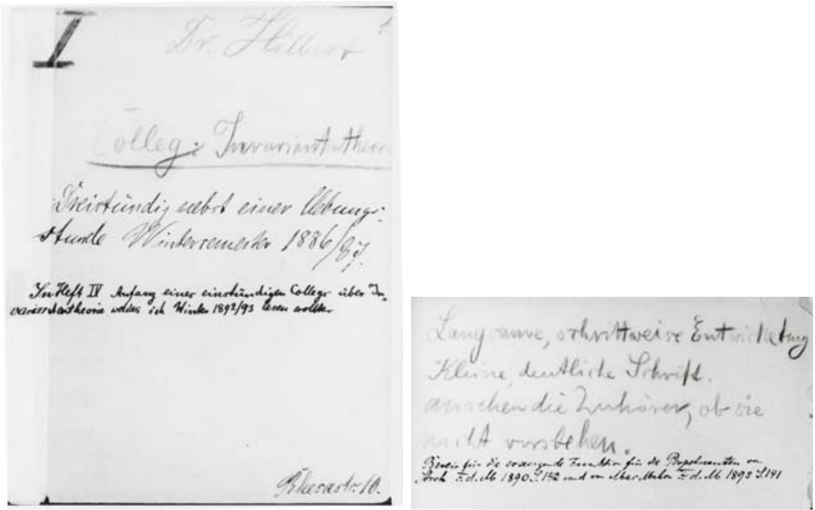


Fig. 10.4. Hilbert's lecture notes for his first lecture course in winter term 1886 in Königsberg. Title page and the facing page with an inscription.
Left: Colleg Invariantentheorie, Dreistündig nebst einer Uebungsstunde (Lecture Course: Theory of Invariants, three hours a week and one hour for exercising).
Right: The large handwriting is a typical example of Hilbert for corrections and insertions, usually done with a thick pencil in blue. The meaning is:
Slow development, step by step,
Small and clear handwriting [on the desk],
Watch the audience to see if they understood.
Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 521.

and abstract thinking was developed in mathematics. Invariants are quantities, entities, relationships, properties, etc., that are unaltered by particular transformations. For example, all geometric facts that are independent of the coordinate system are invariants in Euclidean geometry, such as magnitudes that are not altered by geometric transformations like rotation, dilatation, and reflection. A more general geometric invariant property is that of a cross ratio in projective geometry (with no metrical properties) which is invariant under projectivities. In each geometry every invariant property is based on a certain group of transformations. The task of a geometry in question consists in setting up the invariants of this group. In the language of modern algebra, the theory of invariants deals with linear groups G which act on an n-space (field) K and the polynomials $p(x) = p(x_1, x_2, \dots, x_n)$ on K that are invariant under G.

On the other hand – analytically expressed – such invariants are invariants of tensors: in the terminology of Hilbert's time, invariants of an n-ary form F of degree m under transformations. For example, this is a binary form in x_1

and x_2 of degree three:

$$ax_1^3 + bx_1^2x_2 + cx_1x_2^2 + dx_2^3.$$

The central questions in the theory of invariants are whether there is a system of invariants in which each invariant can be represented rationally and whether there is a finite system of this type.

Fig. 10.5. David and Käthe (née Jerosch) Hilbert, in the year of their marriage, 1892. Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Voit Collection.

Invariant theory was the main research field of Paul Gordan (1837–1912), who was regarded as the greatest expert in the field ("the king of invariants"). Gordan developed many constructive techniques for the representation and generation of invariants, among them his finite basis theorem saying that the invariants of a system of binary forms (with arbitrary many variables) possess a finite basis (1868). Despite the efforts of Gordan it was an open question whether such finite bases exist for forms of arbitrary order. In 1888 the solution came with Hilbert.

From his summer resort on August 28, 1888, Hilbert wrote a letter to Hurwitz.[8] After reporting on a boating trip with four young ladies, among them Miss Käthe Jerosch (1864–1945) (later Mrs. Hilbert), David Hilbert said: "Nevertheless my algebraic-arithmetic questions don't sleep."[9] He informed Hurwitz that when he was finishing a paper on invariant theory, he had found a new approach to Gordan's problem, considering its pure algebraic kernel. Already in the first days of September Felix Klein (1849–1925) received the improved paper, and Hilbert asked for a speedy publication in the *Göttinger Nachrichten* to secure his priority of the new "powerful methods".[10] For any form F with arbitrary degree Hilbert showed that its invariants form a ring $R = K[x]$ with a finite basis $i_1, i_2, \ldots, i_k$; i.e., the ring is generated by this basis or, in other words, any invariant i belonging to the ring R is expressible as a polynomial in this basis: $i = P_1 i_1 + \cdots + P_k i_k$, where P_s $(s = 1, \ldots, k)$ are polynomials with a degree lower than i (Hilbert's basis theorem).

The numerous papers on invariant theory of the time consisted of masses of endless algorithmic calculations, and the authors concerned with these calculations had not recognized the general law. They were, so to speak, unable to see the wood for the trees. Through Hilbert's insight this presentation of the theory of invariants was modified in an essential way. He proved its general theorems in a few pages, especially the most important theorem that every invariant of a given configuration can be expressed by a rational combination of a finite number of them. Hilbert's approach to invariant theory was quite different; the theory was transformed from what it had been in the hand of Gordan. His proof was based on existence procedures and therefore provided no method for constructing the basis in a given case. In resolving the principal problems of invariant theory in his own way Hilbert "had dealt it[11] a mortal blow", as Jean Dieudonné (1906–1992) said. However, in doing so he had laid the foundations of polynomial ideals and, moreover, prepared the way for modern algebra as developed later by, for instance, Emmy Noether (1882–1935), who was Gordan's only doctoral student, and Emil Artin (1898–1962), who studied with Gustav Herglotz (1881–1953) in Leipzig. Gian-Carlo

[8] Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. Math.-Archiv 76, no. 229.

[9] Letter to Hurwitz from August 28, 1888. Ibid., no. 229. "Trotzdem schlafen meine mathematischen Ideen nicht."

[10] Hilbert 1888, 1889; English translations Hilbert 1970. *Göttinger Nachrichten* = *Nachrichten der königlichen Akademie der Wissenschaften in Göttingen.*

[11] "It" means the symbolic methods of invariant theory, i.e. proofs by calculation. Dieudonné 1971 begins the preface with the line: "Invariant theory has already been pronounced dead several times, and like the phoenix it has been again and again rising from its ashes." Already Weierstraß told Hilbert in 1888: "In invariant theory many will go to ruin but not of it alone (Untergehen werde auch vieles in der Invariantentheorie, aber nicht von ihr allein)", Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Bericht über meine Reise (1888), Cod. Ms. D. Hilbert 741, p. 1/7.

Rota (1932–1999) even regarded Lie theory and algebraic geometry as offsprings of invariant theory.[12] From that time on, about the mid-1880's, in Hilbert's research discoveries of the first order were to follow one another in rapid succession for more than twenty years.

Hilbert concluded a paper, a brief résumé of his papers in invariant theory, read in his absence at the International Mathematical Congress in Chicago[13] on August 22, 1893, with historical remarks:

> In the history of a mathematical theory three periods can easily and clearly be distinguished: the naive, the formal, and the critical. As to the theory of algebraic invariants, its founders Cayley and Sylvester are both representatives of the naive period. [...] The discoverers and perfecters of the symbolic calculus Clebsch and Gordan are the representatives of the second period, whereas the critical period has found its expression in the above mentioned theorems.[14]

The theorems Hilbert referred to were his own. Until then the criterion of mathematical existence had been constructibility. Hilbert's revolutionary approach, consisting of pure existence proof, ignored this criterion and perplexed his colleagues, above all the great algorithmician Paul Gordan, who exclaimed: "That is not mathematics, that is theology (Das ist nicht Mathematik, das ist Theologie!)." Felix Klein, however, appreciated Hilbert's approach at once: "the matter is obviously very important."[15]

Hilbert, from 1892 an appointed professor at Königsberg, was offered a chair at the University of Göttingen in 1895, exactly one hundred years after Carl Friedrich Gauss (1777–1855) had enrolled in Göttingen, mainly because of his results in the theory of invariants. Hilbert accepted and remained there

[12] Rota, 1999. Two Turning Points in Invariant Theory. *The Mathematical Intelligencer* 21, 1 (1999) 22-28; a modern presentation for example is V.L. Popov 1992

[13] This Congress was part of the World's Columbian Exposition which was held on the occasion of the 400th anniversary of Columbus's discovery of America in 1893. It was attended by 45 mathematicians; four were from abroad, among them Felix Klein, who brought a number of European contributions. The Congress in Zurich in 1897 is regarded as the first International Mathematical Congress; the last but one took place in Berlin 1998, the last in Beijing 2002.

[14] Hilbert 1893, p. 124 (the given translation is based on that of Reid 1996, p. 34). "In der Geschichte einer mathematischen Theorie lassen sich meist 3 Entwicklungsperioden leicht und deutlich unterschieden: Die naive, die formale und die kritische. Was die Theorie der algebraischen Invarianten anbetrifft so sind die ersten Begründer derselben, Cayley und Sylvester, zugleich auch als die Vertreter der naiven Periode anzusehen. [...] Die Erfinder und Vervollkommener der symbolischen Rechnung Clebsch und Gordan sind die Vertreter der zweiten Periode, während die kritische Periode in den oben genannten Sätzen [...] ihren Ausdruck findet."

[15] Frei 1985. Letter to Hilbert from October 1, 1888., p. 43 (no. 32). "Die Sache ist offenbar sehr wesentlich."

Fig. 10.6. Hilbert's home in Göttingen, 29 Wilhelm-Weber-Straße, from 1897. The lecture halls and the Mathematical Institute were in walking distance of the house. Photo R. Thiele.

until his death in 1943 – almost half a century. Before Klein and Hilbert there had been brilliant mathematicians at Göttingen, for instance Carl Friedrich Gauss, Johann Peter Dirichlet (1805–1859), and Bernhard Riemann (1826–1866); but after Riemann's death in 1866 Göttingen had become a backwater compared to Prussian Berlin, which housed such luminaries as Jakob Steiner (1796–1863), Eduard Kummer (1810–1893), Leopold Kronecker (1823–1891), and Karl Weierstrass (1815–1897). However, Göttingen's bygone mathematical tradition was restored. The second flower prepared by Felix Klein achieved even greater eminence, largely because of Hilbert, who made Göttingen the leading center of mathematics in Germany.

In his Obituary on Hilbert, Hermann Weyl remarked that Hilbert concentrated his energies and focused them on a new area.[16] According to the years of publication of Hilbert's research we have six (almost) sharply distinguished periods. We give some examples:

[16]Weyl 1944, p. 617.

until 1893 algebraic forms:
Über die Theorie der algebraischen Formen.
Mathematische Annalen 36 (1890), 473–531.
Über die vollen Invariantensysteme.
Mathematische Annalen 42 (1893), 313–370.
Über die Theorie der algebraischen Invarianten, paper read at the International Mathematical Congress in Chicago 1893. New York: Macmillan 1896, pp. 116–124.

1894–1899 algebraic number theory:
Die Theorie der algebraischen Zahlkörper [=Zahlbericht].
Jahresbericht der Deutschen Mathematiker Vereinigung 4 (1897), 157–546.

1899–1903 foundations of geometry (axiomatic method):
Grundlagen der Geometrie.
Leipzig: Teubner 1899, 14th ed. 1999.

1903–1912 analysis:
Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, 6 Mitteilungen, in: *Nachrichten der königlichen Akademie der Wissenschaften in Göttingen* 1904–1910; as a book Leipzig: Teubner 1912.

1912–1928 mathematical physics:
Begründung der kinetischen Gastheorie.
Mathematische Annalen 71 (1912), 562–577.
Die Grundlagen der Physik.
Nachrichten der Akademie der Wissenschaften in Göttingen 1915 and 1916, again in *Mathematische Annalen* 92 (1924), 1–32.
R. Courant/D. Hilbert [Hilbert only pro forma],
Methoden der mathematischen Physik. 2 vols.
Berlin: Springer 1924 and 1937.
Über die Grundlagen der Quantenmechanik
(with J. v. Neumann, L. Nordheim),
Mathematische Annalen 98 (1928), 1–30.

after 1918 foundations of mathematics:
Axiomatisches Denken.
Mathematische Annalen 78 (1918), 405–415.
Neubegründung der Mathematik [1. Mitteilung].
Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg 1 (1922), 157–177.
Die Grundlagen der Mathematik [2. Mitteilung].
Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg 6 (1928), 64–85.
Die Grundlegung der elementaren Zahlenlehre.
Mathematische Annalen 104 (1931), 485-494.
Grundzüge der theoretischen Logik
(with W. Ackermann). Berlin: Springer 1928.
Grundlagen der Mathematik
(with P. Bernays). 2 vols. Berlin: Springer 1934 and 1939.

However, if we take into account that Hilbert's choice in his famous speech surveyed nearly all the mathematics of his day, our division into certain special fields simply shows the printed results of Hilbert's scientific activities during these periods, not the enormous diversity of his actual interests. A tolerable overview of Hilbert's interests cannot be obtained from the published sources alone; we must study his drafts of lectures, corresponding lecture notes, his notebooks, and correspondence. Fortunately the Hilbert Nachlass is almost complete. The Nachlass and many other sources are held in the Library of the University of Göttingen (Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung), in the Mathematical Institute of the University of Göttingen, and in the Staatsbibliothek Berlin (Nachlässe Born and Hückel). Among the items of the Nachlass there are three mathematical notebooks[17] containing his handwritten notices from 1886. From our point of view, however, rather surprisingly we find (almost) no references to any of the 23 problems in these notebooks.

10.1 How did Hilbert's Paris Talk Come About?

Like any other science, mathematics is an international matter. To put it in Hilbert's own words: "Mathematics knows no races or geographic boundaries; for mathematics, the whole cultural world is a single country."[18] However, it was necessary to wait for the development of means of transportation in the second half of the 19th century in order to be able to organize international meetings where mathematicians were brought together to communicate in

[17] Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung. Cod. Ms. D. Hilbert 600:1-3.

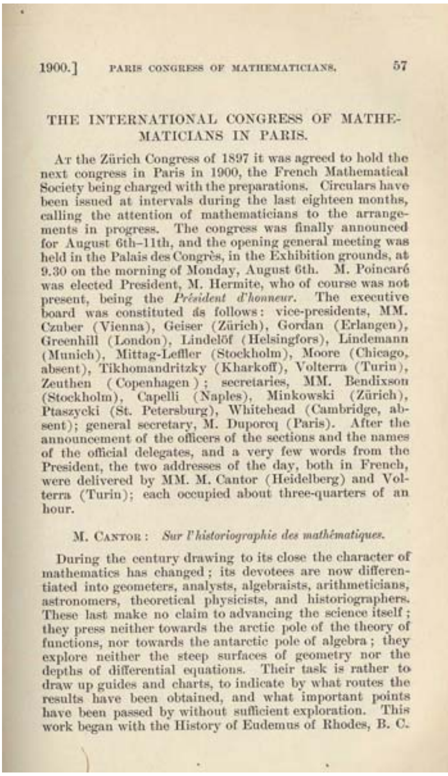
[18] Reid, 1996, p. 188. "Die Mathematik kennt keine Rassen oder geographische Grenzen, denn für sie ist die gesamte kulturelle Welt ein einziges Land."

person rather than writing letters (which indeed had been effective in the days of Father Mersenne (1588–1647)).

Fig. 10.7. Map of Europe before World War I. (In this projection the three towns Königsberg, Göttingen, and Paris related to Hilbert are collinear.) Based on a map of the Karl-Sudhoff-Institut, Universität Leipzig.

In August 1897, 209 mathematicians gathered in Zurich for the first International Congress of Mathematicians; the second ICM met in Paris in 1900, with 262 participants attending. For each ICM it has been customary to invite some mathematicians to deliver lectures on special topics. At Zurich Poincaré delivered his speech "Sur les rapports de l'analyse pure et de la physique mathématique (On the relations between pure mathematics and mathematical physics)".[19] In winter 1899-1900 Hilbert, one of the most respected German mathematicians of the day and nearly 38 years old, was invited to make one of the major addresses in the opening session of the coming ICM in Paris. Hilbert hesitated whether he should reply to Poincaré's Zurich lecture or choose another subject and he asked for Minkowski's opinion on a report on individual problems. On January 5, 1900, his friend wrote:

[19] Rudio 1898, pp. 81–90.



1900.] PARIS CONGRESS OF MATHEMATICIANS. 57

THE INTERNATIONAL CONGRESS OF MATHEMATICIANS IN PARIS.

At the Zürich Congress of 1897 it was agreed to hold the next congress in Paris in 1900, the French Mathematical Society being charged with the preparations. Circulars have been issued at intervals during the last eighteen months, calling the attention of mathematicians to the arrangements in progress. The congress was finally announced for August 6th–11th, and the opening general meeting was held in the Palais des Congrès, in the Exhibition grounds, at 9.30 on the morning of Monday, August 6th. M. Poincaré was elected President, M. Hermite, who of course was not present, being the *Président d'honneur*. The executive board was constituted as follows: vice-presidents, MM. Czuber (Vienna), Geiser (Zürich), Gordan (Erlangen), Greenhill (London), Lindelöf (Helsingfors), Lindemann (Munich), Mittag-Leffler (Stockholm), Moore (Chicago, absent), Tikhomandritzky (Kharkoff), Volterra (Turin), Zeuthen (Copenhagen); secretaries, MM. Bendixson (Stockholm), Capelli (Naples), Minkowski (Zürich), Ptaszycki (St. Petersburg), Whitehead (Cambridge, absent); general secretary, M. Duporcq (Paris). After the announcement of the officers of the sections and the names of the official delegates, and a very few words from the President, the two addresses of the day, both in French, were delivered by MM. M. Cantor (Heidelberg) and Volterra (Turin); each occupied about three-quarters of an hour.

M. Cantor: *Sur l'historiographie des mathématiques.*

During the century drawing to its close the character of mathematics has changed; its devotees are now differentiated into geometers, analysts, algebraists, arithmeticians, astronomers, theoretical physicists, and historiographers. These last make no claim to advancing the science itself; they press neither towards the arctic pole of the theory of functions, nor towards the antarctic pole of algebra; they explore neither the steep surfaces of geometry nor the depths of differential equations. Their task is rather to draw up guides and charts, to indicate by what routes the results have been obtained, and what important points have been passed by without sufficient exploration. This work began with the History of Eudemus of Rhodes, B. C.

Fig. 10.8. Report on the Paris Congress by A.S. Scott, *Bulletin of the American Mathematical Society* 7 (1900). Courtesy Mathematisches Institut, Universität Leipzig.

> Most alluring would be the attempt at a look into the future and a listing of the problems which mathematicians should try themselves during the coming century. With such a subject you could have people talking about your lecture decades later.[20]

In the end Minkowski was right, but Hilbert was still wavering, so he consulted Hurwitz on March 29:

> I must start preparing for a major talk at Paris, and I am hesitating about a subject. [...] The best would be a view into the future. What

[20]Letter to Hilbert from January 5, 1900. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 258: 76; English translation in Reid 1996, p. 69. "Am anziehendsten würde der Versuch eines Vorblicks auf die Zukunft sein, also eine Bezeichnung der Probleme, an welche sich die künftigen Mathematiker machen sollten. Hier könntest Du unter Umständen erreichen, dass man von Deiner Rede noch nach Jahrzehnten spricht."

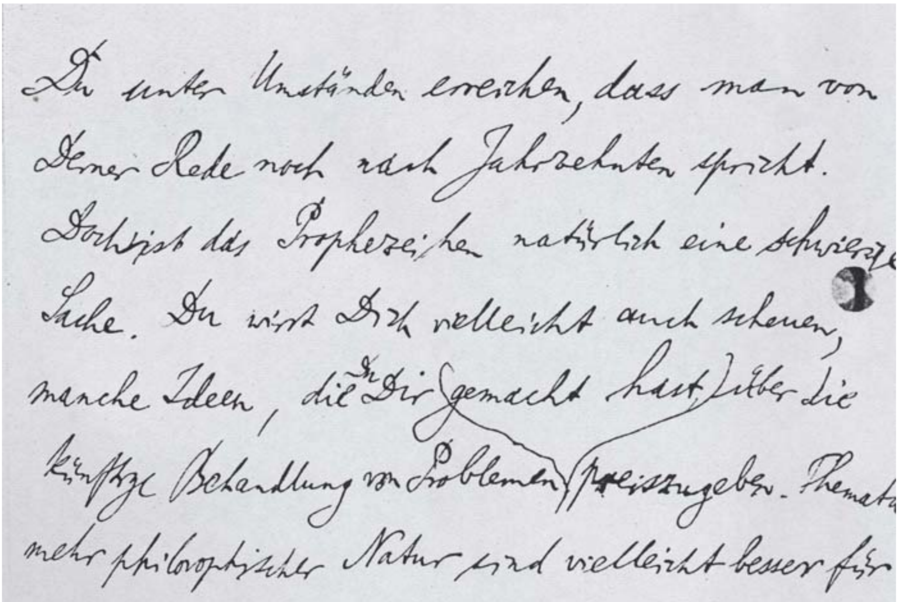

Du unter Umständen erreichen, dass man von
Deiner Rede noch nach Jahrzehnten spricht.
Doch ist das Prophezeihen natürlich eine schwierige
Sache. Du wirst Dich vielleicht auch scheuen,
manche Ideen, die Du Dir gemacht hast, über die
künftige Behandlung von Problemen preiszugeben. Themata
mehr philosophischer Natur sind vielleicht besser für

Fig. 10.9. Letter from Minkowski (Zurich) to Hilbert from January 5, 1900. Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. 258, no. 76.

> do you think about the likely direction in which mathematics will develop during the next century? It would be very interesting and instructive to hear your opinion about that.[21]

We have no record of Hurwitz's reply. In fact Hilbert hesitated rather long and did not decide before the deadline. The mailed program for the Congress included neither an announcement of a major lecture nor any other contribution from Hilbert. Minkowski was disappointed: "Without your lecture the program of the Paris Congress was a great disappointment. The desire on my part to travel to the congress is now almost gone."[22]

In the middle of July, Hilbert surprised his friends Minkowski and Hurwitz (both professors at Zurich) with proofs of a paper entitled "Mathematische Probleme" – the complete version of his Paris talk worked out for publication in the *Göttinger Nachrichten*. Both friends read the proofs carefully and made

[21] I did not find this letter to Hurwitz in the libraries of either the University of Göttingen or of ETH Zürich; I quote the English translation in Reid 1996, p. 70.

[22] Letter to Hilbert, June 22, 1900. Niedersächsische Staats- und Universit"atsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 258: 81. English translation partly by Reid 1996, p. 70. "Das Programm des Pariser Congresses ohne Deinen Vortrag war für mich eine grosse Enttäuschung. Fast ist mir überhaupt die Lust, zum Congress hinzugehen, vergangen."

suggestions for the presentation. Later in a letter to Hurwitz in view of the invariants Hilbert noticed that in his haste he had failed to give credit to Hurwitz (in problem 14) in the paper and begged Hurwitz for pardon. He then mentioned, amused, that Minkowski had even used the opportunity of proof-reading to insert his own results (probably into problem 5).[23]

Fig. 10.10. Hermann Minkowski (1883) in the year he was awarded the prize of the Paris Academy. Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Voit Collection.

On one point both colleagues agreed: the lecture was too long. Minkowski wrote on July 17 and 28 respectively:

> The section on the calculus of variations, to wit: the formulas might be better placed in a note at the end of the lecture.

[23]Letter to Hurwitz from November 21, 1900. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. Math.-Archiv 76, no. 278. "Minkowski hat sich sogar selbst beim Correcturlesen an einer Stelle hineinverbessert."

> Hurwitz and I started under the erroneous impression that a stop was to be put to Ignorabimus, with the calculus of variations neatly dealt with. And now you go and take a lease on mathematics for the twentieth century, and will go down as Managing Director.
> Actually I believe that through this lecture, which indeed every mathematician in the world without exception will be sure to read, your attractiveness for young mathematicians will increase – if that is possible.[24]

A week later Hilbert met Minkowski in Paris; the disappointed Hurwitz did not attend. In the morning of August 8 Hilbert's lecture was delivered at a joint session of two sections "Bibliographie et Histoire" and "Enseignement et Méthodes" chaired by the German historian of mathematics Moritz Cantor (1829-1920).[25] Hilbert introduced his 23 problems with a longer essay, but he followed the advice to shorten the lecture, and presented a selected list of only 10 problems. Moreover, he canceled a 24th problem, which has not been published until now. In September 1900, Hilbert's talk, including the complete list of 23 problems, was published in the *Göttinger Nachrichten* and later in slightly revised versions (mainly for the 23rd problem), and in French and English translations in 1902.[26]

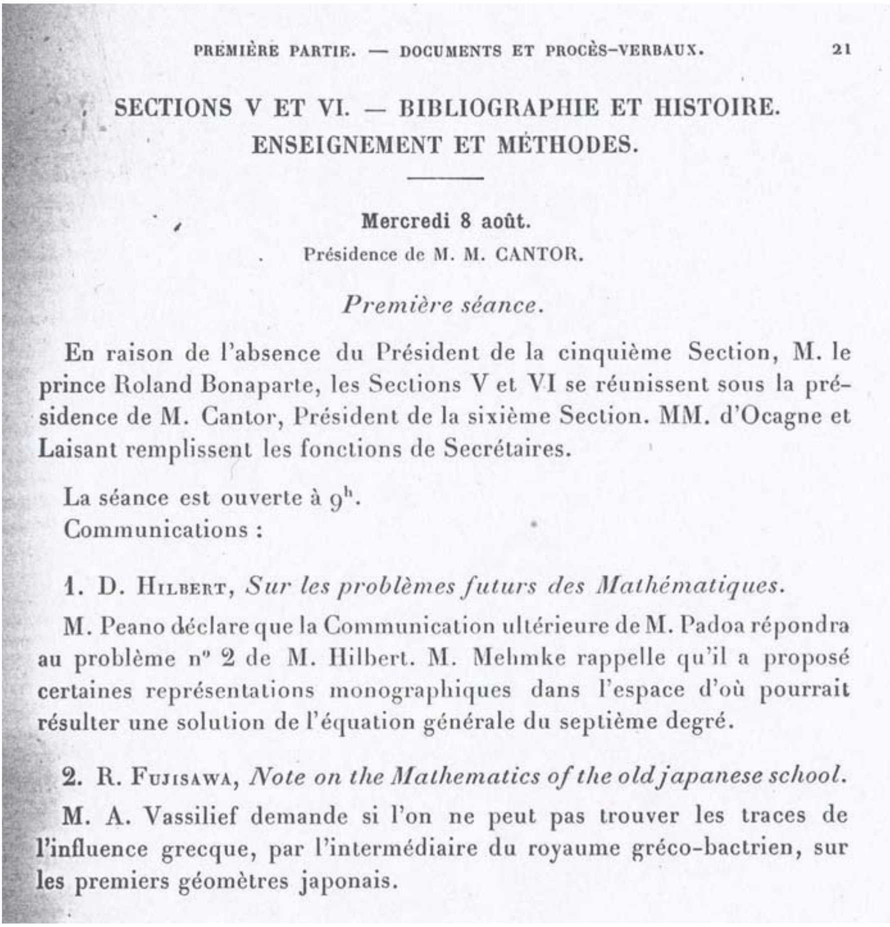
After the Congress Hilbert took a holiday and traveled to the Baltic Sea near Königsberg. He felt somewhat unsatisfied with the Paris Congress for two reasons. The discussion after his lecture was rather disappointing: Giuseppe Peano (1858–1932) declared that a future paper of Alessandro Padoa (1868–1937) would give an answer to the second problem (consistency of arithmetic axioms), and Rudolf Mehmke (1857–1944) obscurely insisted he had already proposed some monographs concerning the solution of equations of the seventh degree (problem 13).

---

[24] Letters to Hilbert. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 258, nos. 83 & 84. "Der Abschnitt über Variationsrechnung, namentlich die Formeln sind wohl besser in eine Anmerkung hinter den Vortrag zu verweisen." (17. 7. 1900) – "Hurwitz und ich hatten uns zunächst ein ganz falsches Bild gemacht, indem wir dachten, mit dem Ignorabimus sollte Schluss gemacht werden, namentlich da die Variationsrechnung schon so genau abgehandelt war. Nunmehr hast Du wirklich die Mathematik für das 20te Jahrhundert in Generalpacht genommen und man wird Dich allgemein gern als Generaldirector anerkennen. - Namentlich glaube ich, dass Deine Anziehungskraft auf junge Mathematiker durch diese Rede, die wohl jeder Mathematiker ohne Ausnahme lesen wird, wenn überhaupt möglich noch wachsen wird." (28. 7. 1900) English translation of the last quotation in Reid 1996, p. 72.

[25] Moritz Cantor, Germany's leading historian of mathematics and no relative of Georg Cantor (the creator of set theory), is best known for his *History of Mathematics* (Leipzig: B.G. Teubner 1880-1908) in four volumes.

[26] Cf. Grattan-Guinness 2000.



PREMIÈRE PARTIE. — DOCUMENTS ET PROCÈS-VERBAUX. 21

SECTIONS V ET VI. — BIBLIOGRAPHIE ET HISTOIRE.
ENSEIGNEMENT ET MÉTHODES.

Mercredi 8 août.

Présidence de M. M. CANTOR.

Première séance.

En raison de l'absence du Président de la cinquième Section, M. le prince Roland Bonaparte, les Sections V et VI se réunissent sous la présidence de M. Cantor, Président de la sixième Section. MM. d'Ocagne et Laisant remplissent les fonctions de Secrétaires.

La séance est ouverte à 9h.
Communications :

1. D. Hilbert, *Sur les problèmes futurs des Mathématiques.*

M. Peano déclare que la Communication ultérieure de M. Padoa répondra au problème n° 2 de M. Hilbert. M. Mehmke rappelle qu'il a proposé certaines représentations monographiques dans l'espace d'où pourrait résulter une solution de l'équation générale du septième degré.

2. R. Fujisawa, *Note on the Mathematics of the old japanese school.*

M. A. Vassilief demande si l'on ne peut pas trouver les traces de l'influence grecque, par l'intermédiaire du royaume gréco-bactrien, sur les premiers géomètres japonais.

Fig. 10.11. The program of Paris Congress for the sections "Bibliographie et Histoire" and "Enseignement et Méthodes" on August 8, 1900 with the announcement of Hilbert's lecture at 9 a.m. In: *Compte Rendu deuxième Congrès International des Mathématiciens.* Paris: Gauthier 1902. Courtesy Mathematisches Institut, Universität Leipzig.

Hilbert believed that science is also to be propagated orally; books alone are infertile.[27] This is why in a letter[28] to Hurwitz written immediately after

[27] Mathematische Notizhefte. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:2, p. 99. "Die Wissenschaft wird auch mündlich übertragen, nur aus Büchern ist unfruchtbar – so etwa."

[28] Letter to Hurwitz from August 25, 1900. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. Math.-Archiv 76, no. 272.

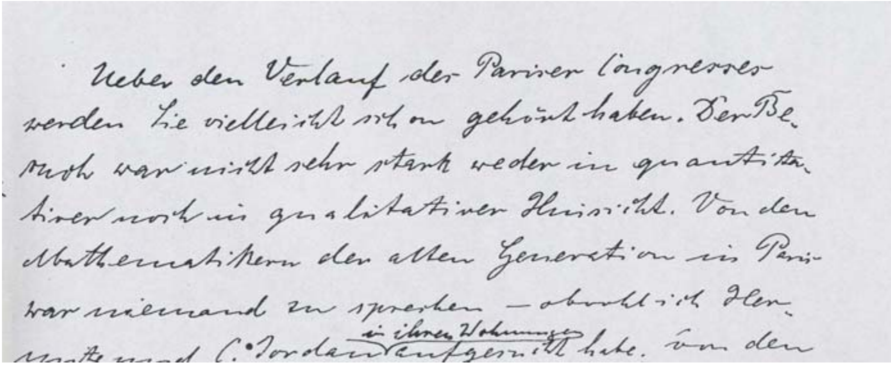

Ueber den Verlauf der Pariser Congresses
werden Sie vielleicht schon gehört haben. Der Be-
such war nicht sehr stark weder in quantita-
tiver noch in qualitativer Hinsicht. Von den
Mathematikern der alten Generation in Paris
war niemand zu sprechen – obwohl ich Her-
mite und C. Jordan in ihren Wohnungen aufgesucht habe. von den

Fig. 10.12. Letter from Hilbert to Hurwitz (Zurich) written from Hilbert's holiday place at the Baltic Sea near Königsberg on August 25, 1900. The Paris Congress ended on August 11, 1900. Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Math-Archiv 76, no. 272.

the Conference, Hilbert regretted the attendance was poor in quantity and in quality because for some reason important French mathematicians were absent. One evening Poincaré even disappeared instead of chairing a banquet. Hilbert also complained that there were not enough rooms for informal meetings. As for the upcoming Congress at Heidelberg, he remarked, at any rate the Germans must make a more efficient organization.[29]

The Paris address intertwined concrete but important problems with the theoretical context. It is this interlocking character which made the collection so fruitful. Hilbert's "leitstern" (lode-star) in research was to find that special case which contains all the germs of generality. His leitmotif was to start investigations (or lectures) with an elementary but instructive example. One outstanding feature of Hilbert's works is that he liked to explain general methods through examples, leaving enough questions for other researches. These are the deeper reasons why Hilbert built up an important school and why Hilbert's list of problems charted the course of mathematics during the 20th century.

Noteworthy, however, is the fact that neither Hilbert himself nor any of his disciples did work exclusively on the 23 problems. Among those who contributed to solutions of the 23 problems we find the following students of Hilbert: Max Dehn (1878–1952, Ph. D. with Hilbert 1899), Teiji Takagi (1875–1960), Georg Hamel (1877–1954, Ph. D. with Hilbert 1901), Paul Funk (1886–1969, Ph. D. with Hilbert 1911), Erich Hecke (1887–1947, Ph. D. with Hilbert 1910), Richard Courant (1888–1972, Ph. D. with Hilbert 1910), Emil Artin (1898–1962), Herbert Busemann (1905–1994), Gerhard Gentzen (1909–

[29]Letter to Hurwitz from August 25, 1900. Ibid., no. 272. "Wir müssen uns mit der Vorbereitung jedenfalls mehr Mühe geben und eine bessere und einheitlichere Organisation ins Werk setzen."

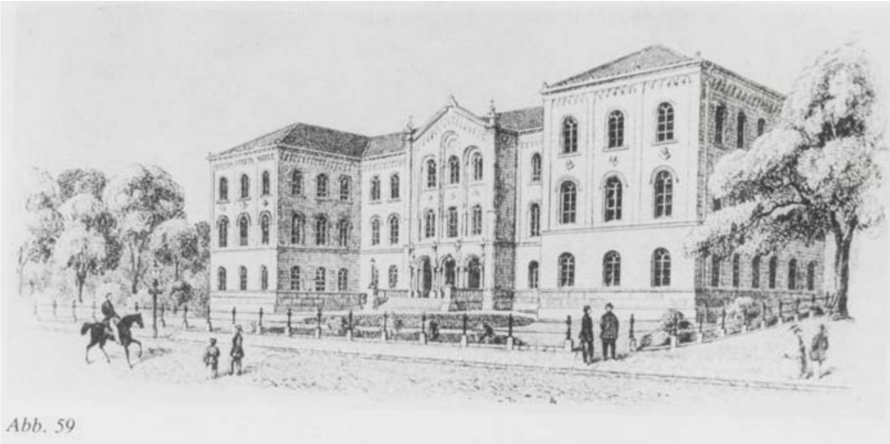


Fig. 10.13. Lecture hall (Auditorium) of the University of Göttingen (at the cross roads Weender Straße and Berliner Straße) in which H.A. Schwarz, F. Klein, D. Hilbert, H. Minkowski, C. Runge, E. Landau, and others lectured. The new Mathematical Institute, Bunsenstraße, was completed in 1929 with the support of the Rockefeller Foundation. Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung.

1945), and others. Three months after the Paris lecture Hilbert reported the first solution of one problem (number 3) by Max Dehn who—so he wrote to Hurwitz—is "one of my best disciples" and "with his results [in his Ph. D. in 1899] I am totally delighted".[30] Dehn was the first to solve a Hilbert problem. Already one year later in Münster he wrote his Habilitationsschrift and became a Privatdozent.

10.2 On the Problems

Hilbert had not had time to deal with all 23 problems, nor did his list cover all branches of mathematics.[31] Moreover, in each survey, complete or incomplete, we can regard the problems from different points of view. So it becomes quite clear that it will be impossible to give an adequate summary or an appreciation of the problems and their sequels here. Hilbert told his audience:

> The problems mentioned are merely samples of problems; yet they are sufficient to show how rich, how manifold and how extensive mathematical science is today, and the question is urged upon us whether

[30]Letter to Hurwitz from November 5–12, 1899. Ibid. no. 275 (sheet 588); "einer meiner besten Schüler", "über dessen Resultate ich ganz entzückt bin". In the letter from November 21, 1900 this appreciation is repeated (ibid., no. 278).

[31]See Grattan-Guinness 2000; for the sequels see Yandell 2002.

Fig. 10.14. David Hilbert about 1900. Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Voit Collection.

> mathematics is doomed to the fate of those other sciences that have split up into separate branches. [...] I do not believe this nor wish it. Mathematical science is in my opinion an indivisible whole. [...] The organic unity of mathematics is inherent in the nature of this science, for mathematics is the foundation of all exact knowledge of natural phenomena. [...] May the new century bring it gifted prophets and many zealous and enthusiastic disciples.[32]

[32]The last sentences of the Paris lecture, English translation Newson 1902, p. 479. "Die genannten Probleme sind nur Proben von Problemen; sie genügen jedoch, um uns vor Augen zu führen, wie reich, wie mannigfach und wie ausgedehnt die mathematische Wissenschaft schon heute ist, und es drängt sich uns die Frage auf, ob der Mathematik einst bevorsteht, was anderen Wissenschaften schon längst widerfahren ist, nämlich daß sie in einzelne Teilwissenschaften zerfällt. [...] Ich glaube und wünsche dies nicht; die mathematische Wissenschaft ist meiner Ansicht nach ein unteilbares Ganzes. [...] Der einheitliche Charakter der Mathematik liegt im inneren Wesen dieser Wissenschaft begründet; denn die Mathematik ist die Grundlage alles exacten naturwissenschaftlichen Erkennens. [...] Mögen ihr [der Mathe-

Furthermore, Hilbert told his audience of what nature problems should be. Instead of the lecture I quote a similar remark from his notebook:

> The problems must be difficult and clear – but not easy and complicated, because confronted with them we would be helpless or we would need some exertion of our memory to bear all the assumptions and conditions in mind.[33]

Hilbert told a related story:

> An old French mathematician [probably Hermite] said: "A mathematical theory is not to be considered complete until you have made it so clear that you can explain it to the first man whom you meet on the street."[34]

I would like to update this statement: you should be content with the fact there is a person on the same floor of your institute who possibly understands you.

Today, with insight, we know that no problem was trivial; all were interesting and fertile. Even so, despite the great impact of Hilbert's problems, we should not regard him as a prophet for the future of mathematics. Hilbert himself regarded the statement "absolutely accurate prophecies are impossible"[35] as an axiom. Indeed, the problems do not indicate that Hilbert would have foreseen the rapid development of functional analysis in the following decade, to which he himself contributed the theory of integral equations.

The difficulty as well as the length of the problems varies. The shortest consists of only six lines, but it took six decades to find its solution; some of the problems are as yet unsolved (nos. 8, 12, 13; no. 8 includes the Riemann

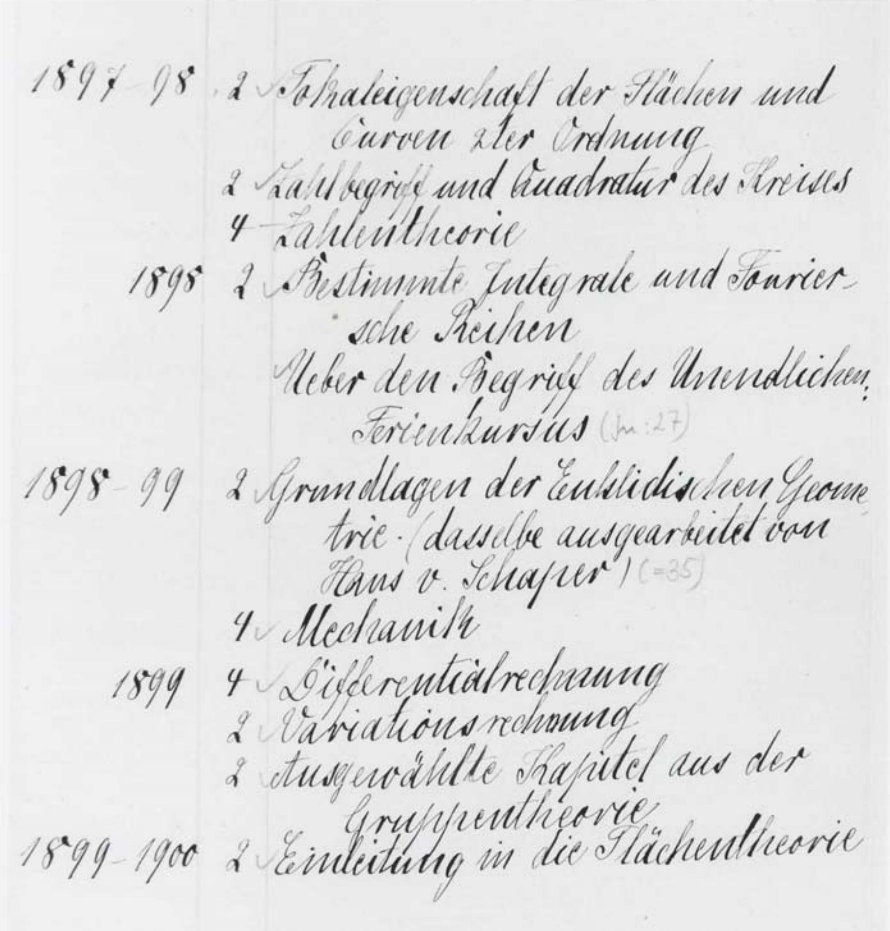
---

matik] im neuen Jahrhundert geniale Meister erstehen und zahlreiche in edlem Eifer erglühende Jünger!"

[33]Mathematische Notizhefte. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:1, p. 55. "Die Probleme müssen schwierig und einfach - nicht leicht und kompliziert sein, so dass man zunächst rathlos vor ihnen steht - nicht so, dass man schon das Gedächtnis anstrengen muss, um bloss alle Voraussetzungen und Bedingungen zu behalten."

[34]At the beginning of his Paris lecture "Mathematische Probleme", Hilbert 1900, p. 254. English translation by Newson 1901, p. 479. "Ein alter französischer Mathematiker hat gesagt: Eine mathematische Theorie ist nicht eher als vollkommen anzusehen, als bis du sie so klar gemacht hast, daß du sie dem ersten Mann erklären könntest, den du auf der Straße triffst."

[35]Mathematische Notizhefte, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:3, inserted pages. "Nimm das Axiom: Absolut richtige Prophezeiungen sind unmöglich (auch für den Laplaceschen Weltgeist?) (Postulate the axiom: absolutely accurate prophecies are impossible (also for Laplace's demon?))."



1897-98 2 Totaleigenschaft der Flächen und Curven 2ter Ordnung
2 Zahlbegriff und Quadratur des Kreises
4 Zahlentheorie
1898 2 Bestimmte Integrale und Fourier-sche Reihen
Ueber den Begriff des Unendlichen; Ferienkursus (Jm: 27)
1898-99 2 Grundlagen der Euklidischen Geometrie. (dasselbe ausgearbeitet von Hans v. Schaper) (=35)
4 Mechanik
1899 4 Differentialrechnung
2 Variationsrechnung
2 Ausgewählte Kapitel aus der Gruppentheorie
1899-1900 2 Einleitung in die Flächentheorie

Fig. 10.15. List of the lectures Hilbert delivered from winter term 1897 until winter term 1899, written by Hilbert's wife Käthe. The numbers before the lectures indicate the number of hours per week. Part of Hilbert's "Verzeichnis meiner Vorlesungen, 1886-1930 (List of my lectures)"; Hilbert's list is incomplete. Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 520.

hypothesis); and at least two problems are so general that they do not have an ultimate solution:[36]

[36]Hilbert 1900, pp. 253-297, and 1935, vol. 3, pp. 290-329; cf. Bieberbach 1930, Aleksandrov 1969, Fang 1970, Browder 1976, Gray 2000, Grattan-Guinness 2000, Yandell 2002.

no. 6: the mathematical treatment of the axioms of physics, and
no. 23: the further development of the methods of the calculus of variations.

We can (roughly) divide the problems into four groups:

	Actual Lecture	Published Version	Canceled
Problems of foundation		1, 2, 6	24
Problems of analysis	19, 21, 22	20, 23	
Problems of geometry		3, 4, 15, 18	
Problems of arithmetic and algebra	7, 8, 13, 16	5, 9, 10, 11, 12, 14, 17	

Once again, Hilbert's rapidly growing international fame was caused largely by his lucky choice of mathematical problems, not by the more usual practice of presenting new solutions, methods, or results. Of course, problems are the lifeblood of science, but in general answers and not questions are more estimated, expected and above all honored. That is our attitude not only in science but also in everyday life. Let me give as a simple example a fictitious but typical dialogue between a father and his child:

> Dad, why is the grass green?
> Mm, no idea.
> Dad, why is the sky blue?
> I don't know that either.
> Dad, do I disturb you? You don't mind me asking all those questions, do you?
> Son, if you don't ask questions how are you going to learn anything?[37]

You will probably be amused by Dad's last answer because you have not expected a logical, but a psychological response like: "Be quiet." We are, however, in a logical context, not in real life.

At the fourth ICM in Rome in 1908 Poincaré said in his lecture "L'avenir des mathématiques (The future of mathematics)": "At one time there were prophets of misfortune; they reiterated that all problems had been solved, that after them there would be nothing but gleanings left. [...] But," he added, "the pessimists have always been compelled to retreat, so that I believe there are none left today."[38] Here Poincaré echoes precisely Hilbert's conviction: "A branch of science is full of life as long as it offers an abundance of problems;

[37]English saying, cf. Vollmer 1993, p. 192.

[38]Poincaré 1908, p. 167. "Il y a eu autrefois des prophètes de malheur. Il répétaient volontiers que tous les problèmes susceptibles d'être résolus l'avaient été déjà, et qu'après eux il n'y aurait plus qu'à glaner. [...] Les pessimistes se trouvaient ainsi toujours débordés, toujours forcés de reculer, de sort qu'à présent je crois bien qu'il n'y en a plus.". English translation in A. Weil 1971, p. 321.

a lack of problems is a sign of death."[39] And yet in his swan song at the Königsberg Meeting in 1930 Hilbert mentioned with satisfaction the failure of Auguste Comte (1798-1857) to pose an unsolvable problem (see next section). André Weil (1906–1998) explained: "If logic is the hygiene of the mathematician, it is not his source of food; the great problems furnish the daily bread on which he thrives."[40] At any rate to pose a problem is no small feat, above all to pose an interesting and important one. In a speech "Die Naturgesetze und die Struktur der Materie (Laws of nature and structure of matter)"[41] in 1961, Werner Heisenberg (1901–1976) said of the ancient Greek philosophers that, disregarding their insufficient and highly speculative answers, above all their attainments in posing the right questions were incredible. Among the philosophers above all it was Karl Popper (1902–1994) who emphasized that progress depends on questions.

There is almost nothing to be compared with that which Hilbert had undertaken: Hilbert's choice of problems is unique, at least as the product of a single mind. There have since been other compilations of problems in books and papers, especially in single branches of mathematics, and some problem columns in journals, among them:

H. T. Croft, *Unsolved Problems in Geometry.* New York: Springer 1991,
R. K. Guy, *Unsolved Problems in Number Theory.* New York: Springer 1981,
D. Mauldin, *The Scottish Book.* Basel: Birkhäuser 1989,
C. S. Ogilvy, *Tomorrow's Math. Unsolved Problems for the Amateur.* New York: OUP 1962,
G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, 2 vols. Berlin: Springer 1925; English translation: *Problems and Theorems of Analysis*, 2 vols. New York: Springer 1998.
P. de Souza, *Berkeley Problems in Mathematics.* New York: Springer 1998,
W. Sierpinski, *A Selection of Problems in the Theory of Numbers.* New York: Macmillan 1964,
D. Shanks, *Solved and Unsolved Problems in Number Theory.* Washington: Spartan Books 1962.
H. Tietze, *Gelöste und ungelöste mathematische Probleme (Solved and Unsolved Problems).* München: Beck 1949,
S. Ulam, *A Collection of Mathematical Problems.* New York: Interscience 1960.

In 1976 Jean Dieudonné inspired Felix Earl Browder (born 1927) to ask a number of mathematicians to describe some unsolved problems in their fields.

[39] Hilbert 1900/1901, p. 254/438. "Solange ein Wissenszweig überfluß an Problemen bietet, ist er lebenskräftig; Mangel an Problemen bedeutet Absterben oder Aufhören der selbständigen Entwicklung."

[40] Weil 1971, p. 324.

[41] Heisenberg 1971, p. 237.

The result, published under the title "Problems of Present Day Mathematics" in the two-volume book *Mathematical Developments Arising from Hilbert Problems*,[42] is likely the largest collection of important problems, at least with respect to the number of represented branches. There is also an earlier Russian edition, *Problemy Gil'berta (Hilbert's Problems)*,[43] edited by Pavel Sergeevich Aleksandrov (1896–1982) with comments by competent Russian mathematicians. Both editions show the influence of Hilbert's problems even after seven decades. I restrict myself to quoting James Serrin (born 1926) on the "The solvability of boundary value problems" in the AMS collection *Mathematical Developments Arising from Hilbert Problems* (on problem 19, which was omitted in the actual lecture):

> Among the prophetic problems in Hilbert's famous list one must surely include the 20th, the general problem of boundary values for elliptic partial differential equations. This subject, only a seedling in the year 1900, has burst into flower during our century, has developed in directions Hilbert never imagined, and today encompasses a vast area of work which to a mathematician of 75 years ago would seem little short of astonishing.[44]

In the same collection Enrico Bombieri (born 1940) concluded his paper "Variational problems and elliptic equations" in this way: "In this sense, it can be said that Hilbert's 19th problem has opened one of the most interesting chapters in mathematics" (p. 434). James Serrin further remarked: "The 20th, like so many of the others in Hilbert's list, consisted as much in a program as in a specific problem requiring some definite answer, and in just this fact we can see one facet of Hilbert's genius and breadth" (p. 507). Indeed, it was the proof of Dirichlet's principle for a specific problem that led Hilbert to the general questions finally arising in the 19th and 20th problems. By the concepts of existence in a generalized sense and of regularity he pointed out two very important issues in the modern theory of partial differential equations.

Inspired in part by Hilbert's list and on behalf of the International Mathematical Union, Vladimir Igorovic Arnol'd (born 1937) recently wrote a letter to some mathematicians asking for the description of great problems for the next century. In response, on the occasion of Arnol'd's 60th birthday, Steve Smale (born 1930) gave a lecture on a "Conference in Honor of Arnol'd" at the Fields Institute in Toronto in June 1997.[45] His talk "Great Problems" is published under the promising title "Mathematical Problems for the Next

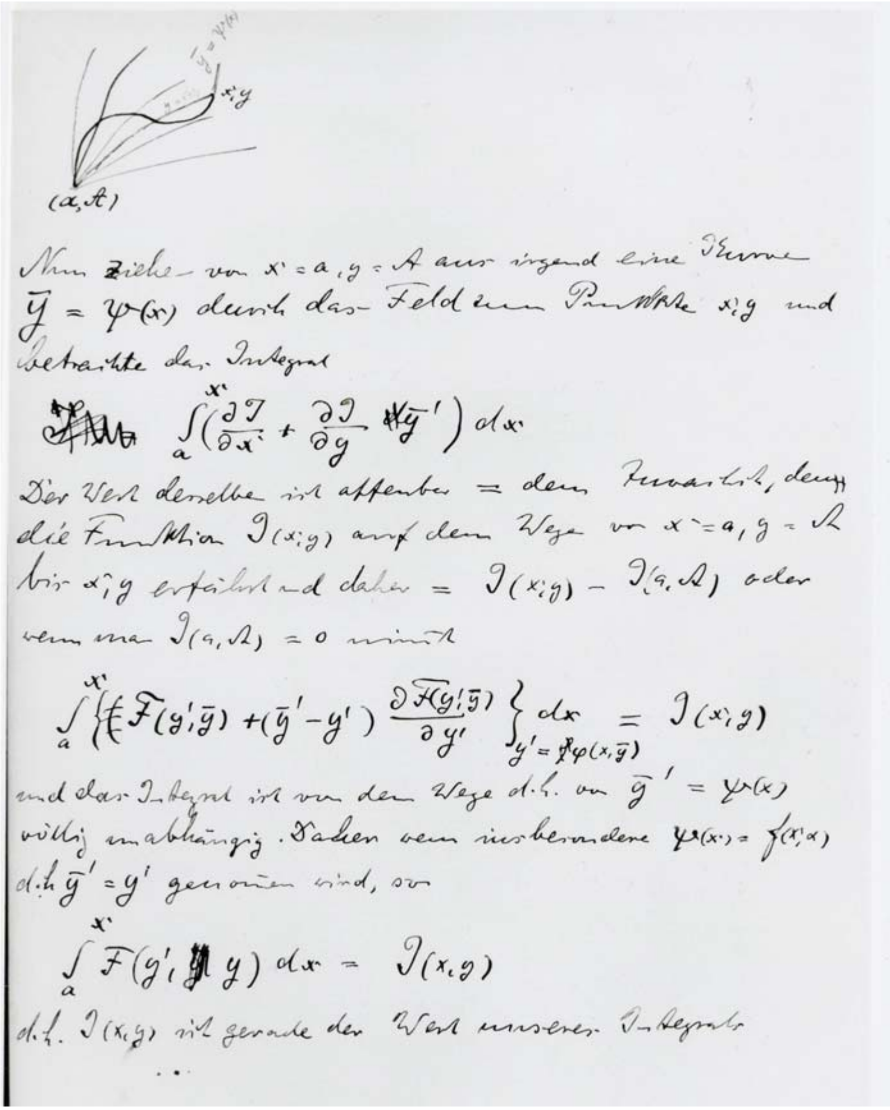
---

[42] Browder 1976.

[43] Aleksandrov 1969.

[44] Browder 1976, p. 507.

[45] Cf. also Atiyah's Fields lecture "Mathematics in the 20th Century", delivered in Toronto, Ont., in 2000. Video Tape of the Fields Institute, Toronto; printed in *N.T.M (N.S.)* 10 (2002), 25-39.



Nun ziehe von $x = a, y = A$ aus irgend eine Kurve
$\bar y = \psi(x)$ durch das Feld zum Punkte x, y und
betrachte das Integral

$$\int_a^{x} \left(\frac{\partial J}{\partial x} + \frac{\partial J}{\partial y}\bar y'\right) dx$$

Der Wert desselben ist offenbar = dem Zuwachs, den
die Funktion $J(x,y)$ auf dem Wege von $x = a, y = A$
bis x, y erfährt und daher $= J(x,y) - J(a,A)$ oder
wenn man $J(a,A) = 0$ nimmt

$$\int_a^{x} \left\{ F(y',\bar y) + (\bar y' - y')\frac{\partial F(y',\bar y)}{\partial y'}\right\}_{y'=\varphi(x,\bar y)} dx = J(x,y)$$

und das Integral ist von dem Wege d.h. von $\bar y' = \psi(x)$
völlig unabhängig. Daher wenn insbesondere $\psi(x) = f(x,\alpha)$
d.h. $\bar y' = y'$ genommen wird, so

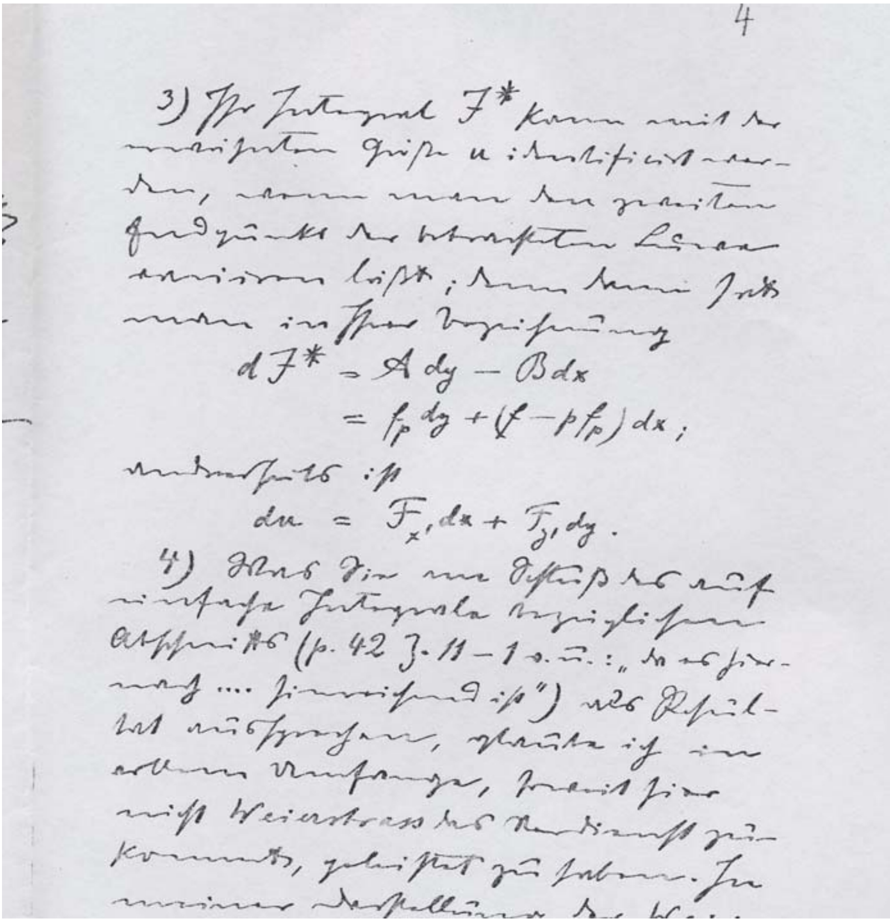
$$\int_a^{x} F(y', y)\, dx = J(x,y)$$

d.h. $J(x,y)$ ist gerade der Wert unseres Integrals

...

Fig. 10.16. First introduction of Hilbert's independent integral in the lecture "Flächentheorie, II (Theory of Surfaces, II)", summer term 1900, only a few weeks before the Paris talk was delivered in which the integral is part of the 23rd problem. Hilbert's lecture note. Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 552, p. 12.

Century"[46] and lists 18 problems, among them two of the Hilbert problems. In Smale's opinion the Riemann hypothesis, the Poincaré Conjecture, and the question "Does P = NP?" are the three greatest open problems. The first two belong completely to classical mathematics; the last also concerns computer science.



$$dJ^* = A\,dy - B\,dx$$
$$= f_p\,dy + (f - pf_p)\,dx;$$
$$du = F_{x'}\,dx + F_{y'}\,dy.$$

Fig. 10.17. Kneser's letter to Hilbert probably from September 2, 1900 (no exact date is available) in which Kneser pointed out his own introduction of the independent integral and desired to be mentioned in an appropriate way in the final version of Hilbert's paper "Mathematical Problems". Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 180, no. 4.

[46] Smale 1998.

Incidentally, the problem of the zeroes of Riemann's zeta function is usually considered to be the most important unsolved problem in mathematics, and so thought Hilbert. Moreover, Hilbert went so far as to maintain that it was absolutely the most important question for humanity. I cannot resist telling you a story I found somewhere in the writings of George Pólya (1887–1985), although the underlying German myth needs some initial explanation. Frederick I (1152–1190), named Barbarossa (i.e. Redbeard), was a German king and Roman emperor of the 12th century who died while on Crusade to the Holy Land. Since the 14th century popular legend has it that he is only sleeping in an imperial castle near Göttingen and every five hundred years looks to see what is going on in Germany, emerging in case Germany needs him.[47] Returning to Hilbert, somebody allegedly asked him: "If you came back, like Barbarossa, after five hundred years, what would you do?" Hilbert replied at once: "I would ask, has somebody proved the Riemann hypothesis?"

Furthermore, consider the role of problems in the work of the four extraordinary Hungarian mathematicians John von Neumann (1903–1957), Paul Erdös (1913–1999), György Pólya, and Gabor Szegö (1895–1985), who at the end of their careers taught in the USA. In 1954 Von Neumann lectured at the Amsterdam ICM on "Unsolved Problems", but he pointed out:

> The total subject of mathematics is clearly too broad for any of us. I do not think that any mathematician since Gauss has covered it uniformly and fully; even Hilbert did not and all of us are of considerably lesser width quite apart from the question of depth than Hilbert.[48]

That is why he restricted himself to a particular area of mathematics: operator theory viewed in its connections with other subjects. For the outstanding mathematician Erdös, problems were a downright passion. He solved problems, posed problems, and in the end he paid for problems. He paid for solutions or refutations – sometimes, for a solution or a refutation of the same problem, he even offered different amounts. Problem books are relatively recent in mathematics. Among them an eminent and early example is the book *Aufgaben und Lehrsätze aus der Analysis* (Springer's Yellow Series, vols. 19 and 20; 1925) of Pólya and Szegö included in the above list. Furthermore, Pólya wrote books on problem solving, among them *How to Solve It?*, and there is also a film *Let Us Teach It* on a class he gave on this topic.[49]

Recent works on problems include Jean-Michel Kantor's (born 1946) "Hilbert problems and their sequels" and Shiing-Shen Chern's (born 1911) complement "Remarks on Hilbert's 23rd Problem"[50] , as well as Jeremy Gray's

[47] The story is like many others, among them that of England's hero Francis Drake (1540?–1596) and the legendary fame of Drake's drum, described in a poem of Henry Newbolt (1862–1938).

[48] Redei 1999.

[49] Pólya 1957; the film is distributed by the MAA.

[50] Kantor 1996, Chern 1996.

Fig. 10.18. The hills of the Kyffhäuser Mountains are crowned by a ruined castle which was one of the largest German castles, destroyed in 1178. The ruins are surmounted by a monument (total height 64 m) showing Kaiser Wilhelm I (1797–1888), Prussian King since 1861 and German Emperor since 1871, as well as the German Emperor Frederick I (1122–1190) asleep within the mountain. The Kyffhäuser lies in Thuringia on the South-East side of the Harz Mountains, whereas Göttingen lies on the North-West side of the Harz Mountains, about 100 km in air distance. Photos R. Thiele.

(born 1947) *The Hilbert Challenge* and Benjamin Yandell's (born 1951) *The Honors Class*.[51] As Victor Katz (born 1943) wrote in 1993: "Hilbert's problems have in fact proved to be central in twentieth-century mathematics. Many have been solved, and significant progress has been achieved in the remainder. Perhaps a late-twentieth-century mathematician will present a new list of problems at the International Congress of Mathematicians in 1988 in Berlin." In Katz's last sentence resound the concluding words of Hilbert's Paris lecture which expressed the hope that well-posed problems will evoke enthusiasm and inspiration among mathematicians.[52] Of course, in Berlin there was no new list presented.[53] In 2000, however, the Clay Institute posed seven problems for our century; for each solution the Institute will award $1,000,000.[54]

In concluding this report on collections of problems let us look back at the 17th and 18th centuries with their rich heritage of important problems. Famous problems of this period are Kepler's sphere problem, Fermat's last theorem, Pascal's cycloid problem, de Beaune's problem, Viviani's Florentine enigma, Goldbach's conjecture, Mascheroni's constructions, etc. In these centuries it was quite common to pose problems to the mathematical community in public, so-called "provocationes" (provocations). Each mathematician who solved such a provocation was authorized to pose another one, and so on, bringing a cascade of questions into existence. Some of these problems became famous: Leibniz's problem of isochronous curves, Jakob Bernoulli's isoperimetric problems, Johann Bernoulli's problem of the shortest line on a surface, and others.

In 1696 John (Johann) Bernoulli (1667–1748) challenged "the most ingenious mathematicians of the whole terrestrial globe"[55] with a new problem, the problem of quickest descent, or, the Brachistochrone Problem.[56] Guillaume François Antoine de l'Hôpital (1661–1701) declared the problem to be one of "the most curious and most beautiful that has ever been proposed."[57] In another announcement John Bernoulli pointed out that nothing encourages noble minds more than the praise of later ages, and that fame and glory is all that a noble expects for his efforts.[58] At this time John Bernoulli and

[51] Gray 2000; Yandell 2002.

[52] Hilbert 1900, p. 297. "Mögen ihr [Mathematik] im neuen Jahrhundert geniale Meister erstehen und zahlreiche in edlem Eifer erglühende Jünger." (cf. footnote 32).

[53] Katz 1993/1998, p. 729/808.

[54] The Clay Mathematics Institute, Millennium Prize Problems, announced May 4, 2000 at the Collège de France, Paris. Presented by J. Tate and Sir Michael Atiyah. Cf. `http://www.claymath.org/millennium/`

[55] "Acutissimis qui toto Orbe florent Mathematicus", headline of a broadsheet (Programm Editum Groningae), distributed in 1697; also in: Speiser 1991, p. 259.

[56] "Problema novum", added at the end of a paper in *Acta eruditorum*, June 1696, p. 269; Speiser 1991, p. 212. Cf. Thiele 2002.

[57] Spiess 1955, p. 319. "Ce probleme [me] paroist des plus curieux et des plus jolis [que] l'on ait encore proposé." Letter to Joh. Bernoulli from June 15, 1696.

[58] "Programm Editum Groningae 1697", also in Speiser 1991, p. 259.

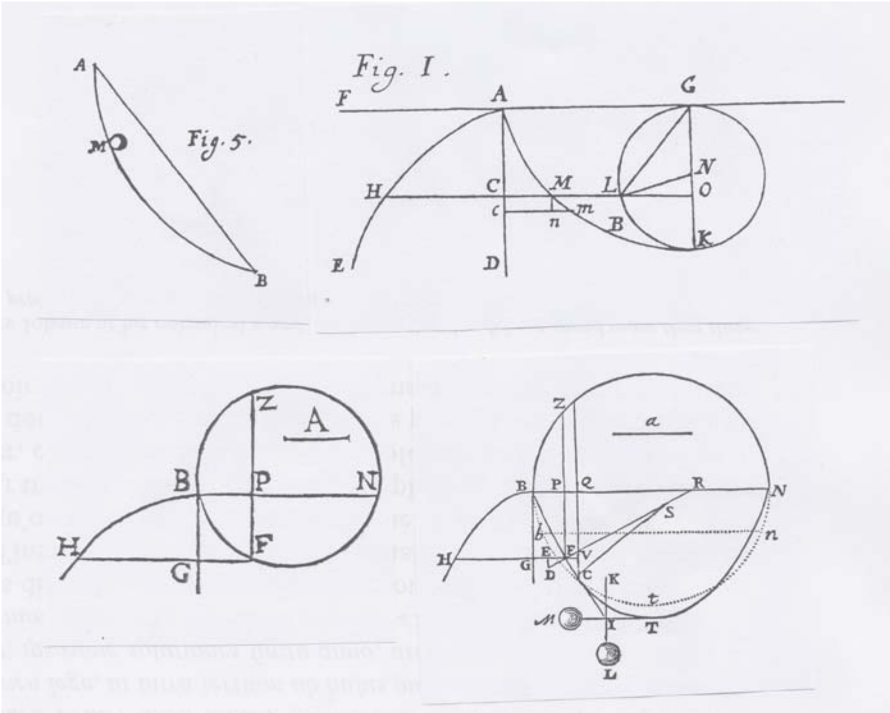


Fig. 10.19. Figures illustrating the influential Brachistochrone Problem (1696) and the Isoperimetric Problem (1697) of John and James Bernoulli respectively. Courtesy Deutsche Akademie der Naturforscher. Leopoldina. Halle. *Acta eruditorum*, June 1696 and May 1697, Journal de Sçavans, August 1698.

his brother James (Jacob) (1654–1705) began to quarrel. The challenge of the Brachistochrone Problem and above all the statement about reward by the argumentative John Bernoulli were aimed especially at his brother. Also quarrelsome, James (who solved the problem by a fundamental technique) took his revenge for the provocation by introducing three further challenges, among them the famous "isoperimetric problems."[59] In order to humble his brother he publicly offered John the amount of 50 imperial ducats which—in case John would solve the problems—were to be paid by an unnamed gentleman.[60] John for his part goaded and sneered at his brother. The situation escalated to the well-known quarrel, ended in 1705 by James's death. John, thirsting for honor and glory, regarded offering payment to a mathematician

[59] Speiser 1991. Solutio Problematum Fraternorum . . . cum Propositione reciproca aliorum. *Acta eruditorum*, May 1697, pp. 211–217; also Speiser 1991, pp. 271–282.

[60] Ibid., Speiser 1991, p. 276; cf. Thiele 1997a.

for the solution of a problem as a slander and wicked defamation.[61] Bygone times, tempi passati.

Just as the problems posed by the Bernoullis were significant for later developments, although it is very difficult to select such problems, Hilbert expected his chosen problems to play an important role in the future development of mathematics. Hermann Weyl aptly compared Hilbert's insight into the future of mathematics with that of politics:

> How much better he predicted the future of mathematics than any politician foresaw the gifts of war and terror that the new century was about to lavish upon mankind.[62]

10.3 Remarks on Hilbert's Philosophy of Mathematics

On one hand Hilbert demanded certainty by formalizing (axiomatic theory), but on the other hand he believed that mathematics advances by solving problems. For this reason it is all too easy to regard Hilbert exclusively as a pure formalist. Some of his remarks taken from his unpublished notebooks reinforce this: "Where does mathematics begin? As soon as concepts are fitted together, and only those facts that are contained in the concepts may be employed further."[63] For that reason axiomatics play an essential role: "The criterion for scientific method (truth) is axiomatizability. The axiomatic is the rhythm that makes music of the method, the magic wand that directs all the individual efforts to a common goal."[64] Incidentally, Hilbert answered the old enigma "Why can the world be described by mathematics? Why is mathematical science possible?" with: "Between thought and action there is no fundamental and no quantitative difference. This explains the pre-established harmony [understood in the sense of Leibniz (1646–1716)] and that simple

[61] In a letter to Legendre, July 2, 1830 C.G. Jacobi held a similar view: "Mais un philosophe comme lui [Fourier] aurait dû savoir que le but unique de la science, c'est l'honneur de l'esprit humain, et que sous ce titre, une question des nombres vaut autant qu'une question du système du monde." (But a philosopher like him ought to have known that the sole aim of knowledge is the honor of the human mind and that from this viewpoint a problem of number theory is as valuable as a problem of the system of the universe.) Jacobi 1830/1875, p. 272f.; also in Jacobi's *Werke*, vol. 1, p. 454f. Hilbert quoted this letter in his 1930 speech "The Knowledge of Nature" (Hilbert 1930, in the last paragraph).

[62] Weyl 1951/1968, p. 466.

[63] Mathematische Notizhefte. Library of the University of Göttingen. Cod. Ms. D. Hilbert 600:3, p. 116. "Wo fängt Mathematik an? Sobald Begriffe zusammengefügt werden, und nur das in den Begriffen liegende weiterhin benutzt werden darf."

[64] Ibid., Cod. Ms. D. Hilbert 600:2, p. 45. "Kriterium für Wissenschaftlichkeit (Wahrheit) ist die Axiomatisierbarkeit. Axiomatik ist der Rhythmus, der die Methode zur Musik macht - ist der Zauberstab, der alle die Einzelbestrebungen auf ein gemeinsames Ziel richtet."

experimental laws generate ever simpler theories."[65] So we live in the best of all possible worlds, at least in the sense that the world has the simplest (mathematical) description. Hilbert was convinced of this metaphysical principle: there is a realm behind phenomena, and the universe is governed in such a way that a maximum of simplicity and perfection is realized. In a notice we read: "Thus pre-established harmony also, because nature does not make such complicated things that cannot solved by mathematicians."[66]

We have many sources for Hilbert's deep belief that each well-posed problem can be solved. In opposition to the pessimism of the *fin de siècle* about 1900, Hilbert maintained his inspiring optimism which played a crucial role in his research and philosophy. Hilbert proudly noticed in 1930 that Auguste Comte's effort to pose an unsolvable celestial problem was met with a solution a few years later. And in conclusion he added the self-confident words:

> The true reason, according to my thinking, why Comte could not find an unsolvable problem lies in the fact that there is no such thing as an unsolvable problem.[67]

In 1922 he concluded his lecture "Wissen und mathematisches Denken (Knowledge and Mathematical Thinking)" with the famous words taken from his Paris lecture in 1900: "Wir müssen wissen, wir werden wissen (We must know, we shall know)" and again in a speech at Königsberg in 1930, partly repeated on air by the local radio station. In the end this optimistic axiom of the solvability of every problem was engraved on his tombstone in Göttingen.[68]

To the end of his career Hilbert repeatedly denied the "foolish ignorabimus"[69] of Emil du Bois-Reymond (1818–1896) and his successors, which he regarded as the battle-cry of slaves and reactionists, even as an intellectual sadism.[70] Du Bois-Reymond concerned himself with the limits of knowledge

---

[65]Ibid., Cod. Ms. D. Hilbert 600:3, p. 95. "Zwischen Denken und Geschehen ist kein prinzipieller und kein qualitativer Unterschied! Dadurch erklärt sich die praestabilierte Harmonie und die Tatsache, dass einfache experimentelle Gesetze auch immer einfachere Theorien ermöglichen."

[66]Ibid., Folder Quantentheorie, Cod. Ms. D. Hilbert 666. "Praestabilierte Harmonie auch darum [,] dass [weil] die Natur so komplizierte Sachen nicht macht, wie auch der Math. sie nicht lösen kann." (cryptic German).

[67]Reid 1996, p. 196.

[68]If you arrive at Göttingen by train from south you pass very close to Hilbert's grave, as well as those of Max Born (1882–1970), Otto Hahn (1879–1968), and Max Planck (1858–1947). The embankment and the cemetery, especially the line of these graves, are separated only by hedges.

[69]Keyword in the famous and widespread speech "Über die Grenzen der Naturerkenntnis (On the Limitations of Knowledge in Natural Sciences)" by the physiologist E. du Bois-Reymond (brother of the mathematician Paul du Bois-Reymond, 1831-1889), delivered in Leipzig in 1872.

[70]Mathematische Notizhefte, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:3, inserted pages.

Fig. 10.20. Hilbert's gravestone in Göttingen, cemetery at the Groner Landstraße. The optimistic lines "Wir müssen wissen. Wir werden wissen (We must know. We shall know)" are engraved in the pedestal and barely legible. Photo R. Thiele.

of nature – a widespread question around 1900. In accordance with the pessimistic spirit of the "fin de siècle (end of the century)" he maintained that there are problems we cannot and shall not solve. His well-known catchphrase was "Ignoramus et ignorabimus (We are ignorant and we shall remain ignorant)". For Hilbert such fruitless scepticism and prophesy of a downfall of culture was unacceptable. Against the general belief of the time Hilbert's device was "Noscemus",[71] which means "We can and shall know." In his 1930 speech he declared: "For the mathematicians there is no ignorabimus, nor, in my opinion, for any part of natural science."[72] In his notebook he wrote: "That there is no ignorabimus in mathematics can probably be proved by my

[71]Ibid., Cod. Ms. D. Hilbert, 600:1, p. 72.

[72]Hilbert 1930, p. 963; also in Hilbert 1932–1935, vol. 3, p. 387. "Für den Mathematiker gibt es kein Ignorabimus, und meiner Meinung nach auch für die Naturwissenschaften nicht." English translation from Ewald 1996, p. 1165.

theory of logical arithmetic", and "Perhaps it will turn out that there is no purport in saying there are insoluble problems."[73]

However, in one of life's little ironies, at almost the same time as Hilbert's great optimistic speech in Königsberg, a 25-year-old Kurt Gödel (1906–1978) proved striking results in 25 pages in a way that Hilbert had not anticipated. As a consequence of Gödel's and Paul Cohen's (born 1924) results, from any systems of axioms the continuum hypothesis[74] can neither be proved nor disproved (assuming the standard axioms of set theory are consistent, Zermelo-Fraenkel axioms plus the axiom of choice).

Despite Gödel's incompleteness theorem of 1931, Hilbert, at this stage 69 years old, continued his work to lay the foundations of mathematics. In his program, proposed in 1905 and more specifically after 1917, Hilbert intended to justify all of mathematics on the basis of elementary methods of finite reasoning. In a lecture on the infinite he declared: "Our thinking is finite; as we are thinking a finite process is going on. [...] The infinite is nowhere realized; it does not exist in nature, nor it is an admissible basis of our thinking - a remarkable harmony between being and thinking."[75] He repeated this conviction in his Königsberg speech in 1930, and in more detail he said: "We must be clear to ourselves that 'infinite' has no intuitive meaning and that without more detailed investigation it has absolutely no sense. For everywhere there are only finite things. There is no infinite speed, and no force or effect that propagates itself infinitely quickly. Moreover the effect itself is of a discrete nature and exists only in quanta. There is absolutely nothing continuous that can be divided infinitely often. Even light has atomic structure just like the quanta of action. I firmly believe that even space is of finite extent. [. . .] Infinity, because it is the negation of a condition that prevails everywhere, is a gigantic abstraction."[76]

---

[73]Mathematische Notizhefte, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert, 600:3, pp. 104, 98. "Dass es kein Ignorabimus in der Mathematik giebt, ist wahrscheinlich durch meine Theorie der Logik-Arithmetik beweisbar." – "Vielleicht stellt sich auch heraus: es hat keinen Sinn zu sagen, es gäbe unlösbare Probleme."

[74]Hilbert's first problem and described by him as a "very plausible theorem (einen sehr wahrscheinlichen Satz)", Hilbert 1900, p. 263.

[75]Hilbert 1924, p. 134. "Unser Denken ist finit, indem wir denken, geschieht ein finiter Prozeß [. . .] Das Unendliche findet sich nirgends realisiert; es ist weder in der Natur vorhanden, noch als Grundlage unseres Denkens zulässig - eine bemerkenswerte Harmonie zwischen Sein und Denken."

[76]Hilbert 1930, "[Wir müssen] uns klarmachen, daß 'Unendlich' keine anschauliche Bedeutung und ohne nähere Untersuchung überhaupt keinen Sinn hat. Denn es gibt überall nur endliche Dinge. Es gibt keine unendliche Geschwindigkeit und keine sich unendlich rasch sich fortpflanzende Kraft oder Wirkung. Zudem ist die Wirkung selbst diskreter Natur und existiert nur quantenhaft. Es gibt überhaupt nichts Kontinuierliches, was unendlich oft geteilt werden könnte. Sogar das Licht hat atomistische Struktur, ebenso wie die Wirkungsgröße. Selbst der Weltraum ist, wie ich sicher glaube, nur von endlicher Ausdehnung. [. . .] Die Unendlichkeit, weil sie eben

Gödel's results clarified the reach of the program and marked its limits. Admittedly, to a certain extent Hilbert reacted to Gödel's results and accepted stronger means in his proof theory such as transfinite induction.[77] Another essential roadblock was the 1936 result of Alan Turing (1912–1954) that the decision problem is unsolvable. To a certain extent Hilbert had dealt with this difficulty when he admitted that there may be no algorithm to find a certain element in an infinite set: "In my proof theory I do not maintain that among infinite objects the discovery of an object can be effected, but we can imagine that the choice is made."[78] We could continue here with the names of Ernst Zermelo (1871–1953), Bertrand Russell (1872–1970), Luitzen Egbert Brouwer (1881–1966), Rudolf Carnap (1891–1970), Alonzo Church (1903–1995), Stephen Cole Kleene, and others. However, that is another story. Nevertheless, taken all in all Hilbert's program proved valuable for the study of formal systems.

However, Hilbert was not only an old but also a sick man; his career was almost over. Furthermore, in 1933 the Nazis came to power and started to rule over Germany and over the German universities. A number of the best-known German mathematicians, including a great many of Hilbert's friends, colleagues and coworkers, became refugees or were murdered. The Hilbert circle in Göttingen was destroyed by the Nazis. The years to follow became for Hilbert years of tragic loneliness. We have a report on those days which shows the old and lonely Hilbert still had his sharp tongue. Asked by the Nazi minister of education whether mathematics in Göttingen suffered by eliminating the Jewish influence, Hilbert sadly answered: "There is really none any more."[79]

10.4 The 24th Problem

It is widely believed among mathematicians that simplicity is a reliable guide to the beauty or elegance of proofs, but like all aesthetic value judgments, such statements are highly subjective. The physicist Ludwig Boltzmann (1844–1906) once declared that we should cede elegance to the tailors. Hilbert admitted that despite the fact that (mathematical) beauty is highly satisfying,

---

die Negation eines überall herrschenden Zustands ist, eine ungeheuerliche Abstraktion." Translation by Ewald 1996, p. 1159 (§10).

[77] In the Preface of Hilbert 1934.

[78] Hilbert 1924, p. 134. "In meiner Beweistheorie wird [...] nicht behauptet, dass die Auffindung eines Gegenstandes unter den unendlich vielen Dingen stets [tatsächlich] bewirkt werden kann, wohl aber, dass man [...] stets so tun kann, als wäre die Auswahl getroffen."

[79] Cf. Reid 1996, p. 205. In oral history there is a report using Hilbert's East Prussian dialect: "Jelitten? Dat hat nich jelitten, dat jibt es doch janich mehr" (Suffered? It [mathematics] did not suffer, it does not exist anymore).

it cannot used in argumentation.[80] Can one really say that certain mathematical proofs are simpler than others? “Mathematical formalism is the base of all natural laws because with its help it is possible to say what simple is; for example, addition is necessarily the ~~first~~ simplest mathematical operation,”[81] wrote Hilbert in his notebooks. In the field which Hilbert later called proof theory and metamathematics, as early as 1900 he wanted a detailed investigation of the question of simplicity; of course, the needed formalism was yet to develop. But he believed that such an investigation carried out by his axiomatic method would be not only promising but also necessarily successful, especially by means of the reduction of proofs to an algebraic calculus, and this remained his belief in the 1920’s. Of course, it is an open problem whether or how such a reduction can be carried out.

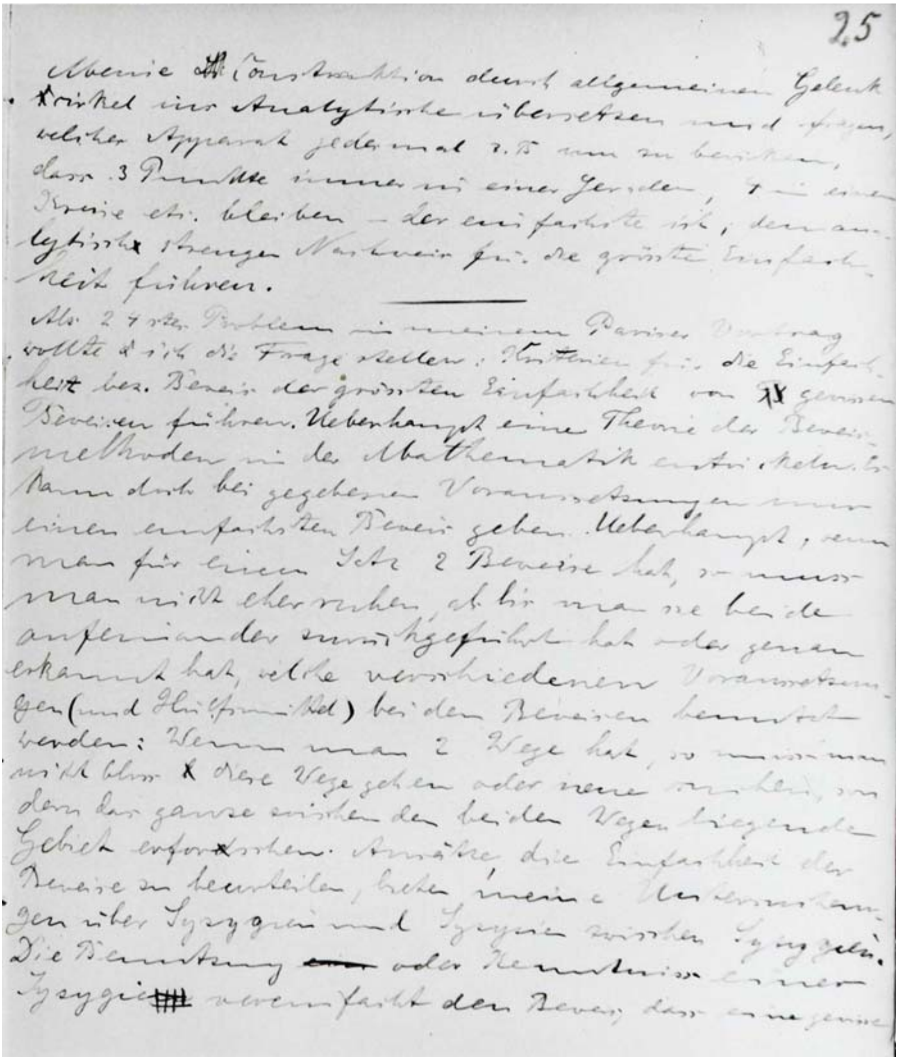
In his Mathematische Notizhefte Hilbert made a note of such a question he intended to include in his Paris lecture as a 24th problem; however, he then canceled it. As far as I know the canceled 24th problem has remained unpublished until now and I am not aware of any responses or references to the challenge, with the exception of Hilbert himself. I came across the problem in the *Hilbert Nachlass* at the Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, while I was studying Hilbert’s notices on the 23rd problem. Let me quote the problem itself:

> The 24th problem in my Paris lecture was to be: Criteria of simplicity, or proof of the greatest simplicity of certain proofs. Develop a theory of the method of proof in mathematics in general. Under a given set of conditions there can be but one simplest proof. Quite generally, if there are two proofs for a theorem, you must keep going until you have derived each from the other, or until it becomes quite evident what variant conditions (and aids) have been used in the two proofs. Given two routes, it is not right to either take these two or look for a third; it is necessary to investigate the area lying between the two routes. Attempts at judging the simplicity of a proof are in my examination of syzygies and syzygies [Hilbert made a slip in writing] between syzygies. The use or the knowledge of a syzygy essentially simplifies a proof that a certain identity is true.[82]

[80]Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 657, folder Physik, sheet 37. “Die Schönheit [... ist] ein wunderbares und den menschlichen Geist hoch befriedigendes Accedenz, aber kein Beweismittel.”

[81]Mathematische Notizhefte. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:3, inserted pages. “Der mathematische Formalismus ist die Grundlage aller Naturgesetze weil durch ihn möglich ist, zu sagen, was einfach ist. z.B. Addieren ist notwendig die erste einfachste mathematische Operation.”

[82]Mathematische Notizhefte. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung. Cod. Ms. D. Hilbert 600:3, p. 25. “Als 24stes


25

Theorie der Construktion durch allgemeinen Gelenk-
[illegible] ins Analytische übersetzen und fragen,
welcher Apparat jedesmal z.B. nur zu beweisen,
dass 3 Punkte immer in einer Geraden, 4 in einem
Kreise etc. bleiben – der einfachste ist, dann ana-
lytische strengen Nachweis für die grösste Einfach-
heit führen.

Als 24stes Problem in meinem Pariser Vortrag
wollte ich die Frage stellen: Kriterien für die Einfach-
heit bez. Beweis der grössten Einfachheit von gewissen
Beweisen führen. Ueberhaupt eine Theorie der Beweis-
methoden in der Mathematik entwickeln. Es
kann doch bei gegebenen Voraussetzungen nur
einen einfachsten Beweis geben. Ueberhaupt, wenn
man für einen Satz 2 Beweise hat, so muss
man nicht eher ruhen, als bis man sie beide
aufeinander zurückgeführt hat oder genau
erkannt hat, welche verschiedenen Voraussetzun-
gen (und Hilfsmittel) bei den Beweisen benutzt
werden: Wenn man 2 Wege hat, so muss man
nicht bloss diese Wege gehen oder neue suchen, son-
dern das ganze zwischen den beiden Wegen liegende
Gebiet erforschen. Ansätze, die Einfachheit der
Beweise zu beurteilen, bieten meine Untersuchun-
gen über Syzygien und Syzygien zwischen Syzygien.
Die Benutzung oder Kenntnis einer
Syzygie vereinfacht den Beweis, dass eine gewisse

Fig. 10.21. Entry in Hilbert's notebook on a canceled 24th problem demanding the "simplest proof", beginning with the eighth line from top. Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:3, p. 25f.

The entry stems from between autumn 1900 and summer 1901. Hilbert might have been inspired to remember the canceled problem again by the notice he made just before this entry. In the preceding entry he dealt with how to transform a geometrical construction into an analytical formalism with the help of a mechanism (linkage), and he asked for the simplest apparatus. Moreover, he had the idea to demonstrate the greatest possible simplicity ("grösste Einfachheit") by an analytic proof.[83]

I will not risk more than a few disjointed remarks.[84] Firstly, I have remarked on the desired proof theory. Secondly, with respect to the mentioned diversity of proofs, I refer to the history of the Fundamental Theorem of Algebra as an example. There are two ideas: one for an algebraic proof (Euler, 1707–1783) and one for an analytic one (d'Alembert, 1717–1783); furthermore, there is the research of Gauss between these ideas. Despite his "Pauca, sed matura (Few but ripe)", Gauss returned to this matter at various times and gave four proofs in total.

Thirdly, a proof is the most straightforward way to justify mathematical reasoning. To quote Godfrey Harold Hardy (1877–1947): "A mathematical

---

Problem in meinem Pariser Vortrag wollte ich die Frage stellen: Kriterien für die Einfachheit bez. Beweis der grössten Einfachheit von gewissen Beweisen führen. Ueberhaupt eine Theorie der Beweismethoden in der Mathematik entwickeln. Es kann doch bei gegebenen Voraussetzungen nur einen einfachsten Beweis geben. Ueberhaupt, wenn man für einen Satz 2 Beweise hat, so muss man nicht eher ruhen, als bis man sie beide aufeinander zurückgeführt hat oder genau erkannt hat, welche verschiedenen Voraussetzungen (und Hülfsmittel) bei den Beweisen benutzt werden: Wenn man 2 Wege hat, so muss man nicht bloss diese Wege gehen oder neue suchen, sondern das ganze zwischen den beiden Wegen liegende Gebiet erforschen. Ansätze, die Einfachheit der Beweise zu beurteilen, haben meine Untersuchungen über Syzygien und Syzyzien [sic] zwischen Syzygien. Die Benutzung ~~einer~~ oder Kenntnisse einer Syzygie~~en~~ vereinfacht den Beweis, dass eine gewisse Identität wahr ist." – The slips and corrections in writing show that Hilbert wrote in haste.

[83]Mathematische Notizhefte. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung. Cod. Ms. D. Hilbert 600:3, p. 25. "Meine Construktion durch allgemeine Gelenkcirkel ins Analytische übersetzen und fragen, welcher Apparat jedesmal [-] z.B. um zu bewirken, dass 3 Punkte immer in einer Geraden, 4 in einem Kreis etc. bleiben - der einfachste ist, den analytische strengen Nachweis für die grösste Einfachheit führen." In the lecture Hilbert 1910 (Elements and Essential Questions in Mathematics) Hilbert explained (p. 42): A Gelenkcirkel or a Gelenkmechanismus (linkage) is a generalization of a circle. In the xy-plane any system of rigid rods the endpoints of which are fixed, either in certain points of the plane or in certain points of the rods, and able to rotate about these points, forms a linkage (joint mechanism) which is assumed to have one degree of freedom. A point of such linkage describes an algebraic curve with the equation $f(x, y) = 0$. Working backward it is possible (at least in principle) to invent such a mechanism for each algebraic curve. Cf. Hilbert 1932; Courant 1953, chap. 3, esp. pp. 155-160; Bieberbach 1952, esp. §§6, 7, and 12.

[84]For more details see Thiele 2003.

proof should resemble a simple and clear-cut constellation, not a scattered cluster in the Milky Way."[85] Allow me to vary another famous bon mot of Hardy: simplicity is a first test; there is no permanent place for other mathematical patterns.[86] Hilbert declared: "The mathematician's function should be to simplify the intricate; instead they do just the opposite, and complicate what is simple, and call it 'generalizing'."[87] Hilbert went on:

> Besides, it is an error to believe that rigor in proof is the enemy of simplicity. On the contrary we find it confirmed ... that the rigorous method at the same time is the simpler and the more easily comprehended. The very effort for rigor forces us to find simpler methods of proof. [...] The most striking example is the calculus of variations.[88]

In general, a mathematical theorem is regarded as "deep" if its proof is difficult. The opposite of deep is "trivial". Nevertheless, there is a constant effort towards simplification, towards the finding of ways of looking at the matter from an easier, more trivial point of view. There is no question that simplicity of proof depends on the length of its presentation, on the method employed, on our familiarity with the used concepts, on its abstract generality, the novelty of ideas, and so on.

Having all such viewpoints in mind, what kinds of simplicity can we define at all precisely? Instead of proving mathematical theorems, how does one examine the deductive systems themselves and prove theses about them?

It is remarkable that, as always, Hilbert started the 24th problem with examples, and that specific results led him to the general idea. For Hilbert the simplest mathematical operation is addition, and to each addition there is a corresponding geometric or logical process. Hilbert investigated the possibility of such corresponding constructions.[89] "The geometrical figures are graphic formulas",[90] he said in his Paris talk. Without doubt, certain constructions or proofs rest on countable processes; that is, by an examination of the number of

---

[85]Hardy 1992, p. 113.

[86]In Hardy 1992 we find "beauty" instead of "simplicity", p. 85.

[87]Mathematische Notizhefte. Library of the University of Göttingen, Cod. Ms. D. Hilbert 600:1, p. 45. "Die Thätigkeit der Mathematiker sollte darin bestehen, das Verwickelte einfach zu machen. Statt dessen machen sie umgekehrt das Einfache verwickelt und nennen das Verallgemeinern." See also Hardy 1992, p. 105: "Generality is an ambiguous and rather dangerous word."

[88]Hilbert 1900, pp. 257, 258. "Zudem ist es ein Irrtum zu glauben, daß die Strenge in der Beweisführung die Feindin der Einfachheit wäre. An zahlreichen Beispielen finden wir im Gegenteil bestätigt, daß die strenge Methode auch zugleich die einfachere und leicht faßlichere ist. Das Streben nach Strenge zwingt uns eben zur Auffindung einfacherer Schlußweisen. [...] Das schlagendste Beispiel aber für meine Behauptung ist die Variationsrechnung."

[89]For example, Hilbert investigated the possibility of such constructions using only a ruler with a given unit in chapter 7 of Hilbert 1899.

[90]Hilbert 1900, p. 259. "Die geometrischen Figuren sind gezeichnete Formeln."

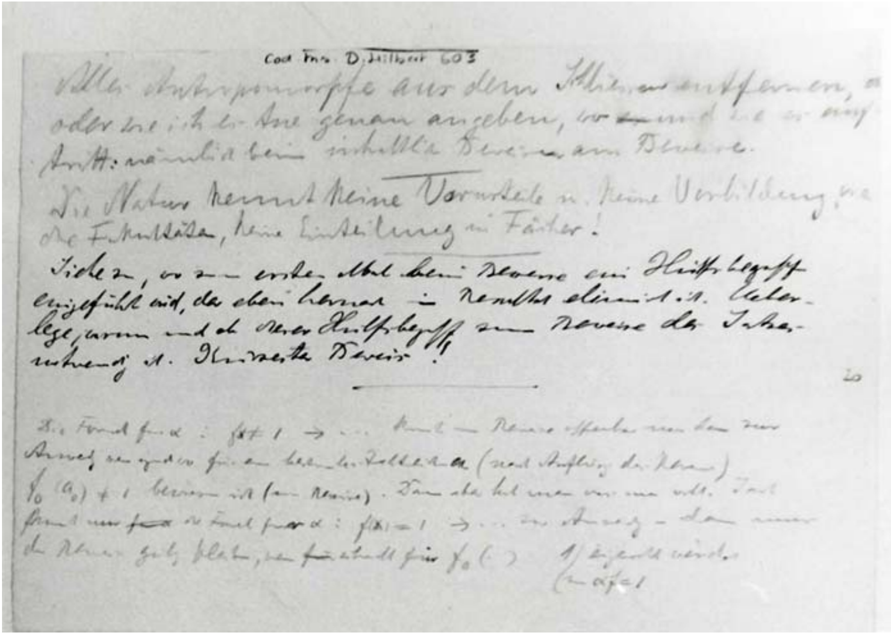


Fig. 10.22. Notice of Hilbert which considers proofs (third entry). "Watch where in a proof for the first time an auxiliary concept is introduced which is later eliminated in the result. Consider why and whether this auxiliary concept is needed for the proof. Shortest proof!" Courtesy Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 603, sheet 20.

operations involved in the proof we can decide which of two proofs is simpler. In this view it seems possible to arrange mathematical proofs somehow in strata. Mathematics appears as a well-ordered stock, as a hierarchical ordering of formulas, and the task under consideration is to examine the proofs or the corresponding chains of formulas respectively with respect to simplicity.

The French mathematician Émile Lemoine (1840–1912) proposed a criterion for simplicity of geometric constructions.[91] In 1888 in *La Géométrographie ou l'Art des Constructions Géometriques* (*The Geometrography or the Art of Geometric Constructions*), Lemoine reduced all geometric constructions by ruler and compass to only five basic constructions. One of them is the placing of a compass end at a given point. Lemoine called the total number of times any of these basic operations was used the simplicity of the construction ("coëfficient de simplicité", or just "la simplicité"). This measure of the complexity of a geometric construction produced some unexpected and surprising results with respect to standard constructions. For example: the standard con-

[91] É.M.H. Lemoine, *La Géométrographie ou l'Art des Constructions Géometriques.* Paris: Naud 1902. There is a summary by Van der Waerden 1937/38.

struction of four tangents to two circles has degree of simplicity 92; Lemoine gave another construction with a reduced degree of only 34.[92]

Hilbert similarly wanted to make proofs a measurable object of another theory by his "logical arithmetic", in which only finite methods should be used. Of course, the discovery of such a measure is a delicate business. Invariant theory is a bridge between geometry and algebra. Hilbert's fundamental theorem reads that we can pick a finite number of invariants $i_1, i_2, \ldots, i_k$ by which we can express any invariant as a polynomial in these basic invariants. In his aim Hilbert was probably guided by his investigations on invariants, especially on "syzygies". What are syzygies? The basic invariants $i_1, i_2, \ldots, i_k$ are not algebraic independently but themselves fulfill a homogeneous polynomial relation $f(i_1, i_2, \ldots, i_k) = 0$. Such an identity is called a syzygy.

What is the role of such syzygies in geometry? As an example consider any geometry with inhomogeneous Cartesian coordinates and all linear homogeneous transformations (affine transformations). In this affine geometry geometric magnitudes are invariants which under linear homogeneous transformation are altered only by a factor (such as the area of an triangle, which can be expressed analytically by certain determinants of the coordinates). Such invariant magnitudes have a geometric meaning in the affine geometry. Moreover, in affine geometry the mentioned determinants make the full system of invariants. In other words, every invariant can be expressed as a polynomial in these determinants and because of the geometric meaning of an invariant, this meaning is expressed analytically by the polynomial relation (syzygy). Conversely, to each theorem (on invariants) of the affine geometry corresponds a syzygy. This means that by determining the full system of syzygies the theory of invariants allows one to describe all theorems of affine geometry.

Can we extend such considerations to other geometries? The polynomial relators, syzygies, can be added and multiplied; they are closed under the operations of addition and multiplication and form a subring of R, where R denotes the ring formed by the invariants (see above, p. 250). Moreover, they even form an ideal I of R which has a finite basis (Hilbert). That means we can pick a finite number of relators, $f_1, f_2, \ldots, f_k$, and express every polynomial relator f in the form $f = Q_1 f_1 + Q_2 f_2 + \cdots + Q_k f_k$, the Q_i being polynomials (Hilbert's basis theorem). This finite basis $f_1, f_2, \ldots, f_k$ is not algebraically independent. Thus one obtains new relators (second-order syzygies) again with a finite basis; these relators are not algebraically independent, and so on. However, Hilbert proved that the cascade of syzygies stops in at most $r + 1$ steps, where $r = \xi(m)$ is the number of invariants of the full invariant system of the n-ary forms F of degree m; $\xi(m)$ is called the characteristic function and is a rational function of m. It was in the years up to 1892 that

[92] Cf. also Lemoine 1893, submitted to the Chicago Congress 1893. Hilbert contributed to the same volume with Hilbert 1893 (On the Theory of Algebraic Invariants), pp. 116–124, cf. footnote 14. See also Hilbert 1890.

Hilbert proved the fundamental finiteness theorem of the theory of invariants for the full projective group, the Hilbert Syzygy Theorem: "The chain of syzygies terminates after finitely many steps."[93] Furthermore, Hilbert's basis theorem reads: "If every ideal in a commutative ring R is finitely generated, then so is $R[x_1, \ldots, x_n]$." Taken all in all, we stay in a *finite* system.

Suppose we can eliminate complex geometric conceptions, i.e., express them in a finite form. Then it should be possible to exploit this finiteness to establish the simplicity of geometric proofs. More generally Hilbert pointed out: "With each mathematical theorem [...] one can ask whether there is any way to determine how many operations are needed at most to carry out the assertion of the theorem. Kronecker has particularly emphasized the question of whether one can carry it out in a finite number of steps."[94]

As to the complexity of technical details for proofs consider the well-known Four-Color Problem.[95] Aside from the possibility of a computer-aided proof which we cannot survey step by step, there remains Hilbert's practical question "whether in mathematics problems exist that cannot be dealt with in a prescribed short time?"[96] His example: calculate the n-th digit in the decimal expansion of π, where n is equal to $(10^{10})^{10}$. A simple calculation on an envelope shows that an ideal computer of the largest size working since the Big Bang would have been able to carry out only a finite number of operations somewhere between 10^{120} and 10^{160}. Nevertheless, we read in Hilbert's notebook: "All our effort, investigation, and thinking is based on the belief that there can be but one valid opinion", or: "The proof of proofs: that it must

[93]Hilbert 1897. Lecture *Theorie der algebraischen Invarianten nebst Anwendungen auf Geometrie*, delivered in Göttingen summer term 1897. Lecture notes taken by S. Marxsen, Mathematisches Institut, Universität Göttingen. English translation: *Theory of Algebraic Invariants* by R. Laubenbach, Cambridge: University Press 1994, p. 173. "Die Kette der Syzygien bricht nach einer endlichen Anzahl von Schritten ab." Bl. 773.

[94]Ibid., p. 133. "Man kann bei jedem mathematischen Satze [...] fragen, ob man auf irgendeine Weise darüber Aufschluß geben kann, wie viele Operationen man höchstens gebraucht, um das im Satze Gesagte auszuführen. Mit besonderem Nachdruck hat Kronecker die Frage nach der Ausführbarkeit durch ein endliche Anzahl von Schritten betont." Bl. 563f.

[95]Using this example, Hilbert pointed out that a peculiarity of mathematics consists in the state of affairs that despite the fact that problems can be very simply formulated, their solutions can be very hard. ("Es ist eine Eigentümlichkeit der Math.[ematik], dass einfachste spezielle Probleme (z.B. das 4-Farbenproblem) sehr schwierig zu lösen sind. Erkläre dies Phänomen." Mathematische Notizhefte. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600: 3, inserted pages.)

[96]Mathematische Notizhefte. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung. Cod. Ms. D. Hilbert 600:1, p. 53. "Ob es Probleme in der Mathematik giebt, deren Erledigung nicht in einer vorgeschriebenen kleinen Zeit möglich ist?"

always be possible to arrive at a proof."[97] Thus there must be a simplest proof for any given proposition. Colloquially he added: "always apply the strictest proof! Philological-historic import must be wiped out. Given a 15-inch gun, we don't shoot with the crossbow."[98]

But if we are willing to remain in the sphere of the finite, can we actually justify all the needed mathematical conclusions? Let us consider the proofs needed in mathematics and all proofs we can actually execute. Maybe a gap appears: with the help of finite methods we are only able to deduce a countable set of proofs (Gromov's Hilbert tree), while the set of possible proofs is probably uncountable.[99] Was Hilbert's belief in the power of thinking somewhat naive? Was it indebted to the widespread belief of progress at the turn of the century, part of the zeitgeist? Consider this quotation from his Paris lecture:

> Occasionally it happens that we seek the solution under insufficient presuppositions or in an incorrect sense, and for this reason we do not succeed. The problem then arises: to show the impossibility of the solution under the given hypothesis.[100]

Nevertheless, despite an increasing diversity of mathematical branches and their problems by the simplification of proofs due to axiomatic methods, very much in the spirit of Hilbert, mathematics as a whole has become more economically designed, has been increasingly unified and has widened our horizons. Hilbert's spirit and influence has been vivid.

This is a good point to stop. The last word is Hilbert's: "Mathematics stalks on earthly ground and at the same time touches the divine firmament."[101]

---

[97]Ibid. 600:3, p. 96. "Allem unserem Streben, Forschen und Denken liegt doch die Meinung zu Grunde, dass es nur eine richtige Meinung geben kann (Maximum)" - "Beweis aller Beweise: dass man den Beweis immer muss finden können!"

[98]Ibid., 600:3, inserted pages. "Immer das schärfste Mittel anwenden! philologisch-historische Sinn muss ausgerottet werden. Wenn wir 42cm-Kanone haben, schiessen wir doch nicht mit Armbrust."

[99]For details see Gromov 2000, p. 1213f.

[100]Hilbert 1900, p. 261. "Mitunter kommt es vor, daß wir die Beantwortung unter ungenügenden Voraussetzungen oder in unrichtigem Sinne erstreben und infolgedessen nicht zum Ziele gelangen. Es entsteht dann die Aufgabe, die Unmöglichkeit der Lösung des Problems unter den gegebenen Voraussetzungen [...] nachzuweisen."

[101]Mathematische Notizbücher. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:3, p. 104. "Die Math.[ematik] schreitet auf der Erde und berührt zugleich den Himmel."

Hilbert's Mathematical Problems[102]

This is the problem, find its solution. You can find it by pure thinking since in mathematics there is no Ignorabimus.[103]

1. The cardinality of the continuum, including well-ordering (Cantor's problem, 1878),
Gödel 1940, Cohen 1963 (awarded Fields Medal 1966), Vopenka 1975.
2. The consistency of the axioms of arithmetic,
Gödel 1931, Gentzen 1936, Novikov 1941.
3. The equality of the volumes of two tetrahedra of equal bases and equal altitudes,
Dehn 1902, Kagan 1903.
4. The straight line as shortest connection between two points,
Hamel 1901, Funk 1929, Busemann 1955.
5. Lie's concept of a continuous group of transformations without the assumption of the differentiability of the functions defining a group,
Kolmogorov 1930, E. Cartan 1930, Pontrjagin 1932,
von Neumann 1933, Chevalley 1941, Malcev 1946, Gleason 1952,
Montgomery 1952, Zippin 1952.
6. The axioms of physics,
It is not yet clear what axiomatizing physics really means;
not a problem but a program of research.
Particular fields were axiomatized:
Classical mechanics by Hamel 1906,
Thermodynamics by Carathéodory 1909,
and others.
7. Irrationality and transcendence of certain numbers,
Siegel 1921, Gelfond 1929, 1934, Schneider 1934, Baker 1966 et al.
8. Prime number theorems (including the Riemann hypothesis),
Hecke 1917, Schnirelman 1930, Vinogradov 1937.
9. The proof of the most general reciprocity law in arbitrary number fields,
Hilbert 1897, Takagi 1920, Artin 1928, Hasse 1935, Shafarevich 1950.
10. Decision on the solvability of a Diophantine equation,
no solution: Thue 1908, Siegel 1928, Robinson 1969,
Matijasevich 1970.
11. Quadratic forms with any algebraic coefficients,
Hasse 1929, Siegel 1936, 1951, Weil 1964, Ono 1964.

[102] Problems with a number in bold print were presented in the historic speech in Paris. For more details cf. Bieberbach 1930, Aleksandrov 1969, Fang 1970, Browder 1976, Gray 2000, Yandell 2002.

[103] "Da ist das Problem, suche die Lösung. Du kannst sie durch reines Denken finden; denn in der Mathematik gibt es kein Ignorabimus." End of the introductory essay in Hilbert 1900, p. 262.

12. The extension of Kronecker's theorem on Abelian fields to arbitrary algebraic fields,
    unsolved.

13. Impossibility of solving the general seventh degree equation by means of functions of only two variables,
    Arnold 1957; unsolved if analyticity is required.

14. Finiteness of systems of relative integral functions,
    no solution: Nagata 1959.

15. A rigorous foundation of Schubert's enumerative calculus,
    van der Waerden 1930f., in general unsolved.

16. Topology of real algebraic curves and surfaces,
    partial results.

17. Representation of definite forms by squares,
    for real-closed fields Artin 1926,
    negative solution in general Du Bois 1967.

18. The building up of space from congruent polyhedra,
    Bieberbach 1908, Hajos 1941.

19. The analytic character of solutions of variation problems,
    special results.

20. General boundary value problems,
    Hilbert dealt with the special case of Dirichlet's principle 1900, 1904 and Plateau's problem, Bernstein 1904, Douglas 1939, Courant 1950,
    Morrey 1966, and many others.

21. Linear differential equations with a given monodromy group,
    Hilbert 1905, negative solution in general Anasov 1994,
    Bolibruch 1994.

22. Uniformization of analytic relations by means of automorphic functions,
    Koebe 1907, Poincaré 1907.

23. The further development of the methods of the calculus of variations,
    Kneser 1900, Hilbert 1905, Carathéodory 1935, Weyl 1936,
    Boerner 1936, Lepage 1936, and others.

[24.] The simplicity of proofs,
    canceled.

References

Sources

Hilbert, D., *Nachlass.* Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung (Special Collections). Cod. Ms. D. Hilbert and Cod. Ms. Math.-Archiv 76 (some letters): among them
    Mathematische Notizhefte, 3 vols, some inserted pages.
    Cod. Ms. D. Hilbert 600:1–3.

Fimpel, M., 2002. *Spezialinventar zur Geschichte der Mathematik und Naturwissenschaften an der Universität Göttingen von 1880-1933.* Schriften des Universitätsarchivs Göttingen, Bd. 1 (U. Hunger and H. Wellenreuter, eds.). Göttingen: Universitätsarchiv.
II. B.4 Nachlaß Hilbert, Cod. Ms. D. Hilbert, pp. 438-485,
II. B. 15 Mathematik-Archiv, Math-Arch., pp. 605-616.

Hilbert, D., 1886. Invariantentheorie. Wintersemester 1886. (Theory of invariants.) Lecture notes from Winter 1886 by Hilbert. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 521.

Hilbert, D., 1897. Theorie der algebraischen Invarianten nebst Anwendungen auf Geometrie. Sommersemster 1897. (Theory of algebraic invariants including applications to geometry.) Lecture notes taken by S. Marxsen. University of Göttingen, Library of the Mathematical Institut.

Hilbert, D., 1905. Logische Prinzipien des mathematischen Denkens. Sommersemester 1905. (Logical principles of mathematical thinking.) Lecture notes taken by M. Born. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 558; Lecture notes taken by E. Hellinger. University of Göttingen, Library of the Mathematical Institut.

Hilbert, D., 1910. Elemente und Principienfragen der Mathematik. Sommersemester 1910. (Elements and principal questions of mathematics.) Lecture delivered in the summer term 1910 and recorded by R. Courant. Library of the Mathematical Institute, University of Göttingen.

Hilbert, D., 1917. Principien der Mathematik. Wintersemester 1917. (Principles of Mathematics, winter term 1917/18.) Lecture notes by P. Bernays. University of Göttingen, Library of the Mathematical Institut.

Hilbert, D., 1919. Natur und mathematisches Erkennen. Wintersemester 1919. (The mathematical knowledge of nature.) Lecture notes taken by P. Bernays, winter term 1919/20. University of Göttingen, Library of the Mathematical Institut.

Hilbert, D., 1921. Geometrie und Physik. Sommersemester 1921. (Geometry and Physics.) Type-written manuscript. Staatsbibliothek Berlin, Handschriftenabteilung, Nachlaß Hückel 2.14.

Hilbert, D., 1921a. Kopenhagener Vorträge, 1921. Lectures delivered in Copenhagen 1921, 2 copy books. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 589.

Hilbert, D., 1921b. Hamburger Vorträge, 1921. Lectures delivered in Hamburg 1921, 3 copy books with an insertion. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 596.

Hilbert, D., 1924. Über das Unendliche. Wintersemester 1924/25. (On the Infinite, winter term 1924/25.) Lecture notes, taken by L. Nordheim. University of Göttingen, Library of the Mathematical Institut.

Hilbert's printed works

Hilbert, D., 1932-1935. *Gesammelte Abhandlungen*. 3 vols. Berlin: Springer.

Hilbert, D., 1900/1902. "Mathematische Probleme". Vortrag, gehalten auf dem internationalen Mathematiker-Kongreß. *Nachrichten der Akademie der Wissenschaften in Göttingen* (1900), 253–297; English translation by Mary Winston Newson with the permission of Hilbert in: *Bulletin of the American Mathematical Society* 8 (1901/02), 437–479; reprinted in the same *Bulletin* 37 (2000), 407–436.

Hilbert, D., 1888. Zur Theorie der algebraischen Gebilde. *Nachrichten der Akademie der Wissenschaften in Göttingen*, p. 450–457.

Hilbert, D., 1889. Über die Endlichkeit des Invariantensystems für binäre Grundformen. *Mathematische Annalen* 33, 223–226.

Hilbert, D., 1890. Über die Theorie der algebraischen Formen. *Mathematische Annalen* 36, 473–534.

Hilbert, D., 1896. Über die Theorie der algebraischen Invarianten. In: *Mathematical Papers read at the International Congress held in Chicago 1893*. E.H. Moore et al., eds. New York: Macmillan 1896, pp. 116–124.

Hilbert, D., 1897/1993. *Theory of algebraic invariants*. Translation of a lecture delivered in 1897. Eds. R.C. Laubenbacher and B. Sturmfels. New York: Cambridge University Press 1993.

Hilbert, D., 1899. *Grundlagen der Geometrie*. Leipzig: B.G. Teubner.

Hilbert, D., 1930. "Naturerkennen und Logik" speech delivered in Königsberg. *Die Naturwissenschaften* 18 (1930), 959–963; also in Hilbert 1932–1935, vol. 3, pp. 378–387; English translation in Ewald 1996. pp. 1157–1165.

Hilbert, D. and S. Cohn-Vossen, 1932. *Anschauliche Geometrie*. Berlin: Springer.

Hilbert, D. and P. Bernays, 1934. *Grundlagen der Mathematik*, I. Berlin: Springer.

Hilbert, D., 1970. *Hilbert's Invariant Theory Papers*. Translated by M. Ackerman. Brookline (MA): Math Sci Press.

Frei, G. (ed.), 1985. *Der Briefwechsel David Hilbert - Felix Klein (1886-1918)*. Göttingen: Vandenhoeck & Ruprecht.

Selected references

Aleksandrov, P.S., 1969/1976. *Problemy Gilberta*. Moscow: Nauka 1969; German translation *Die Hilbertschen Probleme*. Leipzig: Geest & Portig 1976 (reprints).

Bernays, P., 1922. Über Hilberts Gedanken zur Grundlegung der Arithmetik. *Jahresbericht der Deutschen Mathematiker Vereinigung* 31, 65–65.

Bernays, P., 1935. Hilberts Untersuchungen über die Grundlagen der Arithmetik. In: Hilbert 1932–1935, vol. 3, pp. 196–216.

Bieberbach, L., 1930. Über den Einfluß von Hilberts Pariser Vortrag über "Mathematische Probleme" auf die Entwicklung der Mathematik in den letzten dreißig Jahren. *Die Naturwissenschaften* 18 (51), 1101–1111. B

Bieberbach, L., 1952. *Theorie der geometrischen Konstruktionen*. Basel: Birkhäuser.

Blumenthal, O., 1922. *David Hilbert. Die Naturwissenschaften* 10 (4), 67–72.

Browder, F.E. ed., 1976. *Mathematical Development Arising from Hilbert Problems*. Proceedings of Symposia in Pure Mathematics, vol. 28 (2 parts). Providence: AMS.

Carathéodory, C., 1943. David Hilbert. *Sitzungsberichte der Bayerischen Akademie der Wissenschaften. Mathematisch-naturwissenschaftliche Klasse*, pp. 350–354.

Chern, S.-S., 1996. Remarks on Hilbert's 23rd Problem. *The Mathematical Intelligencer* 18 (4), 7–9.

Courant, R. and H. Robbins, 1953. *What is Mathematics?* New York: Oxford University Press.

Dieudonné, J., 1971. *Invariant Theory, Old and New*. New York: Academic Press.

Ewald, W., 1996. *From Kant to Hilbert*, 2 vols. Oxford: Claredon Press.

Fang, J., 1970. *Hilbert*. Towards a Philosophy of Modern Mathematics, II. Hauppauge: Paideai.

Freudenthal, H., 1970. "David Hilbert". In: *Dictionary of Scientific Biography* (C. Gillispie, ed.), vol. VI. New York: Scribner's, pp. 388–395.

Grattan-Guinness, I., 2000/2001. A Sideway Look at Hilbert's Twenty-three Problems of 1900. *Notices of the American Mathematical Society* 47, 752–757; corrections by G.H. Moore and response of Grattan-Guinness, in same *Notices* 48 (2001), 167.

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

Gray, J., 2000a. The Hilbert Problems 1900-2000. *Newsletter European Mathematical Society* 36, 10–13.

Gromov, M., 2000. Interview with Gromov. In: *Development of Mathematics, 1950-2000* (J.-P. Pier, ed.). Basel: Birkhäuser, pp. 1213–1227.

Halsted, G.B., 1900. The International Congress of Mathematicians. *The American Mathematical Monthly* 7, 188–189.

Hardy, G.H., 1992. *A Mathematician's Apology*. Foreword by C.P. Snow. Cambridge: Cambridge University Press. (First edition 1940).

Heisenberg, W., 1971. Talk "Die Naturgesetze und die Struktur der Materie", delivered in Athens 1964. In: W. Heisenberg, *Schritte über Grenzen*. München: Piper. English translation: *Across the Frontiers*. Oxbow Press 1990 (Reprint).

Hellinger, E., 1935. Hilberts Arbeiten über Integralgleichungen und unendliche Gleichungssysteme. In: Hilbert 1932–1935, pp. 94–145.

Jacobi, C.G.J., 1830/1875. Correspondance Mathématique. *Journal für die reine und angewandte Mathematik* 80, 205–279.

Kantor, J.-M., 1996. Hilbert Problems and Their Sequels. *The Mathematical Intelligencer* 18 (1), 21–30.

Kaplansky, I., 1977. *Hilbert's Problems*. Lecture notes. University of Chicago.

Katz, V.J., 1993/1998. *A History of Mathematics*. New York: Harper Collins 1993; 2nd ed. Reading (MA): Addison-Wesley 1998.

Koetsier, T., 2001. Een verrassende vondst: Duits historicus ontdekt Hilberts 24ste probleem. *Nieuw Archief voor Wiskunde* 5 (2), 65–67.

Lemoine, É.M.H., 1893. Considérations Générales sur la Mesure de la Simplicité ... , in: *Mathematical Papers Submitted to the International Mathematical Congress in Chicago 1893*, published by the American Mathematical Society, vol. 1. E.H. Moore et al., eds. New York: Macmillan 1896, pp. 143–154.

Lemoine, É.M.H., 1902. *La Géométrographie ou l'Art des Constructions Géometriques*. Paris: Naud.

Parshall, K.V.H., 1990. The One-Hundredth Anniversary of the Death of Invariant Theory? *The Mathematical Intelligencer* 12 (4), 10–16.

Pier, J.-P. (ed.), 1994/2000. *Development of Mathematics*. 1900-1950 and 1950-2000 (2 vols.) Basel: Birkhäuser 1994 and 2000.

Poincaré, H., 1908. L'Avenir des Mathématiques. In: *Atti del IV Congresso Internazionale Mathematici, Roma 6-11 Aprile 1908* (G. Castelnuovo, ed.), vol. I. Roma: Accademia dei Lincei 1909, pp. 1167–182.

Pólya, G., 1954. *Mathematics and Plausible Reasoning*, 2 vols. Princeton: Princeton University Press.

Pólya, G., 1957. *How to Solve It?* Princeton: University Press.

Popov, V.L., 1992. *Groups, Generators, Syzygies, and Orbits in Invariant Theory*. Translated from the Russian. Providence: American Mathematical Society.

Reid, C., 1970. *Hilbert*. New York: Springer 1970 and Copernicus 1996.

Redei, M., 1999. Unsolved Problems in Mathematics. John von Neumann's Address to the International Congress of Mathematicians, Amsterdam, 1954. *The Mathematical Intelligencer* 21 (4), 7–12.

Rota, G.C., 1999. Two Turning Points in Invariant Theory. *The Mathematical Intelligencer* 21 (1), 22–28.

Rudio, F. (ed.), 1898. *Verhandlungen des ersten Internationalen Mathematiker-Kongresses, Zürich 9.-11. August 1897*. B.G. Teubner: Leipzig.

Schmidt, A., 1933. Zu Hilberts Grundlagen der Geometrie. In: Hilbert 1932–1935, vol. 2, pp. 404–414.

Scott, C.A., 1900. The International Congress of Mathematicians in Paris. *Bulletin of the American Mathematical Society* 7, 57–79.

Smale, S., 1998. Mathematical Problems for the Next Century. *The Mathematical Intelligencer* 20 (2), 7–15.

Speiser, D. (ed.), 1991. *Die Streitschriften von Jacob und Johann Bernoulli* (commented by H.-H. Goldstine). Birkhäuser: Basel.

Spiess, O. (ed.), 1955. *Der Briefwechsel von Johann Bernoulli*, vol. I. Basel: Birkhäuser.

Thiele, R., 1997. Über die Variationsrechnung in Hilberts Werken zur Analysis. *N.T.M. (N.S.)* 5, 23–42.

Thiele, R., 1997a. Das Zerwürfnis Johann Bernoullis mit seinem Bruder Jakob. *Acta Historica Leopoldina* (Volume in Honor of K.-R. Biermann) 27, 257–276.

Thiele, R., 2002. 300 Jahre Brachistochronenproblem. In: *Medium Mathematik* (Proceedings of an International Leibniz Conference at Altdorf in 1996). G. Löffladt and M. Toepell, eds. Hildesheim: Franzbecker 2002, pp. 76–99.

Thiele, R., 2003. Hilbert's Twenty-fourth Problem. *The American Mathematical Monthly* 110, 1–24.

Thiele, R., 2003. Hilbert und Hamburg. *Mitteilungen der Mathematischen Gesellschaft in Hamburg* 22, 99–126.

Thiele, R. 2004. *Von der Bernoullischen Brachistochrone zum Kalibratorkonzept.* Turnhout: Brepols.

Thiele, R., 2004a. Variationsrechnung und Wirkungsprinzipien. In: *Formale Teleologie und Kausalität.* Proceedings of a Symposia held in Salzburg 2002, ed. by M. Stöltzner and P. Weingartner. Paderborn: Mentis Verlag. In print.

Thiele, R., 2004b. Van der Waerdens Leipziger Jahre. *Mitteilungen der DMV.* In print.

Van der Waerden, B.L. 1933. Bemerkungen zu Hilberts algebraischen Arbeiten. In: Hilbert 1932–1935, vol. 2, pp. 401–403.

Van der Waerden, B.L., 1936/37. Die Geometrographie von Lemoine. *Semester-Berichte zur Pflege des Zusammenhangs von Universität und Schule*, Mathematisches Seminar Münster. 11. Semester 1937/38, pp. 86–93.

Vinnikov, V., 1999. We Shall Know: Hilbert's Apology. *The Mathematical Intelligencer* 21 (1), 42–46.

Vollmer, G., 1993. Gelöste, ungelöste und unlösbare Probleme. In: G. Vollmer, *Wissenschaftstheorie im Einsatz.* Stuttgart: S. Hirzel 1993, pp. 183–210.

Weil, A., 1971. The Future of Mathematics. In: *Great Currents of Mathematical Thought* (ed. F. le Lionnais), New York: Dover, vol. 1. Translation of the French edition *Les Grands Courants de la Pensée Mathématique*, Paris 1962.

Weyl, H., 1944. David Hilbert and his Mathematical Work. *Bulletin of the American Mathematical Society* 50, 612–654.

Weyl, H. 1951/1968. A Half-century of Mathematics. *The American Mathematical Monthly* 58, 523–553; also in: *Gesammelte Abhandlungen*. Heidelberg: Springer 1968 (I quote this edition).

Yandell, B.H. 2002. *The Honors Class. Hilbert's Problems and Their Solvers*. Natick: Peters.

11

Turing and the Origins of AI*

Stuart Shanker

Distinguished Research Professor, Departments of Philosophy and Psychology,
Atkinson College, York University

11.1 Turing's Two Questions

The publication of *Mechanical Intelligence*[1] provides us with an invaluable opportunity to survey Turing's contributions to the origins of AI. Reprinted here are all of the major articles which Turing published on the Mechanist Thesis.[2] Reading through these papers gives us a chance to appreciate anew the sparkling freshness of Turing's thought, but also, to reassess the profound and in many ways problematic influence which Turing had on the foundations of AI. This exercise takes on a special significance in light of the fact that, in the past few years, some of the most influential figures in the cognitive revolution have repudiated AI (see Bruner 1990; Shanker 1992). *Mechanical Intelligence* enables us to see quite clearly, both why the fledgling cognitive revolution so quickly adopted AI as its paradigm, and the reasons for this inevitable rupture.

The book is dominated by Turing's celebrated 1950 paper, 'Computing machinery and intelligence', which in turn is dominated by the *philosophical* question that Turing asks at the outset: *'Can machines think?'* It is a question whose import cannot be divorced from the resounding success of Turing's version of Church's Thesis (*viz.*, that mechanically calculable functions are Turing-machine computable). Moreover, if judged by the amount of interest which it has aroused, this question surely stands unrivalled in post-war analytic philosophy. And yet, as far as the foundations of Cognitive Science are concerned, there is a definite sense in which this question places the emphasis on the wrong issue. Indeed, there is even a sense in which it places the emphasis on the wrong issue as far as Turing's own interests and contributions are

*First published in *Philosophia Mathematica* **3** (1) (1995), 52–85.

[1]Volume III (D. C. Ince, ed., 1992, ISBN 0-444-88058-5, pp. xix + 227) of the *Collected Works of A. M. Turing*, Amsterdam: North-Holland.

[2]The two most notable omissions are 'Intelligent Machinery: A Heretical View' and Turing's 1951 BBC lecture 'Can digital computers think?'.

concerned. For over and over again we find him returning to the *psychological* question: *Do thinkers compute?* This is a different, and in many ways, a much more significant matter.

These two questions belong to very different traditions. The former was a central concern of English mathematicians in the nineteenth century (*e.g.*, Babbage, Jevons and Marquand); the latter a mainstay of empiricist psychology in Germany, England, and America. But Turing not only regarded these two questions as intimately connected: in fact, he thought they were internally related—that in answering one you would *ipso facto* be answering the other. The result was a remarkable synthesis: not only did Turing succeed in merging recursive function theory and cognitive psychology, but within psychology itself, he brought together two distinct—and even hostile—schools of thought under the banner of *post-computational mechanism.* But while Turing's conception of automata may have been strikingly original, his approach to the analysis of thought pursues themes that can be traced back to the Greeks.

Admittedly, it is difficult to view the question 'Can machines think?' as anything other than a modern phenomenon (which, in philosophical terms, means post-Cartesian). But the question 'Do thinkers compute?' is another matter. The succession of mechanical metaphors of mind—*qua* hydraulic pipes, clock, telegraph system, telephone exchange, feedback circuit, serial and parallel computer—are part of a tradition that stems from a persisting picture of a *mechanist continuum* (see Shanker [forthcoming]). Locke's argument that 'in all the visible corporeal world we see no chasms or gaps. All quite down from us the descent is by easy steps, and a continued series that in each remove differ very little one from the other' (Locke 1690: III vi §12); Whytt's claim that 'in all the works of nature, there is a beautiful gradation, and a kind of link, as it were, betwixt each species of animals, the lowest of the immediately superior class, different little from the highest in the next succeeding order' (Fearing 1930: 75); Herrick's premise that there is 'an unbroken graded series from the lowest to the highest animal species' (*Ibid*: 179); and George's insistence that there is a cognitive continuum, 'with simple negative adaptation (habituation, or accommodation, and tropisms, which are orientating responses and are known to be mediated by fairly simple physico-chemical means) at one end, and maze-learning, puzzle-box learning... and ape-learning... in stages of increasing complexity, leading to human learning at the other end' (George 1962: 180), all spring from the same source as that which led to Aristotle's maxim in *De Generatione Animalium* that 'Nature orders generation in regular gradation' (Aristotle 1938: 186).

It is precisely this psychological issue, however, from which the question of whether machines can think inadvertently serves to deflect attention. Turing repeatedly insists that his sole concern is with 'the meaning of the words "machine" and "think"', and that it is his faith in 'semantic progress' which leads him to express his belief 'that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted'

(Turing 1950: 133, 142). Most of the papers he wrote towards the end of his life begin with a defence of 'machine intelligence'. His 1947 'Lecture to the London Mathematical Society', 'Intelligent Machinery' (1948), 'Computing Machinery and Intelligence' (1950), and his 1951 BBC lecture 'Can Digital Computers Think?' all begin by making the same point: Turing's repeated claim that he was 'not interested in developing a *powerful* brain... just a mediocre brain' (Hodges 1983: 251), and that 'if a machine is expected to be infallible, it cannot also be intelligent' (Turing 1947: 105).

The concept of *machine* had already undergone radical changes. At the beginning of the nineteenth century it had been confined to the static motions dictated by Newtonian mechanics, but by the 1870s it had evolved into the homeostatic systems conceived by Claude Bernard. These developments were essential to the transition in the mechanist/vitalist debate from the Life Sciences to the Human Sciences: from the question whether the body could be explained in mechanical terms to the question whether the mind could be so explained.

The problem faced by both physiological and psychological mechanists at the turn of the century was the same: it stemmed from the widespread doubt that machines would ever approximate the self-regulating adaptative and selective behaviour which characterizes physiological and mental phenomena. The key word here is 'ever', which signifies that the issue was regarded as empirical. The obvious solution would be, in G. H. Lewes's words, to 'think through the essentials of such a mechanism'. But 'An automaton that will learn by experience, and adapt itself to conditions not calculated for in its construction, has yet to be made; till it is made, we must deny that organisms are machines' (Lewes 1877: 436).

This is precisely the problem which was continuing to preoccupy and frustrate mechanists fifty years on; and indeed, might have remained beyond the reach of their ambitions had Turing not completed the mathematical transformation of the concept of *machine*. What Turing had proved in 'On Computable Numbers' is that an 'effective function' is an algorithm that can be so encoded (*e.g.*, in binary terms) as to be machine-executable. But for the advocates of strong AI, Turing had proved far more than this: what he had really accomplished was to transform machines into a species of 'rule-following beasts' (as Hofstadter describes computers). And the manner in which he achieved this feat was by postulating a category of meaningless (sub-)rules which could guide the operations of a machine (and/or the brain), thereby providing the rudiments for a new understanding of 'machine' and thence the creation of artificial intelligence (see Shanker 1987).

For almost three decades, philosophical discussions of Turing's contributions to the origins of AI centred on his preoccupation with strong AI, the nature of consciousness, and in particular, with the significance of the Turing Test. It is not difficult to account for the overwhelming response which these issues elicited. Turing's argument spoke directly to the anxieties of a society that had just lived through a terrifying war, only to find itself in a world that had not just been transformed, but was continuing to change at a rate never before experienced in ways that few had envisaged. Turing's was the bold and, to many, the reassuring voice of the new vanguard; but the message was as old as the Renaissance: technological advances cannot be halted, man must adapt to the inexorable march of progress. And so there followed a flurry of articles and books in which AI-theorists rhapsodized and humanists anguished, all of them mesmerized by the debate over man's position on the *Scala Naturae*. Archetypal issues clearly die hard.

With all of the rhetoric about the computational possibilities being opened up, or the singular phenomenological characteristics of human experience, it was easy to overlook the fact that, in order to defend his philosophical thesis—*viz.*, his proof that, if not quite yet, at some point in the future machines will indeed be capable of thought—Turing was led deeper and deeper into the development of an appropriate psychological theory: *viz.*, that thinkers do indeed compute. By the time he came to write 'Computing Machinery and Intelligence' he was explaining how his real goal was that of 'trying to imitate an adult human mind' (Turing 1950: 155). The result was a sublimely simple mechanist theory whose appeal lay in its claimed ability 'to resolve complex psychical phenomena into more elementary process'—a sentiment which, significantly, was expressed fourteen years before the publication of 'On Computable Numbers' (Rignano 1923).

Turing's psychological theory represents a marriage of *Denkpsychologie* and behaviourism. His basic idea is that thinking is an effective procedure (no doubt because the brain is a digital computer): *i.e.*, the mind proceeds, via an unbroken chain of mechanical steps, from α to ω, even though the subject himself may only be aware of α, δ, ξ, and ω. By mapping the subject's thought-processes onto a program designed to solve the same problem, we can thus fill in the intervening—*subconscious*—steps. This is the thesis underlying Turing's observation in 'Can digital computers think?' that 'The whole thinking process is still rather mysterious to us, but I believe that the attempt to make a thinking machine will help us greatly in finding out how we think ourselves' (Hodges 1983: 442).

To be sure, there is nothing particularly novel about this picture of the unconscious mind. As a matter of fact, Hadamard's *The Psychology of Mathematical Invention* is about little else; Hadamard cites story after story to establish how all the important creative work in mathematical discovery is unconscious (see Hadamard 1947). But Hadamard's book also vividly illustrates why Turing's psychological thesis was to have such a profound and widespread impact. For on traditional theories of the *cognitive unconscious*, the problem of 'insight' was simply shifted down a level, where any amount of extraordinary cognitive abilities that escaped the conscious mind could be attributed to the unconscious. But Turing's argument claims to take no such metaphysical step; for the processes occurring 'beneath the threshold of consciousness' are all said to be Turing-machine computable. But then, by placing the emphasis on the question of whether machines can think, Turing only manages to blur the crucial demarcation lines between materialism and mechanism: between breaking the machine down into its component parts, and using the machine as a psychological paradigm for understanding how the mind works.

Newell and Simon were perhaps the first to realize the significance of this point. Indeed, it was partly for that reason that they were initially opposed to the name 'artificial intelligence'.[3] For they were not interested in the philosophical question of whether or not machines can be said to think; rather, they wanted to place the emphasis firmly on psychological explanation. As they explained in 'GPS: A program which simulates human thought':

> We may then conceive of an intelligent program that manipulates symbols in the same way that our subject does—by taking as inputs the symbolic logic expressions, and producing as outputs a sequence of rule applications that coincides with the subject's. If we observed this program in operation, it would be considering various rules and evaluating various expressions, the same sorts of things we see expressed in the protocol of the subject. If the fit of such a program were close enough to the overt behaviour of our human subject—*i.e.*, to the protocol—then it would constitute a good theory of the subject's problem-solving (Newell & Simon 1961: 283).

It is easy to see how this could have been read as an expression of strong AI: if the computer can solve the same problems as man—and what's more, do so in exactly the same steps—then it would have satisfied the demands for the attribution of 'intelligence'. But the psychological significance of the argument is contained in the last line: in the claim that such a program would constitute a *theory* of problem-solving. The reasoning here is straightforward: given that thinking is an effective procedure, then, if a program simulates a subject's 'overt behaviour'—both what the agent did and the 'fragments' of his thinking-process which he observed and reported—that program will explain the 'hidden processes' that brought about that subject's behaviour.

The essential details of Newell and Simon's argument had already been sketched by Turing. This is the reason why we find Turing spending so much time during this period on the development of chess programs. Chess served as the ideal medium for the computational explanation/simulation of those processes which a subject('s mind) employs when solving problems, where this was seen as a matter of grasping a problem, tentatively experimenting with various methods of solution, selecting a favoured route, and then verifying the success of the chosen means. What made chess so suitable for this purpose was the fact that the decision-making procedures leading up to a move are highly amenable to hierarchical recursive analysis: a player thinks through a problem by shuttling back and forth between sub- and primary-goals as he

[3]In fact, they refused to use the term for several years, preferring instead to describe their work as 'complex information processing'.

first formulates a plan, calculates the permutations, deduces whether any of these will implement the strategy he has in mind, and if none of these succeeds in realizing the overall goal, formulates a new plan and repeats the process.[4] Most important of all was the fact that chess afforded the perfect environment in which to map out the goal-directed pattern of the mechanical procedures that constitute problem solving in a fixed domain: given the prior assumption that these are moves within a calculus. The resulting models may of course be far removed from anything one might encounter in ordinary chess play; but the fact that we may not be aware of such processes is merely a confirmation of the limits of introspection.

Significantly, Turing was not the first to seize on chess as a paradigm for studying thought processes. Interest in the psychology of chess at the turn of the century was largely confined to the study of prodigies. The subjects were primarily masters, and the methodology strictly introspectionist (see Binet 1894, Cleveland 1907). The shift in attitudes towards the psychology of chess had begun prior to and entirely independent of Turing: largely as a result of Selz's work on the 'laws of cognitive activity'. Proceeding on the basis of the Würzburg School's approach to the psychology of thinking, Selz undertook to confirm the existence of directed associations (*i.e.*, of the manner in which an *Aufgabe* influences a subject's response to a stimulus). Where Selz's particular importance for the foundations of AI lies is in his insistence that

> The individual analysis of task-conditioned thought processes always shows an uninterrupted chain of both general and specific partial operations which at times cumulatively ($A + B + C$) and at times in a stepwise fashion (B after failure of A) impel the solution of the task. These operations are continued until a solution is found or up to a momentary or lasting renunciation of the solution (Simon 1982: 153).

Although Selz's writings were relatively unknown amongst English speaking psychologists, the work of one of his followers, Adriaan De Groot, had an immediate impact on the evolution of GPS.[5] De Groot sought to implement Selz's ideas in an exhaustive investigation of how a broad spectrum of chess players set about solving board problems. His primary result (presupposition?) was that, as Selz had outlined, such problem-solving processes

[4] This theme was clearly expressed by Shannon in 'A Chess-playing Machine': The thinking process is considered by some psychologists to be essentially characterized by the following steps: various possible solutions of a problem are tried out mentally or symbolically without actually being carried out physically; the best solution is selected by a mental evaluation of the results of these trials; and the solution found in this way is then acted upon. It will be seen that this is almost an exact description of how a chess-playing computer operates, provided we substitute 'within the machine' for 'mentally' (Shannon 1956: 2132–3).

[5] Although *Thought and Choice in Chess* was not translated into English until 1965, Simon recounts how, in 1954, he had taught himself Dutch solely in order to read it (see Simon 1982: 149).

must be based on a linear chain of operations: a point which, as De Groot noted in the Epilogue to the later English translation of *Thought and Choice in Chess*, rendered his findings highly compatible with the requirements of chess programming.

The picture of thinking which guides De Groot is the same as what we find in *Mechanical Intelligence*: it postulates that we cannot—where the 'cannot' is thought to be psychological—hope to capture the full range of our thoughts in the net of language, either because so much of the thinking process is subliminal, too rapid or too far removed for our powers of introspection, or simply of a nature that outstrips the present possibilities of linguistic expression. With training it might be possible to ameliorate some of these deficiencies, but no amount of laboratory experience can enable a subject to discern the 'elementary information processes' out of which human problem solving is compounded. Computer models thus provide the cognitive psychologist, not just with a valuable, but in fact, with an essential adjunct to thinking-aloud experiments. For without them we could never hope to overcome the inherent limitations of introspection. Moreover, it is difficult to see how we could otherwise hope to explain such phenomena as 'moments of illumination' or the mind's ability to solve problems of enormous computational complexity in what, from the vantage point of current technology, seems like an astonishingly small amount of time.

It is important to remember that De Groot laid the foundation for this argument long before he became familiar with computer models of chess thinking. At the outset of his argument he approvingly cites Selz's dictum that the psychologist's goal must be to deliver "a complete (literally: 'gapless') description of the causal connections that govern the total course of intellectual and/or motor processes" in problem solving (De Groot 1965: 13). De Groot's major task was then to explain the phenomenon of pauses in a player's reports followed by a new approach; for 'It is often during these very pauses that the most important problem transformations appear: the subject takes a "fresh look" at the entire problem' (184). Given that a 'subject's thinking is considered one continuous activity that can be described as a *linear chain of operations*' (54), 'transitional phases have to be assumed in order to understand the progress of the thought process even though the written protocol gives no indication at all' (113). Thus, we can hypothesize from the program what *must* have been going on during the pauses in the player's mind. For there can be no such *lacunae* in the computer model: a program with such 'gaps' would literally grind to a halt. And the same must hold true for thinking itself, *given that it too is an effective procedure.*

Reading through *Mechanical Intelligence*, one begins to appreciate just how propitious Turing's timing was. If Turing's major accomplishment in 'On Computable Numbers' was to expose the epistemological premises built into formalism, his main achievement in the 1940s was to recognize the extent to which this outlook both harmonised with and extended contemporary psychological thought. Turing sought to synthesize these diverse mathematical and psychological elements so as to forge a union between ‘embodied rules’ and ‘learning programs’. Through their joint service in the Mechanist Thesis each would validate the other, and the frameworks from whence each derived. What is all too often overlooked by AI theorists, however, is that by providing the computational means for overcoming the impasse in which mechanism found itself before the war, Turing was committed to the very framework—as defined by its set of assumptions—which had created it. As important as Turing’s version of Church’s Thesis was for the foundations of AI, no less significant was the psychological thesis which provided the means for the transformation of Turing’s ‘slave machines’ into ‘intelligent automatons’. It is to the latter that we must look, therefore, in order to understand, not simply the genesis, but more importantly, the presuppositions of AI. For it suggests that the gulf between pre- and post-computational mechanism may not be nearly so great as has commonly been assumed.

11.2 Turing’s Behaviourist Ambitions

The main reason why it is so tempting to speak of Turing’s ‘computational revolution’ is because of the contrast between behaviourist and AI attitudes towards the role of ‘mental states’ in the explanation of actions. But before we accept the radical divergence commonly postulated between ‘pre-’ and ‘post-computational’ mechanism, we should consider the extent to which Turing saw himself as working within the framework of behaviourism: as taking the behaviourist account of problem solving a step further by treating ‘mental states’ as ‘machine-state configurations’, thereby allowing for a reductionist explanation of such higher-level activities as chess playing and theorem proving. This enables us to see that the route leading from Huxley’s ‘sentient automatons’ through Jevons and Marquand’s ‘reasoning machines’ to Hull’s and Turing’s ‘learning systems’ displays far more continuity than is commonly acknowledged.

In order to understand the significance of Turing’s contribution to the evolution of behaviourism, it is important to be aware that the thought-experiment machines portrayed in ‘On Computable Numbers’ are not credited with cognitive abilities as such; on the contrary, they are explicitly referred to as ‘devoid of intelligence’. The routines that they execute are described as ‘brute force’: a reminder, not just of the repetitive strategy they use to solve computational problems, but also, that they belong to the intellectual level of the brutes—with all the Cartesian overtones which this carries. In Turing’s words, these machines ‘should be treated as entirely without intelligence’; but, he continues, ‘There are indications. . . that it is possible to make the machine display intelligence at the risk of its making occasional serious mistakes’ (Turing 1946: 41). Just as we can say of a student exposed to ‘teachers [who] have been intentionally trying to modify’ his behaviour that, ‘at the end of the period a large number of standard routines will have been superimposed on the original pattern of his brain', so too, 'by applying appropriate interference, mimicking education, we should hope to modify the machine until it could be relied on to produce definite reactions to certain commands' (Turing 1959: 14).

The key to accomplishing this feat lies in the introduction of 'learning programs': self-modifying algorithms that revise their rules in order to improve the range and sophistication of the tasks they can execute, thereby satisfying Lewes's demand for an automaton capable of adapting to conditions not calculated for in its construction. The obvious problem which this argument raises, however, is whether, or in what sense, such programs can be described as 'learning'; and should this be deemed inappropriate, how Turing could have assumed that 'self-modifying' means the same thing as 'learning'. But before this issue can be explored, there lies the prior question of why Turing should have seized on this particular notion in order to implement his mechanist ideas.

There are several reasons why *learning* assumed such importance in mechanist thought *vis-à-vis* both of the 'traditions' outlined in §1. In mythopaeic terms, the automaton only springs to life once it displays the ability to recognize and master its environment (at which point humanist anxieties invariably surface in the form of the creator's loss of control over this now autonomous creature). In both physiological and psychological terms, questions about the nature of learning dominated the mechanist/vitalist debates during the nineteenth century. And in terms of the history of AI, the first and in some ways most potent objection raised against the Mechanist Thesis was voiced nearly a century before the invention of computers.

In her Notes on Menabrea's 'Sketch', Ada Lovelace cautioned that Babbage's 'Analytical Engine has no pretensions whatever to *originate* anything. It can do whatever we *know how to order it* to perform. It can *follow* analysis; but it has no power of *anticipating* any analytical relations or truths. Its province is to assist us in making available what we are already acquainted with' (Lovelace 1842: 284).[6] As he makes clear in 'Computing Machinery and Intelligence', the crux of Turing's version of the Mechanist Thesis turns on the very premise which Lovelace denies in this passage. 'Who can be certain,' Turing asks, 'that "original work" that he has done was not simply the growth of the seed planted in him by teaching, or the effect of following well known general principles' (Turing 1950: 150). The important point is that, granted that the operations of a machine can be guided by rules (however simple these might be), it should be possible to develop programs of sufficient complexity to warrant the attribution of intelligence. It was this argument which was to have so dramatic an effect upon mechanist thought. For Turing was to insist that the essence of a 'learning program' is its ability to simulate the creative aspect of human learning (see Turing 1947: 103-4).

[6]Peirce was to develop this theme in several important papers on 'logical machines'. I am indebted to Kenneth Ketner for drawing this to my attention.

To serve as a defence of machine intelligence, this argument must assume that *learning* 'denotes changes in the system that are adaptive in the sense that they enable the system to do the same task or tasks drawn from the same population more efficiently and more effectively the next time' (Simon 1983: 28). On first reading, this statement looks like little more than a strained attempt to tailor the concept of *learning* so as to mesh with that of *mechanical rules*. For if all that learning amounted to were the adaptation of something to its environment we should be forced to conclude, not just that machines, but indeed, the simplest of organisms is capable of some primitive form of learning. But far from seeing this as an *objection*, the mechanist will respond: Exactly, that is the whole point of the theory!

One cannot simply assume, therefore, that Turing only succeeded in subverting the concept of learning in his zeal to reduce it to a level commensurate with the minimal 'cognitive abilities' possessed by his machines. For such a charge would fail to do justice to the manner in which AI evolved from the union of mathematical and mechanist thought, and the extent to which the latter had come to dominate learning theory. Moreover, it would ignore the conceptual evolution of *machine* which underpins this outcome, and the bearing which this had, not just on Turing's psychological programme, but as a result of his influence, on automata theory and thence AI. But most serious of all, it would obscure the extent to which behaviourist presuppositions were absorbed into the fabric of AI.

This behaviourist orientation is particularly evident in 'Intelligent Machinery', the report which Turing completed for the National Physical Laboratory in the summer of 1948. The purpose of this paper was to defend the claim that self-modifying algorithms can legitimately be described as 'learning programs'. The opening premise recalls Pavlov and Lashley's theory that what we call 'learning' is the result of new neural pathways brought about by conditioning (see Pavlov 1927: 4ff). According to Turing, 'the cortex of the infant is an unorganized machine, which can be organized by suitable interfering training' (Turing 1948: 120). By enabling the system to modify its own rules, Turing thought he had demonstrated how Turing Machines could in principle simulate the formation of neural reflex arcs that take place during conditioning. The ensuing argument then expands on this notion of conditioning in terms of the 'Spread of Effect' experiments inspired by Thorndike.

In Turing's eyes, the most important element in this behaviourist theory is that *learning* consists in neural stimulus-response connections. He takes it as given that 'The training of the human child depends largely on a system of rewards and punishments.... Pleasure interference has a tendency to fix the character, *i.e.*, towards preventing it changing, whereas pain stimuli tend to disrupt the character, causing features which had become fixed to change, or to become again subject to random variation'. Accordingly, 'It is intended that pain stimuli occur when the machine's behaviour is wrong, pleasure stimuli when it is particularly right. With appropriate stimuli on these lines, judiciously operated by the "teacher", one may hope that the "character" will converge towards the one desired, *i.e.*, that wrong behaviour will tend to become rare' (*Ibid*: 121).

The concepts of extinction and positive reinforcement on which Turing placed so much emphasis in his 'learning'-based version of the Mechanist Thesis were thus directly culled from behaviourist writings: it was by employing 'analogues' of pleasure and pain stimuli that he hoped 'to give the desired modification' to a machine's 'character' (*Ibid*: 124). As he put it in 'Intelligent Machinery, A Heretical Theory':

> Without some... idea, corresponding to the 'pleasure principle' of the psychologists, it is very difficult to see how to proceed. Certainly it would be most natural to introduce some such thing into the machine. I suggest that there should be two keys which can be manipulated by the schoolmaster, and which can represent the ideas of pleasure and pain. At later stages in education the machine would recognize certain other conditions as desirable owing to their having been constantly associated in the past with pleasure, and likewise certain others as undesirable (Turing 1959: 132).

The mechanist metaphor would now appear to be twice removed from the established meaning of 'learning'. Whereas behaviourists had taken the liberty of depicting *habituation* as a lower form of *learning*, Turing went a step further and added the premise that machines display 'behaviour' which as such can be 'conditioned' by 'analogues' of 'pleasure and pain stimuli' in what can reasonably be described as 'training'. Yet Turing was not alone in this move; at much the same time Hull was arguing that 'an automaton might be constructed on the analogy of the nervous system which could learn and through experience acquire a considerable degree of intelligence by just coming in contact with an environment' (Hull 1962: 820).

Far from being a coincidence, the affinity between Turing's and Hull's thinking is a consequence of their shared outlook towards the nature of problem-solving. This thinking is exemplified–and was to some extent inspired–by Thorndike's experiments on the 'learning curve'. Thorndike designed a puzzle box to measure the number of times a cat placed inside would randomly pull on chains and levers to escape. He found that when practice days were plotted against the amount of time required to free itself, a 'learning curve' emerged which fell rapidly at first and then gradually until it approached a horizontal line which signified the point at which the cat had 'mastered the task'. According to Thorndike, his results showed how animal learning at its most basic level breaks down into a series of brute repetitions which gradually 'stamp' the correct response into the animal's behaviour by creating 'neuro-causal connections'. But repetition alone does not suffice for the reinforcement of these connections; without the concomitant effects produced by punishment and reward new connections would not be stamped in.

One's immediate reaction to Thorndike's experiment might be that this manifestly constitutes an example of anything but *learning*. But to appreciate the full force of Thorndike's argument, you have to imagine that you were shown a cat, that had already been conditioned, quickly freeing itself from a puzzle box. Let us suppose that one's natural inclination here would be to say that the cat had clearly learnt how to free itself. The whole point of the experiment is to show us, in an artificial condition, exactly what processes had led up to this outcome. To speak of 'insight' here is completely vacuous; and so too, Thorndike wants us to conclude, must it be in all other cases. That is, we only speak of 'insight' when we are unfamiliar with the processes that have led up to the results we have witnessed.

Herein lies the thrust of the behaviourist version of the *continuum picture* outlined in §1. The crucial point is the idea that *learning* consists in the formation of stimulus-response connections which themselves require no intelligence. A plant that turns its leaves towards the sun can be said to have *learned* how to maximize its photosynthesis; a dog that is conditioned to salivate at the sound of a bell can be said to have *learned* that it is about to be fed. The 'higher' forms of learning—*e.g.*, learning how to speak, how to count, how to solve logical problems—are distinguished from these lower forms by the complexity of the stimulus-response connections forged in the organism's brain. But the mechanical nature of the atomic associations which form the basis for all levels of learning remain identical. This provides the rationale for describing what had hitherto been regarded as disparate phenomena—as reflexive *as opposed to* cognitive phenomena—as constituting a *learning continuum* ranging from simple negative adaptation, habituation, accommodation, and tropisms, through animal and infant learning, to the highest reaches of education and scholarship.

Turing's Thesis was thus tailor-made for behaviourism. For as we saw in §1, the epistemological significance of his computational version of Church's Thesis was said to consist in the demonstration that algorithms are complex systems of meaningless sub-rules each of which can as such be applied purely mechanically. The essence of Turing's version of strong AI is that *machine intelligence* is a function of the complexity of the program which the computer follows rather than the individual steps of the algorithm. The difference between 'slave' and 'learning' programs lies in the shift from fixed to self-modifying algorithms. In the former, the Turing Machine repeatedly performs the same elementary steps; in the latter it alters its program, using heuristics which enable it to augment its knowledge base and/or store of rules, and thence the range and sophistication of the tasks it can execute.

It is this argument which, according to the standard mechanist interpretation of Turing's Thesis, enables us 'to face the fact that a "human computer" does need intelligence—to follow rules formulated in a language he must understand' (Webb 1980: 220). In order to provide a 'non-question begging' analysis of computation, 'the smallest, or most fundamental, or least sophisticated parts must not be supposed to perform tasks or follow procedures requiring intelligence' (Dennett 1978: 83). Thus, 'Turing's analysis [of computation] succeeded just because it *circumvented* the problem of how a computer can understand ordinary language' (Webb 1980: 225). Without 'meanings' to deal with, 'these atomic tasks presuppose no intelligence', from which 'it follows that a non-circular psychology of computation is possible' (*Ibid*: 220). But not just a psychology of computation; rather, a psychology of thinking *simpliciter*. For provided the lowest level of the 'learning continuum' can be simulated, there is no *a priori* reason why machines should not be capable of ascending this cognitive hierarchy. And this was exactly the theme which Turing exploited in the 1940s in his defence of the Mechanist Thesis. In so doing he saw himself as providing the crucial definition of 'mechanical' on which the behaviourist theory of learning rests, and as a result, opening the door to a future populated with thinking machines.

11.3 The Continuum Picture

One of the major philosophical questions that emerges as one reads through *Mechanical Intelligence* is whether, by forging a union between recursive function theory and psychology, and within psychology, between *Denkpsychologie* and behaviourism, Turing was surmounting or subsuming the conceptual problems that afflicted the latter, and to what extent this can be said to have impinged on AI. The mechanist picture of a *learning continuum* which Turing embraces is putatively one in which the higher forms of learning are built up out of simpler components. But in actual fact, the theory proceeds in the opposite direction: it is only by presupposing that the network of learning concepts can be applied in an ever-diminishing state to the descending levels of the 'continuum' that the converse reductionist analysis of these same concepts can then be instituted.

The key to this move lies in treating the criteria—what an agent says or does—that govern the application of *learning* as evidence of some hidden transformation in an organism's CNS. That is, being able to multiply sums correctly is a criterion for saying 'S has learnt the multiplication tables'. But the mechanist continuum presupposes that being able to multiply sums correctly is only evidence of the formation of the synaptic connections that constitute *learning how to multiply*. Hence, an organism undergoing cellular division also provides us with evidence that it has learnt how to multiply (albeit at a much lower stage on the multiplication continuum).

The criteria governing the application of *learning* can thus be hived off from declining applications of 'learning' until the atomic level is reached, where learning is said to be purely a function of assimilation and accommodation to stimuli. In other words, the theory postulates that it makes sense to speak of 'learning' in the absence of what we would normally regard as rule-following behaviour: *i.e.*, in the absence of the very possibility of explaining, teaching, correcting, or justifying the use of a concept. Given that a plant, or a dog, no less than a child, can respond to discriminably different stimuli, a child's ability to learn how to speak, to count, to apply colour or shape or psychological concepts, must simply be the end result of a more complicated cerebral nexus of exactly the same sort as guides the plant's or the dog's behaviour. (And note that it does indeed make sense on this picture to speak of the plant's 'behaviour'; to resist this neologism would be to court 'semantic conservatism'.) The prior question of whether, or in what sense, we can speak of a plant or a dog as *learning how to distinguish* between discriminably different stimuli—as opposed to *reacting* to these stimuli—is an issue which the continuum picture dismisses with the simple expedient of placing problematic uses of the concept in question in inverted commas.

'On Computable Numbers' is filled with such inverted commas. Turing tells us that a Turing Machine is 'so to speak, "directly aware"' of the scanned symbol. 'By altering its m-configuration the machine can effectively remember [the inverted commas are missing here—SS] some of the symbols which it has "seen" (scanned) previously' (Turing 1936: 117). Here Turing conforms to the precedent established by Jacques Loeb, who demonstrated that when Porthesia Chrysorrhoea are exposed to a source of light coming from the opposite direction to a supply of food, they invariably move towards the former (and perish as a result). In Loeb's words, 'Heliotropic animals are... in reality photometric machines': not that far removed, as it happens, from Turing Machines (Loeb 1912: 41). Loeb designed his experiments to undermine the vitalist thesis that all creatures are governed by an unanalyzable instinct for self-preservation. His conclusion was that, 'In this instance the light is the "will" of the animal which determines the direction of its movement, just as it is gravity in the case of a falling stone or the movement of a planet" (*Ibid*: 40–1). Since it is possible to explain, 'on a purely physicochemical basis', a group of 'animal reactions... which the metaphysician would classify under the term of animal "will"', the answer to no less than the 'riddle of life'—*viz.*, the nature of free will—must lie in the fact that 'We eat, drink, and reproduce not because mankind has reached an agreement that this is desirable, but because, machine-like, we are compelled to do so' (*Ibid*: 35, 33).

Loeb treats tropisms as quantitatively, not qualitatively, different from the mastery of a concept: they are simply a more primitive form of 'equivalence response'. For on the continuum picture, concepts are 'the mediating linkage between the input side (stimuli) and the output side (response). In operating as a system of ordering, a concept may be viewed as a categorical schema, an intervening medium, or program through which impinging stimuli are coded, passed, or evaluated on their way to response evocation' (Harvey, Hunt, Schroder 1961: 1). To be sure, on the folk theoretical level of psychological discourse, to say that 'S possesses the concept ϕ' may be to say that S has mastered the rules for the use of 'ϕ'; but this is seen as nothing more than the name of a problem, and psychology's task is to explain in satisfactory casual terms what this 'normativity' consists in.

The reason why the mechanist has no qualms about, *e.g.*, describing caterpillars as 'learning' where the light was coming from is because of his fundamental assumption that when we describe an organism as 'acquiring knowledge', 'we suppose that the organism had some specific experience which caused or was in some way related to the change in its knowledge state' (Bower & Hilgard 1981: 13). And this is where Turing's claim that 'To each state of mind of the [human] computer corresponds an "*m*-configuration" of the machine' (Turing 1936: 137) stepped in. With the benefit of this reduction of *mental* to *machine states*, and the step-by-step interaction of these 'internal configurations' with external input, AI was in a position to reintroduce the various cognitive concepts, hierarchically arranged, at each successive stage on the phylogenetic continuum.

Turing could thus defend his speaking of a chess program as 'learning' from its past 'mistakes' on the grounds that this use of 'learning' is no different from the sense in which Loeb, Thorndike, and Pavlov employed the term. What we have to remember here is the reductionist *animus* of these early behaviourists. Thorndike and Pavlov were not guilty of inadvertently failing to distinguish between *learning* and *conditioning*. The whole point of their theories is that what we refer to in ordinary language as 'animal learning' is in fact merely a species of conditioning: *i.e.*, that an animal's behaviour can be explained without appealing to any 'mentalist' concepts. And whatever is true of their behaviour applies, in virtue of the continuum picture, to human learning.

It is precisely this reductionism—as it applies to the explanation of human behaviour—which Bruner is reacting to in his repudiation of AI (see Bruner 1990: 4ff). But if one is to challenge the eliminative results of the continuum picture, then it is the picture itself that must be scrutinized (see Shanker 1992, 1993a). For it is not AI *per se* that is the problem here, it is the epistemological framework which underpins the succession of mechanist models described in §1: a framework which sees adaptation and accommodation as simpler versions of the same neural phenomena that constitute human learning.

What if we should approach this issue from the opposite starting-point? What if we should take as the paradigm for applying the concept of *learning* the mastery of the rules governing the use of a concept? The application of 'learning' in more primitive contexts would then be limited by the extent to which these resemble standard normative practices: *i.e.*, that doing such-and-such in appropriate circumstances constitute criteria for what we call 'learning how to ϕ'. In some cases, these criteria can be relaxed to a point that allows us to speak of 'primitive' applications of 'learning how to ϕ'. For example, the very concept of *language* is such that it makes sense to speak of 'primitive linguistic practices' (see Shanker 1994).

The 'continuum' which results from this perspective is *grammatical*, not mechanical. The question of whether the behaviour of any particular organism can be described as a primitive form of 'learning' depends on whether its actions can be described as benefitting from training or experience, as responding to correction and instruction, as acquiring simple and complex skills. Has Kanzi *learned* the meaning of 'bubbles': can he use and respond to the use of words to initiate or participate in the appropriate interindividual routines (see Savage-Rumbaugh *et al.* 1993)? Did Pavlov's dogs *learn* that food was imminent when a bell sounded? By extinguishing a conditioned stimulus did Pavlov *teach* a dog something else, or did he demonstrate how misleading it is in such circumstances to describe the dog as having *learned* something? Is the ability of lichen to survive in the most inhospitable of environments a *skill*? By correcting a program have we *trained* it to ϕ? Is this really comparable, as Turing so readily assumes, to the use of punishment and rewards in the training of a human child?

In all of these cases, the issue is decided by clarifying the logical grammar of *learning*. Typically, it is a subject's ability to respond appropriately to the use of 'ϕ', to master the rules for the use of 'ϕ', and to explain the meaning of 'ϕ' (even if only by ostension) that licenses our describing him as having 'learned how to ϕ'. These rules can clearly be stretched to a point that allows us to speak of primitive instances, but not so far as to regard any form of causal regularity (*e.g.*, a tropism, or a thermostat) as displaying evidence of having *learned* x. It is precisely because it makes no sense to speak of a dog that salivates at the sound of a bell as responding 'correctly' or 'incorrectly' to the signal that it makes no sense to try to reduce *learning* to the same terms. A dog which, bell or no bell, incessantly salivates, has no more made a *mistake* than has an electric door that fails to open when its photo-electric beam is crossed.

The central issue underlying all these questions is whether the behaviourists established what *learning* consists in at the primitive level, or rather demonstrated that the primitive behavioral responses that they studied should be explained in mechanical as opposed to normative terms. Indeed, were they really seeking to explain the concept of *learning*, or were they simply attacking its indiscriminate use by comparative psychologists at the turn of the century? By conditioning an animal to ϕ, and then extinguishing that conditioned response, they showed how inappropriate it would be to describe such behaviour in normative terms. But that hardly entails that there are *no* circumstances in which it makes sense to describe an animal as learning something: that the concept of animal learning—indeed, *learning simpliciter*—reduces to the concept of conditioning. If anything, the upshot of their experiments is that they sharpened the criteria for speaking of 'conditioning'. But that only serves to reinforce the categorial distinction between *learning* and *responding to a stimulus* by reminding us that the difference between a tropism and, *e.g.*, learning how to count is one of kind, not degree, insofar as it only makes sense to speak of explaining the latter in normative terms—*viz.*, of mastering the rules for the use of the concept—but a tropism or a conditioned response can only be explained in terms of internal mechanisms.

That is certainly not to deny that reflexes shade into skills, but instead, to see this observation as drawing attention to the grammatical continua which frame any psychological investigations into the possible existence of a phylogenetic/ontogenetic continuum: *scalae linguisticae* which range from primitive to rarefied uses of psychological concept words. Adult human beings serve as the paradigm subjects for our theoretical discussions about the use of psychological concepts: any question about the psychological capacities of animals or infants demands that we compare their behaviour with the relevant adult human actions which we regard as licensing the use of the concepts in question. These grammatical continua are counterpoised against behaviour and demand ever more complex actions to license the attribution of ever more complex cognitive skills and abilities.

It is on the basis of the grammatical continuum governing the application of *learning*, therefore, that we question whether it makes any sense to describe Turing's machines as 'learning how to play chess': *i.e.*, the question is whether there are any justificational criteria—whether there are sufficient grounds in the machine's behaviour—for the attribution of 'learning' to get a foothold. That does not mean that we should expect there to be a hard and fast line between 'reacting' and 'learning' in all contexts: between the use of causal and the use of normative terms. When studying a child's development, for example, it is often difficult to judge where exactly reacting ends and learning begins. But to speak of 'learning' nonetheless presupposes that it makes sense to describe the child as mastering the rules for the use of '*p*', as opposed to responding in a consistent manner to the sound of '*p*'. That is, if we are to make the transition from describing a child's (or an animal's) behaviour in causal to describing it in normative terms, his (or its) actions and reactions must warrant our speaking of him (or it) as *e.g.*, 'beginning to use or to respond to the use of "*p*" correctly'.

This argument illustrates the full gravity of the charge being levelled here that the continuum picture does not genuinely attempt to 'build up' from bottom to top but rather, proceeds in the opposite direction. By presupposing that *learning* can be treated as the same sort of process as *assimilation* or *accommodation*—that *habituation* constitutes a species of *cognition*—the mechanist has already assumed that *e.g.*, when caterpillars are attracted to a light this signifies a change in their 'knowledge state'. But is knowledge a state, much less one that is caused by external stimuli and whose inception and duration can be measured? Can knowledge be equated with a change in behaviour? After all, one can behave in a certain way without knowing what one is doing, or conversely, conceal what one knows. And is it really the case that 'we infer someone's knowledge from inputs to him and outputs from him, and we *infer* learning caused by an experience because of before-to-after changes in his inferred knowledge' (Bower & Hilgard 1981: 14)? Do we *infer* that a child who is able to answer all the adding and subtracting questions that we put to him knows how to count? Isn't this just what we call 'knowing how to count', and when a child can perform these sums we *see* that he has learnt how to count? Aren't these all criterial *as opposed to* evidential issues: grammatical as opposed to empirical questions?

Like *understanding*, to which it is so intimately connected, the concept of *learning* embraces a wide spectrum of activities that are loosely based on the development of skills, not 'cortical connections' (assuming that there even is such a thing). To learn how to speak is different from learning a second language; to learn how to tie one's shoes is different from learning how to play a Bach fugue. But this grammatical continuum, which progresses from simple to compound skills, from practice and drill to experience and understanding, does not 'descend' still further, to the point where we can treat the statements 'The thermostat clicked on', 'The leaves of the plant turned towards the sun', 'The pigeon pecked the yellow key in order to get a food pellet', 'Kanzi pressed the "drink" lexigram in order to get a drink', and '*S* has learned how to play the Toccata and Fugue in D minor' as all similar in kind: as distinguished only by their mounting internal complexity.

Just as our descriptions of reflexive movements shade into our descriptions of automatic behaviour (in the way *e.g.*, that red—or rather, crimson—can be said to shade into purple), so too causal descriptions of training (of reacting, associating, repeating or memorizing words) shade into normative descriptions of learning and understanding: of following, appealing to, explaining, teaching the rules for the use of words. Different kinds of training lead into different kinds of normative practice (*e.g.*, how we are taught to count, to describe colours, or objects, or pains, or actions). The different kinds of practice result in different kinds of concepts, and to seek to reduce this heterogeneity to a common paradigm—*e.g.*, the so-called 'functional' definition of concepts—is to embrace the presuppositions which ultimately result in eliminative materialism or mechanist reductionism.

Behaviourism and AI are converse expressions of this point. Both proceed from the assumption that, since there is a continuum leading from reflexes to reactions to concept acquisition, the only way for psychology to explain the mechanics of learning is by having a uniform grammar of description: in the case of behaviourism, by reducing all explanations of behaviour to the terms that apply to reflexes, and in the case of AI, by reading in the attributes which apply to higher-level cognitive abilities and skills into the lowest levels of reflex actions (*e.g.*, the brain is said to make hypotheses, the nervous system to make inferences, the thermostat to possess knowledge and beliefs). This is why it is so often argued these days that AI is just a form of neobehaviourism; for the fact is that, although the AI scientist may approach learning theory from the opposite point of view from the behaviourist, he does so because he shares the same framework as the behaviourist. Thus, on both theories we end up with caterpillars and even thermostats that think (or perhaps one should say 'think').

This brings us back to the different problems involved in the philosophical question 'Can machines think?' and the psychological question 'Do thinkers compute?' To talk about the 'mechanics of learning' is to refer to the importance of drill, repetition, a systematic training, of allowing scope for creativity and inculcating intrinsic motivation. We are no more interested in neurophysiology when concerned with the mechanics of learning arithmetic than we are interested in kinesiology when we speak of mastering the mechanics of golf. It might be possible to design a program that scores the same success rates on a series of standardized tests as a young child, but could a program be designed 'to simulate' all those things that a child does when learning how to count: *e.g.*, to find the same problems difficult or easy, to make the same sorts of excuses, to have its attention wander? Merely designing a program that matches and even predicts the mistakes a child will make hardly satisfies our criteria for describing that program as having learned how to count. And it in no way *entails* that the operations of that program must shed light on how a child learns how to count.

The problem here has nothing to do with semantic conservatism. It lies rather in the fact that, far from explaining, the continuum picture only serves to undermine the normative foundation on which the concept of *learning* rests: the criteria which govern our application of the concept. This is what Turing's version of the Mechanist Thesis is all about. Turing's Thesis only works as a psychological programme—a 'model of how the mind works'—on the basis of the reductionism inherent in the continuum picture. But if the latter should be a product of unwarranted epistemological assumptions, what are we to say of the models which it yields? The problems that arise here have nothing to do with any technical 'shortcomings' of computationalism. They are concerned rather with the question of whether, because of its grounding in the continuum picture, AI distorts the nature of the cognitive phenomena which it seeks to explain. There is no need to delve into social or cultural psychology to see the force of this point. The core of Turing's argument—his 'analysis' of computation—amply demonstrates the manner in which the framework drives the theory, rather than the other way round.

11.4 Wittgenstein versus Turing on the Nature of Computation

On the argument being sketched here, the union which Turing forged between recursive function theory and behaviourism was grounded in an archetypal epistemological picture. Furthermore, it was this epistemological picture which served to establish AI as the paradigm for the burgeoning field of Cognitive Science. For the cognitive revolution was driven by the premise that 'Good correspondence between a formal model and a process—between theory and observables, for that matter—presupposes that the model will, by appropriate manipulation, yield descriptions (or predictions) of how behaviour will occur and will even suggest forms of behaviour to look for that have not yet been observed—that are merely possible' (Bruner 1959: 368). The big appeal of AI lay in the hope that 'perhaps the new science of programming will help free us from our tendency to force nature to imitate the models we have constructed for her' (Bruner 1960: 23).

Turing was responding, in effect, to a challenge that had been laid down by Kant, and which had preoccupied psychologists throughout the latter half of the nineteenth century. Kant had insisted that psychology could 'become nothing more than a systematic art'. It could 'never [be] a science proper', since it is 'merely empirical' (Kant 1970: 7). What he meant by this last remark is that, because psychology is based on evidence yielded by the 'inner sense', it could not (*pace* taxonomic botany) become a true science—*i.e.*, a science on the paradigm of mechanics—since the latter requires the mathematization of its subject matter. What Turing was doing in 'On Computable Numbers'—even if he did not become fully aware of the fact until about five years later—was attempting to lay the foundation for just such a *science* by showing how, through the use of recursive functions, mental processes could indeed be mathematized. More is involved here, however, than simply establishing the *bona fides* of psychology. More to the point, Turing was responding—using the tools which Hilbert had developed (see Detlefsen 1993)—to the epistemological problems that result from Kant's picture of the mind confronted with and forced to make sense of reality.

Kant saw his basic task as that of explaining the rule-governed manner in which the mind imposes order on the flux of sensations that it receives (see the opening pages of the *Logic*). The power of Turing's Thesis lay in the fact that he was able to analyse these rules in such a way as to dispell the air of metaphysical speculation surrounding Kant's argument. For Turing's 'embodied rules' seemed to open up the prospect of a mechanist explanation of the connection between perception (input) and behaviour (output), between forming and systematising concepts, mastering and being guided by rules, acquiring and exercising an ability, hearing a sound or seeing a mark and grasping its meaning, forming an intention to ϕ and ϕing. But these are classic epistemological, not psychological problems; and the thought-experiment whereby Turing sought their solution was no more a piece of psychology than the Transcendental Aesthetic.

Turing discusses the nature of Turing ('computing') Machines at two different places in 'On Computable Numbers': §§1–2 and §9. In the first instance he defines the terms of his thought-experiment; in the second he takes up his promise to defend these definitions. The crucial part of his argument is that

> The behaviour of the computer at any moment is determined by the symbols which he is observing, and his 'state of mind' at that moment.... Let us imagine the operations performed by the computer to be split up into 'simple operations' which are so elementary that it is not easy to imagine them further divided. Every such operation consists of some change of the physical system consisting of the computer and his tape. We know the state of the system if we know the sequence of symbols on the tape, which of these are observed by the computer... and the state of mind of the computer.... The simple operations must therefore include:
> (a) Changes of the symbol on one of the observed squares.
> (b) Changes of one of the squares observed to another square within L squares of one of the previously observed squares....
> The operation actually performed is determined... by the state of

> mind of the computer and the observed symbols. In particular, they determine the state of mind of the computer after the operation is carried out.
>
> We may now construct a machine to do the work of this computer. To each state of mind of the computer corresponds an 'm-configuration' of the machine... (Turing 1936: 136–7).

Davis remarks of this argument that 'What Turing did around 1936 was to give a cogent and complete logical analysis of the notion of "computation". ... Thus it was that although people have been computing for centuries, it has only been since 1936 that we have possessed a satisfactory answer to the question: "What is a computation?"' (Davis 1978: 241). Ignoring for the moment the validity of this claim, it is important to see how accurately Davis has represented the tenor of Turing's argument. It was never intended to be read as a *psychological thesis*. It is rather a *reductive conceptual analysis* whose epistemological significance is said to lie in the demonstration that algorithms can be defined as complex systems of meaningless sub-rules, each of which can as such be applied purely mechanically, from which Turing's psychological thesis is said to follow. Davis goes on to explain how

> Turing based his precise definition of computation on an analysis of what a human being actually does when he computes. Such a person is following a set of rules which must be carried out in a completely mechanical manner. Ingenuity may well be involved in setting up these rules so that a computation may be carried out efficiently, but once the rules are laid down, they must be carried out in a mercilessly exact way.

Even before one gets to the problems involved in this use of 'mechanical' (see Shanker 1987), it is important to consider the accuracy of this first line. Far from basing his 'precise definition of computation on an analysis of what a human being actually does when he computes', Turing based his 'analysis' on a Kantian epistemological picture of what the human being's mind is doing when he computes. Computation, according to Turing, is a cognitive process, large parts of which are hidden from introspection; we can only infer the nature of these pre-conscious operations from an agent's observed behaviour and self-observations. (Kant presents a similar conception of *inference* in the *Logic*.) Given that the phenomena to be explained in psychology are so strongly analogous to those with which physics deals, is it any wonder that both disciplines should be seeking for good correspondence between a formal model and a process—between theory and observables?

It is precisely because his thought-experiment represents an attempt at conceptual analysis that Turing's Thesis has been the source of so much philosophical discussion. And perhaps the most pertinent such investigation is that conducted by Wittgenstein. Interestingly, Turing's thought-experiment shares some affinities with Wittgenstein's use of artificial language-games to clarify a concept: particularly Wittgenstein's discussion of 'reading-machines' at

§§156ff in *Philosophical Investigations*, which bears an uncanny resemblance to Turing's argument.[7] Moreover, Turing's reductionist thesis exemplifies, not just the target which Wittgenstein is attacking, but the larger point that he is making at §§122–33 of the *Investigations* about the dangers of philosophical theorizing.

The hardest thing to get clear about in Wittgenstein's discussion of *reading* is its purpose. Wittgenstein introduces this theme in the midst of his investigation into the nature of understanding. If the discussion of *reading* is to help clarify the nature of *understanding* it obviously cannot presuppose understanding. Thus Wittgenstein dwells on reading at its most mechanical: *viz.*, 'reading out loud what is written or printed'. But this is not an early version of the call for a 'presuppositionless psychology' (see Dennett 1978). Wittgenstein was not looking for the 'atomic units' out of which understanding is 'composed'. Rather, he uses a form of primitive language-game in order to clarify some of the problems involved in reductionist analyses of psychological concepts.

The parallel between Wittgenstein's 'reading-machines' and Turing's computing machines is striking. *Reading* (*counting* at *PI* §§143ff) can serve as an example of a rule-governed procedure at its most mechanical. The case Wittgenstein wants to consider is where the agent 'function[s] as a mere reading-machine: I mean, read[s] aloud and correctly without attending to what he is reading.' Here is a possible key to Wittgenstein's remark at §1096 of *Remarks on the Philosophy of Psychology* (volume I): 'Turing's "Machines". These machines are *humans* who calculate.' That is, what Turing is doing in his thought-experiment is imagining people performing the most routine of calculating tasks *in order to analyse the concept of calculation.* But whereas Turing's goal was indeed 'to break calculation down into its elementary psychic units', Wittgenstein was looking to clarify the criteria which license us in speaking of possessing a cognitive ability at its most primitive level, and the bearing which this has on reductionism.

The target in Wittgenstein's discussion is the continuum picture. Wittgenstein distinguishes between three broad uses of the term 'read':

(1) Causal, as this applies to a machine reading signs;
(2) Primitive, as this applies to someone who does not understand a word of what he is reading;
(3) Pardigmatic uses of 'read' (*i.e.*, understanding what one is reading).

Wittgenstein argues that there are many situations in which it is difficult to distinguish between (2) and (3) (*e.g.*, a child learning how to read, a politician reading from a cue card). And now the interesting question is: does (or could) a

[7]The overlap appears to be coincidental, although that in no way diminishes its significance. Wittgenstein had already formulated his argument by 1934 (see Wittgenstein 1960: 117ff), and it seems unlikely that Turing had read this before he composed 'On Computable Numbers'.

similar ambiguity exist between (1) and (2)? Is the machine-use just a further extension of the primitive use—as the continuum picture stipulates—and if so, is this because the internal mechanisms guiding the machine's operations are a simpler version of the internal mechanisms guiding the organism? Could we therefore learn about the 'processes of reading' involved in (2) and (3) by studying the machine's processes?

Wittgenstein's response to this problem is to clarify the criteria which govern the application of 'reading' in primitive contexts. He looks at the case in which we would say that 'S read p' even though he had no idea as to the meaning of p. Even in a limiting case like this, we still demand some criterion to distinguish between 'S read p' and 'It only seemed as if S read p' (*e.g.*, 'The parrot read the sign' and 'It only seemed as if the parrot read the sign'). That is, no matter how primitive the context, we insist on being able to distinguish between *ϕing* and *seeming to ϕ*. The question then is, on what basis do we draw this distinction. Do (or could) neural considerations play any role here? Could our judgment that someone is reading be overturned by neurological evidence? And could the neurophysiologist tell us what the subject is doing when he is reading?

Wittgenstein makes the distinction between *reading* and seeming to *read* (pretending to read, memorizing the words, repeating sounds) as hazy as possible. We are at that indeterminate point in a pupil's training where no one—not the pupil, not his teacher—can say exactly when he started to read. The argument is intended to take up the point made in §149. The question Wittgenstein posed there is: what is the relation between what we call 'the possession of an ability' ('the acquisition of a concept') and what we call 'the exercise of that ability' ('application of that concept'). If we treat this relation as causal—*i.e.*, possession causes exercise—then we shall be drawn into treating *understanding* (*reading, calculating*) as a mental state or process, and thence, into the construction of 'hypothetical mind-models' to explain how that hidden state or process causes the agent's observed behaviour:

> [A] state of the mind in this sense is the state of a hypothetical mechanism, a mind model meant to explain the conscious mental phenomena. (Such things as unconscious or subconscious mental states are features of the mind *model*.) In this way also we can hardly help conceiving of memory as of a kind of storehouse. Note also how sure people are that to the ability to add or to multiply... there *must* correspond a peculiar state of the person's brain, although on the other hand they know next to nothing about such psycho-physiological correspondences. We regard these phenomena as manifestations of this mechanism, and their possibility is the particular construction of the mechanism itself (Wittgenstein 1960: 117–18).

The conclusion which Wittgenstein draws in §149 is that the pair of concepts conscious/unconscious has no bearing on the relation between possession/exercise of an ability (acquisition/application of a concept). Wittgenstein returns to this theme in §156. We look for the difference between possessing and not possessing the ability to read in 'a special conscious mental activity', and because we cannot find any unique phenomenon to distinguish between *reading* and *seeming to read*, we conclude that the difference must be unconscious: where this is construed as either a mental or a cerebral activity. ('If there is no difference in what they happen to be conscious of there must be one in the unconscious workings of their minds, or, again, in the brain' (*PI*: §156).) This, according to Wittgenstein, is a paradigmatic example of the 'metaphysical "must"'. For if we construe the relation between 'knowing how to read' and 'reading' as *causal*—*i.e.*, if we construe the *grammatical relation* between 'possessing the ability to ϕ' and 'ϕing' (*viz.*, that doing x, y, z constitute the criteria for what we call 'ϕing') as causal (*viz.*, the ability to ϕ consists in a state-configuration which causes a system to x, y, z)—then we shall be compelled to identify the 'two different mechanisms' which *must* distinguish reading from seeming to read.

In the *Brown Book* Wittgenstein remarks how the failure to observe this logical distinction proceeds from a picture of the agent as being 'guided by the signs'. From this it seems to follow that we can only understand the nature of the mental activity in which reading consists 'if we could look into the actual mechanism connecting seeing the signs with acting according to them. For we have a definite picture of what in a mechanism we should call certain parts being guided by others. In fact, the mechanism which immediately suggests itself... is a mechanism of the type of a pianola' (Wittgenstein 1960: 118). The problem with this picture has nothing to do with machines lacking consciousness, or with an illicit use of 'reads'. For, regardless of whether or not it makes any sense to speak of machines as being conscious, there is indeed a sense in which we can say that, *e.g.*, a pianola is 'guided by' the notes that it 'reads'. But then, is the pianola 'guided by' the notes that it 'reads' in the same way that a child is guided by the rules he has been taught, (*i.e.*, the way the child goes through these rules in his mind, checks to see if this instance applies, automatically repeats certain mnemonics, *etc.*)?[8]

[8] In §162 of the *Investigations* Wittgenstein advances a definition of 'reading' as 'deriving a reproduction from the original'. The argument is about whether the pupil's behaviour warrants saying that he derived his action from a rule we have taught him: *e.g.*, if he can be seen to check the rule in the process of reading a text, if he cites the rule to justify what he has read. But the 'definition' proposed in the first line of §162 is after something much stronger: *viz.*, you can only be said to be 'reading' if this intermediary 'inferential' process occurs, and if you are not conscious of this process then it *must* have occurred unconsciously. The thesis under attack here is that of 'secondarily automatic actions'. Since learning how to read demands that we be conscious of such intermediary processes, they must have become automated in the adult reader. Hence, we can deploy a protocol of the child's behaviour as the paradigm for the 'unconscious processes' that must occur in a skilled reader. Wittgenstein often touches on this theme in his discussions of 'calculating in the head': *e.g.*, we use the long-hand calculations that would be done

The continuum picture presupposes that these two uses of 'being guided' are indeed the same: but only because it presupposes that the sense in which it applies to the child can be *reduced* to the sense in which it applies to the machine. On the mechanist argument which Turing embraces, to say that the child learning how to read is guided by the rules he has been taught is to say that these rules have been 'embodied' in his brain. That is, the child is guided *in exactly the same way* as the machine is guided by its program. Only the differing complexity of their internal state-configurations accounts for their divergent reading abilities.

Nowhere does Wittgenstein suggest that the kind of use which Turing makes of inverted commas should be deemed illicit. For who made the philosopher the custodian of good grammar? Can we not understand what Turing is saying? The point is not that it makes no sense to speak of a Turing Machine as 'reading' or 'calculating'. It is that this machine use of 'reading' or 'calculating' is categorially different from the normative sense in which we use it for children and adults: even in its most primitive applications. *Reading* is indeed a family-resemblance concept: a family which ranges from the infant responding to flash cards to a philosophy class reading the *Investigations* or a navigator reading the stars. But when we speak, *e.g.*, of a scanner 'reading' a bar code, this is not a *still more primitive extension* of an infant reacting to signs.

The mechanist insists that the only difference between the scanner 'reading' a bar code and someone reading something mechanically lies in the operations of the internal mechanisms guiding their respective actions. Accordingly, it should be possible to design a system that would perfectly simulate—and thus explain—the operations of the mechanism guiding the human reading-machine. Wittgenstein responds: 'By calling certain creatures "reading machines" we meant only that they react in a particular way to seeing printed signs. No connection between seeing and reacting, no internal mechanism enters here' (Wittgenstein 1960: 121). Neurophysiological evidence has no bearing on the application of 'reading mechanically': can neither corroborate nor undermine its use. What misleads us into thinking otherwise is 'the idea of a mechanism that works in special media and so can explain special movements.... But what is of interest to us in reading can't be essentially something *internal*' (Wittgenstein 1974: 99). That is, 'The distinction between "inner" and "outer" does not interest us' when clarifying the concept of *reading* in paradigmatic or in primitive contexts. The criteria for describing an agent as 'reading' or as 'reading mechanically' lie in his behaviour and the situation in which this occurs: *i.e.*, we call acting thus-and-so in such-and-such circumstances 'reading' or 'reading mechanically'.

In §157 Wittgenstein returns to the theme that the description of behaviour in causal terms shades into the description of behaviour in normative

on paper as the paradigm for what *must have gone on in an agent's mind* when the answer to a problem suddenly occurs to him.

terms.[9] Causal descriptions of behaviour—of conditioned responses, associations, repeating sounds—merge into normative descriptions of behaviour: of reading or calculating (of being able to cite the rules one has followed in reading or calculating). But despite the graduated nature of applying psychological concepts in primitive contexts, the transition from 'reacting' to 'reading' involves a fundamental *categorial* shift: the terms that apply to conditioned responses do not carry over into the description of rule-following behaviour, even though it may at times be difficult to identify where 'reacting' ends off and 'reading' begins. The crucial point here is that this grammatical continuum does not license the shift to speaking of a mechanist continuum, such that the sole difference between machine uses, primitive uses, and paradigmatic uses, is one of degree, not of kind. We do indeed speak of an agent 'reading like a machine', but only when we want to signal his failure to attend to what he was reading, or to describe the manner in which he was reading (in a monotone, without pauses or inflections).

Wittgenstein does not present us with an empirical generalization here; rather, with the grammatical observation that 'it makes no sense' to ask a question like 'What was the first word S read?'. This distinguishes the two 'mechanical' uses of 'reading' described above: the primitive (2) and the machine use (1). For the question 'What was the first symbol the scanner read?' is the easiest thing in the world to answer. Here we have another useful reminder of the different uses of 'reading', and more importantly, of the fact that the relation between possession and exercise is not causal. If it were, then the question would make perfectly good sense in both normative and causal contexts.[10]

§158 ties this argument in to the attack on causal theories of meaning and understanding initiated at §6 (cf. Wittgenstein 1974: 190, 117–18). The mechanist misconstrues the grammatical proposition that we cannot speak of 'a first word in S's new state' (*i.e.*, 'it makes no sense') as an empirical claim; maybe, he suggests, this is only a temporary difficulty. But the mechanist is only drawn into this position because he is already captivated by the idea that the connection forged in training (between S's possessing the ability to ϕ and his ϕing) is causal rather than grammatical. And if the changes in S's behaviour are brought about by changes in his internal mechanism, then it must be possible to say or predict what he has read/learnt if only we could read the brain's secrets: *i.e.*, the changes occurring in 'learning'.[11]

[9] But note that not all causal uses imply a lack of understanding. Cf. 'He reacted with alarm when he saw the sign of the four'.

[10] Note also the subtle distinction drawn in the last line of §157 between 'change in his behaviour' and 'state-transformations'. If there were no difference between these two uses then 'change in behaviour' would be taken as a sign of an underlying change in the hypothetical internal mechanism.

[11] Note the picture of learning here. Cf. the argument presented at p. 118 of Wittgenstein 1960 with Turing's conception of 'learning programs'.

Wittgenstein's response to this argument is grounded in his remarks on psychophysical parallelism (see Shanker, 1993b). His concern here is with the source of the 'must' underpinning the mechanist argument: *viz.*, the misconception of the relation between possession and exercise of an ability (concept). Neurophysiological knowledge would not explain but rather presupposes independent criteria for using 'reading' ('calculating'). That is, we could only look for the neural events that take place during the exercise of an ability—*i.e.*, we can only speak of mapping neural events onto behavioural acts—if we have independent criteria for saying 'Now he is ϕing'. The mistake here is that of supposing that neurophysiological models of 'internal mechanisms' (based on whatever paradigm: be it that of dynamics, physiological, computational, connectionist) are essential to the 'analysis' of a psychological concept.

The standard mechanist response to this argument is to query: Can we not at least assume that it is possible to map what an agent does when he is reading something onto his neural processes? So much is involved in this seemingly innocent question. Every time I read the *Investigations* I understand something different: is the same thing happening in my brain every time I read the book, or something different? And what about everyone else's brain? Is it conceivable that different agents might experience different neural processes while reading the *Investigations*? If you answer yes, then does that mean that a concept must be defined according to the brain that is using it? And if you say no—as Turing does—then what is the source of the 'must' driving your conceptual analysis? But most important of all, even if everyone did experience the same neural processes when reading, say, §571 of the *Investigations*, would this constitute an *analysis* of what reading that passage consists in, or just a *correlation*?

Suppose a Turing Machine (or an 'android') were indeed able to 'pass' the Turing Test: this would only be the case on the basis of its satisfying the criteria that govern the application of 'reading' or 'calculating'—as Turing himself, and countless mechanists following in his footsteps, have endlessly insisted. The machine's internal operations would be completely irrelevant to this issue: could neither undermine nor corroborate such a judgment. But in that case, why should we suppose that these same operations could serve as a paradigm of what *reading* or *calculating* consist in? That is, how does Turing pass from his mechanist to his psychological thesis?

The answer lies in the use which he makes of the continuum picture: *i.e.*, in the presupposition that *learning* consists in the formation of synaptic connections. For Turing could then assume that the machine's 'behaviour' only satisfies our criteria for saying that it is 'calculating' because its internal operations are isomorphic with those guiding the human computer when he passes the Turing Test. Thus, given that the human computer's behaviour is the end result of mechanical processes, and that any machine can be simulated by a Turing Machine, it follows that mechanist and psychological theses are internally related: that machines can be said to think precisely because thinkers compute. But then, these very 'processes' said to be guiding the human subject are themselves the product of the assumption driving this argument: *viz.*, that the *criteria* governing the use of 'read' and 'calculate' can be transformed into *evidence* of (embodied) effective procedures (which as such must be isomorphic with those guiding the machine).

The picture operating here is that of *reading* or *calculating* as consisting in the mental rule-governed transformation of input data. The connections forged in teaching are these 'embodied' rules. In §163 Wittgenstein argues that this notion of 'rule-governed' is empty. These are 'hidden' rules: rules that are inaccessible to both observer and agent. Wittgenstein's point is not simply that any number of possible rules could be formulated to satisfy the required transformation: more importantly, it is that this indeterminacy can be stretched to a point where it is impossible to distinguish between rule-governed and random behaviour. His intention is not to force us into the sceptical conclusion that 'reading' or 'calculating' are not rule-governed procedures; rather, it is to emphasize that these rules, *qua* rules, must be public. Hence the question raised by the passage is: what misleads us into postulating these 'hidden rules'. And the answer is: the epistemological framework underpinning Turing's Thesis. The premise that there is a *gap* between input and action in the exercise of an ability which must be bridged by a series of internal operations.

§164 brings us to the heart of this issue. The passage ties in to the family-resemblance argument at §§65ff. Here the mechanist is accused of looking for the 'hidden essence' of 'deriving' or 'reading'. But the 'essence' of these concepts is that there is no 'hidden essence'. Hence the mechanist is guilty of trying to define what can only be explained. (Bear in mind that the point of these discussions is to clarify the concept of *understanding*, which is *a fortiori* a family-resemblance concept.) Here is the source of Davis' claim that Turing succeeded in presenting 'a cogent and complete logical analysis of the notion of "computation"'. *I.e.*, the idea that Turing succeeded in defining the essence of computation.

Instead of asking 'What does computation consist in?', we might begin our investigation of the nature of computation by considering 'What justifies our use of "calculating" in such-and-such contexts?'. Waismann has the latter question in mind when he asks in *Principles of Linguistic Philosophy*: 'What, then, is the difference between a causal connection and a transition in a calculus? What is the difference between a calculation formed by a machine, and that made by a person? Or else between a musical box playing a tune and a person playing it?' The answer lies in a statement of the categorial distinction between reasons and causes *vis-à-vis* the use of 'calculation' or 'reading':

> the playing or calculation of a person can be justified by rules which he gives us when asked; not so the achievements of a machine, where the question 'Why do these keys spring out?' can *only* be answered by describing the mechanism, that is by describing a *causal nexus*. On the other hand, if we ask the calculator how he comes to his results, he

> will explain to us what kind of calculation he is doing and then adduce certain laws of arithmetic. He will not reply by describing the mode of action of a hidden machine, say, a machine in his brain (Waismann 1965: 122).

The way that Waismann has phrased this argument invites the mechanist response that a machine could be programmed to perform exactly the same linguistic acts; but as we have just seen, such an objection is beside the point. For what matters here are the criteria for distinguishing between justification and causal explanation. That is, the important distinction here is between applying the rules which logically determine whether we can describe a subject as having *read* or *calculated* x, and explaining the causes (if there are any) of a subject or a machine saying or writing 'x'.

The idea that Turing succeeded in giving the first cogent and complete logical analysis of *computation* turns, not just on his success in resolving Church's Thesis, but at an even more fundamental level, on the persisting strength of Kant's epistemological picture of the hidden rule-governed processes whereby the mind makes sense of reality. It is precisely this epistemological picture which Wittgenstein's argument is designed to subvert. This is why he insists that it is 'necessary to look and see how we carry out [calculations] in the practice of language; what kind of procedure in the language-game [calculating] is' (Wittgenstein 1956: I §17). That is, to paraphrase what Waismann says about *concepts* at pp. 227–8 of *Principles of Linguistic Philosophy*, we need to recognize that 'what we call "calculating" comes into existence only by its incorporation in language; that it is recognizable not by one feature but by a number, which, as it were, constitute its facets'. Hence, to get clear about what *calculating* 'consists in' we need to clarify the rules governing the use of 'calculate'. This is the reason why Wittgenstein remarks, *vis-à-vis* the clarification of psychological concepts, that 'It is very noteworthy that *what goes on* in thinking practically never interests us' (Wittgenstein 1967: §88). And that is exactly the area where Turing's 'analysis' has had its greatest influence.

Not only does Wittgenstein's argument lead us to query from yet another angle the manner in which Turing sought to analyse the concept of calculation (*viz.*, as recursive 'mental process'), but it also brings into question the whole basis of Turing's interpretation of the logical relation in which algorithms stand to *inferring, reasoning, calculating*, and *thinking*. To be sure, this barely touches the surface of the issues inspiring the information-processing approach to thinking: in particular, the basic epistemological picture of the mind sundered from and forced to make sense of reality. But it does reflect some of the philosophical problems involved in interpreting the relation in which programs stand to speaking-aloud protocols.

It also sheds light on the import and relevance of the puzzling argument in *Philosophical Grammar* that begins with the question 'if thinking consists only in writing or speaking, why shouldn't a machine do it?' and concludes with the warning that 'It is a travesty of the truth to say "Thinking is an activity of our mind, as writing is an activity of the hand"' (*PG*: 105–6). As the foregoing makes clear, this passage bears directly on Turing's initial 'analysis' of *calculation* and his subsequent attempt to base an 'analysis' of *thinking* on this flawed foundation. Moreover, the very fact that Wittgenstein broached this topic five–six years before the publication of 'On Computable Numbers' provides an important insight into the conceptual environment which shaped Turing's attitude towards the Mechanist Thesis; for the themes which most concerned Wittgenstein in the early 1930s also served as the spring-board for Turing's 'post-computational' entry into the philosophy of psychology.

Here is the reason why Turing's writings continue to attract such deep philosophical interest. Turing brought to the fore the consequences of the epistemological framework that has virtually governed psychology from its inception. But that framework is itself the product of an archetypal picture to which Kant, no less than Aristotle, was responding—of the relation between thought and reality. Turing's mechanist thesis can either be seen as the crowning achievement of that framework, or as a *reductio*, forcing us to reassess these epistemological premises. Similarly, Turing's psychological thesis can either be seen as one more step in mechanism's technological advance (one that has already been displaced), or as forcing us to reassess the consequences of seeking to establish psychology on this epistemological foundation: of trying to resolve epistemological problems psychologically. Is it any wonder that *Mechanical Intelligence* makes for such stimulating and timely reading?

References

Aristotle [1938]: *Selections*, W. D. Ross (trans. and ed.). New York: Scribner.

Binet, Alfred [1894]: *Psychologie des grands calculateurs et joueurs d'échecs*. Paris: Deuxième Partie.

Bower, Gordon H. and Ernest R. Hilgard [1981]: *Theories of learning*. (5th ed.). Englewood Cliffs: Prentice-Hall.

Bruner, J. S. [1959]: 'Inhelder and Piaget's *The growth of logical thinking*', *General Psychology* **50**, 363–70.

——— [1960]: 'Individual and collective problems in the study of thinking', *Annals of the New York Academy of Science* **91**, 22–37.

——— [1990]: *Acts of meaning*. Cambridge, Mass.: Harvard University Press.

Carpenter, B. E., and R. W. Doran, eds. *A. M. Turing's ACE report of 1946 and other papers*. Cambridge, Mass.: MIT Press, 1986.

Cleveland, Alfred A. [1907]: 'The psychology of chess and of learning to play it', *The American Journal of Psychology* **XVIII**.

Davis, Martin [1978]: 'What is a computation?' in Steen, Lynn Arthur (ed.), *Mathematics today*. New York: Springer Verlag.

De Groot, Adriann D. [1965]: *Thought and choice in chess*. The Hague: Mouton Publishers.

Dennett, Daniel [1978]: 'Why the law of effect will not go away', in *Brainstorms*, Sussex: The Harvester Press.

DETLEFSEN, MICHAEL [1993]: 'Hilbert's formalism', *Revue internationale de philosophie* **47**, no. 4.
FEARING, FREDERICK [1930]: *Reflex action: A study in the history of physiological psychology*. New York: Hafner.
GEORGE, F. H. [1962]: *Cognition*. London, Methuen.
HADAMARD, JACQUES [1945]: *The psychology of invention in the mathematical field*. New York: Dover.
HARVEY, O. J., D. E. HUNT, and H. M. SCHRODER [1961]: *Conceptual systems and peronality organization*. Wiley: New York.
HODGES, ANDREW [1983]: *Alan Turing*. London: Burnett Books.
HOFSTADTER, DOUGLAS [1980]: *Gödel, Escher, Bach: An eternal golden braid*. London: Penguin Books.
HULL, CLARK [1962]: 'Psychology of the Scientist, VI', *Perceptual and Motor Skills* **15**.
KANT, IMMANUEL [1970]: *Logic*, Robert S. Hartman & W. Schwarz (trans). Indianapolis: Bobbs-Merill.
LEWES, G. H. [1877]: *The physical basis of mind*. Boston.
LOCKE, J. [1690]: *An essay concerning human understanding*, (John Yolton, ed.). London: Dent, 1961.
LOEB, JACQUES [1912]: 'The significance of tropisms for psychology' in Fleming, Donald (ed.), *The mechanistic conception of life*. Cambridge, Mass.: The Belknap Press of Harvard University Press, 1964.
LOVELACE, ADA AUGUSTA [1842]: 'Sketch of the analytical engine invented by Charles Babbage by L. F. Menabrea', in Babbage, Charles, *On the principles and development of the calculator*. New York: Dover Publications, 1961.
NEWELL, ALLEN and HERBERT A. SIMON [1961]: 'The simulation of human thought', in *Current trends in psychological theory*. Pittsburgh: University of Pittsburgh Press.
PAVLOV, I. P. [1927]: *Conditioned reflexes* (B. V. Anrep, trans. and ed.). New York: Dover.
RIGNANO, E. [1923]: *Psychology of reasoning* (W. A. Holl, trans.). London: Kegan Paul, Trench, Trubner.
SAVAGE-RUMBAUGH, E. SUE, JEANNINE MURPHY, ROSE A. SEVCIK, KAREN E. BRAKKE, SHELLY L. WILLIAM, DUANE M. RUMBAUGH [1993]: 'Language comprehension in ape and child', *Monographs of the Society for Research in Child Development* **58**, Nos. 3–4.
SHANKER, S. G. [1987]: 'Wittgenstein *versus* Turing on the nature of Church's Thesis', *Notre Dame Journal of Formal Logic* **28**, 615–649.
——— [1992]: 'In Search of Bruner', *Language & Communication* **12**, no. 1, 53–74.
——— [1993a]: 'Locating Bruner', *Language & Communication* **13**, no. 4, 239–263.
——— [1993b]: 'Wittgenstein *versus* Russell on the analysis of mind', in Irvine, Andrew and Gary Wedeking (eds.), *Bertrand Russell and the rise of analytic philosophy*. Toronto: University of Toronto Press.
——— [1994]: 'Ape language in a new light', *Language & Communication* **14**, no. 1, 59–85.
——— [forthcoming]: 'The mechanist continuum', in Shanker, S. G. (ed.), *Philosophy of the English speaking world in the twentieth century 1: Logic, mathematics and*

science. Routledge History of Philosophy. (G. H. R. Parkinson and S. G. Shanker, eds.). London: Routledge.
SHANNON, CLAUDE E. [1956]: 'A chess-playing machine', in Newman, James R. (ed.), *The world of mathematics*. New York: Simon and Schuster.
SIMON, HERBERT A. [1982]: 'Otto Selz and information-processing psychology', in Frijda, Nico H. and Adriaan D. De Groot (eds.) *Otto Selz: His contribution to psychology*. The Hague: Mouton Publishers.
________ [1983]: 'Why should machines learn?', in Michalski, R. S., *et. al.* (eds.), *Machine learning*. Berlin: Springer-Verlag.
TURING, ALAN [1936–37]: 'On computable numbers, with an application to the *Entscheidungsproblem*', *Proceedings of the London Mathematical Society* (2) **42**, 230–265.
________ [1946]: 'Proposals for development in the mathematics division of an automatic computing engine (ACE)', in Carpenter 1986. Reprinted in A. M. Turing 1992, pp. 1–86.
________ [1947]: 'Lecture to the London Mathematical Society on 20 February, 1947', in Carpenter 1986. Reprinted in A. M. Turing 1992, pp. 87–106.
________ [1969]: 'Intelligent machinery', in B. Meltzer and D. Michie (eds.), *Machine intelligence 5*, Edinburgh University Press: Edinburgh. Reprinted in A. M. Turing 1992, pp.107–128.
________ [1950]: 'Computing machinery and intelligence', *Mind* **59**, 433–460. Reprinted in A. M. Turing 1992, pp. 133–160.
________ [1951]: 'Can digital computers think?'
________ [1959]: 'Intelligent machinery: A heretical view', in Sarah Turing 1959.
TURING, A. M. [1992]: *Collected works of A. M. Turing*. Vol III, *Machine Intelligence* (D. C. Ince, ed.). Amsterdam: North-Holland.
TURING, SARAH [1959]: *Alan Matheson Turing*, Cambridge: Heffers.
WEBB, JUDSON [1980]: *Mechanism, mentalism, and metamathematics*. Dordrecht: D. Reidel Publishing Co.
WAISMANN, F. [1968]: *The principles of linguistic philosophy* (R. Harré, ed.). London: Macmillan.
WITTGENSTEIN, L. [1953]: *Philosophical investigations* (3rd ed.) (G. E. M. Anscombe, trans.). Oxford: Blackwell.
________ [1956]: *Remarks on the foundations of mathematics* (G. H. von Wright, R. Rhees, and G. E. M. Anscombe, eds., G. E. M. Anscombe, trans.). Oxford: Blackwell.
________ [1960]: *The blue and brown books*. Oxford: Blackwell.
________ [1967]: *Zettel* (G. E. M. Anscombe and G. H. von Wright, eds., G. E. M. Anscombe, trans.). Oxford: Blackwell.
________ [1974]: *Philosophical grammar* (Rush Rhees, ed., Anthony Kenny, trans.). Oxford: Blackwell.

12

Mathematics and Gender: Some Cross-Cultural Observations*

Ann Hibner Koblitz

Professor of Women's Studies, Arizona State University

According to folklore, the eminent mathematician Hermann Weyl occasionally declaimed: "There have been only two women in the history of mathematics. One of them wasn't a mathematician [Sofia Kovalevskaia], while the other wasn't a woman [Emmy Noether]." Most people in the mathematical community today would not venture so far as to endorse this pronouncement. Yet an astonishing number of mathematicians, mathematics educators, and historians of mathematics subscribe to the belief that there have been only a handful of women in the history of mathematics and that women's numbers in mathematics-related fields today are uniformly low everywhere in the world. Moreover, many people would agree with some form of the statement that the female nature and mathematical thought are incompatible. This paper challenges these beliefs and argues that the historical legacy and current status of women in mathematics are complicated and often contradictory. In the interests of brevity, I have organized what follows as a series of statements.

STATEMENT 1:

The historical legacy of women in mathematics has been mixed.

There have been women mathematicians since classical times. During at least the second half of the 19th century as well as the whole of the 20th century, many women have been active participants in and contributors to the mathematical community (Grinstein & Campbell, 1987). Seven women are generally cited in superficial overviews (Coolidge, 1951; Mozans, 1974; Osen, 1974): (1) Hypatia (370?–415? AD), who according to legend lived in Alexandria, did work on conic sections, and was martyred by Christians; (2) Emilie du Châtelet (1706–1749), courtier and *philosophe*, who translated Newton into French and extensively commented on his work; (3) Maria Gaetana Agnesi (1718–1799), of "witch" of Agnesi fame, who occasionally lectured at the University of Bologna and turned down an appointment to the Academy of

*First published in Gila Hanna, ed., *Towards Gender Equity in Mathematics Education*, Dordrecht: Kluwer, 1996, pp. 93–109.

Sciences to devote herself to a religious life; (4) Sophie Germain (1776–1831), who proved an important case of Fermat's Last Theorem and whose work on elasticity won her the Grand Prix of the French Academy of Sciences in 1816; (5) Mary Fairfax Somerville (1780–1872), an English polymath and popularizer whose *On the Connexion of Mathematics to the Physical Sciences* went through numerous editions in the 19th century; (6) Sofia Kovalevskaia (1850–1891), the first woman to receive a doctorate in mathematics (in the modem sense of the term) and winner of the Prix Bordin of the French Academy of Sciences; and (7) Emmy Noether (1882–1935), one of the founders of modem algebra.

These women are the most famous. But it would not be difficult to compose a list of a hundred or so prominent women mathematicians from many time periods and diverse cultures.

Women mathematicians have often been the ones to break down educational barriers and open up professional opportunities for all women. Some examples:

- The first woman in modern times to be fully integrated into professional, academic life at the university level in Europe was a mathematician. Sofia Kovalevskaia joined the faculty of Stockholm University in 1884 and became an ordinary (full) professor there in 1889. Kovalevskaia was also the first woman to be elected corresponding member of the Russian Imperial Academy of Sciences; the rules were changed to permit her membership (Koblitz, 1993).
- The first PhD granted to a woman in any field by Columbia University (and the first doctorate in mathematics given to any woman by a US university) was awarded to a mathematician, Winifred Haring Edgerton, in 1886 (Green & LaDuke, 1990).
- In 1920, Emmy Noether became the first woman in any field to obtain full qualifications to teach in German universities (Junginger, 1993).
- The first Nigerian woman to obtain a doctorate in any field was a mathematician; Grace Alele Williams received her PhD from the University of Chicago in 1963 (American Association for the Advancement of Science, 1993).
- The first woman to be awarded a full professorship in a scientific or technical field in Vietnam (where full professor is very rare as a title) was a mathematician, Hoang Xuan Sinh. Sinh, a former student of A. Grothendieck and L. Schwarz, has several times been the only female professor involved in the International Mathematics Olympiads (in her capacity as a coach of the Vietnamese team).[1]

[1] Much of my information on women mathematicians of Asia, Africa, and Latin America does not come from published sources, which are generally conspicuous by their absence. Rather, the data emerge from extensive personal interviews with the women themselves, their male colleagues, officials in women's organizations and ministries, and so on.

Moreover, there are a couple of instances in which the attitude of male mathematicians toward their female colleagues has been particularly impressive:

- In 1896, a survey of the German professoriate was taken on the question of whether women should be admitted to universities with the same rights as men. Mathematicians were unanimously in favor, physicists only slightly less so. Historians, however, were almost all opposed to the entry of women (Koblitz, 1993).
- Carl Friedrich Gauss, whom many mathematicians consider the greatest number theorist who ever lived, attempted to arrange for Sophie Germain to obtain a doctorate from Göttingen University. Gauss had tremendous admiration for what Germain had achieved in mathematics. He wrote:

  > But when a person of the sex which, according to our customs and prejudices, must encounter infinitely more difficulties than men to familiarize herself with these thorny researches, succeeds nevertheless in surmounting these obstacles and penetrating the most obscure parts of them, then without doubt she must have the noblest courage, quite extraordinary talents, and a superior genius. (Quoted in Edwards, 1977, p. 61)

On the other hand, one can chronicle enough cases of blatantly discriminatory conduct of male mathematicians towards their female colleagues to remove any doubt that male mathematicians are as much a product of their culture as any other occupational group. Some examples:

- Sophie Germain was a mathematical correspondent of Gauss, Lagrange, and others, and, as mentioned above, her work on the elasticity of metals won the Grand Prix of the French Academy of Sciences. Nevertheless, her name was not on the original list of prize winners on the Eiffel Tower, even though her work contributed to making the tower itself possible (Mozans, 1974, p. 156).
- Christine Ladd-Franklin, a student of J. J. Sylvester at Johns Hopkins University, completed all work for the PhD, up to and including her soon-to-be published dissertation, in 1882. The university did not award her a degree at the time, and in fact did not do so until 44 years later, in 1926. Ladd-Franklin's work earned her one of the very few stars given to women in early editions of *American Men* [sic] *of Science* (Green & LaDuke, 1990, p. 123).
- D. E. Smith and J. Ginsburg note in their *A History of Mathematics in America Before 1900* that *Mathematische Annalen* published 15 articles by U.S. mathematicians in the period 1893 to 1897. They list 14 authors, all male. The only one they omit is the one woman – Mary Frances Winston, a student of Felix Klein at Göttingen (Green & LaDuke, 1990, p.117).
- Emmy Noether was an editor of *Mathematische Annalen* in the late teens and twenties, as the whole of the European mathematical community knew

well. She was not listed on the masthead of the journal (in contrast to Sofia Kovalevskaia, who 30 years earlier was listed as editor of *Acta Mathematica*). Moreover, she was sometimes disrespectfully referred to by her colleagues as "Der Noether."

- In the course of several years of giving talks in mathematics departments on Sofia Kovalevskaia, I have sometimes encountered mathematicians who assume demeaning stereotypes about her. They say, for example, that she must have been Weierstrass's mistress, or that the French mathematicians gave her the Prix Bordin out of gallantry, or (my personal favorite) "no, no, no, dear, you've got it all wrong. Kovalevskaia was an *amateur* mathematician; her *husband* was responsible for the Cauchy-Kovalevskaia Theorem." (That particular piece of nastiness has several layers. There was a male mathematician named Kovalevskii, slightly younger than Kovalevskaia and not in her field; the Cauchy-Kovalevskaia Theorem is sometimes called Cauchy-Kovalevskii, due to differing transliterations of her name; her husband was also a well-known scientist – a paleontologist!)

Sometimes, people have a tendency to get rather smug and complacent when reading about examples of historical discrimination against women or about the situation in countries far away. They congratulate themselves that they are not as stupid as the Johns Hopkins professors were when they refused Christine Ladd-Franklin her degree, or as whoever it was who kept Emmy Noether off the masthead of *Mathematische Annalen*. It is well to remember, though, that just because certain formal prohibitions and rules barring women have been abolished in academia today, that does not mean that discrimination against women in mathematical fields no longer exists. It behooves us to be especially alert to contemporary methods of discrimination as well as their historical analogs. Although discriminatory practices are in no sense intrinsic to mathematics, nor necessarily worse in mathematics than in other fields, nevertheless there are many ways in which women mathematicians can be the victims of unfairness.

It seems obvious, for example, that a person who can cavalierly and ignorantly propagate sexist myths about historical women mathematicians (such as Hermann Weyl's insults or the nonsense about Kovalevskaia cited previously) will not be above concocting analogously scurrilous stories about present or prospective female colleagues. And that, of course, could be quite harmful. I shall return to this point in Statement 8.

STATEMENT 2:

There is immense variability in the position of women in mathematics both historically and cross-culturally. One cannot necessarily predict what the status of women will be even across ostensibly similar cultural and economic settings.

The percentage of women receiving PhDs at U.S. universities, for example, has varied by over 300 percent; I use "varied" as opposed to "increased" because the percentage has by no means increased steadily. Women begin at a small non-zero percentage of PhDs awarded,[2] and constituted 14.3 percent of mathematics doctorates in the period 1900 to 1940. This compares satisfactorily to the 14.1 percent of doctorates received by women overall in the same period. In the 1940s through the 1960s the percentage of women doctorates in mathematics declined severely, as it did in all fields, so that during the period 1950 to 1969 the figure was always under 6 percent. Not until the 1980s did the percentage of US women receiving doctoral degrees climb past the earlier figure of 14.3 percent, and in the late 1980s and early 1990s it is between 18 and 28 percent, depending on whether one counts allied fields like statistics and computer science. No longer is this close to the percentage of women PhD recipients overall, however. That has peaked at approximately 36 percent (Almanac, 1991; Green & LaDuke, 1990; Keith, 1991; National Research Council, 1979). One can see similar up and down trends in other countries. For example, pre-revolutionary Russia to about 1900 had more women receiving doctorates in mathematics than in the next 20 years. The numbers start to rise again in the experimental 1920s (Koblitz, 1988a; Lapidus, 1978).

Although one might assume that the percentages of female mathematicians in countries with similar cultures or with similar economic indices might be comparable, in fact there are many obvious (and not so obvious) counterexamples. Among the obvious are Great Britain and northern Europe in general as opposed to the US and southern and eastern Europe. The US, France, Italy, Portugal, Turkey, and Spain have significantly higher proportions of female mathematicians than England, western Germany, Sweden, and so on (Burton, 1990; Lovegrove & Segal, 1991; Ruivo, 1987; Stolte-Heiskanen, 1991). Nor can we assume that level of material advancement necessarily correlates positively with percentage of women in mathematics. Japan's and Singapore's numbers are relatively low, while Mexico's, China's, Brazil's, Cuba's, and Costa Rica's are higher (Azevêdo et al., 1989; Burton, 1990; Faruqui, Hassan, & Sandri, 1991; Graf & Gomez, 1990; Koblitz, 1988b; Ruivo, 1987).

Analogously, neighboring countries with similar cultures can also have different percentages of women in mathematical areas. Costa Rica's national university has 12.5 percent women on its mathematics faculty, including a woman chair. Nicaragua's national university, by contrast, has virtually no female professors in mathematics or the physical sciences (though the statistics department at the Leon branch of the university is chaired by a woman).

In so-called developing countries women appear to be leaders in their departments no less often – and frequently more often – than in the US and Canada. In preparing this paper, I came up with a list of women who were or had recently been heads of mathematics, statistics, or computer science departments in India (Madras Christian College, Lucknow, Mangalore), the

[2] Graduate education in the US and Canada did not really get started until the late 1870s, so there were some women involved in the enterprise almost from the beginning.

Philippines (University of the Philippines and De La Salle University), the University of Costa Rica, University of Nicaragua-León, Mexico, Ivory Coast, Peru, Hong Kong. This does not necessarily mean that the situation for women in mathematics is better in these countries, merely that certain stereotypes need to be treated with caution. One cannot assume that academic communities in Asia, Africa, and Latin America will be more backward on women's issues than those of North America and Europe. For example, the Third World Academy of Sciences membership is 15 percent female, as compared to approximately 5 percent for the academies of science of the US and the former USSR (Salam, 1988).

At the student level, too, the picture can be varied and can run counter to some of our stereotypes and the images put forward in the popular press. Women's participation in post-secondary level programs in mathematics and computer science in some countries (Liberia, Cuba, Indonesia, South Korea, Kuwait, Saudi Arabia, Turkey, Albania, and Italy) approaches or even exceeds their percentage in all programs at that level (UNESCO, 1993, Table 3-14). Other countries (the US, Canada, and most countries of northern and western Europe) have far lower percentages of women students in mathematics and computer science (Chipman, Brush, & Wilson, 1985; Stolte-Heiskanen, 1991).

STATEMENT 3:

Periods of social pressure and reaction can have special consequences for women in "non-traditional" fields like mathematics.

These effects can be either positive or negative (or mixed). In fact, sometimes "good things" can happen for unusual, or even for the wrong reasons. Take, for example, the situation in Mexico. Like many countries all over the world, Mexico is experiencing economic difficulties. The national universities are in fiscal crisis, and (disproportionately male) professors are leaving in droves for the private universities and the business sector. At a roundtable discussion of women mathematicians and scientists in Mexico City in 1991,[3] the question was raised of the impact of the crisis on university women, whether it might not be a blessing in disguise that the men are abandoning the university to women. True, the percentage of women on the mathematics faculty is rising. But prestige and salary are falling, and many women have the feeling that the best students are not going to be attracted to mathematics anymore.

This is a complex situation, and one that has analogs in many other countries and fields. If a specialty becomes "saturated" (that is, too many professionals to make good salaries) or unattractive for some other reason, opportunities for women can increase.[4] Also, a field can become stratified. That is, the teaching of mathematics (and university teaching in general) can become more

[3]The discussion was part of a week of activities commemorating the centenary of Sofia Kovalevskaia's death.

[4]Veterinary medicine in the US, pharmacy in El Salvador, medicine in the USSR and the Philippines are examples of this phenomenon.

or less the province of women, while the academies of sciences and research institutes can remain largely or entirely male. One sees this in many countries that overall have decent percentages of women in mathematics, such as Mexico, India, China, Nigeria, and Costa Rica. The most prestigious positions and places become or remain largely male preserves.

STATEMENT 4:

One must be very careful about making generalizations from historical and cross-cultural comparisons.

Mindless use of one or another indicator to make a sweeping generalization about women's status in mathematics can give misleading results. It is not possible to use the same indicators to determine the situation in every country. The significant statistic might be the percentage of women teaching at the university level. But it might also be the proportion of women at research institutes and academies of sciences (and at what level), or the percentage of women who publish (or who publish in foreign as opposed to domestic journals), or the proportion of women who go abroad for conferences, postgraduate study, and so on, or the percentage of women awarded grants by national and international funding agencies. Indices can have different meanings in different countries, and the prestige of various positions and honors can vary considerably. This is not to say that it is unimportant that women constitute a large percentage of professors on a mathematics faculty in a certain country. But this measure might not be the unique indicator of women's success or status in the mathematical world.

STATEMENT 5:

Despite the problems and societal obstacles, women can fare reasonably well in mathematics for a number of reasons.

There are, after all, some recognized and relatively objective standards in the mathematical community. Women might have to be better than their male counterparts to be judged equal, but the standard is not impossible to achieve. Sofia Kovalevskaia, for example, offered *three* works to Göttingen in fulfillment of her degree requirements. She and her adviser Karl Weierstrass reasoned that as the first woman applying for the doctorate, her case would have to be especially strong. Three *were* sufficient, however.

Moreover, there is perhaps not so great an "Old Boy Network" in mathematics as in other fields and more of a tradition of tolerance and eccentricity. There is also a certain pride in traditions of internationalism – in not wanting to allow political, ethnic, or ideological considerations to interfere with one's mathematical judgments. These factors have worked in women's favor in a large variety of contexts.

Women can also fare well in mathematics because of a relatively small number of men who have served as excellent mentors, in a couple of cases despite otherwise not terribly progressive politics or overall commitment to women's rights. Some random examples:

- Karl Weierstrass performed valuable mentoring functions not only for Sofia Kovalevskaia but also for her friend Iulia Lermontova. It was through Weierstrass's intervention with university officials at Göttingen that Lermontova was able to become the first woman in the world to receive her doctorate in chemistry (Koblitz, 1993).
- Felix Klein supported Mary Winston, Grace Chisholm, and Margaret Maitby in their admission to graduate studies in mathematics at Göttingen University in 1893. He later said that the women were fully comparable to his male students.
- Of the seven PhDs in mathematics awarded to women by Johns Hopkins University before 1940, five were students of one professor (Morley). Eleven of the 13 women who completed PhDs at Catholic University were students of Landry. The University of Chicago granted 46 of the 229 doctorates given to women in the US through 1939; of these, 30 were students of either Leonard Eugene Dickson (18 of his 67 students were women) or Gilbert Ames Bliss (12 of his 52 students were women) (Green & LaDuke, 1990).
- Lee Lorch, now emeritus at York University, performed mentoring functions for many women – including several of the few black women who later received PhDs in mathematics – during the time he taught at Fisk University.

STATEMENT 6:

Cross-cultural disparities in female educational and employment patterns in mathematics and computer science raise serious questions about theories that claim innate gender differences in mathematical ability.

We need to ask ourselves whether there are other explanations: faulty test design, socio-cultural factors, and so on. Theories must be examined critically and tested historically and cross-culturally. Several generalizations have become quite popular in late-20th century US society, yet they rest on dubious foundation and would fail under historical and cross-cultural scrutiny.

For at least three decades the received wisdom – and the line being pushed by "objective scientists" like Camilla Benbow and Julian Stanley – was that girls are better at verbal tests and boys at mathematical ones (Benbow & Stanley, 1980). Now, though, that picture is breaking down in several important ways. The so-called gender gap on US standardized tests narrowed considerably during the 1980s, to the point that specialists at the Educational Testing Services (ETS) in Princeton are now saying the differences are not statistically significant. Also, several recent studies have shown that even in mixed groups where males had performed noticeably better than females on mathematics Scholastic Aptitude Tests (SATs), on other tests, including ETS's Mathematics Achievement Test itself, there were no significant gender differences. (For a reasonably current review of the literature see Kenschaft, 199lb.)

More importantly, the supposed gender differences are constant neither across ethnic groups within the US and Canada nor across cultures. A 1987 study noted that even within the mathematics SAT, the gender gap varies considerably. It is largest for Hispanics and smallest for Afro-Americans (Ruskai, 1991). Moreover, Gila Hanna has pointed out in her comparative studies that differences *between countries* are much larger than those between boys and girls, and that the gender gap is larger in countries with low scores (like the US) than in those with high scores (like Hungary and Japan). On the geometry test, for example, US males scored 39.7 while US females scored 37.9 – a mere 1.8 points. Meanwhile, the advantage over the US of Hungary and Japan was relatively massive, since all subgroups in both countries scored between 55 and 60 points (Hanna, 1989).

Some countries (Thailand and South Korea, for example) do not exhibit any statistically significant differences between male and female performance on mathematics achievement tests (Hanna, 1989; Hanna, Kündiger, & Larouche, 1990; Kwon, unpublished paper). As Hanna and her collaborators note, the one clear conclusion that emerges from all the statistics is that one must doubt the biological explanations of male/female difference in mathematical ability, since it is "very unlikely" to vary between countries (Hanna, Kündiger, & Larouche, 1990, p. 96).

Other studies have also pointed to the culture-bound nature of many of our notions of gender difference. For example, spatial ability tests given to Native American children in Alaska and to central African children show either no difference or one favoring the females (Fausto-Sterling, 1985; Kenschaft, 1991a; Koblitz, 1987; Ruskai, 1991). Needless to say, these tests are not the ones reported with fanfare in the *New York Times*.

Benbow's and Stanley's studies have immense cultural problems. Gender differences are not consistent across ethnic groups, and even girls who did worse than boys on the Benbow and Stanley test outperform the boys in school. Moreover, before they test, Stanley's research centre apparently sends parents a pamphlet noting that boys do better than girls on the test. Such a message to parents reveals the researchers' blindness to issues of bias-free experimental design (Jackson, 1990; Ruskai, 1991).

Also dubious are generalizations about a fundamental physiological difference which affects the way men and women reason about mathematics or their interests in mathematics. The latest manifestation of this type of pseudo-scientific study is the brain lateralization "research." This has a long, inglorious history dating back to early 19th-century phrenology (Alper, 1985; Bleier, 1984; Dumdell, 1991; Fausto-Sterling, 1985).

STATEMENT 7:

There are fundamental problems with anything that smacks of essentialism; this includes so-called feminist gender-and-science theory.

"Essentialism" encompasses any theory that attributes gender differences to biological, genetic, psychosocial, or other immutable factors. One problem with such theories is that, as indicated above, women's position in mathematics and the natural sciences is constant neither across cultures nor across time periods. What in one country or time might be considered unfeminine might in another country or time not be so. Specifically with regard to the mathematical sciences, there are several countries, including the Philippines, Turkey, Kuwait, and Mexico, for example, in which women constitute a rather high percentage of mathematics-related professions, especially when compared with the numbers in northern Europe, Canada, or the United States (Faruqui, Hassan, & Sandri, 1991; Lovegrove & Segal, 1991; Stolte-Heiskanen, 1991; UNESCO, 1993; United Nations Statistical Division, 1992). Yet most feminist theorizing about gender and mathematics assumes that women's participation in these fields is uniformly low.

The theorists also assume that Victorian-era bourgeois stereotypes concerning femininity and gender polarities are uniform across all cultures, classes, and historical periods. This is far from being the case, even within western Europe. In Italy, for instance, the stereotype is different from that in the US or Sweden or the UK. Women are purported to be "natural" theoreticians (hence the relatively large number of Italian women mathematicians and computer scientists), while men are supposed to be more practical by nature (and thus become engineers rather than theoretical scientists).

Moreover, the position of women in mathematics and the sciences can change quite rapidly for the better (or worse). The changes are far too rapid to be explainable by biological theories of difference or by psychosocial theories such as Nancy Chodorow's (1978).[5] Gender and science theorists have the unfortunate tendency to make generalizations despite clear historical and cross-cultural counterexamples. For example, it is simply not true that women's status in the sciences has remained unchanged since the Scientific Revolution, though that is a claim often made by theorists like Evelyn Fox Keller (1985), Sandra Harding (1986, 1991), and their imitators.

In like manner, I am disturbed by certain aspects of work by Belenky and her collaborators (1986), Gilligan (1982), and others on women's purportedly different ways of knowing. Belenky and her collaborators, for example, assert that the gender differences they describe are independent of culture, ethnicity, and class. This is highly questionable. A colleague of mine at Hartwick College, Katherine O'Donnell, has conducted years of field research with poor migrant farm women in New York state. At the beginning she had the expectation of

[5]In general terms, object-relations theory says that a girl infant never has to disidentify herself from her mother. Therefore, she never sees the world as alienated from herself in the same way a boy infant does. A boy baby, on the other hand, realizes very early that he is not the same gender as his mother. He has to distance himself from her and thus begins to objectify the world. Because object-relations theory attributes gender differences in intellectual outlook to an immutable mechanism of early childhood, it is almost indistinguishable from a genetic or biological theory. Moreover, there is no way the theory can account for differences between individuals or for change over time or across cultures.

finding support for the writings of Belenky and Gilligan (O'Donnell, unpublished paper a) Now, however, she is firmly convinced that, on the contrary, the interactions of gender, culture, race, and class are far too complex to be encompassed by a simplistic gender-polarity theory (O'Donnell, unpublished paper b).

To bring this discussion closer to mathematics, let us take the work of Sherry Turkle (1984) on children and computers. Under the influence of object-relations theory and gender-and-science theory, Turkle has created the concepts of "soft mastery" and "hard mastery" to categorize her analysis of the use of computers by school-age boys and girls. For Turkle, "hard mastery is the imposition of will over the machine through the implementation of a plan ... the hard masters tend to see the world as something to be brought under control." Soft mastery, on the other hand, is more interactive – "the soft masters are more likely to see the world as something they need to accommodate to, something beyond their direct control." Though Turkle gives examples of soft masters of both sexes, she says that girls tend to be soft masters, while hard masters are "overwhelmingly male."

Large parts of the theory sound quite plausible at first. Certainly it is true that in most societies males and females are socialized differently, with female socialization tending more towards valuing qualities like accommodation and interaction, and male socialization tending more towards control. Upon reflection, however, some people have raised questions about the validity of the theory and voiced concerns about its social implications.

It has been pointed out (by Beth Ruskai [1990], for example) that the theory contains certain assumptions about the nature of computer science that are rather far off the mark. What Turkle dichotomizes as "hard" and "soft" mastery might be better characterized as two inextricably interwoven parts of the creative scientific process. The attempt to label people as being *either* hard *or* soft, therefore, misses a crucial point of what it is to do science.

Turkle's theory makes no allowances for historical and cross-cultural variation. It ignores the evidence that women's participation in the sciences including computer science – has varied widely from one decade to the next and even between neighboring countries (Koblitz, 1991; Lovegrove & Segal, 1991). Moreover, stereotypes regarding women's innate capacities and their relation to mathematics and computer science vary from culture to culture.

There is also reason to be uncomfortable with the way Turkle's theory dovetails with current western European and North American stereotypes about women's intrinsic nature. In practice, these stereotypes can contribute to discrimination against women in the workplace, and to the segregation of women in so-called pink-collar ghettoes like data processing.

Sherry Turkle is fairly skillful in inserting caveats in her generalizations and in reminding us that the picture is complex. (Harding, Belenky, and Gilligan are not nearly so careful.) Unfortunately, however, the caveats and reservations rarely make it into popular accounts of the work. What gets picked up in the media is the rather simplistic idea that girls cannot be attracted into computer-related fields unless the machine can be portrayed as artistic, relational, and "soft." We do not work in a vacuum. If the media can distort a theory, they will.[6]

In the US, one of the most common manifestations of sexism and racism in the classroom is a refusal to intellectually challenge girls and members of minority groups. They are condescended to and patronized and do not receive adequate exposure to the more rigorous, thought-provoking, and elegant aspects of mathematics. Their understanding thus rarely attains the level of the systematic and the structural; they seldom arrive at the stage where they can see much point in doing mathematics. Unfortunately, this phenomenon can be exacerbated by overzealous followers of the ideas of Turkle, Gilligan, and Belenky.

STATEMENT 8:

Any discrimination against younger faculty automatically falls disproportionately on women because of the demographics of the profession. The injustice is increased because of certain usually unconscious but quite pervasive attitudes about women's "natural" roles.

In the US, for example, women are often assigned heavier teaching loads than men and more courses at the introductory level. This is a triple blow: more work is devoted to teaching as opposed to research; teaching evaluations are automatically worse because of the nature of the course (introductory courses virtually always receive lower evaluations than upper division courses); and any deviation from stereotypically feminine behavior (such as attempting to enforce high academic standards) is met with displeasure (and low ratings) by students. Numerous studies indicate differential treatment of women faculty on evaluations. For example, students expect to be "nurtured" by women and punish them for deviations from the ideal "feminine" standard (N. Koblitz, 1990).

Women faculty are caught in a bind. Either they devote tremendous time to teaching, in which case their research suffers, or they devote as much time to their research as their male colleagues do, remaining aloof from students, in which case they are penalized more heavily than men on evaluations.

Moreover, the evaluatory process for granting tenure is in essence a blackballing system.[7] Even if 80 percent of the department have not a sexist bone

---

[6]We had a distressing illustration of exactly this point in Swedish newspaper coverage of the 1993 ICMI Study Conference "Gender and Mathematics Education" in Höör, Sweden. The headline of the 9 October 1993 *Dagens Nyheter* story was "Mathematicians disagree whether biology makes a difference," and ICMI President Miguel de Guzmán was misquoted as saying that gender differences manifest themselves from day one in the classroom!

[7]I realize that the concept of tenure (the right to relative job security, granted after a probationary period of some years) is not institutionalized in all countries. But I believe that the following discussion is applicable to most other kinds of hiring and promotion processes as well.

in their bodies, the opposition of a relatively few curmudgeons can sink a woman's chances in a variety of ways. It is rather like the old anti-Semitic blackballing system for admission to US country clubs – the system persisted for so long because all that was needed was one person in opposition.

There is an analogous blackballing system in place for hiring and tenure in academic departments today. A couple of people can skew the process and, in fact, wreck it. Consider the following situations:

- Say a woman has children and resumes mathematical activity after a couple of years hiatus. How does one interpret this? One could talk about the fact of her return to active research as indicating resolve and high mathematical ability and dedication. But one typically hears the diehard sexists referring instead to the "unfortunate gap in her publication record."
- Women are often held up to far higher standards than men on the pretext of not lowering standards. For a woman's appointment, there cannot be the least shadow of a doubt, while men are often given the benefit of the doubt.[8] We all know at least one senior professor who thinks nothing of spreading stories to the effect that it would be lowering the department's standards to hire a woman, even when the woman being considered is a far better researcher than he is himself.

In the US a certain amount of sound and fury signifying little or nothing has sprung up around the issue of Affirmative Action. Departments sometimes have women come on campus for interviews, under pressure from the administration, only to have the appointment sabotaged by a couple of diehard sexists working assiduously to undermine the process. There is a lot of rumormongering that goes on about Affirmative Action. Traditionalists routinely exaggerate the success of the program, telling female graduate students things like, "oh, you'll have no problem getting a job, you're in fashion these days," or jokingly advising their male students to wear a skirt to the job interview. This kind of childish and disingenuous behavior in itself creates a bad atmosphere for women. It conveys the impression that colleagues do not have confidence in the women they have hired and implies that any women in the department are there on sufferance.

Conclusion

My goal here clearly has not been the presentation of a definitive account or a polished treatise. It was my intention to throw out food for thought,

[8] This kind of discrimination includes behavior such as the infamous cases of analogous curricula vitae being sent to chairs with male and female names attached – the chairs routinely recommended the women for lower positions than the men! Bernice Sandler has documented many examples of this sort of (unconscious) prejudice in her "Chilly Climate" series (Sandler, n.d.).

to illustrate some of the curiosities and ironies inherent in women's position in mathematics historically and across cultures. The interactions of gender and culture are never simple or straightforward. The history of women in mathematics has not been some Whiggish triumphal passage from darkness into light, but neither has it been a chronicle exclusively of discrimination and marginalization. The history and present status of women in mathematics are complicated, and often the picture has elements of contradiction. One conclusion seems obvious, however. The complex and multifaceted interactions of gender and mathematics can be understood only if one takes into account historical and cross-cultural perspectives.

References

Almanac. (1991). *The chronicle of higher education*, XXXVIII, No. 1.

Alper, J. S. (1985). Sex differences in brain asymmetry: A critical analysis. *Feminist Studies*, 11(1), 7-37.

American Association for the Advancement of Science [AAAS]. (1993). Science in Africa; Women leading from strength. Washington, DC: Author.

Azevêdo, E. S., et al. (1989). A mulher cientista no Brasil. Dados atuais sobre sua presença e contribuiçao. *Ciência e cultura*, 41(3), 275-283.

Belenky, M., Clinchy, B., Goldberger, N., & Tarule, J. (1986). *Women's ways of knowing: The development of self, voice and mind.* New York: Basic Books.

Benbow, C. P., & Stanley, J. C, (1980). Sex differences in mathematical ability: Fact or artifact? *Science*, 210, 4475, 1262-1264.

Bleier, R. (1984). *Science and gender.* New York: Pergamon.

Burton, L. (Ed.). (1990). *Gender and mathematics: An international perspective.* London: Cassell.

Chipman, S. F., Brush, L. R., & Wilson, D. M. (Eds.). (1985). *Women and mathematics: Balancing the equation.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Chodorow, N. (1978). *The reproduction of mothering: Psychoanalysis and the sociology of gender.* Berkeley; University of California Press.

Coolidge, J. L. (1951). Six female mathematicians. *Scripta Mathematica*, 17, 20-31.

Durndell, A. (1991). Paradox and practice: Gender in computing and engineering in Eastern Europe. In G. Lovegrove & B. Segal (Eds.), *Women into computing.* London: Springer-Verlag.

Edwards, H. M. (1977). *Fermat's last theorem: A genetic introduction to algebraic number theory.* New York: Springer-Verlag.

Faruqui, A. M., Hassan, M.H.A., & Sandri, G. (Eds.). (1991) *The role of women in the development of science and technology in the third world.* Singapore: World Scientific Publishing.

Fausto-Sterling, A. (1985). *Myths of gender: Biological theories about women and men.* New York: Basic Books.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Harvard University Press.

Graf, H. B., & Gomez, H. G. (1990). Acerca de las cientfficas en Ia UNAM. Unpublished paper delivered at the National Autonomous University of Mexico. (I am indebted to the authors for giving me a copy of this paper.)

Green, J., & LaDuke, 1. (1990). Contributors to American mathematics. In G. Kass-Simon & P. Farnes (Eds.), *Women of science: Righting the record.* Bloomington: Indiana University Press.

Grinstein, L. S., & Campbell, P. J. (1987). *Women of mathematics: A biobibliographical sourcebook.* New York: Greenwood Press.

Hanna, G. (1989). Mathematics achievement of girls and boys in grade eight: Results from twenty countries. *Educational Studies in Mathematics*, 20, 225-232.

Hanna, G., Kündiger, E., & Larouche, C. (1990). Mathematical achievement of grade 12 girls in fifteen countries. In L. Burton (Ed.), *Gender and mathematics* (pp. 87-97). London: Cassell.

Harding, S. (1986). *The science question in feminism.* Ithaca, NY: Cornell University Press.

Harding, S. (1991). *Whose science? Whose knowledge? Thinking from women's lives.* Ithaca, NY: Cornell University Press.

Jackson, M. V. (1990). SATs ratify white male privilege. Association for Women in Mathematics *Newsletter*, 20(5), 9-10.

Junginger, G. (1993). A woman [sic] career in veterinary medicine. Lecture given at the XIXth International Congress of History of Science, August 1993.

Keith, S. Z. (1991). A statistical overview of American women doctorates, 1988–1989. in P. C. Kenschaft (Ed.), *Winning women into mathematics* (pp. 59-60). Mathematics Association of America.

Keller, E. F. (1985). *Reflections on gender and science.* New Haven: Yale University Press.

Kenschaft, P. C. (1991a). Fifty-five cultural reasons why too few women win at mathematics. In P. C. Kenschaft (Ed.), *Winning women into mathematics* (pp. 11-18). Mathematics Association of America.

Kenschaft, P. C. (Ed.). (1991b). *Winning women into mathematics.* Mathematics Association of America.

Koblitz, A. H. (1987). A historian looks at gender and science. *International Journal of Science Education*, 9(3), 399-407.

Koblitz, A. H. (1988a). Science, women, and the Russian intelligentsia. *Isis*, 79, 208-226.

Koblitz, A. H. (Ed.). (1988b). *La mujer en la ciencia, la tecnología y la medicina.* Seattle: Kovalevskaia Fund.

Koblitz, A. H. (1991). Women in computer science: The 'soft mastery' controversy. *Kovalevskaia fund newsletter*, Vl(2), 5-6.

Koblitz, A. H. (1993). *A convergence of lives. Sofia Kovalevskaia: Scientist, writer, revolutionary* (2nd. ed). New Brunswick, NJ: Rutgers University Press.

Koblitz, N. (1990). *Are student ratings unfair to women?* Association for Women in Mathematics *Newsletter*, 20(5), 17-19.

Kwon, O. (unpublished paper). U.S.-Korea cross-national studies on college entrance exams.

Lapidus, G. W. (1978), *Women in Soviet society.* Berkeley: University of California Press.

Lovegrove, G., & Segal, B. (Eds.). (1991) *Women into computing.* London: Springer-Verlag.

Mozans, H. J. (1974). *Woman in science.* Cambridge, MA: MIT Press. [Original publication date, 1913.]

National Research Council, Commission on Human Resources. (1979). *Climbing the academic ladder: Doctoral women scientists in academe.* Washington, D.C.: National Academy of Sciences.

O'Donnell, K. (unpublished paper a). A class act. *Phoebe.*

O'Donnell, K. (unpublished paper b). *Got nowhere to go and no way to get there: Migrant women's struggle for dignity.*

Osen, L. M. (1974). *Women in mathematics.* Cambridge, MA: MIT Press.

Ruivo, B. (1987). The intellectual labour market in developed and developing countries: Women's representation in scientific research. *International Journal of Science Education*, 9(3), 385-391.

Ruskai, M. B. (1990). Why women are discouraged from studying science. *The Scientist*, 4(5), 17, 19.

Ruskai, M. B. (1991). Are there innate cognitive gender differences? Some comments on the evidence in response to a letter from M. Levin. *American Journal of Physics*, 59(1), 11-14.

Salam, A. (1988). Speech given at the Third World Academy of Sciences, Trieste, October 3.

Sandler, B. R. (n.d.) The campus climate revisited: Chilly for women faculty, administrators, and graduate students. *Project on the status and education of women.* Association of American Colleges.

Stolte-Heiskanen, V. (Ed.) (1991). *Women in science: Token women or gender equality?* Oxford: Berg.

Turkle, S. (1984). *The second self: Computers and the human spirit.* New York: Simon and Schuster.

UNESCO. (1993). *1992 Statistical yearbook.* New York: United Nations.

United Nations Statistical Office. (1992). *Statistical yearbook 1990-1991.* New York: United Nations.

Index