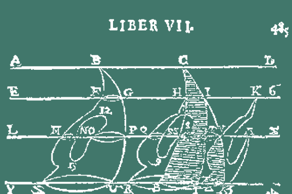
Annals of the Canadian Society for History and Philosophy of Mathematics/
Société canadienne d'histoire et de philosophie des mathématiques

Maria Zack
Dirk Schlimm
Editors

Research in History and Philosophy of Mathematics

The CSHPM 2019-2020 Volume



Birkhäuser

Birkhäuser

Annals of the Canadian Society for History and Philosophy of Mathematics/ Société canadienne d'histoire et de philosophie des mathématiques

Series Editor

Maria Zack, Dept of Mathematical Info & Comp Sci, Point Loma Nazarene Univ, San Diego, CA, USA

The books in the series contain selected papers writttten by members of the Canadian Society for History and Philosophy of Mathematics. Founded in 1974, this society promotes research and teaching in the history and philosophy of mathematics, as well as in the connection between the two. Volumes in this series cover a broad range of topics from a variety of time periods and cultures. They will be accessible to anyone who has had exposure to mathematics at the university level and will appeal to scholars of the history and/or philosophy of mathematics, graduate and undergraduate students undertaking research projects, and anyone with a general interest in mathematics.

More information about this series at https://link.springer.com/bookseries/16576

Maria Zack • Dirk Schlimm
Editors

Research in History and Philosophy of Mathematics

The CSHPM 2019-2020 Volume

Birkhäuser

Editors
Maria Zack
Point Loma Nazarene University
San Diego, CA, USA

Dirk Schlimm
McGill University
Montreal, QC, Canada

ISSN 2662-8503 ISSN 2662-8511 (electronic)
Annals of the Canadian Society for History and Philosophy of Mathematics/ Société canadienne d'histoire et de philosophie des mathématiques
ISBN 978-3-030-95200-6 ISBN 978-3-030-95201-3 (eBook)
https://doi.org/10.1007/978-3-030-95201-3

This book is published under the imprint Birkhäuser, www.birkhauser-science.com by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume contains 11 papers that have been complied by the Canadian Society for History and Philosophy of Mathematics. These papers provide some interesting insights into contemporary scholarship in the history and philosophy of mathematics, and the teaching of the history of mathematics.

The volume begins with Joel Silverberg's paper on Nathaniel Torporley (1564–1632), "The Most Obscure and Inconvenient Tables Ever Constructed?" In this paper, Silverberg describes both the astronomy and mathematics contained in Torporley's only published work *Diclides Coelometricae, seu valvae astronomicae universal*. Joel Silverberg died in August of 2019, making this paper the last in a long list of publications. Joel was an excellent scholar who was devoted to his family and his friends, and he is deeply missed.

The volume continues its exploration of seventeenth century mathematics with Duncan Melville's "Commercializing Arithmetic: The Case of Edward Hatton." For a period of roughly 40 years from 1695 onwards, more than 40,000 copies of Hatton's (ca. 1664–1733) arithmetic books were sold. In this paper, Melville offers a survey of Hatton's works with an emphasis on those that were key to his commercial success.

The next two papers focus on French mathematics in the nineteenth century. In "Leading to Poncelet: A Story of Collinear Points," Christopher Baltus discusses Jean-Victor Poncelet's (1788–1867) *Traité des propriétés projectives des figures*. Baltus examines the wide variety of mathematical work on collinear points that provided the foundation for Poncelet's 1822 treatise. This is followed by Roger Godard's paper "Cauchy, Le Verrier et Jacobi sur le problème algébrique des valeurs propres et les inégalités séculaires des mouvements des planets." Godard looks at the connections between work of three mathematicians Urbain Le Verrier (1811–1877), Carl Gustav Jacob Jacobi (1804–1851), and Augustin-Louis Cauchy (1789–1857) on the algebraic eigenvalue problem.

Amy Ackerberg-Hastings continues the focus on nineteenth century mathematics in the paper "Mathematics in Astronomy at Harvard College Before 1839 as a Case Study for Teaching Historical Writing in Mathematics Courses." In this paper, Ackerberg-Hastings considers how mathematics was employed in the teaching of astronomy at Harvard before the college established an observatory in 1839. The author also uses this history at Harvard as a framework for considering how to engage current undergraduates in historical research. In "Lectures for Women and the Founding of Newnham College, Cambridge," James J. Tattersall and Shawnee L. McMurran study the opportunities for women to study mathematics at Cambridge on the late 1800s. In 1869, the Cambridge Examination for Women was established with the intent to certify a candidate's qualifications for teaching. To help women prepare for this exam, a lecture series was established and many of Cambridge's leading mathematicians participated. This paper highlights the first 10 years of those lectures.

Then next two papers focus on the philosophy of mathematics. David Waszek's paper "Are Euclid's Diagrams 'Representations'? On an Argument by Ken Manders," considers Ken Manders' (2008) argument against conceiving of the diagrams in Euclid's *Elements* in "semantic" terms (as representations). Waszek shows some of the limitations of Manders' argument while also suggesting that Manders makes a compelling case that semantic analyses ought to be relegated to a secondary role for the study of mathematical practices. Bernd Bult's paper "Abstraction by Embedding and Constraint Based Design," considers the traditional approach to concept formation via abstraction arguing that this approach does not provide an adequate model for concept formation in mathematics. Bult's paper contributes to the ongoing discussion about how the notion of concept formation can be brought into alignment with mathematical practice and experience.

The final three papers in the volume examine various aspects of undergraduate mathematical education. Walter Meyer's "The Birth of Undergraduate Modern Algebra in the United States," charts how courses in modern algebra slowly moved into the curriculum of American universities in the twentieth century. Meyer's paper considers some of the historical factors that may have contributed to slowing down the development of these courses.

In "History as a Source of Mathematical Narrative," Po-Hung Liu describes some of the ways that mathematical narrative can be used to help students understand mathematics. In this paper, Liu provides a detailed example of how mathematical narrative writing has been used in a mathematics course for students who were not earning a degree in mathematics. In "Thoughts on Using the History of Mathematics to Teach the Foundation of Mathematical Analysis," Fairouz Kamareddine and Jonathan Seldin discuss a strategy for avoiding the use of "formal definitions" in a transitional analysis course until after students have seen a number of examples. The paper is rich in examples that illuminate the authors' suggested approach to teaching analysis to third-year undergraduates.

This collection of papers contains several gems from the history and philosophy of mathematics, which will be enjoyed by a wide mathematical audience. Many of these papers were written after the global coronavirus pandemic began in late

2019. Each of the authors was dedicated to completing their work and providing some scholarly light in a time of global darkness. As editors, we are grateful to the authors and the referees who brought volume in being.

San Diego, CA, USA Maria Zack
Montreal, QC, Canada Dirk Schlimm

Editorial Board

Contents

Contributors

Amy Ackerberg-Hastings Rockville, MD, USA

Christopher Baltus Department of Mathematics, SUNY College at Oswego, Oswego, NY, USA

Bernd Buldt Department of Mathematical Sciences, Purdue University, Fort Wayne, IN, USA

Roger Godard Department of Mathematics and Computer Science, Royal Military College of Canada, Kingston, ON, Canada

Fairouz Kamareddine School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK

Po-Hung Liu National Chin-Yi University of Technology, Taichung, Taiwan (R.O.C.)

Shawnee L. McMurran Department of Mathematics, California State University San Bernardino, San Bernardino, CA, USA

Duncan J. Melville St. Lawrence University, Canton, NY, USA

Walter Meyer Adelphi University, Garden City, NY, USA

Jonathan P. Seldin Department of Mathematics and Computer Science, University of Lethbridge, Lethbridge, AB, Canada

Joel S. Silverberg Roger Williams University, Bristol, RI, USA

James J. Tattersall Department of Mathematics, Providence College, Providence, RI, USA

David Waszek Department of Philosophy, McGill University, Montreal, QC, Canada

The Most Obscure and Inconvenient Tables Ever Constructed?

Joel S. Silverberg

Abstract Nathanial Torporley (1564–1632) published only one work, the *Diclides Coelometricae, seu valvae astronomicae universale*. In it he laid out a theory of astrology, then turned to spherical trigonometry problems that arise in that theory. The two tables he designed to solve these problems were denigrated by those who came after him and were even called "the most obscure and the most inconvenient tables ever constructed." I will outline his astronomical work, then set the scene for the spherical problems his tables were designed to solve, showing how these tables were to be used, with examples drawn from his text. I conclude that while the tables were quickly superseded, they did, in fact, serve Torporley's expressed purpose.

1 Introduction

With an Oxford degree and close ties to both François Viète and Thomas Harriot, Nathaniel Torporley (1564–1632) was one of the most enigmatic contributors to the mathematics of the late sixteenth and early seventeenth centuries. Only a generation

---

Joel S. Silverberg is deceased at time of publication.
Note: Joel Silverberg died on August 11, 2019, at the age of 73, making this paper the last in a long list of publications. Joel earned four degrees from Brown University, in music and engineering, and was a faculty member of Vassar College, Boston University, Brown University, and Roger Williams University. He took pleasure in being a Mainer, from Bangor, and in his Jewish heritage. His wide range of pursuits included music, sailing, birding, and Jewish studies, and he was a valued consultant on technical aspects of George Washington's papers. Much of his work addressed the history of navigational and practical mathematics, interests evident in this paper, with its roots in astrological calculations and the mathematics of the celestial sphere. Joel was devoted to his family and his friends, and he is deeply missed.

J. S. Silverberg (deceased)

M. Zack, D. Schlimm (eds.), *Research in History and Philosophy of Mathematics*,
Annals of the Canadian Society for History and Philosophy of Mathematics/Société canadienne d'histoire et de philosophie des mathématiques,
https://doi.org/10.1007/978-3-030-95201-3_1

later, John Aubrey, in his *Brief Lives*, (Aubrey, 1898) conveyed how little was known about him, reporting:

> Mr. Hooke affirms to me that Torporley was amanuensis to Vieta; but from whom he had that information he haz now forgot, but he had good and credible authority for it, and bids me tell you that he was certainly so . . . [and that] Mr. Nicholas Mercator assures me that the earle of Northumberland who was prisoner in the Tower gave also a pension to one Mr. Torporley, Salopiensis, a learned man; and that in the library of that family at Petworth, are some papers of his.

Torporley lay at the crossroads between British and Continental mathematics, between scholars who enjoyed noble or royal patronage and scholars who trained and functioned within the universities, between those who studied mathematics for its own sake and those who sought practical applications in the real world, between the Renaissance mathematics of the fifteenth and early sixteenth centuries and the mathematics of the early modern period.

Torporley published only one mathematical work, the *Diclides Coelometricae, seu valvae astronomicae universale* (Torporley, 1602), whose double-barrelled title essentially repeats itself. *Diclides Coelometricae* is Latinized Greek for "Double-folding doors of the measurement of the heavens," while *valvae astronomicae universale* is Latin for "bifold doors having general applicability to astronomy." In this 1602 book he presented his two "doors," the *Quadrans* and the *Quincunx*, two lengthy sets of tables relating to spherical trigonometry and astrology.

Several writers thought the work was essentially incomprehensible. For instance Jean Baptiste Joseph Delambre wrote in his *Histoire de l'astronomie moderne* (Delambre, 1821) of the *Diclides*:

> Cette dernière est composée de cinq parties différentes. Ces deux tables peuvent passer à bon droit pour les plus obscures and les plus incommodes qui aient jamais été construites.[1]

In his article on Viète in the *Penny Cyclopedia* (De Morgan 1837a), Augustus De Morgan mentioned that Torporley had preceeded by twelve years Napier's well-known circular parts mnemonic for spherical triangles, but called Torperley's version, "the greatest burlesque on mnemonics we ever saw." The title page of the *Diclides*, Fig. 1, includes several of Torporley's whimsical figures representing spherical triangles, both inside the miter-like diagram the two men are holding, and in the triangles the men are forming with their bodies. Some of these will be discussed below, a fuller discussion appeared in Silverberg (2009).

In a later publication (De Morgan, 1843b) De Morgan gave a cursory description of the nature of Torporley's *Quincunx*, but then threw up his hands, saying, "Those who like such questions may find out the meaning of the other table." Details of De Morgan's commentary may be found in Silverberg (2017).

To my knowledge, no one till now has explored these tables, nor has anyone commented on or explained the theorems Torporley developed. I will discuss the

[1] This work is made up of five separate parts. These two tables may rightly seem to be the most obscure and the most inconvenient that have ever been constructed.

DICLIDES COELOMETRICÆ
seu
VALVÆ ASTRONOMICÆ
VNIVERSALES

Omnia artis totius munera Psephophoretica in sat modicis finibus duarum Tabularum Methodo noua, generali, & facilima continentes.

Praeunte Directionis accurata consumata Doctrina, Astrologiæ hactenus plurimùm desiderata.

Authore NATHALE TORPORLAEO Salopiensi in secessu Philotheoro.

Fragilis & laboriosa mortalitas in partes ista digessit, vt portionibus quisque coleret, quo maximè indigeret; PLIN.
Sed quæ carptim singula constant
Eadem nos hic mista iugamus. BOET.

LONDINI
Excudebat FELIX KINGSTON. 1602.

Fig. 1 Title page of the *Diclides*

astrological and mathematical settings for the problems these tables are designed to solve, then show how Torporley used them.

The *Diclides* developed a novel and byzantine scheme for applying plane and spherical trigonometry to refine methods of natal or "genethliac" astrology—the preparation and interpretation of a natal chart intended to reflect the influence of planets, signs, and "houses" upon the subject's life and character. Torporley included his reinterpretation of Ptolemaic house boundaries as suggested by Johannes Müller (also known as Regiomontanus), and the calculation of Circles of Position and Arcs of Direction to help predict the timing of possible or likely events, life span, and time of death.

2 Overview of Torporley's Astrology

While modern mathematicians and scientists no longer consider astrology a valid subject for study, it did in fact serve as a major historical impetus for much of the astronomical calculations and related mathematics still in use today. It certainly was the major reason Torporley studied the mathematics that he did. Those who are more interested in the part of Torporley's work that bears directly on spherical trigonometry may prefer to skip this section, or give it a superficial glance.

Nathaniel Torporley's views on astrology were likely developed between his years at Oxford (1581–1584) and his writing of the *Diclides* (1602). They were based on the classical four-part model of Ptolemy described in the *Tetrabiblos* (Ptolemy c 160), as modified by Regiomontanus in his 1496 *Epitome of the Almagest* (Regiomontanus, 1533). Within the *Diclides* Torporley frequently referred to both sources as well as the contributions of Georg Joachim Rheticus, Bartholomeus Pitiscus, Jean-Baptist Morin, Thomas Fincke, Nicolaus Copernicus, and Erasmus Reinhold.

The natal chart, as drawn by modern astrologers, is a circle of the heavens, representing the path of the sun against the fixed stars over the course of one year. The chart is divided into 12 signs of 30 degrees each, thus 360 equal parts. The first sign begins at the point of the ecliptic, the great circle that is the apparent path of the sun through a year, marking the spring equinoctial point or zero point of Aries (the vernal equinox). When the sun is in that position, the length of day and night is equal at all points on the Earth. Successive signs are named after the well-known constellations of the zodiac. In Fig. 2 the signs appear in the outer, tinted wheel.

Figure 2 also includes the boundaries of 12 "houses," based upon divisions of a single day. The movement of the sun along the ecliptic takes one year; the movement of the ecliptic through the sky takes one day. Over the course of a day, the planets (except the moon, which astrologers sometimes considered a planet) will not move perceptibly with respect to each other, but the entire ecliptic and all of the planets placed within it will revolve 360 degrees with respect to these houses, which are defined in terms of the horizon.

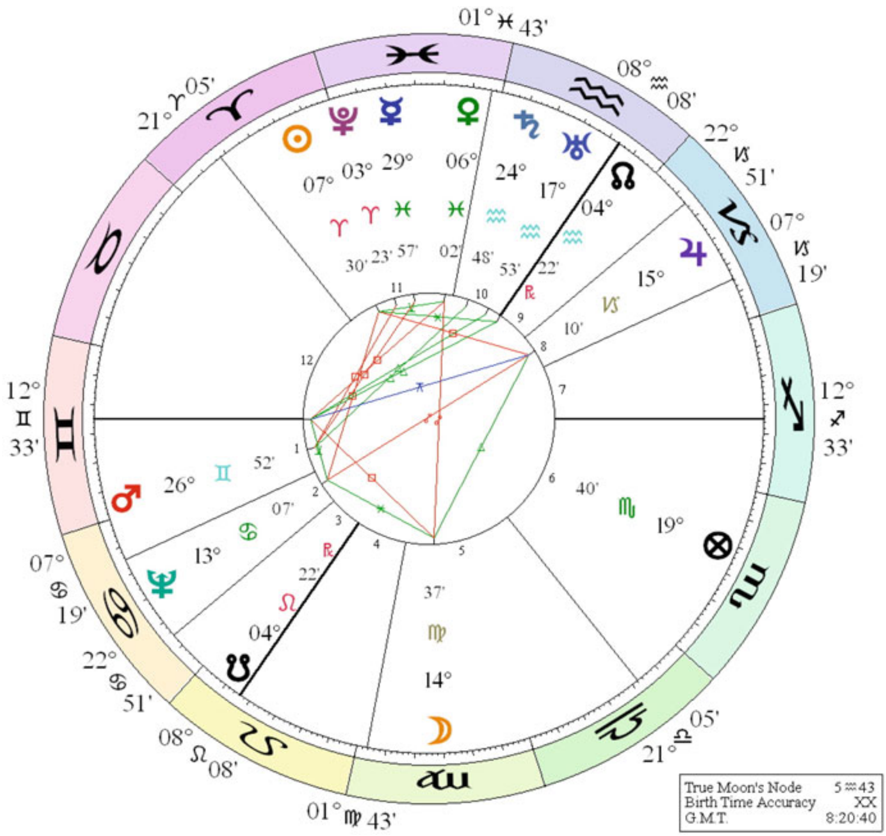


Fig. 2 March 28, 1581, 8:20 am, latitude 51° N. House divisions by the method of Regiomontanus

The first house starts at the position in the zodiac rising over the eastern horizon at the specific moment for which the chart is cast, or from the particular latitude and longitude of the birth or other event. How far (i.e., how many degrees) the ecliptic rises over the horizon in a two-hour period of time determines the number of degrees within the house. Since the ecliptic is oblique to the equator some signs will rise more quickly than others, the various houses will differ in their degree, measure. The houses in Fig. 2 are numbered from 1 to 12 on the inner wheel.

The sun, moon, and planets are charted in accordance with their position on the ecliptic (their ecliptic longitude), based on the sign on the outside wheel.

The four cardinal points on the wheel, called the *angles,* comprise the ascendant, the midheaven (*medium coeli*), the descendant, and the lowest heaven (*immum coeli*.) The line segment between the twelfth and the first houses marks the ascendant. The line segment between the sixth and seventh houses marks the descendant. The horizontal diameter that these segments form is the horizon, with the sky in the upper half-circle visible and that in the lower half-circle invisible. The ascendant marks the spot on the ecliptic just rising over the horizon at the time for which the chart is cast. The descendant marks the spot on the ecliptic setting below the horizon at that same time.

The line segment between the ninth and tenth houses marks the midheaven or *medium coeli*. The line segment between the third and fourth houses marks the lowest heaven or *immum coeli*. These points mark the intersection of the ecliptic with the meridian, the *medium coeli* being the culmination or highest point of the ecliptic above the horizon, and the *immum coeli* being the anticulmination or lowest point. The position of the *medium coeli* and *immum coeli* will shift over the course of the day, but the horizon will not. This will have a marked effect on the size of the houses, depending on how fast or slow the signs are to rise and set at the particular time of day. Now, when we say the houses are given by the number of degrees of the ecliptic rising above the horizon in two hours, we must use seasonal hours, where the length of a day is divided into 12 seasonal hours and the length of a night is also divided into 12 seasonal hours, regardless of the latitude of the observer or the season of the year.

In Torporley's day the natal charts were the same as given by Regiomontanus, except that they were placed in a square divided into 12 equal triangles, with an inner square containing details, such as the date, time, location of birth, and name of the subject. See Fig. 3.

Torporley's interest in astrology led to his desire to find effective solutions of trigonometric problems.

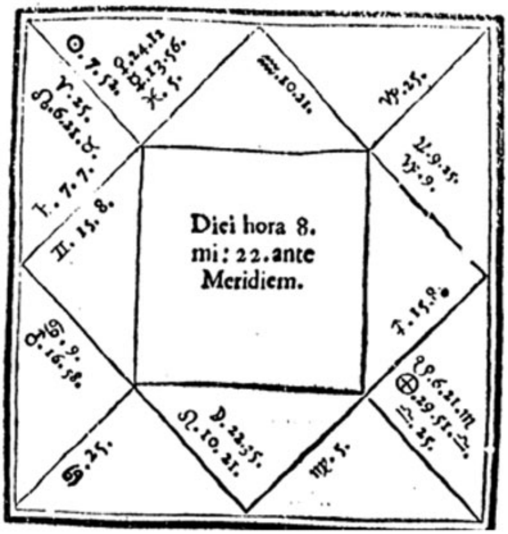


Fig. 3 Natal chart, date unspecified. Sun at 7°51′ of Gemini, 8:22 am, latitude 51° N

3 All Circles Great and Small

Torporley's astrology drew heavily upon Regiomontanus, but on his astronomical work, not his astrology. Regiomontanus was primarily an astronomer and a mathematician with only peripheral works and tables dealing with astrological considerations. His masterpiece, *De triangulis omnimodis* (Regiomontanus, 1533), provided a carefully built foundation for spherical trigonometry, commencing with Book III, which concerns spherical geometry. Regiomontanus stated all of the theorems and conditions needed for the various coordinate systems used to specify the position of heavenly objects in the celestial sphere.

Here are several important theorems from that book.

- Theorem 1. If a sphere is intersected by a plane, the line of intersection of the spherical surface and the intersecting plane will be the circumference of a circle. The intersecting plane may pass through the center of the sphere or it may not.
- Theorem 16. A great circle passing through one pole of another circle (great or small) will also pass through the other pole of that circle.
- Theorem 17. A great circle in a sphere, passing through the pole of another circle (great or small), will bisect the other orthogonally.
- Theorem 19. All great circles in a sphere bisect each other.
- Theorem 22. Circles in a sphere that have the same poles are parallel to each other, and parallel circles will have the same two poles.
- Theorem 23. In a sphere any two great circles intercept similar arcs from parallel circles through whose poles they pass.

3.1 Coordinate Systems

Theorem 17 of Regiomantus allows the astronomer to choose a pair of great circles that pass through the poles of each creating a coordinate system. It is often useful to convert the position of a celestial object from one system to another. The three coordinate systems most commonly used in astrology are listed below and illustrated in Fig. 4.

- The *horizon* (with poles at zenith and nadir) can be paired with the *prime vertical* (with poles at the east and west points of the horizon) to define position in terms of azimuth and altitude.
- The *equator* (with poles at the north and south celestial poles) can be paired with the *meridian* (with poles at the vernal and autumnal equinox, which lie upon the equator and also the ecliptic) to define position in terms of right ascension and declination.
- The *ecliptic* (with poles at the north and south ecliptic poles) can be paired with the great circle that passes through the ecliptic poles and through the vernal

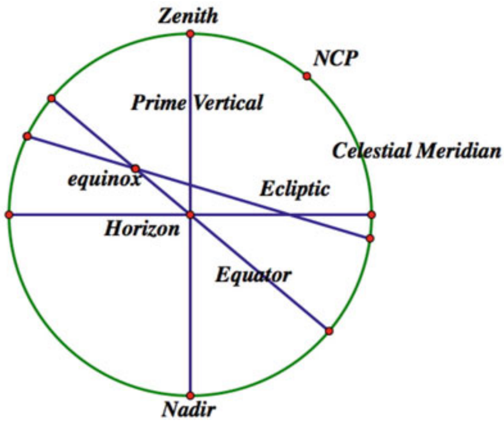


Fig. 4 Celestial meridian, celestial equator, and prime vertical

and autumnal equinoxes (which lie upon the ecliptic and the equator) to define position in terms of ecliptic longitude and ecliptic latitude.

4 The Law of Tangents

The unifying principle of the doctrine of triangles is that in any right-angled triangle the radius may be chosen to be any of the three sides, and that in any similar triangles the ratio of corresponding sides will be equal (Fincke, 1583; Pitiscus, 1600 [2]).

In the left triangle of Fig. 5, the radius is the hypotenuse, BC is the sine of A. In the middle triangle, the base is the radius, the perpendicular is the tangent of A, and the hypotenuse is the secant of A. In the right-most triangle the perpendicular is the radius, the base is the tangent of the complement of A, and the hypotenuse is the secant of the complement of A.

The next consequence of the doctrine of triangles is the key to Torporley's *Quincunx* table. It is the law of tangents. In modern form, given sides a and b with corresponding opposite angles A and B the law is

$$\frac{a+b}{a-b} = \frac{\tan\frac{A+B}{2}}{\tan\frac{A-B}{2}}.$$

The evolution of the statement of this theorem indicates some of the difficulties of reading old texts. Here are three versions, the third being a good translation of either Latin version. First, Fincke wrote:

[2] This work introduced the word trigonometry.

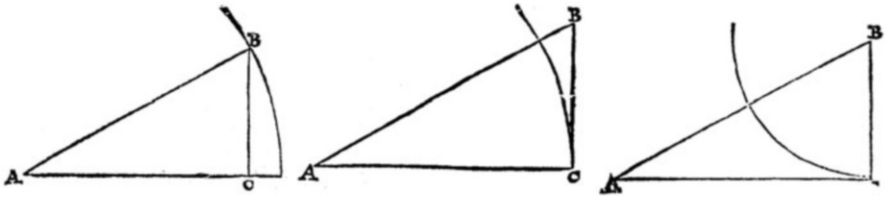


Fig. 5 Three possibilities for the radius

> Et ut semissis summae crurum ad differentiam summa semissis alteriusquae cruris, sic tangens semissis anguli crurum exterioris ad tangentem anguli quo minor interiorum semissee dicti reliqui minor est, aut major, major. (Fincke, 1583, Book X proposition 15)

Fincke seems to have been the first to use the word tangent in its trigonometric sense in print.

The same theorem appears in Book III of the *Trigonometria* (Pitiscus:1600, Axiom V), in slightly more compact form:

> *In Triangulis planis universis:*
> Ut summa duorum laterum ad differentiam eorundem: ita tangens dimidii summae duorum angulorum oppositorum, ad tangentem differentiae infra vel supra dimidium.

This said the same thing as Fincke, but used different terminology.

Finally comes R. A. Handson's English version of Pitiscus's Axiom V *Trigonometria* (Handson 1614), which he renumbered as axiom 3:

> In all plaine triangles. As the somme of the two sides is to their difference; So is the Tangent of halfe the somme of the two angles opposite, to the Tangent of the difference, lesse or more then the halfe.

The only part that might cause us difficulty is the phrase "tangent of the difference, less or more than the half." In all three books the examples immediately following the statement of the theorem make it clear the "half" is $(A + B)/2$. Then, supposing $A > B$, the "difference" is A minus the half or the half minus B, in either case $A - B$. The choice of naming the angles avoids the problem of negatives.

This theorem may be used to solve a triangle given two sides and the included angle, to find the other angles. The sum of the opposite angles is thus 180° minus the given angle. Now three parts of the law of tangents have been given. Solving the proportion and taking the arctan gives a numerical value for $(A - B)/2$, which is then added to the half $(A + B)/2$ to get A and B. Torperley would have known this from his reading of Fincke and Pitiscus.

4.1 Spherical Trigonometry

Spherical triangles are formed by the intersection of the arcs of three great circles. Since any arc of ecliptic longitude will intersect the ecliptic at a right angle, any line of declination will intersect the equator at a right angle, and any vertical circle will intersect the horizon at a right angle. Thus the spherical triangles involved usually have at least one right angle.

Spherical trigonometry cannot be used directly to solve problems involving small or lesser circles. Since all diurnal arcs, with the sole exception of the equator, will be small circles, astrologers developed a variety of methods for projecting positions on these diurnal arcs onto a great circle. They used the equator if seeking the time of an event and the ecliptic if seeking the sign. A common method had been to project the desired object onto the plane of the ecliptic along an arc passing through the object and the north and south celestial poles.

Regiomontanus instead suggested identifying the position of an object by an arc passing through the object and the north and south points of the horizon. This system rests on a brief passage in the *Tetrabiblos* in which Ptolemy defined his understanding of "similar places" in the arcs of the different planets:

> For a place is similar and the same if it has the same position in the same direction with reference both to the horizon and to the meridian. This is most nearly true of those which lie upon one of these semicircles which are described through the sections of the meridian and the horizon, each of which at the same position makes nearly the same temporal hour. (Robbins, 1940, p. 290)

What these procedures mean for astrology lies beyond the scope of this paper, which is concerned with Torporley's spherical triangle tables. Although Torporley explained at length how one might use these tables, he did not explain how he derived them. His work contained no syncopated or symbolic algebra of any kind. All explanations were made through either rhetorical algebra or, more often, a detailed verbal discussion of applications illustrated by numerical examples of their solutions. Full mastery of the *Diclides* would require a deep knowledge of astrological, astronomical, and mathematical language of the era, as well as the Latin language.

5 Torporley's Miter and Menelaus's Figure

The arrangement now called Menelaus's figure (Fig. 6) is present in Ptolemy's *Almagest*. On his title page, Torporley has a drawing composed of two overlapping Menelaus figures (Fig. 7). He calls this a miter for its resemblance to a bishop's hat. It is intended as mnemonic for his spherical triangle terminology, discussed below.

The Menelaus figure led to two theorems, both present in the *Almagest* Book I, Section 13 (Ptolemy c. 150). Neugebauer (Neugebauer, 1975) named them Menelaus Theorem I and Menelaus Theorem II, which we abbreviate MT I and MT II. Symbolically they are:

$$\frac{\sin(a+b)}{\sin a} = \frac{\sin(g+h)}{\sin g} \cdot \frac{\sin f}{\sin(e+f)} \qquad \text{(MT I)}$$

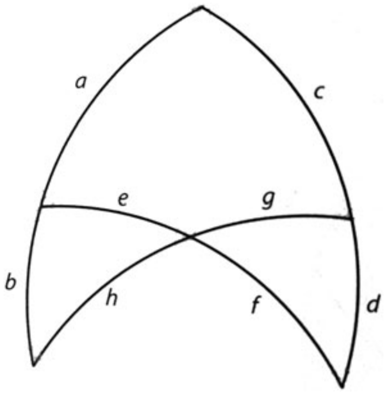


Fig. 6 The Menelaus figure. Four quadrantal great circle arcs: $a + b,\ c + d,\ e + f,\ g + h$. The vertex at the intersection of b and h is the pole of arc $c + d$. The vertex at the intersection of arcs d and f is the pole of arc $a + b$. Hence $e + f$ and $a + b$ are orthogonal as are $h + g$ and $c + d$

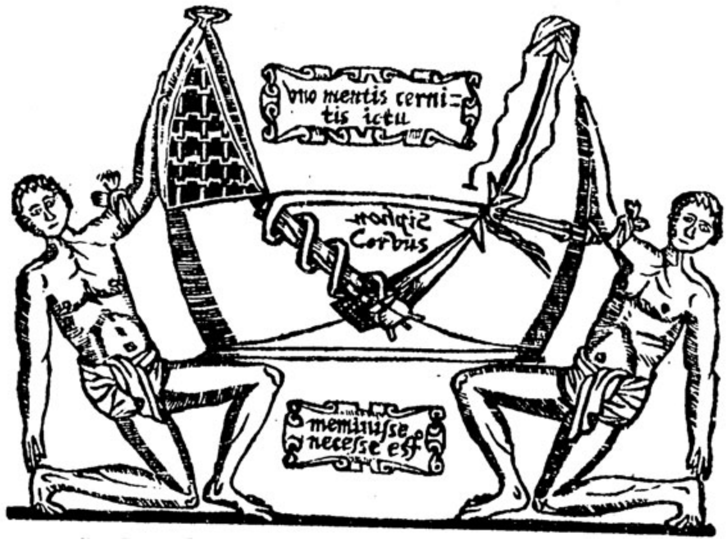


Fig. 7 The Miter from the title page of the *Diclides*. In the middle are the words *Corvus* and (read from the back) *Siphon*. The upper and lower boxes read, "By one stroke of the mind you see how to remember what you need." This refers to the diagram itself. Note that the two men form several triangles with their limbs

$$\frac{\sin a}{\sin b} = \frac{\sin(c + d)}{\sin d} \cdot \frac{\sin g}{\sin h} \tag{MT II}$$

As Torporley would have known, any right-angled spherical triangle for which two parts are known in addition to the right angle can be solved for the other parts by using one of these two relationships.

5.1 Considering MT I

The Menelaus figure (Fig. 6) described a spherical triangle with $e + f$ orthogonal to $a + b$ and $h + g$ orthogonal to $c + d$, as discussed in the caption. For his miter, Torporley further assumed that $a + b$ is a quadrant, as are $g + h$ and $e + f$. The sine of a quadrant is the radius R, also called the *total sine*. In earlier times R was taken however was convenient for a given table, so the sines were large whole numbers. Today we take $R = 1$, so all our sines lie between -1 and 1. With $R = 1$ MT I can be expressed as

$$\sin g = \sin a \cdot \sin f.$$

Knowing values for any two of a, f, and g, the third can be determined.

Sides f and g appear in Fig. 6 as the hypotenuse (f) and an adjacent side (g) of the triangle gfd. Side a in the upper part of the figure is equal to the angle in triangle gfd opposite side g. This is the figure that Torporley called the *Siphon*: hypotenuse, adjacent side, and the angle opposite that side. The equation above is the generator of his *Quincunx* table.

5.2 Considering MT II

In Ptolemy's time, the only trigonometric function was the chord, and every arc was associated with a particular chord. The chord and the sine are much the same function, the sine being half of the chord of double the arc.

From Ptolemy on, calculations frequently required dividing the chord of an arc by the chord of its supplement—or in Renaissance terms, dividing the sine of an arc by the sine of its complement—yet neither era thought of this quotient as a different function, nor gave it a different name. It was simply an often-needed computation. To eliminate repetitive calculations, eventually auxiliary tables were constructed relating an arc to the quotient of the sine of the arc divided by the sine of its complement. The concept of the tangent was eventually introduced with the development of the doctrine of triangles, as discussed above.

Sides a and b in Fig. 6 are complements, sides g and h are also complements, and $c + d$ is a quadrant. Thus, again taking $R = 1$, MT II can be written as

$$\tan a = \frac{1}{\sin d} \cdot \tan g$$

or, more simply,

$$\tan g = \sin d \cdot \tan a.$$

Sides d and g appear in the lower left part of Fig. 6, but side a does not. Since the angle between sides f and d is a pole of side a, the two are equal.

In the triangle gdf, we have isolated two legs (sides other than the hypotenuse) and one of the acute angles. The tangent of the side opposite an acute angle is equal to the product of the sine of the angle and the tangent of the adjacent side. This is precisely the rule of Torporley's figure of the *Corvus*, the crowbar, the basis for the *Quadrans* table.

6 The Spherical Right Triangle

The *Quadrans* and the *Quincunx* are Torporley's two sets of tables for solving right spherical triangles. In following his approach, I will use modern notation when convenient, as illustrated in Fig. 8, with c for hypotenuse opposite the right angle C; A and B for the other angles; and a and b the respective opposite sides.

Torporley gave names to each of the triplicities, or sets of three parts of the triangle, along with sometimes amusing illustrative diagrams. For example the triplicity of the three sides (a, b, and c) he called the *Carcer* or prison (Fig. 9). In each case given two of the items the third could be found using his tables.

7 The *Quadrans* or Right Gateway

The *Quadrans* is the simpler table. It is 12 double pages with 35 rows and 20 columns on each double page. The row number he called *latus* (side), while the column number was the *caput* (head); both are in degrees. The cell numbers are written in degrees and minutes and are given different names depending on the triplicity, aligned under G for *Gradus* (degree) and M for minutes. If the degree in a cell is missing, it is taken to be the same as that in the row above it. In using the

Fig. 8 Spherical triangle

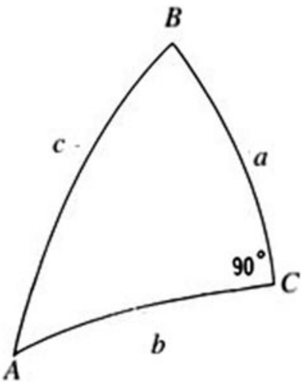
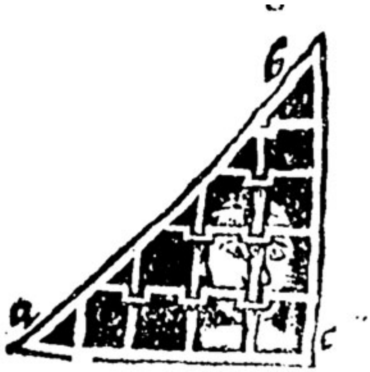


Fig. 9 The *Carcer* (Prison)



cells, I will identify them by the left side row number and the top column number. The defining rule for the table is

$$tan(row\ number) \cdot sin(column\ number) = tan(cell\ number).$$

We saw a version of this with MT I.

The pages shown in Fig. 10 show this array, with a typical cell in Fig. 12. The row number gives one angle, perhaps A. Then the cell number is the side opposite that angle a, and the column number is the other side b. (Or, starting with the row as B, the cell would have b and the column a.) Thus from the second row of Fig. 10, $A = 2°$, $b = 41°$, and $a = 1°19'$. In practice, given any two of these parts, the *Quadrans* can be used to solve for the third.

This triplicity, two sides other than the hypotenuse, and an angle opposite one of those sides, is the *Corvus* (Fig. 11). The drawing shows not the crow but a crowbar.

A useful computational feature is built into the table. The left and right labels of each row are complementary, as are the top and bottom labels of each column. If the right row number is used for B, and the bottom column number for a, then the cell number is the complement of the hypotenuse. This allows us to solve other configurations.

A *Corvus* example (Torporley, 1602, p. 105) gives one side as 20° and the other side as 50°, to find the angle opposite the first side. In the columns headed 50 we seek 20 among the cells below. Awkwardly, due to the size of the pages there are several columns headed 50, so we must search a while only to find 20 is not in the table. However, row 25 has 19 39 and row 26 has 20 29 (Fig. 12), and a bit of implicit interpolation gets us to 25 24 49, i.e., $25°24'49''$. Torporley was doubtless and comfortable navigating the 12-page *Quincunx*, but we can see why others might have been reluctant to follow him.

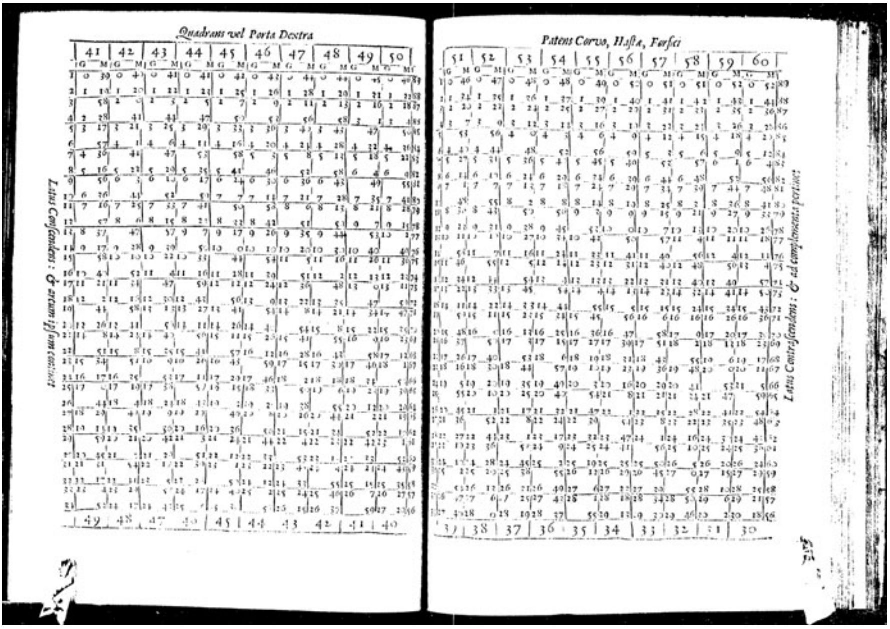
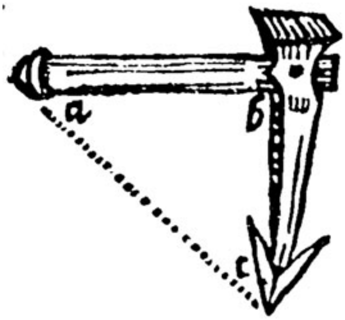


Fig. 10 A sample page from the *Quadrans*

Fig. 11 The *Corvus*



8 The *Quincunx* or Left Gateway

The *Quincunx* table (Fig. 13) is made up of 45 double pages, each with 18 rows and 15 columns. As in the *Quadrans*, row and column numbers are in degrees, and the cells are aligned under degrees, G, and minutes, M. The cells in this table contain three numbers. For example, Fig. 14 shows a small but typical section of the table where the first entry in the first cell is 25 25, meaning $25°25'$. The name of the table comes from the five numbers involved: row, column and three in the cell.

Note the cell entries in row 56, column 46. As mentioned before, with blank positions we carry down values from the corresponding place in the cell above.

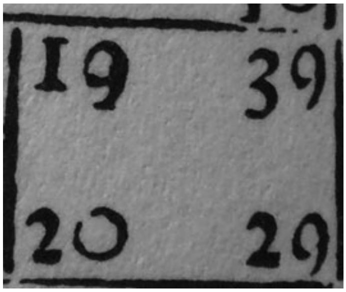


Fig. 12 *Quadrans* entry rows 25 and 26, column 50

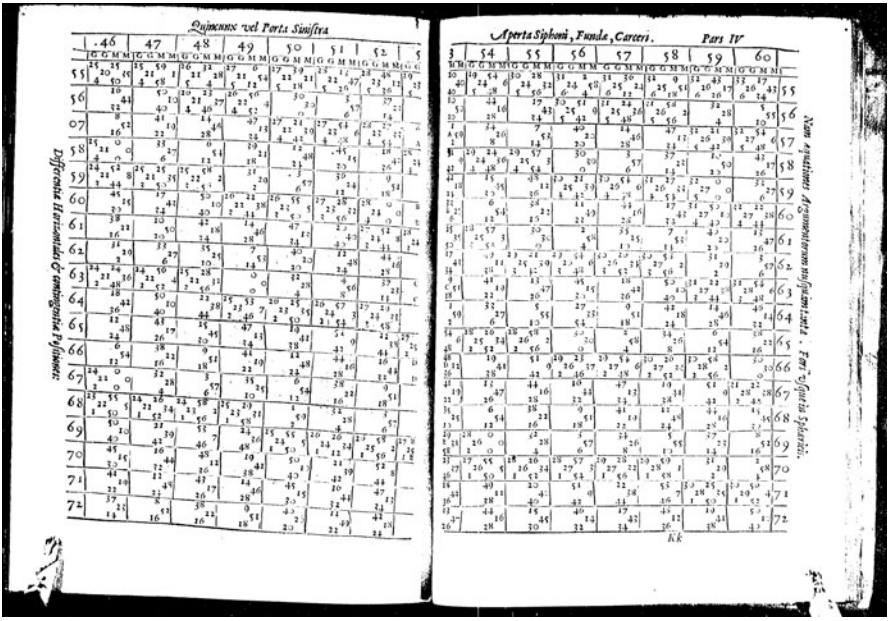


Fig. 13 A sample page from the *Quincunx*

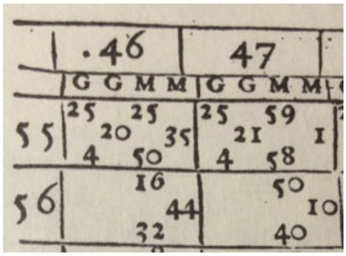


Fig. 14 Typical *Quincunx* entry from the upper left of Fig. 13

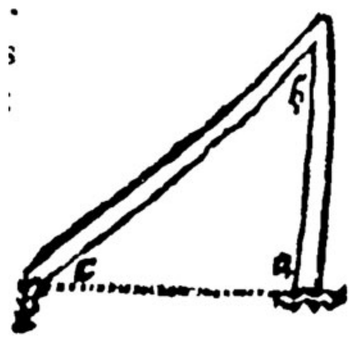


Fig. 15 The *Siphon*

Thus the first cell of row 56 represents (25) 16, above (20) 44, and at the bottom (4) 32. I will use parentheses to indicate values supplied in this way.

As in the *Quadrans*, Torporley called the row number the *latus* and the column number the *caput*. The three numbers in the cell were identified by their position: the topmost being the *supremum*, with the *medium* in the middle and the *infimum* below. From his text and examples, we can see what the table numbers mean. The *latus* is the hypotenuse c, the *supremum* is one of the angles A (or B), and the *medium* is the side opposite that angle a (or b).

Torporley gave the rule for the table as

$$\sin(latus) \cdot \sin(supremum) = \sin(medium), \text{ i.e.,}$$
$$\sin c \cdot \sin A = \sin a.$$

We have seen this equation derived from MT II. He also pointed out the identities:

$$medium + supremum = caput$$
$$supremum - medium = \mathit{infimum},$$

i.e.,

$$a + A = \text{column number}$$
$$A - a = \text{number at the bottom of the cell.}$$

Despite noting these identities, which are reminiscent of the law of tangents, Torporley did not explain their significance. I speculate they may be a clue to how the table was made.

The *Quincunx* is easiest to use if we are given the hypotenuse and a side to solve for the opposite angle, or given the hypotenuse and a given angle to solve for the opposite side. The triplicity of hypotenuse, side, and angle opposite forms what he called the *Siphon* (Fig.15), whose figure reminded him of a bent hydraulic pipe.

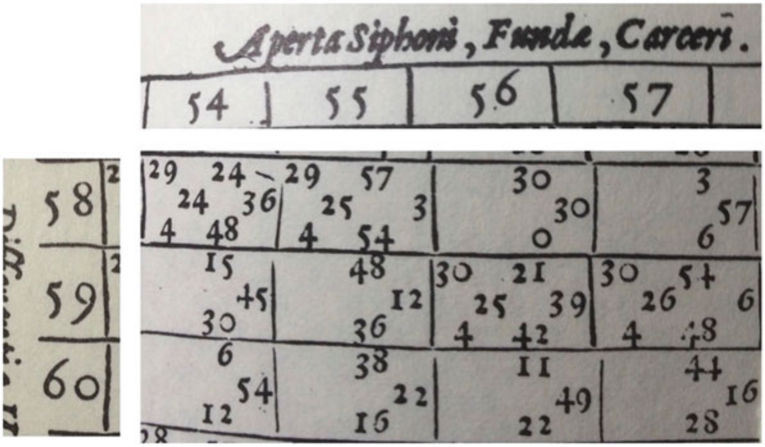


Fig. 16 Relevant parts of the *Quincunx* for the *Siphon* Example

For other cases, the *Quincunx* takes a bit more work, but despite Delambre's sneer, the table is actually fairly easy to use, even if the terminology is a bit clumsy. Let us consider Torporley's first example of the *Siphon* (Torporley, 1602, p. 106).

> It is proposed: That from the given hypotenuse 60 & the other angle 30 the side opposite the given angle is sought. Therefore this is a *Siphon*: Mother of the *Quincunx* & the first case of it. Thus the calf [*sura*] [3] is sought, and knowing what we do about the *Siphon* etc., that which was sought will be found in the *medium* to be 25 39 32, occupying the *latus* 60 & the supreme class 30. If vice versa we seek 30 in the *latus* and 60 in the *supremum* there will be found 25 39 30.

What is he up to? Using our terminology we are given $c = 60°$ and $A = 30°$ to find a. You can look along the table displayed in Fig. 13 if your eyes are excellent, otherwise consult Fig. 16. In row 60 (the *latus*) we want to find the *supremum* value 30, remembering that missing G values mean the value is copied from above. You will not find 30, but in column 55 the *supremum* is (29) 38, and in column 56 the *supremum* is (30) 11. Torporley must have interpolated the *medium* values in those cells, (25) 22 and (25) 49, to get 25 39 32. We are not sure how. Using simple linear interpolation, however, since 30 is 1/3 of the way from the column 55 value to the column 56 value in the *supremum* place, the *medium* would be 25 40 12, just 80 seconds difference.

The only slightly tedious part was scanning along the 60 row until finding the *supremum* values around 30.

Here is another example (Torporley, 1602, p. 105), involving the configuration, mentioned above that he called *Carcer*, the prison (Fig. 9). This triplicity is the three

[3] He often used anatomical terms like calf and forearm for the sides.

sides of the right triangle. In the figure he envisions some poor soul imprisoned by the three sides.

> Given one of the right sides 20 gr. [*gradus*, degrees] and the other side 36, the hypotenuse is sought.

We need not trace every comment in his solution. He took 70 (the complement of 20) to be the *latus* and 54 (the complement of 36) to be the *supremum*. The *supremum* value lies between the entry (53) 44 in column 103 and 54 17 in column 104. The *medium* values are (49) 16 and 49 43. He interpolated to get 49 29 1, thus the hypotenuse is the complement of this: 40 39 59.

Why use the complements? That comes from the cosine rule, known to Torporley,

$$\cos c = \cos a \cos b + \sin a \sin b \cos C,$$

with

$$\cos C = \cos 90^\circ = 0.$$

Torporley gives many more examples of the cases for the *Quincunx*, as well as many for the *Quadrans*. Using both tables we can solve all right spherical triangles and then even general spherical triangles, after splitting the triangle with a perpendicular from some angle to the side opposite it.

9 Conclusion

Industrious future workers may yet discover how Torporley created his mysterious tables. His awkward approach to solving right triangles was quickly superseded, and later writers poured scorn on his "doors," but as we have seen they actually work well enough for their intended purposes. Moderns often overlook the important role that astrology played in motivating the development of mathematical and astronomical tools. Moderns may also learn a lesson from Torporley's whimsical diagrams and terminology, for they remind us that while mathematics is serious business, it also can be great fun.

References

Aubrey John (1898) 'Brief Lives,' chiefly of Contemporaries, set down by John Aubrey, between the Years 1669 & 1696. Clark Andrew (ed) Clarendon Press, Oxford

Delambre Jean Baptister Joseph (1821) Histoire de l'Astronomie Moderne. Courcier, Paris

De Morgan Augustus (1843a) On the invention of the circular parts. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 22 350–353

De Morgan Augustus (1843b) Vieta. In: The Penny Cyclopedia of the Society for the Diffusion of Useful Knowledge 26. Charles Knight and Co, London

Fincke Thomas (1583) *Geometriærotundi libri XIIII.* Henricpetri, Basel

Neugebauer Otto (1975) A History of Ancient Mathematical Astronomy. Springer, Berlin

Pitiscus Bartolomeus (1600) *Trigonometriæ sive De dimensione Triangulæ.* Augsburg

Pitiscus Bartolomeus (1614) Trigonometrie or the Doctrine of Triangles, etc. Handson, R A (tr) Jo. Tappe, London.

Ptolemy, Claudius (c. 150) Almagest. Ptolemy's Almagest. See: Toomer G (tr), Princeton University Press, Princeton, NJ, 1998

Ptolemy, Claudius (c. 160) Tetrabiblos. See: Robbins

Regiomontanus (1533) (Johannes Müller von Königsberg) *De triangulis omnimodis.* Nuremburg

Robbins F E (tr) (1940) Ptolemy Tetrabiblos. Loeb Classical Library 435, Harvard University Press, Cambridge, MA

Silverberg Joel (2009) Nathaniel Torporley and his *Diclides Coelometricæ* (1602)—A Preliminary Investigation. In: Cupillari A (ed) Proceedings of the Canadian Society for the History and Philosophy of Mathematics 22 142–154

Silverberg Joel (2017) Napier, Torporley, Menelaus, and Ptolemy: Delambre and De Morgan's Observations on Seventeenth-Century Restructuring of Spherical Trigonometry. In: Zack M and Schlimm D (eds) Proceedings of the Canadian Society for the History and Philosophy of Mathematics, 149–168. Birkhäuser/ Springer, Berlin

Torporley, Nathaniel (1602) *Diclides Coelometricæ seu Valvæ Astronomicæ Universales.* Kingston, London

Commercializing Arithmetic: The Case of Edward Hatton

Duncan J. Melville

Abstract For a period of some 40 years from 1695 onwards, Edward Hatton was a dominant author in the burgeoning world of commercial arithmetic publishing in London. Hatton was a prolific as well as successful author, selling some 40,000 copies across a dozen titles, some of which went through many editions. In this paper, we offer a survey of his works with emphasis on those that were key to his commercial success.

1 Introduction

At the dawn of the eighteenth century, increasing trade and a growing mercantile class required an expansion of numeracy and mathematical education in England, and especially in London. England had far-flung trade networks, the best-known companies of this period being the East India Company and the Hudson Bay Company, but there were many smaller merchants, both sole proprietors and companies or syndicates, trading across Europe, the Near East, Africa, and North America.[1] Tracking and accounting for the flows of goods and money demanded the training of ever-increasing numbers of clerks, bankers, merchants, and accountants. They needed training materials and aids for their daily calculation. The specific, practical needs were for accounting and general commercial arithmetic for merchants and clerks, and, since trade in Britain was carried on the seas, navigation. While the development of the theory and practice of navigation, including the famous

[1] For a general orientation to English history of the period, see, for example, Hoppit (2000).

D. J. Melville (✉)
St. Lawrence University, Canton, NY, USA
e-mail: dmelville@stlawu.edu

M. Zack, D. Schlimm (eds.), *Research in History and Philosophy of Mathematics*, Annals of the Canadian Society for History and Philosophy of Mathematics/Société canadienne d'histoire et de philosophie des mathématiques,
https://doi.org/10.1007/978-3-030-95201-3_2

search for longitude, has been well-studied, the commercial side much less so.[2] Commercial arithmetic training was also a world away from the élite study of Euclidean geometry.

As the demand grew for educating clerks and "young merchants," the publishers and booksellers saw a potentially lucrative market.[3] To serve that market, they needed authors. One such was Edward Hatton, the subject of this study. In this paper, I essentially give a catalog of the wide range of Hatton's publications, with more detailed analysis of his major works, the *Merchant's Magazine*, and the *Comes Commercii*.[4] For reasons of space, it has not been possible to fully contextualize his work, and a detailed comparative analysis of Hatton, his predecessors, and competitors in the various sub-genres in which he worked must be deferred to the future. Hatton wrote in a clear and engaging, if at times loquacious, style. He had a passion for calculation, and a penchant for gathering and organizing information, often in the form of lists or tables, depending on the kind of information. As such, he rose from obscurity to prominence in the mercantile world of commercial arithmetic in London, before returning to obscurity in this somewhat neglected area of the history of mathematics.[5]

2 Edward Hatton and his Publishers

Little is known for certain about the background, origins, and education of Edward Hatton (1664–1737). He appears to have come from a Lancashire family of yeoman farmers.[6] How he came to be educated, we do not know; how he came to London, we do not know, although then as now, London drew in people from all across the

[2] The literature on the development of astronomy, astronomical tables, navigation, and navigational instruments in the seventeenth and eighteenth centuries is voluminous. For a recent access point, see several of the chapters in Flood (2020), and the references therein.

[3] The boundaries between what we might consider as the professions of printer, publisher, and bookseller were much more fluid and complex in the seventeenth and eighteenth centuries. If we think of a printer as someone with a printing press, a publisher as someone who (typically) holds the copyright to a printed work, and a bookseller as someone who sells printed material, it is clear that there is plenty of room for overlap. In particular, booksellers were often publishers in order to have books to sell and to trade to other booksellers in return for more stock. A good guide to navigating the world of London publishing is Raven (2007).

[4] Many of Hatton's publications have now been digitized and are available from *Eighteenth Century Collections Online* (ECCO), HathiTrust, and Google Books, as well as derivative print-on-demand publishers. Others are available on microfilm. The *English Short Title Catalogue* (ESTC) has a nearly complete listing of his works, including information on accessible digital or microfilm versions.

[5] For a recent paper on authors of practical mathematics in London in the period just before Hatton was publishing, see Beeley (2019).

[6] Baptized 19 Jun 1664 in Standish, Lancashire Lancashire Baptisms (2012); Will: PROB 11/686/333.

country. Born in 1664, in the early years of the Restoration, he may have travelled to London as a young man after his father died in 1684. Somehow, he gained experience in arithmetic and geometry, in the intricacies of commercial arithmetic, and double-entry book-keeping. Most textbook writers of the time were teachers of mathematics or accounting, often also advertising themselves as practitioners, but Edward Hatton seems to have been the exception. He was not a school teacher, and never seems to have been a merchant. Instead, he was secretary and surveyor for a fire insurance company.

In the 1660s, London was visited with twin disasters of plague and fire. The great Plague of 1665–1666 killed over 100,000 people, including an estimated 15–25% of the population of London. September 1666 saw the Great Fire of London, which destroyed or damaged the homes of more than 80% of the population of the City of London. Over 13,000 houses were burnt down along with 89 parish churches, St. Paul's Cathedral, the Guildhall, and virtually all the civic buildings of the city.

The subsequent reconstruction was overseen by Robert Hooke and Christopher Wren.[7] New building regulations were promulgated covering widths of streets, set-backs of houses from streets, and building material (stone or brick rather that wood and thatch). As the rebuilt, safer city took shape another economic opportunity arose—that of spreading the risk of the devastating effects of a fire among members of societies or customers of insurance companies.

The first such organizations were founded in the 1680s, the two most successful being the *Amicable Insurers* (or *Hand in Hand*) and the *Friendly Society*.[8] An intense rivalry for trade played out in the newspapers of the day in a way that was neither friendly nor amicable. The *Friendly Society* was founded by Henry Spelman and William Hale in 1684, and was organized along the lines of a mutual aid society. Members paid an escrow deposit, a management fee, and were liable to an annual call that depended on the amount insured. Losses sustained by members were made good by raising funds from the subscribers up to the maximum call. The obvious problem of default was dealt with by moral suasion, having socially prominent members as trustees (Christopher Wren and the Lord Mayor of London, for example), and the fact that anyone in the City of London who was in a position to insure property worth hundreds of pounds was also known to others in a similar position.

As surveyor for the *Friendly Society*, a position he seemingly held for his entire career, Hatton would have been responsible for surveying and assessing property to be insured and property damaged by fire, collecting and disbursing premiums and

[7] Both Hooke and Wren have been the subject of numerous studies and have extensive literature on all aspects of their lives and work. See, for example, the twin biographies Jardine (2002) and Jardine (2004).

[8] For a collection of early documents concerning early fire insurance, including some relating to the *Friendly Society*, see Jenkins and Yoneyama (2000). The business of fire insurance in eighteenth-century London, including the *Hand in Hand*, but not the *Friendly Society*, is addressed in Pearson (2002).

claims, and maintaining the company books. Certainly, his position required a solid ability in arithmetic and accounting; however, he came by his education.

Unusually for the time, Edward Hatton had a close and open relationship with his publishers and booksellers. While he liked to style himself as “Gent.”, when not using philomath or philomercat, the fact was Hatton worked for a living, and his publications were often done at the behest of the publishers. At one time or another, most of the significant booksellers in London handled some of Hatton’s works. He began with Christopher Coningsby, and by 1699 had added Daniel Midwinter and Thomas Leigh as well as John Nicholson. Daniel Midwinter became a life-long friend.[9] Over the course of the next decades, they were joined by Robert Knaplock, Jonas Brown, James Knapton, James Holland, John Osborn, Thomas Longman, and many more.[10] On the publishers side at least, it was a lucrative relationship.

3 A Textbook: The *Merchant’s Magazine*

Edward Hatton’s first publication, *The Merchant’s Magazine or Tradesman’s Treasury* appeared in 1695. It was a general textbook for commercial arithmetic and became one of his most enduring and popular works with a second edition in 1697, a third in 1699, and a ninth impression published in 1734.[11]

From the beginning Hatton was open about his close commercial liaison with his publishers and his desire to educate “young merchants”. Printed for Christopher Coningsby and dedicated to Henry Spelman, the principal of the *Friendly Society*, Hatton opened the book with an epistle “To the Reader” laying out the origins, rationale, and goals of the work.

> Having for many Years past spent some leisure hours in the Study of Arithmetick, Geometry, &c. I was often solicited to write something Mathematical. But considering the many ingenious Tracts of that subject already extant, together with the Censoriousness of the Age, I refused it; knowing that to write what others had done before, without making some improvements, would look like a Transcript, and not be agreeable to the End, which every Author ought to propose of himself. i.e. To make some new Discoveries, and advance the Art he makes his subject a Degree nearer Perfection: However being requested by the Bookseller to publish something of the Use of some Arithmetical Copper-Cutts which he had by him, I complied with his Desire, hoping I had acquired such a competent Knowledge in Arithmetick, as to offer several things I had not observed in Print before. I resolved also to make such Additions, as might render the Book of general Use to Men of Business, especially to young Merchants, for whom it is chiefly intended. And the Practice of Merchandize being of so great Consequence to a Nation; I have endeavoured so

[9] In his will, Hatton left a small legacy to “my very good friend Daniel Midwinter. . .in consideration of his friendship which I have enjoyed near forty years”. PROB 11/686/333.

[10] For a survey of London publishers in the early eighteenth century, see Dugas (2001); for an overview of Hatton’s kind of instructional publishing, see Raven (2014), Chapter 9.

[11] Raven claims that it was, “by far the most successful of the early guides [to accounting]” (Raven 2014, 184), while also noting some of Hatton’s competitors.

> to handle Arithmetick (as a Foundation) and to apply that to Merchants Accompts, and both to Book-keeping, as might be most likely to accomplish those concerned in that honourable Employment; for it would signifie little for a young Merchant to understand Arithmetick without Merchants Accompts, and to know the latter without the former is impossible; and though he should assume a good Knowledge of both; yet if he is ignorant of the Art of Book-keeping, that alone will prove a great Deficiency. (*MM*1: v)

Let us understand how Hatton is positioning himself and his work in the marketplace. Firstly, he claims sufficient expertise in mathematics to justify authorship, but, importantly, for one signing himself on the title page as "Edw. Hattton, Gent.", that expertise was gained during his "leisure hours", not employment. Preparing to "offer several things I had not observed in Print before", he is willing to share his knowledge in order to make "his subject a Degree nearer Perfection". Thus his disinterested gentlemanly *bona fides*.

He was, however, propelled into print by the purely commercial designs of a bookseller who had some attractive arithmetical copper plates and needed a suitable text to surround them in order to be saleable. Hatton accepted the commission, but rather than producing a simple arithmetical textbook he negotiated an expansion of the work to encompass the three main areas of arithmetic, merchant's accounts, and book-keeping so as to "render the Book of general Use to Men of Business, especially to young Merchants, for whom it is chiefly intended". Hatton therefore entered the commercial arena at the behest of his publisher in order to provide a practical, mercantile education while himself remaining somewhat above the fray.

Titling the book, *The Merchant's Magazine or Tradesman's Treasury* sent a clear signal of intended audience and the title page expanded the categories of beneficiaries beyond the "young Merchants" to claim the work was "likewise usefull for Schools, Bankers, Diversion of Gentlemen, the Business of Mechanicks, Land-waiters, and other officers of Their Majesty's Customs and Excise" as well as "those concerned in the Bank of England", (founded just the year before) (*MM*1: title page). Christopher Coningsby, the publisher and presumed owner of the arithmetical copper plates had made a good choice of author. The book was successful and he had a long and doubtless lucrative association with Edward Hatton.[12]

3.1 Commercial Arithmetic

After the preliminary frontmatter, the main body of the text opens with the section on commercial arithmetic, described on the title page as, "Vulgar Arithmetick in Whole Numbers, with the Reason and Demonstration of each Rule, adorn'd with

[12] Christopher Coningsby's last publication of Hatton's was the 7th impression of the *Merchant's Magazine* in 1719; he died in 1720 (London Burials, 2010).

curious Copper-Cutts of the chief Tables and Titles: Also Vulgar and Decimal Fractions, after a New, Easie and Practical Method".

The adorning copper plates that caused the book to come into being also provided its structure. With each chapter except the last headed by the appropriate plate, Hatton proceeded for 80 pages through the standard arithmetic subjects: Numeration; Addition; Subtraction; Multiplication; Division; Reduction; The Golden Rule; and Fractions with little opportunity to influence the topics covered.

Hatton did, presumably, have more flexibility in mode of presentation, and here he did not prove himself particularly innovative, giving many definitions, subdividing topics into cases, and utilizing the standard Case–Rule–Example format. He did, however, have his eye firmly fixed on the commercial, rather than abstract, nature of his arithmetic. A few examples will give the flavour.

In this chapter, on addition of whole numbers, he gives definitions, rules, and then an interesting choice of first worked example.

> Addition is either Simple or Compound.
>
> 1. Simple Addition is when Numbers are to be added that have but one Name or Denomination, as Pounds to Pounds, Feet to Feet &c.
> 2. Compound Addition is when Numbers of divers Denominations are added together, as Pounds, Ounces, and Drams, to Pounds, Ounces, &c. in both which Cases these two Rules are to be considered.
>
> The *First* is for the right placing the Numbers to be added.
> The *Second* is for the adding together those Numbers after they are stated. (*MM*1: 5)

He showed how to place numbers in columns according to their units and then proceeded to the first example for simple addition.

		l.	
Admit I have owing to me for	Holland-Cloath	3794	
	For Thread	896	
	For Cambrick	6285	
	For Latten Wyre	3745	
	For Sugar	2392	
	For Nutmegs	3058	
	Total	20170	(*MM*1 : 6)

Hatton worked through the details of the addition, beginning with the 8 in the lower right corner, and closing with the comment "so you will find your self Creditor by 20170 Pound" (*MM*1: 6). By page 6, the young merchant of the intended audience is already a dealer in diverse products and is owed more than £20,000,[13] surely an incentive for continued study.

A single example sufficed for simple addition, but compound addition, "adding Numbers of divers Denomination together" (*MM*1: 6), confronts the practitioner

[13] The main English currency units were the pound (denoted £, l., or li.), shilling (s.), and penny (d.). There were 12 pence (pennies) in a shilling and 20 shillings in a pound. Additionally, there were 4 farthings in a penny. Actual coinage was much more complicated.

with the rich collection of metrological systems then in use and requires knowledge of the relations between the units in each system. Hatton gave five simple examples and used them as a springboard to indulge in his passion for tables, laying out eleven of them (English coin, Troy-weight, Averdupoize-weight, Apothecary-weight, Wine-Measure, Beer-Measure, Ale-Measure, Dry-Measure, Long-Measure, Square or Superficial-Measure, and Time) with following instructions for their use.

Subtraction paralleled addition, with a single simple example involving money, "Admit I have laid out Cash, the Summ of 4579 pounds, out of 6947 pounds, which I had in bank; what Summ remains yet in my Hands?" (*MM* 1: 18), and several examples of compound subtraction using large quantities of specific merchandize:[14]

	C.	Qrs.	lb	
Bought Cotton-Wool	121	1	20	
Sold out	92	3	27	
Remains	28	1	21	(*MM*1 : 19)

Chapters on multiplication and division follow, with each operation being divided into cases depending on magnitude of various terms. At the end of each chapter on the four basic arithmetical operations, Hatton provided a brief proof of the method.

The next chapter is on reduction, the turning of a quantity in one unit into either smaller or larger units (reduction descending and reduction ascending), and naturally plunged the reader deep into the metrological world, converting pounds to farthings, hundred-weight to pounds, and dealing with problems such as, "In 484 Gross of Tape, each Gross 12 Dozen, each Dozen 2 Pieces, and each Piece 36 Yards, how many Yards?" (418176 Yards) (*MM*1: 42). In this case, the worked example multiplied first $484 \times 12 = 5808$, converting into dozens, and then $45808 \times 72 = 418176$ in a single step to obtain yards, without passing through pieces. The Reduction Ascending examples recycle many of the Descending problems ("In 418176 Yards, how many Gross of Tape?" (*MM*1: 46)) again showing the inverse relationship of multiplication and division. Mixed examples include such problems as "In 46 C. of Cotton-wool, how many Pounds, and what Price, at 15d, a Pound?" (*MM*1: 48) requiring a conversion of hundred-weight into pounds, multiplication by 15 to gain a cost in pence, and then division to convert pence to pounds.

Following arithmetic and the unit conversions of reduction, Hatton turned to the next standard topic, the Golden Rule, or Rule of Three. This was divided into four sections, single and double rules of both direct and indirect proportion, given with assorted subcases and numerous examples. Opening his discussion of the single rule of direct proportion, Hatton stated, not incorrectly, "All the Difficulty in this Rule consisteth in the right stating the three Numbers given; for when you have done that, you have onely Multiplication and Division, and the Work is performed" (*MM*1: 49). The examples proceed through the various cases from, "If 32 Rundlets

[14] The system of (averdupois) weights had 16 ounces in a pound (lb.) and 112 pounds in a hundred-weight (C.). A quarter (Qr.) was a quarter of a hundred-weight, or 28 lbs.

of Brandy cost 96 pounds, what will 4 Rundlets cost at that rate?" (*MM*1: 50) up to "What Principle will raise 20*l*. in Eight Months at 6 *per Cent per Annum*" (*MM* 1: 60). Despite all the cases and the fourteen examples, Hatton disposes of the crucial topic of the Rule of Three in just over 10 pages.

The last chapter in the commercial arithmetic section concerns fractions, both vulgar and decimal. Almost a third of the 23-page chapter is devoted to reduction of fractions, covering conversion of mixed numbers to or from improper fractions, common denominators, reduction of lowest terms, and conversion of units such as the computation of $\frac{124}{146}$ of a hundred-weight as 3 quarters 18 pound 12 ounces and $11\frac{58}{146}$ drams (*MM*1: 69). Each of the arithmetic operations then only requires a one or two page collection of cases, rules, and examples. Division is accomplished by laying out the sum as in division by integers and then cross-multiplying numerators and denominators. A final section disposes of decimal fractions with a few specialized rules for money based upon the relationships of 20 shillings in the pound and 960 farthings to a pound.

All in all, the arithmetic section of *Merchant's Magazine* is a clear and straightforward description of arithmetic aimed at creating practitioners, and relatively unconcerned with theoretical justification or proof. As Hatton commented after his description of the simple way of reducing fractions to lowest terms, "There are other Rules for the performing this, but none so proper for the young Merchant's practice" (*MM*1: 68). There is no great innovation in either content or organization, but compared to some other texts of the period a lack of clutter and a careful focus.

3.2 *Merchant's Accompts*

The second main section of the book, *Merchants Accompts; or, Rules of Practice*, is contained in the 50-page Chap. 9 and covers applications of commercial arithmetic. To begin with, Hatton admonished the reader to gain skill in division by single digits and also by 12, "For the Rules of practice being of daily use with the Merchants, ought to be performed with all imaginable Brevity" (*MM*1: 87). To aid with brevity of computation, Hatton gave a number of specialized shortcut cases for arithmetic, "to shew how the Value of any Quantity of Merchandize may be found with most Dispatch" (*MM*1: 87). As is usual, the techniques are presented in Case–Rule–Example format with little additional explanation. The first is:

CASE 1

When the Price of a Unit or Integer of a Commodity is one Shilling.

RULE

Take $\frac{1}{20}$ of the given Number for the Answer.

EXAMPLE

What is 46743 Pound of Cotton-wool worth at 12d. per Pound? See the Operation.

$\frac{1}{20}$ of 46743

is 2337l. 3s. Ans. (*MM*1 : 88)

The next cases are when the price of the commodity is 2 shillings (divide by 10); any even number of shillings up to 20; odd numbers of shillings (treat as even plus one); one penny or any aliquot part of a shilling; any "Number of pence under 12 that are not an even part of a Shilling" (*MM*1: 91) (this case was subdivided into rules for the various different possibilities); farthings; shillings and pence; pence and farthings; and finally, pounds, shillings, pence, and farthings. The final example of this section is "What cost 276 Hundred, 2 Quarters of Steel at 2 l. 3 s. $8\frac{1}{2}$ d. per Hundred?" (*MM*1: 99). Doubtless a skilled merchant performing these kinds of computations every day would develop a mental bag of tricks for simplifying the work according to various cases, but the beginner would probably be overwhelmed and slowed down trying to remember all the different shortcuts. One can imagine a learner returning to this section and gradually developing familiarity with commonly met special cases rather than absorbing all the material in one gulp.

The second topic in the merchant's accounts chapter concerned tare and trett. "Tare is an Allowance in Merchandize made to the Buyer for the Weight of the Bagg, Cask, Chest, Freal, Hogshead &c. in which any Merchants Goods is put… the Allowance for Tare, is various, as you shall see by and by" (*MM*1: 100). On the other hand, "Trett is an Allowance made for the Waste that may be mixt with thee Commodity, as Dust, Moats &c. which is always 4l. at 104" (*MM*1: 100). The fact that trett was a fixed allowance simplified the calculations, but the way tare varied depending upon how the product was packed opened the way for a large number of cases and Hatton certainly availed himself of the opportunity. Case 1 considered "When the Allowance is 14l per Cent. (as of Almonds, Figgs, Steel or Hemp) how to compute the Nett-weight" (*MM*1: 100). The modern reader is cautioned that "per cent" means "per hundred-weight of 112 pounds". Hatton gave a specific example and showed how to compute the net weight in three different ways and then proceeded through a dozen pages of examples, ending with "What is the Nett-Weight of the 4 Puncheons of Pruons following, Allowance being made for 14 Pound at 112 for Tare, and 4 Pound at 104 for Trett?" (*MM*1: 112). In this example, the puncheons of prunes all had different weights and were subject to both tare and trett.

Turning to barter, Hatton was dismissive, saying, "there is much more difficulty in the Name than the Rule; for that is no other than the Rule of Proportion which has been taught already" (*MM*1: 113). After a few examples, we get to the much more complicated issue of "Exchange of Coin" where the complexity lies in the multitude of currencies and disparity between face value and market value of different coins. Hatton noted, "This is also a kind of Barter; though 'tis not called by that Name, and is a Rule by which Merchants know what Summ in English Coin will answer any Summ of Foreign Coin, paid by their Factor or Correspondent" (*MM*1: 116). English coinage at the time was in disarray, debased and clipped, and the country was chronically short of coins. The Bank of England was set up in 1694, and Isaac Newton would be appointed Warden of the Mint in 1696 with a brief of reforming the currency.[15] Hatton waded into the topic with relish, listing weights, "extrinsick values" and "currant values" (*MM*1: 112) for a bewildering variety of English coinage before launching into the valuation of foreign coinage in English sterling. He gave a number of examples, and a handy conversion table (*MM*1: 122–123).

After dealing with currencies, the natural next step was a consideration of interest, both simple and compound. Interest computations were framed as using the Golden Rule, and Hatton treated compound interest as repeated steps of simple interest. He noted, "There is a much briefer way of finding the Compound Interest, which is done for any Number of Years, at one Operation by Artificial Numbers, called Logarithms; but since that kind of Arithmetick falls not within the Subject of this Book, which tends chiefly to accomplish the Young Merchant; and since Compound Interest is seldom either taken or given by great Traders, I shall therefore omit the former, and say no more of the latter" (*MM*1: 130).

Rounding out the portion of the volume on Rules of Practice was a brief straightforward section on computing gain and loss and Fellowship, as usual by rule of proportion.

3.3 Book-Keeping

Hatton's "Method of Book-keeping by way of Debter and Creditor or (as some call it) after the Italian manner" (*MM*1: 137) did indeed follow the standard approach of three books, the Waste book, the Journal, and the Ledger, going back to Pacioli and even in England well-established by the late seventeenth century. He introduced the three books and explained the basics of double-entry book-keeping and the ledger with debtors on the left-hand pages, creditors on the right. He described various classes of transactions, such as stock debtor and cash creditor and then gave a long list of cases showing how to enter all kinds of accounts into the ledger in both domestic and foreign trade. Domestic trade spanned 29 cases, with a further 18 cases for foreign transactions before he started dealing with factors and company accounts. Despite the brevity of description of each case, the entire collection runs to 17 pages. Certainly this is far more dense material than a young merchant could be expected to absorb on a first reading, but it was the style of the times, and presumably Hatton intended his reader to use this as a reference text while gaining practice with casting up accounts.

[15] On late seventeenth-century coinage and Newton's work at the Mint, see, for example, Craig (1963), Westfall (1980), Belenkiy (2013).

The theoretical material was followed by a practical example showing the use of Waste book, Journal, and Ledger for a fortnight-long set of fictitious accounts "of me C.D. of London, Merchant: Containing all my Dealings from the First day of July 1694" (*MM*1: [159]). The accounting period was extended to a month in the later editions by the simple expedient of changing the dates in the later entries. Hatton began by drawing up an inventory of all cash, goods, and debts as of July 1st for the Waste book. We will trace a small example of his dealings to illustrate not only the techniques of double-entry book-keeping, but also how he chose his examples to engage the reader. Among the inventory he had on hand:

Item	I have Drugs, viz.	li.	s.	d.	
340 l.	of Scammony at 10 s. per l.	170	00	0	
565 l.	Opium at 6 s. per l.	169	10	0	
105 C.	Gallingale at 40 s. per C.	210	00	0	
		549	10	0	(*MM*1 : [159])

The drugs made up a noticeable fraction of his net worth of 3159 li. 10 s. Scammony is a purgative from the root of a flower of the morning glory family, Gallingale (with various spellings) is a plant in the sedge family with roots then used for a variety of medical treatments including for dropsy, and opium is still widely used although not usually so openly accounted for.

Among other transactions, "C.D." recorded in the waste book a sale on July 2 to George Higgs of "300 l. of Scammony for ready Money at 20s. 6d. per lb" (*MM*1: [160]) for a total of £307.10.00, registering a healthy profit. The transaction was then recorded in the Journal as "Cash Debter to Druggs for 300l. of Scammony sold George Higgs for ready Money at 20s. 6d. per l." (*MM*1: [163]). In the ledger, the appropriate entries appeared under the "Cash Debtor" and "Drugs Creditor" sections (MM1: [166 and 168]). On July 9, is recorded "Sold William Short the following Druggs, viz. 40l. of Scammnony at 21s. per l. 42.00.00, 350 l. of Opium at 12s. per l. 210.00.00, 390l. in all for 252li. of which I have received 160l. and the rest, which is 92l. to be paid in 3 months" (*MM*1: [160]). As before, the Journal entry directed how this more complicated transaction was to be parcelled out in the ledger accounts. Finally, on July 10 came a yet more complicated example, "Received from my Factor Gilbert Gainwell at Aleppo by my Order and on my Accompt 8 Chests of Myrrh, containing 30C. Nett, which at 22 Dollars per C. comes to 660 Dollars, the Exchange at 4s. 6d. per Dollar makes Sterling 148l. 10s. 00d." (*MM*1: [161]).

As this subset of the transactions illustrates, Hatton gave a wide variety of entries of different types of sales and purchases and trades in the model accounts, all of which was encompassed in no more than three or four pages each for the waste book, journal, and ledger. It was a very succinct textbook approach and any young merchant copying out and practicing with the accounts would soon become well-versed in the standard daily entries. It would also not escape notice that "C.D." was a very successful merchant, registering a drug profit of £284, out of a total profit of £1138 4s. 6d. in the two-week period covered by the exercise.

Hatton ended the accounts section of *Merchant's Magazine* by showing how to use the accounts to generate an inventory, how to close accounts, and how to balance accounts. There followed a short, four-page chapter on Maxims and Rules for Drawing Bills. The whole treatise on arithmetic, merchants' rules of practice, and book-keeping came in at just 200 pages, quite succinct for the times.

3.4 Later Editions

The success of the first edition of *Merchant's Magazine* called forth a second edition in 1697 and Hatton's enthusiasm poured out in a new preface to the reader, "The First Impression of the Treatise. . .having however for its usefulness and familiarity found Acceptation in the World, exceeding my Expectation: I thought myself obliged in point of Gratitude in this Second. . .to study how. . .I might make it a Book farther usefull to the Publick" (*MM*2: iii). Hatton's gratitude extended to the inclusion of several more tables, explanations of "all those mysterious Rules" (*MM*2: iii) for calculations with fractions, reasons behind the rules of practice in the merchants accounts, extension of book-keeping to include a cash book, household expenses and charges of merchandize, and five whole new chapters. The maxims section was expanded to a chapter, and extended with examples concerning factors, there was a new chapter on "what Commodities are produced by all Countries in the World" (*MM*2: iv), a chapter "concerning the Post and Postage of Foreign Letters" (*MM*2: iv) and the dates of the foreign mail, a dictionary where "I have explained all the mysterious Terms or Words I could think on, that relate to Merchandizing" ((*MM*2: iv) (this glossary of terms does seem to be something that had not been done before, at least in extent), and model examples of merchant paperwork: bills of lading, invoices, bills of exchange, letters of credit, etc. The preface also included an extraordinary paean to the merchant. All the new material came at a price, and the second edition was some 70 pages longer than the first.

The third edition of 1699 differs little in content from the second. Henry Spelman having died, Hatton dedicated the third edition to Sir Francis Child, Lord Mayor of London. On the title page, where the author had signed himself, E. Hatton, Gent., the third edition gave the curious appellation, E. Hatton, Philomercat. While many would identify themselves as philomaths (lovers of mathematics), Hatton appears to be unique in championing his love of trade (he used it on other works as well as *Merchant's Magazine*).

1701 brought the fourth edition, "Corrected and Improv'd", although in fact little changed from the third. Hatton was back to being a "Gent." and this edition was dedicated to Thomas Welham, Deputy Register of the Prerogative Court of Canterbury. A fifth edition followed in 1707 and a sixth in 1712. In the dedication to Colonel James Seamer of the seventh edition of 1719, Hatton noted, "The following Treatise having, in a few Years, passed six Impressions of about Eight Thousand Books" (*MM*7: ii), testifying both to its staying power (the few years being twenty-five) and its popularity.

The eighth edition of 1726 was dedicated “To the Merchants and those Gentlemen who Instruct Youth in their Accounts.” The dedication opened, “Since it is chiefly owing to your Approbation and Recommendation of this Book, that hath encouraged the printing of about Twelve Thousand in Eight Impressions; I could not therefore think on any to whom I now owe so much of Gratitude on the Account; nor of a more ample and just Way of expressing it, than by this publick Acknowledgment” (*MM*8: ii). Hatton’s refreshing mix of candour and boasting indicates that he fully understood his market, and that the last two impressions probably had a print run of about two thousand copies each. The ninth, and last, edition of *Merchant’s Magazine* in 1734 included the same dedication with the sales updated to “about Fourteen Thousand” (*MM*9: ii).

Hatton’s textbook on accounting had a remarkable forty year run and, after the expansion of the second edition, the text remained largely unchanged. Even the 1694 date of the sample accounts was never updated. While, in the nature of publishing at the time, the book may not have made its author much if any money, it doubtless was lucrative for the booksellers. The first edition was published for Christopher Coningsby. In the second edition he was joined by Daniel Midwinter, and by the ninth impression we learn the book was printed for “J. Knapton, R. Knaplock, D. Midwinter, B. Sprint, R. Robinson, W. Innys and R. Manby, J. Osborn and T. Longman, and A. Ward”, quite the roster of 1730s publishers. Of course, by this time, Hatton was a well-known quantity in the London book scene. The *Merchant’s Magazine* was a mainstay from its first edition, but it was only one of his books, and, measuring by copies sold, not even his most popular.

4 Ready Reckoners: *Comes Commercii*

In terms of sales, Hatton’s most successful work was probably the *Comes Commercii: Or, The Trader’s Companion*, the first edition of which appeared in 1699, printed for Christopher Coningsby, J. Nicholson, and Daniel Midwinter. Unlike the comprehensive textbook approach of the *Merchant’s Magazine*, the *Comes Commercii* was a reference work, principally a ready reckoner. A tall, narrow volume, it was designed to be consulted in the office, counting-house, and anywhere a merchant might need to do some quick computations. Thus, it was portable, could be slipped into a pocket, carried around town, opened many times a day, and worn out, necessitating the purchase of a new copy.

The first edition was dedicated to the “Honourable Court of Directors of the English Company Trading to the East-Indies,” the dedication showing once again Hatton’s commercial acumen, “Having at the Request of the Booksellers written, and finished the following Treatise. . .considering it might in some Measure prove usefull to those employed in your momental Concerns both at home and abroad; I hope it might for that Reason gain your Approbation, which I believed would not a little contribute to the Credit of the Book” (*CC*1: ii–iii).

The volume was divided into four main parts plus a supplement. The first and largest part, occupying nearly half the volume was, “An exact and usefull Table, shewing the Value of any Quantity of any Commodity ready cast up, more adapted to Merchants Use than any other extant; which is demonstrated by 14 Examples relating chiefly to Buying and Selling” (*CC*1: title page). Each page of the *Exact and Useful Table* contained a table recording the value of multiples of a commodity of a given price, “The price of the Pound, Ell, Yard, Ounce or other thing, being” [base price]. The base prices start at 1/8th of a Penny or half a farthing and increase through One Farthing, Two Farthings, Three Farthings, One Penny, and so on up to Nineteen Shillings and Six Pence. The quantities or multiples range from 2, 3,. . ., 70, 80, 84, 90, 100, 112, 200,. . ., 1000,. . ., 10000,. . ., 70000. The layout is clear and the reader can easily find the value of any quantity of any item. A brief extract will suffice.

The price of the Pound, Ell, Yard, Ounce or other thing, being Five Pence

Value of	l.	s.	d.
12	0	5	0
13	0	5	5
14	0	5	10
15	0	6	3
16	0	6	8

As noted in the extended title of the Table, the section concludes with a five page, 14 example, guide to using the table for various problems.

The second section was mostly taken up with a large multiplication table, or “A Table Calculated for Universal Use”, given as “The Length or Value of anything being [x]”; “The Breadth or things valued 2,. . ., 99” and the product recorded. The head numbers for the tables run from 2 to 99, with two tables crammed onto a page except for the last. Rounding out the abstract multiplication table is a series of tables converting multiples of one unit into another: square feet into square inches; acres into square perches; gallons of wine (of 231 solid inches) into solid inches; gallons of ale or beer (of 282 solid inches) into solid inches; square of the gauge point for a wine gallon times gallons into Dividends, and solid feet into solid inches. The tables themselves are followed by thirty pages of instruction into their use for multiplication, division, reduction, and “merchandizing”.

Chapter 3 contained a brief practical description of the kinds of units and computations used in the work of various tradesmen: glaziers; joiners; painters; plaisterers; brick-layers; masons; and carpenters. Rounding out the main part of the volume was a 50-page description of “Such business of merchants as is to be done in the Custom House” (*CC*1: 254). While Hatton only pointed to the clarity and accuracy of the preceding tables and instructions, he did claim some originality in collecting the information relevant to the customs. The last section of this chapter was, “Concerning Insuring of Ships, Merchandize, Houses &c.” (*CC*1: 289). After explaining how to insure ships and the merchandize carried on them, Hatton gives a detailed description of the process of insuring a house, with emphasis on the *Friendly Society*, including a table of their rates.

The "Supplement Concerning Simple and Compound Interest" was a 25-page description of calculating simple and compound interest together with tables of both for 6 per cent for numbers of days.

In due course (1706), a second edition appeared. Unlike the case of the *Merchant's Magazine*, the second edition of the *Comes Commercii* was almost identical to the first, although the dedication to the Directors of the East India Company was replaced with a more general dedication to "The Merchants (More particularly Those of the Honourable City of London)". By the third edition, the Booksellers felt a need to insert a note ensuring potential buyers of the usefulness, accuracy, and popularity of Hatton's work compared to competitors, claiming that it was "so approved by Merchants and Traders, that the last Impression of 3000 sold off much sooner than the former of 2000" (*CC*3: iii). This claim by the booksellers was the first indication of the size of the print-runs for the work. Oddly enough, the fourth edition of 1723 contained exactly the same note from the booksellers, although by the 5th edition of 1727, the numbers had been revised upwards to 4000 and 3000. The 6th edition of 1734 was the last to appear in Hatton's lifetime, containing the same 4000 and 3000 claims. Depending on how we interpret these somewhat suspect figures, it does seem likely that *Comes Commercii* outsold *Merchant's Magazine*, despite the popularity of the latter.[16]

As soon as Hatton passed from the scene, a pirated Irish version appeared in Dublin. This volume in turn went through several editions, a 6th edition appearing as late as 1781. Meanwhile, Hatton's regular booksellers and their successors kept his volume going, corrected by one W. Hume, through 7th, 8th, 9th, and 10th editions, the last being published in 1759. Clearly the format of the ready reckoner was a solid seller among merchants and accountants, and Hatton's reputation was sufficient to keep his name on the cover long after his decease.

5 Ready Reckoners: *Index to Interest*

While the *Merchant's Magazine* and *Comes Commercii* were very successful works, and *Comes Commercii* included an interest table supplement, it is not so clear that Hatton's full-scale foray into the production of interest tables was as popular. The first edition of *Index to Interest* appeared in 1711, by which time Hatton was a well-established name in the world of commercial publishing. In his preface to the reader, Hatton says that he "was prevail'd upon by some Gentlemen in the Law, to undertake the following laborious Work" (*II*1: iv). The legal profession was a slightly different market from the merchants and tradesmen that were his primary audience. A ready

[16] On the question of interpreting booksellers' claims of sales, and the related issue of comparing sales and readers, see Glaisyer (2006): 105.

reckoner is intended to substitute for computation, but here Hatton expressed more doubt about the mathematical abilities of lawyers working with interest payments, "a Tenth Part of whom may not perhaps understand Decimals" (*II*1: iv). This remark was by way of justifying his decision to give his tables in terms of currency to farthings rather than decimal parts of a pound, but it does give one pause. The volume was published for a consortium of a dozen booksellers, although only Daniel Midwinter was one of Hatton's regulars.

The *Index to Interest* featured extensive tables of interest, opening with tables of simple interest for each day of a year and multiple years to 19 for five, six, seven, and eight per cent rates for principal amounts of 1 to 1000 pounds. These were followed by a table of discount at 6 per cent and then instructions for the use of the tables. The large tables of simple interest were followed by smaller tables of present value of annuities and reversions at 6 and 10 per cent compound interest, values of leases of church or college lands, and annual tables of compound interest at 6 and 10 per cent.[17] There was a postscript converting guineas into pounds, shillings, and pence (with a guinea at 21s. 6d.) and a "Circle of Time" for computing the number of days between given dates.

The second edition of 1714 gained ringing endorsements from Directors of the South Sea Company, Accountants from the East India Company, lawyers and Writing Masters, as well as certification by William Whiston, Humphrey Ditton, and one Mr. P. Hoet.[18] Adorned with a poem of "Panegyrical Lines", the volume also retained the dedication to Hugh, Lord Willoughby of Parham, although he had died in 1712. However, the list of publishers had shrunk to just two, Jonas Brown and John Richardson. Jonas Brown included an advertisement for other books he sold. None of them was by Hatton and two were accounting texts, competitors to the *Merchant's Magazine*. Apart from the frontmatter, the book was essentially unchanged.

A third edition followed three years later. The third edition lost the poem but retained the endorsements and dedication. This edition was published by John Walthoe, who also included an advertisement for other books he sold, none of which was by Hatton. The tables were also again unchanged. It is the changing and dwindling numbers of publishers that casts doubt on the success of the *Index to Interest*, especially when so many of Hatton's other works were in wide circulation.

[17] For a discussion of annuities, reversions, and interest computations in the eighteenth century and Hatton's place in the developing field, see Bellhouse (2017).

[18] On the subject of endorsements and other techniques of gaining credibility, especially in ready reckoners of interest rate calculations where the reader must implicitly trust the author's accuracy, see Glaisyer (2007).

6 Posters

At the opposite end of the scale from comprehensive commercial textbooks and extensive interest rate tables, Hatton early in his career published a couple of broadsheets squeezing information onto a single sheet of paper, one on money and one on interest rates. In 1696, for Christopher Coningsby, Hatton published, "An Exact Table of the Weight of Gold and Silver". Two-thirds of the page was given over to a table of the value of silver weighed in grains, pennyweights, and ounces, from 1 grain to 200,000 ounces, for silver values of 5s. 2d., 5s. 4d., 5s. 6d., and 5s. 8d. per troy ounce. A second table gave the standard weight of the new milled coinage from 1 penny to 1000 pounds, and a third gave the value of gold from 1 grain to 40 ounces at 4l. 2s. per ounce. At the bottom of the page, in typical Hatton style, were a couple of worked examples of using the tables, and an advertisement for the *Merchant's Magazine*. The sheet was designed to be displayed prominently, probably pasted to the wall, for anyone who dealt in coinage or plate.

The *Exact Table* was followed in 1697 by "The Assessors and Collectors Companion," printed for Robert Vincent, a near neighbour of Christopher Coningsby, giving a large one page table of interest rates together with conversions into quantities of silver money valued at 5s. 8d. an ounce. This poster was a direct response to a recent act of parliament raising a tax on land and various duties to raise money for the war with France, and was intended to be "useful not only for the assessors and collectors, but for all other persons chargeable with the duties of the said act". Whether or not the two posters were successful, they represented the only time Hatton ventured into such brief publications.

7 Other Works

Edward Hatton's publishing career began, at the behest of his publishers, with the publication of the *Merchant's Magazine* in 1695. Above, we have discussed his significant contributions to commercial arithmetic in the late seventeenth and early eighteenth centuries. However, Hatton published more widely, sometimes on his own initiative, and sometimes as a jobbing author for his publishers. In this section, we briefly summarize his other works to give a more rounded picture of his output.

Hatton's early publishing was commercial, the *Merchant's Magazine* in 1695, followed by the *Exact Table* in 1696, and the *Assessors and Collectors Companion* and the second edition of the *Merchant's Magazine* in 1697. In 1699 came the third edition of the *Merchant's Magazine*, now joined by the first edition of *Comes Commercii*, and Hatton's edition of Robert Recorde's *Arithmetick: or Ground of Arts*. Recorde's *Ground of Arts* was by this time 150 years old, and well on its way to being superseded, not least by Cocker's *Arithmetic*.[19] The booksellers thought there was life (or profit) still in it and had Hatton produce his own version (indeed, this was the last edition of Recorde to be published). As Hatton put it in his preface, "the Stile and Phrase growing obsolete,. . ., the Booksellers (not willing so choice a piece of Arithmetick should be lost for want of a little Polishing, the Principal parts being Extraordinary) were pleased to recommend the performance thereof to me" (*Recorde*: vi). Hatton preserved the overall approach, including the presentation in dialogue, while "expunging what is now useless, and substituting in lieu thereof such Expressions, Rules, and Examples as are most agreeable to the present times" (*Recorde*: vi). At the end of the arithmetic, Hatton adjoined an 80-page *New Treatise of Decimals*, including tables of simple and compound interest. This treatise had separate pagination and title page, and may have also been published independently. Hatton stressed the importance of arithmetic to trade, signed himself Philomercat., and dedicated the work to the young Duke of Gloucester, who unfortunately died the next year.

The early 1700s saw new editions of the *Merchant's Magazine* and *Comes Commercii*, and then in 1708, something quite different, a topographical work, *A New View of London*, in 2 volumes, xlii+824 pp.[20] As surveyor for the *Friendly Society* fire insurance company, it was Hatton's job to travel throughout London measuring and valuing houses for insurance purposes. Along the way, he seems to have built up a detailed fund of information, ranging from the objective and useful to the idiosyncratic. In the preface, he explained,

> I must tell the Reader, That I had not undertaken the Compiling this Work had I not had, by virtue of Business, an opportunity for near 20 Years past of being acquainted with the Names and Situation of most Streets, and the Nature of the Building therein contained, which, together with what knowledge I had of Geometry, Heraldry and Architecture and the Request of Friends, induced me to essay the Performance, which I should however have declined had I been aware of the very great deal of Time and Pains which after I was engaged, I found my self obliged to bestow thereon. (*New View*: vi)

The work was published by a consortium of booksellers who ran an advertisement in the first volume for reprinting Stow's *Survey of London*. These were presumably the friends who made the request.

The *New View* is in eight parts, of which the first is an alphabetical list of the "Streets, Squares, Lanes, Markets, Courts, Alleys, Rows, Rents, Yards and Inns in London, Westminster and Southwark." For a single taste, here is Hatton's entry for Fleet Street showing the mixture of information he presented as well as his abbreviations for space (PC is St. Paul's Cathedral):

> *Fleetstreet*, a very publick and spacious str. of excellent Buildings, the 3d and 4th Rates, which fetch great Rents (one House having been Let near *Temple Bar*, for 360 l. Sterling per annum, with 1400 l. Fine, and few or none under 40 or 50 l.) It is between *Temple Bar* VV. (which see in the part of the Gates) and Fleet Bridge E. L. 570 Yds, and from PC

[19] First edition, 1677. For the question of authorship, see Wallis (1997).

[20] For a more detailed analysis of this work, see Cherry (2001).

> W. 440 Yds: In this str. are 19 Taverns, as many Booksellers, and many Linen Drapers. I find it Recorded, that one James Farr, a Barber; who kept the Coffee House which is now the *Rainbow*, by the *Inner Temple Gate*, (one of the first in England) was in the Year 1657 presented by the Inquest of St. *Dunstans*, in the W. for Making and Selling a sort of Liquor, Called *Coffee*, as a great Nusance and Prejudice of the Neighbourhood, &c. And who would then have thought *London* would ever have had near 3000 such Nusances, and that *Coffee* would have been (as now) so much Drank by the best of Quality, and Physicians. (*New View*, I: 29–30.)

After the appearance of the *New View of London*, the 1710s were quite quiet for Hatton, with the only new work being the *Index to Interest* in 1711. In the 1720s, however, roughly the decade between when he was 55 to 65 years old, he had a sudden burst of energy. First upon the scene was *An Intire System of Arithmetic* in 1721. The *Intire System* was a 500-page comprehensive textbook on arithmetic "in all its Parts," that is, "Vulgar, Decimal, Duodecimal, Sexagesimal, Political, Logarithmical, Lineal, Instrumental, Algebraical." In his preface, Hatton noted that, "Having for many Years past spent my Leisure Hours in Mathematical Studys, and not meeting with any Treatise teaching in a Succinct Order, and Good Method, all the Parts of Arithmetic;. . .I supposed the Publication thereof would be accordingly received by the Studious in this Art" (*Intire System*: v). Once again, Edward Hatton, Gent. was giving of the product of his "Leisure Hours."

The book was "sold by" a collection of a dozen booksellers, ten in London plus Mr. Clements at Oxford and Mr. Crownfield in Cambridge, but not published for them. Instead, Hatton rounded up the "Studious in the Art". He dedicated the book to John Keill, who unfortunately died in 1721, stressed his own *bona fides* in the preface, referring particularly to the *Merchant's Magazine* and the *Index to Interest*, and quietly affixed the list of endorsers to the *Index of Interest* to the end of the preface. There followed a list of subscribers. Subscribers typically paid half down and half on arrival for a new work, and helped defray the cost and risk of publication. Hatton had collected a full 300 names ranging from the obscure to the well-known. Among the latter we find John Theophilus Desaguliers, Edmund Halley, Christopher Wren, and Sir Isaac Newton. This is not the venue to enter into a detailed discussion of the contents and presentation of the book; suffice it to say that a second edition duly appeared 10 years later.

The next year, 1722, brought *A Supplement to the Review of London*. In contrast to the bulk of the original *New View of London*, this work was a 40-page pamphlet listing all the churches in London, Westminster and Southwark in alphabetical order. Compressed into a single paragraph, each entry gave a potted history of the church, the times and days of worship, and how much the benefice was worth. Hatton was never one to stray far from the commercial, even in a work aimed at "all Christian Readers, who will hereby have an Opportunity. . . of frequenting such Churches, on such Times and Occasions of performing their Duties to God as shall be most suitable to their Circumstances" (*Supplement*: 4).

In 1724, Hatton was back to working for the publishers, bringing out a third edition of Isaac Keay's *The Practical Measurer his Companion: Containing Tables Ready cast up for the speedy mensuration of Timber, Board, &c*. Isaac Keay is an obscure figure. Apparently the first edition of the *Practical Measurer* was published in 1704, but there do not appear to be any extant copies, and nothing else is known about the author. Presumably he had died by the time Edward Hatton was called in, "This Book being by the Proprietors put into my hands, to peruse and make fit for the Press: I have faithfully examined every Number herein, and have corrected upwards of 3000 Errors" (*Keay*: 1). A ready reckoner of tables of mensuration for merchants was definitely something Hatton could get behind. In this case, the work contained "A Table of Solid Measure, as Timber Stone, &c. Ready cast up; from 3 to 72 Inches the Side of the Square...And From 1 Foot to 43 in Length", as well as a shorter table of flat measure, comments on boards, and an appendix on four ways of measuring timber. In all, the volume was just short of 200 pages, mostly of tables. A 4th edition was published in 1730, and Hatton's name lingered through the 10th edition in 1777.

Next, in 1726, Hatton was drafted in to support Daniel Midwinter by lending his name to the 17th edition of John Ayres' *Arithmetick Made Easie for the Use and Benefit of Trades-Men*. This venerable work (the first edition appeared in 1682) had been published by Thomas Norris and Daniel Midwinter since the 11th edition in 1711. For the 17th edition, the title page proclaimed, "The Whole perused and many Errors corrected by E. Hatton, Gent." The actual extent of Hatton's involvement beyond perusal is not clear. In 1727, Hatton was back as a jobbing editor, this time helping out his friend Daniel Midwinter with a fourth edition of John Hill's *Arithmetick Both in the Theory and Practice...*, another comprehensive arithmetic treatise, running almost 500 pages and covering much the same territory as Hatton's own *Intire System*. The current edition was "accurately Revised, Corrected, and Improved by E. Hatton, Gent."

The following year, Hatton had a new original publication, *A Mathematical Manual: or, Delightful Associate*. This work was mostly a treatise on the use of celestial and terrestrial globes, including list of stars and constellations on the celestial side and cities and countries on the terrestrial side. Tucked away at the back was a collection of "Mysterious Curiosities in Numbers". The first of these delightful propositions was, "There is a Number consisting of nine Digits; which being multiplied by five different Digits, each of the five Products shall have the nine Digits in it, and neither more nor less; and the Sum of the five Products shall contain the 9 Digits and 0, which are the Characters by which all Numbers whatsoever are expressed" (*Mathematical Manual*: 176).

Despite the fact Hatton had said in the preface to *A Mathematical Manual* that "this will be the last [book] that I intend to write" (*Mathematical Manual*: v), he was back the following year with *The Gauger's Guide; or, Excise-Officer Instructed*. This was intended to be a practical manual for determining volumes of (taxable) liquids in casks and barrels of various sorts. Gauging was a complicated business involving a great deal of value and Hatton was concerned that neither party should have an unfair advantage while minimizing the calculations the excise man would need to perform. Hatton was gathering up old material: "It is now near 40 Years since I first apply'd my Thoughts to the Study of what is the Subject of this Treatise" (*Gauger's Guide*: v), and at 65 seems to have been made redundant. Very little is known about the history of the Friendly Society. It was still going strong in 1722, but perhaps it faded out later in the decade (the founder Henry Spelman had died in 1698). As Hatton put it, "having for some Months past had a Relaxation from other Avocations", he "was put upon writing on this Subject" (*Gauger's Guide*: v). The book was published by Daniel Midwinter.

Hatton's last original publication was, *A New Treatise of Geography*, published in 1732 for a wide consortium of booksellers. Weighing in at some 430 pages, this was a large collection of descriptive geographical information, covering "Empires, Kingdoms, Republicks, and Countries...Cities and Towns...Latitude and Longitude...Things most remarkable in each Country...Lakes, Gulphs, Bays,...The whole after a new, easy and comprehensive Method." Hatton had never been able to resist compiling information and tabulating it, from listing major trading cities and their chief commodities in the *Merchant's Magazine*, to tables of latitude and longitude of major cities in the *Mathematical Manual*. Although not strictly a mathematical work, the *Treatise of Geography* was definitely grounded in his concern for commerce and trade.

Hatton's major publications had an extensive after-life, both in pirated editions such as *The Irish Comes Commercii* first published in 1739, two years after Hatton's death, and through his regular publishers. Hatton had become a brand in commercial mathematical publishing and his name went on being used for the rest of the century.

Appendix: A Summary Hatton Bibliography

- *The Merchant's Magazine or Tradesman's Treasury*, 1695; 2nd edition 1697; 3rd edition 1699; 4th edition 1701; 5th edition 1707; 6th impression 1712; 7th impression 1719; 8th impression 1726; 9th impression 1734.
- *An Exact Table of the Weights of Gold and Silver*, 1696.
- *The Assessors and Collectors Companion*, 1697.
- *Arithmetick; or The Ground of Arts*, by Robert Recorde, ed. E. Hatton, 1699.
- *Comes Commercii; or The Trader's Companion*, 1699; 2nd edition 1706; 3rd edition 1716; 4th edition 1723; 5th edition 1727; 6th edition 1734; 7th edition 1740; 8th edition 1747; 9th edition 1754; 10th edition 1759; 12th edition 1766; 13th edition 1783; 14th edition 1794.
- *A New View of London*, 1708.
- *An Index to Interest*, 1711; 2nd edition 1714; 3rd edition 1717.
- *An Intire System of Arithmetic*, 1721; 2nd edition 1731; 3rd edition 1753.
- *A Supplement to the Review of London*, 1722.
- *The Practical Measurer His Pocket Companion*, by Isaac Keay, 3rd edition, ed. E. Hatton, 1724; 4th edition 1730; 5th edition 1736; 6th edition 1750; 9th edition 1764; 10th edition 1777.
- *Arithmetick made Easie, for the Use and Benefit of Trades-men.* By John Ayres. To which is added, A Short and Easie Method ... of Accompts. By Charles Snell. The Whole perused, and many Errors corrected by E. Hatton, Gent., 1726.

- *Arithmetick Both in Theory and Practice*, by John Hill, 4th edition, ed. E. Hatton, 1727; 5th edition 1733; 6th edition 1736; 7th edition 1745; 8th edition 1750; 9th edition 1754; 10th edition 1761; 11th edition 1772.
- *A Mathematical Manual: or, Delightful Associate*, 1728.
- *The Gauger's Guide; or Excise-Officer Instructed*, 1729.
- *A New Treatise of Geography*, 1732.
- *The Irish Comes Commercii: or, The Trader's Companion*, 1739; 3rd edition 1752; 4th edition 1758; 5th edition 1765; 6th edition 1781.
- *The Merchant and Trader's Daily Companion*, 1763; 2nd edition 1765; 4th edition 1765; 5th edition 1783; 6th edition 1790; 12th edition 1799.

References

Ancestry.com (2010) London, England, Baptisms, Marriages and Burials, 1538–1812, Ancestry.com Operations, Inc., Provo, UT.

Ancestry.com (2012) Lancashire, England, Baptisms, Marriages and Burials, 1538–1812, Ancestry.com Operations, Inc., Provo, UT.

Beeley, Philip (2019) Practical mathematicians and mathematical practice in later seventeenth-century London. The British Journal for the History of Science 52:2, 225–248.

Belenkiy, Ari (2013) The Master of the Royal Mint: how much money did Isaac Newton save Britain? Journal of the Royal Statistical Society. Series A (Statistics in Society) 176: 2, 481–498.

Bellhouse, David R. (2017) Leases for Lives: Life Contingent Contracts and the Emergence of Actuarial Science in Eighteenth-Century England, Cambridge University Press, Cambridge.

Cherry, Bridget (2001) Edward Hatton's *New View of London*. Architectural History 44: 96–105.

Craig, John (1963) Isaac Newton and the Counterfeiters. Notes and Records of the Royal Society of London 18:2, 136–145.

Dugas, Don-John (2001) The London Book Trade in 1709 (Part One). The Papers of the Bibliographical Society of America 95:1 31–58.

Flood, Raymond; Mann, Tony; and Croarken, Mary, eds. (2020) Mathematics at the Meridian. The History of Mathematics at Greenwich, CRC Press, Boca Raton, FL.

Glaisyer, Natasha (2006) The Culture of Commerce in England, 1660–1720, Boydell Press, Woodbridge.

Glaisyer, Natasha Calculating Credibility: Print Culture, Trust and Economic Figures in Early Eighteenth-Century England (2007). The Economic History Review, New Series, 60:4, 685–711.

Hoppit, Julian (2000) A Land of Liberty? England 1689–1727, Oxford University Press, Oxford.

Jardine, Lisa (2002) On A Grander Scale: The Outstanding Life Of Sir Christopher Wren, HarperCollins, New York.

Jardine, Lisa (2004) The Curious Life of Robert Hooke, the Man who Measured London, HarperCollins, New York.

Jenkins, David; and Yoneyama, Takau (2000) History of Insurance Volume 1 Fire, Pickering & Chatto, London.

Pearson, Robin (2002) Mutuality Tested: The Rise and Fall of Mutual Fire Insurance Offices in Eighteenth-Century London. Business History 44:4 1–28.

Raven, James (2007) The Business of Books, Yale University Press, New Haven.

Raven, James (2014) Publishing Business in Eighteenth-Century England, Boydell Press, Woodbridge, Suffolk.

Wallis, Ruth (1997) Edward Cocker (1632?–1676) and his arithmetick: De Morgan demolished. Annals of Science 54: 507–522.

Westfall, Richard S (1980) Never At Rest: A Biography Of Isaac Newton, Cambridge University Press, Cambridge.

Leading to Poncelet: A Story of Collinear Points

Christopher Baltus

Abstract Even for a highly original work, such as Jean-Victor Poncelet's *Traité des propriétés projectives des figures* (1822), previous work prepared the ground. The claim of this paper is that the prevalence of problems and propositions in which collinear points (or concurrent lines) are assumed or demonstrated is a good measure of that groundwork. Euclid, Apollonius, Ptolemy, Pappus, Desargues, Monge, L. Carnot, and C. J. Brianchon all have roles in the story, with special attention to the first decade of the nineteenth century.

1 Introduction

As we look for the seeds of the ideas that grew into Jean-Victor Poncelet's pioneering introduction to projective geometry, the *Traité* of 1822 (Poncelet, 1822), we first look where Poncelet tells us to look: the work of Gaspard Monge (1746–1818), Lazare Carnot (1753–1823), and Charles-Julien Brianchon (1783–1864). One notable feature of their work, especially that of Carnot and Brianchon, is the interest in theorems involving collinear points and concurrent lines. (Points lying on one line are *collinear* and lines on one point are *concurrent*). As plane-to-plane projection preserves collinearity and concurrence, these concepts are central to projective geometry, and theorems of this sort will not surprise someone who has taught or taken a College Geometry course. What is surprising is the scarcity of such theorems in mathematics before 1800. A corresponding phenomenon is the disappearance of propositions of this type discovered in the early modern period, propositions now well known that were forgotten soon after they were discovered. It seems reasonable to associate the burst of interest, soon after 1800, in theorems

C. Baltus (✉)
Department of Mathematics, SUNY College at Oswego, Oswego, NY, USA
e-mail: christopher.baltus@oswego.edu

M. Zack, D. Schlimm (eds.), *Research in History and Philosophy of Mathematics*, Annals of the Canadian Society for History and Philosophy of Mathematics/Société canadienne d'histoire et de philosophie des mathématiques,
https://doi.org/10.1007/978-3-030-95201-3_3

involving collinear points and concurrent lines, with the emergence of projective geometry in the 1820s and 1830s.

The contention of this paper is, first, that just before 1800 there was, for active mathematicians, a paucity of propositions involving collinear points and concurrent lines, with much of that paucity due to a broad forgetting of discoveries made after 1600. Second, that there was a burst of interest in that same type of proposition soon after 1800, especially by Lazare Carnot and Charles-Julien Brianchon. Although beyond the scope of this paper, one can trace aspects of Jean-Victor Poncelet's *Traité des Propriétés Projectives des Figures*, of 1822, back to the work of Carnot and Brianchon.

Let us start with a little test: name a proposition from Euclid's *Elements* in which the conclusion is about collinear points or concurrent lines. (Pause.) OK, time is up. Have you found anything? Here are two possibilities, which are exceptions that support our claim that Euclid was not interested in that sort of proposition. The first is Book 3 Prop. 12, the unsurprising observation that with two circles tangent to each other, the point of tangency and the two centers of the circles are collinear points. The second is Book 4 Prop. 4, in a given triangle to inscribe a circle. The solution is the circle whose center is the intersection of two of the angle bisectors. Euclid justifies his answer by showing that perpendiculars from this point to the three sides are of the same length. What is interesting is that Euclid did not find it worth mentioning that the three angle bisectors would, therefore, be concurrent (Euclid, 1998). In a similar way, Archimedes, in *On the Equilibrium of Planes*, Book 1, Prop. 14, shows that the "center of gravity of any triangle is at the intersection of the lines drawn from any two angles to the middle points of the opposite sides, respectively," but does not go on to point out that, therefore, the medians of a triangle are concurrent (Archimedes, 2002).

In later antiquity, we do have theorems involving collinear points and concurrent lines. These have importance in our investigations since they were influential with European mathematicians of the seventeenth century. First, we mention Menelaus's Theorem, which Girard Desargues made great use of in 1639 in proving theorems about involution. He called it Ptolemy's Theorem, since it appeared in Ptolemy's *Almagest* (Ptolemy, 1984, editor Toomer). Here is that theorem, in modern form, with proof, as in the *Almagest*, Book 1 Section 13 (Fig. 1):

Given triangle EGZ, with the three sides cut in A, D, and B, as pictured. Then

$$EB \cdot GA \cdot ZD = EA \cdot GD \cdot ZB.$$

For the proof, draw a line on E a parallel to GD, meeting AB in H. By parallels,

$$\frac{GA}{EA} = \frac{GD}{EH} \text{ and } \frac{ZD}{EH} = \frac{ZB}{EB},$$

from which the theorem follows.

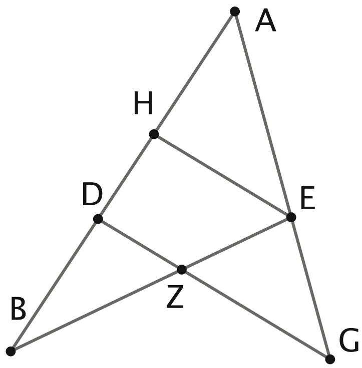


Fig. 1 Menelaus's Theorem in Ptolemy Book 1 Section 13

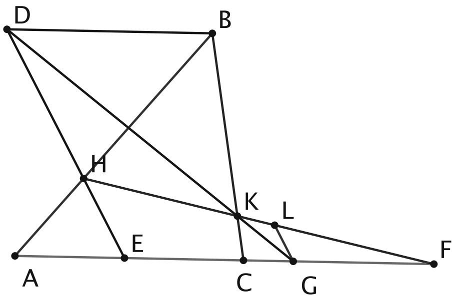


Fig. 2 Pappus, Book 7 Prop. 130

2 Pappus

In Book 7 of Pappus's *Mathematical Collection*, of the fourth century AD, we find a number of problems involving collinear points (Pappus, 1933). Here is Prop. 128: Given figure $ABCDEFGH$ with AF parallel to DB, $AE : EF = CG : GF$, $K = BC \cap GD$, and $H = AB \cap ED$. Then, as pictured in Fig. 2, a line lies on H, K, F. (LG is drawn parallel to DE in the proof.)

Among the theorems that follow in Book 7 are the invariability of the *cross ratio* under projection:

$$\frac{HE \cdot GF}{HG \cdot FE} = \frac{HB \cdot DC}{HD \cdot BC}$$

when a line on H, E, G, and F is projected through a point A onto a second line on H, with E projected to B, F projected to C, and G projected to D (Prop.129).

Making use of the invariability of the cross-ratio, in Prop.143 we have a proof of the Pappus Hexagon Theorem (Pappus, 1933). The following is equivalent to Pappus's formulation: When points A, E, and B are collinear and points Γ, Z, and Δ are collinear, then the opposite sides of hexagon $A\Delta E\Gamma BZ$ meet in collinear points, H, M, K, as in Fig. 3 (where A, H, Z are collinear).

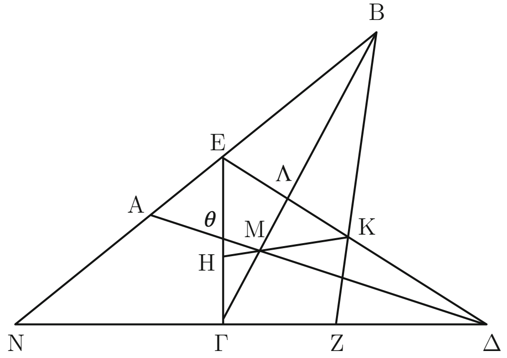


Fig. 3 Pappus Book 7 Prop 143

3 Desargues

Moving to early modern Europe, we have the difficult case of Girard Desargues (1591–1661). He made a step in language and concept toward treatment of collinear points and concurrent lines. Where there had been no term for "concurrent" with earlier writers, in his *Brouillon project* of 1639 (Desargues, 1639), Desargues introduced *ordonnance* to name a set of concurrent lines, where the common point was a *but*. A set of parallel lines forms an *ordonnance* whose *but* is at infinity. And he defined as a *traversal* the line we now call the *polar* of a point with respect to a given conic.

And in 1648, in *La Perspective de Mr Desargues*, (Field and Gray, 1987, pp. 161–169), we have what is now called *Desargues' Theorem*, perhaps the most famous theorem involving concurrent lines and collinear points: When triangles EKD and bla are in perspective from a point H, then "the points c, f, g lie on a straight line cfg," where $c = ab \cap DE$, $f = bl \cap EK$, $g = al \cap DK$ (Desargues, 1648, p 340). See Fig. 4. Desargues followed with two proofs, the first a familiar proof in which the two triangles are in three-dimensional space, the second a proof in the plane, which applies Menelaus's Theorem and its converse. Then Desargues proved the converse theorem.

However, for our story, we note that Desargues' *Brouillon project* disappeared for 200 years. His work with points in involution was known, but little of the material on what we now call the pole and polar of a point with respect to a conic section. And even his theorem of 1648 on triangles in perspective seems to have been unknown. Michel Chasles, in his *Aperçu* of 1837 (Chasles, 1837, p. 82) lists only mathematicians of the early nineteenth century among those who revived that theorem. We see a similar fate for Pascal's Hexagon Theorem of 1640, that for a hexagon inscribed in a conic, the opposite sides meet in collinear points (Pascal, 1640). It was soon forgotten. Although published in the 1779 edition of works of Pascal, by Bossut (Pascal, 1779), it only reappeared with works of Carnot (Carnot, 1806) and Brianchon (Brianchon, 1806) in 1806, and only cited as the work of Pascal in Brianchon (1813). Following the diagram in René Taton's (Taton, 1955,

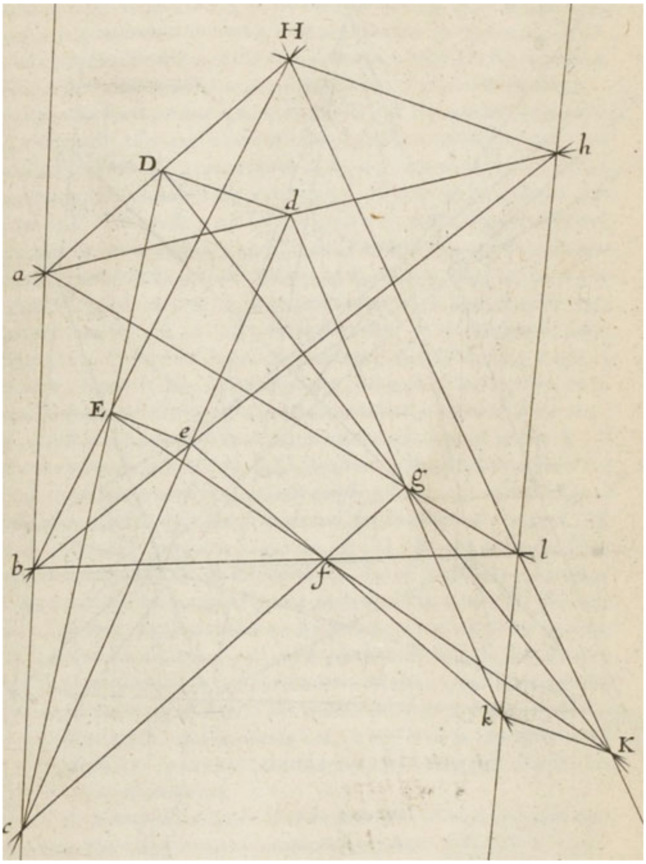


Fig. 4 For Desargues' Theorem on Triangles in Perspective, 1648

p. 12), with Taton's corrections, the claim is in Pascal's Lemma 2, without proof: If hexagon $QPKNOV$ is inscribed in a circle (or conic), with opposite sides PK and VO meeting at M and opposite sides KN and QV meeting at S, then "je dis que les droites MS, NO, PQ sont de même ordre [are concurrent]." NO and PQ are the third pair of opposite sides of the hexagon, and lines *de même ordre* are concurrent (Fig. 5).

The theorem of Giovanni Ceva had a similar fate. It was stated and proved in *De lineis rectis* of 1778 (Ceva, 1678; Oettel and Ceva, 2008); the proof involved statics, weights placed at points in a plane: Given lines on the three vertices of a triangle ABC, meeting opposite sides, respectively in a, b, and c, the three lines are concurrent exactly when

$$aB \cdot bC \cdot cA = aC \cdot bA \cdot cB.$$

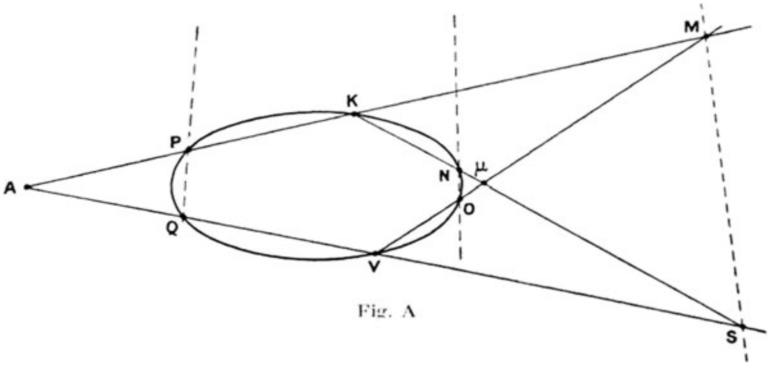


Fig. 5 Taton's diagram for Pascal's Essay, 1640

We do not see the theorem referred to again until Carnot's *Essai sur la théorie des transversales* of 1806, Theorem 5, which did not credit the theorem to Ceva.

In short, by 1800, although much work had been done in early modern Europe involving collinear points and concurrent lines, the only widely known propositions were those from the ancient world, especially Pappus, and what had been recently found by Gaspard Monge.

4 Monge

Gaspard Monge (1746–1818) revived enthusiasm for geometry in France. He taught at the *École royale du génie de Mésièrs* from 1769 to 1784, where he developed "descriptive geometry." As a founder of *l'École Normale* and the *Polytechnique*, in the 1790s, his approach to geometry assumed a great influence when it was incorporated into the curriculum of the *École Polytechnique*. He included only a few propositions about collinear points or concurrent lines, but those few propositions seem to have been widely studied (Fig. 6).

Here is one proposition, in *Géometrie Descriptive*, Articles 38 and 39, with Figures 18 and 20 (Monge, 1799): Monge began with a 3-dimensional argument: Given a sphere, with center A, and an outside line CD, there is one tangent plane to the sphere, which includes CD and meets the sphere above plane ADC. Let N be the tangent point on the sphere. Now, any cone with its vertex at a point G on CD and tangent to the sphere includes a line of the plane CDN, and that line must be on N. The plane ACD cuts the cone and the sphere in a circle with two tangents to the circle from G. When we project down, perpendicular to plane ACD, the intersection of the cone and sphere projects to a line on the projection of N (which I will still call N). For different choices of vertex G on line CD, we get a different line projected onto plan ACD, but always on N. We conclude that for a circle and outside line CD in one plane, for pairs of tangents from all points on line CD, the

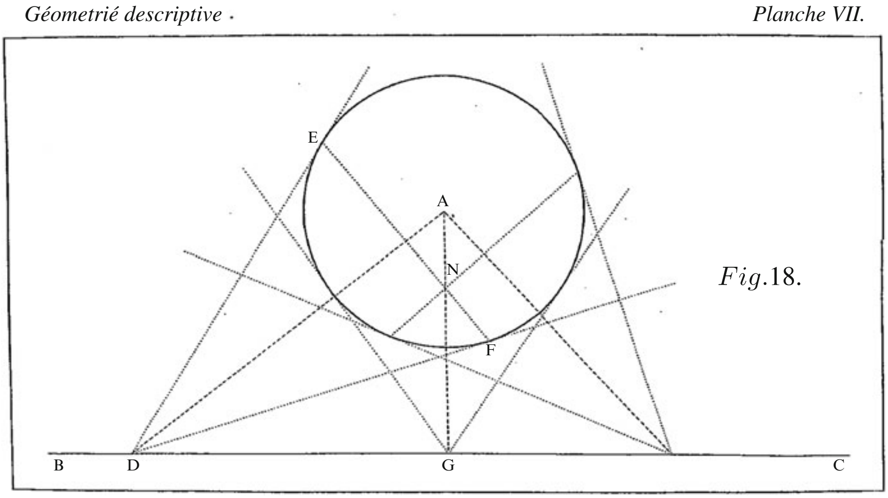


Fig. 6 Monge, *Géometrié Descriptive*, 1798, Fig. 18

chords joining the pairs of points of tangency all lie on one point, N. Two decades later, that point, N would be called the *pole* of line CD, and CD the *polar* of N. Monge pointed out, by similar three-dimensional considerations, that the converse is also true. Monge continued, in his Figure 20, to claim the same result for any conic section and outside line CD. But his argument is only clear for an ellipse, where he introduced the three-dimensional solid formed by revolving the conic about one of its axes. It seems Monge was unaware that Philippe de la Hire had it in his 1685 work on conic sections (La Hire, 1995).

A second proposition, now called *Monge's Theorem*, gives several triples of collinear points associated with three given circles, each outside the others. For each pair of circles, there are four common tangent lines, two internal and two external. We pay attention to the point where the internal common tangents meet, point G for circles A and B in Fig. 7, and the point where the common external tangents meet, point D for circles A and B. The theorem is that for three given circles, centered at A, B, and C, the points where the common exterior tangents of pairs of the circles meet, D, E, and F, are collinear. The theorem holds if we replace two of the points at which exterior common tangents meet, say E and F, by the corresponding points at which the interior common tangents meet, I and H, then those three points are collinear: H, I, and D. The proof is, again, in three dimensions. Take the spheres with centers at A, B, and C and which contain the three given circles, and take a plane tangent to the three spheres. There will be several such planes, each meeting the plane ABC in a line. For each pair of spheres, a cone (or cylinder) will be externally tangent to the two spheres, and the vertex of that cone will be on the line where a plane tangent to the spheres meets plane ABC. That cone's vertex will be the point at which the common external tangents meet. So points D, E, and F will

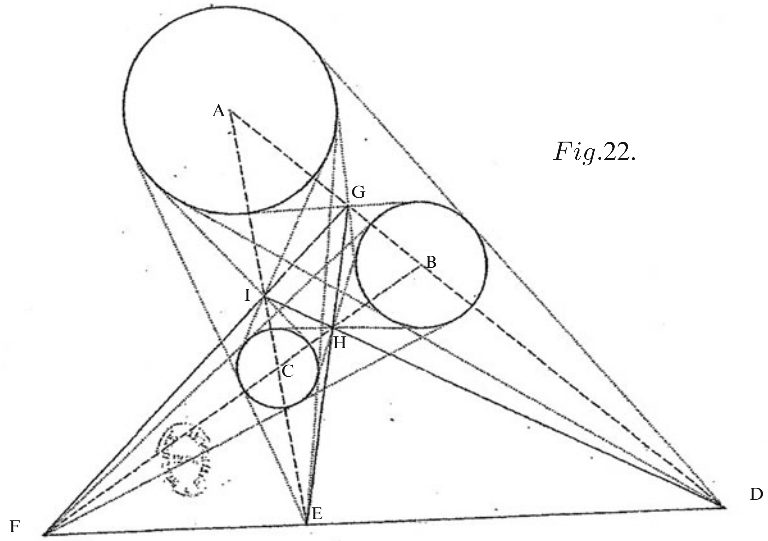


Fig. 7 Monge, *Géometrié Descriptive*, 1798, Art. 44, Fig. 22

be collinear. We get a similar result for other planes tangent to the three spheres, using cones that extend on both sides of the vertex.

5 Carnot

Among the students of Monge at Mésières was Lazare Carnot (1753–1823). Carnot was a physicist and a member of the Committee of Public Safety in the French Revolution, but he was also a mathematician. We see a theorem, apparently new, in his 1803 work (Carnot, 1803, Theorem 36), that if three circles in a plane meet pairwise, then the three common chords (secants) of the pairs of circles are concurrent. Carnot's proof, in the spirit of Monge, considers the three spheres for which the given circles are great circles. The three spheres meet in a single point above the given plane; projecting perpendicular onto the given plane, the three circles in space at which the three spheres meet pairwise become the three common secants, concurrent at a point.

Carnot's most "projective" work was his short *Essai sur la Théorie des transversals* (Carnot, 1806) of 1806. It is largely about concurrent lines and collinear points. He tells us that a *transversal* is a line or curve cutting a system of other lines or curves. (Art. 1) "The simplicity and fecundity of these principles would seem to grant [the theory of transversals] the right to be admitted into the ordinary elements of geometry." Theorem 1 is Menelaus's Theorem, and Carnot wrote that this theorem can be regarded as the "fundamental principle of the entire

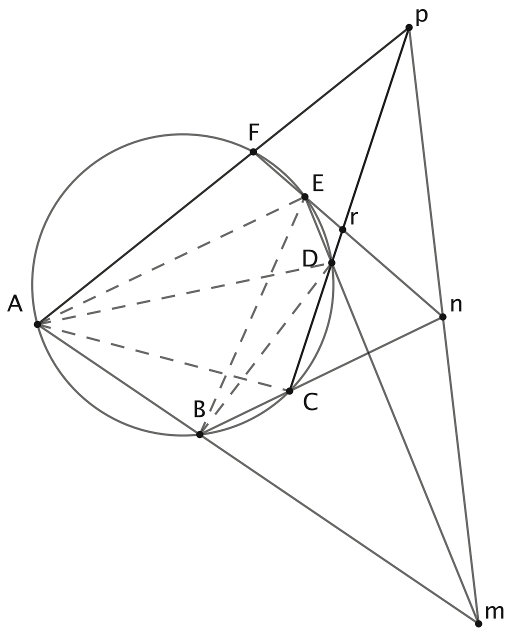


Fig. 8 Carnot, 1806, based on *Fig*.24: Pascal's Hexagon Theorem

theory of transversals."(p 67) There was no attribution to Menelaus or to Ptolemy. After extensions of Menelaus's Theorem to other polygons, Theorem 5 is Ceva's Theorem, without any credit to Ceva, proved using Menelaus's Theorem. Theorems 6 through 9 involve both Ceva's Theorem and Menelaus's Theorem, and from that work comes Theorem 10, which includes Desargues's Theorem. After two more theorems, we have applications. First, Article 27 is Monge's Theorem, now proved in just two dimensions, making use of the converse of Menelaus's Theorem. In Article 31, we have Pascal's Hexagon Theorem, proved using the Law of Sines and the converse of Menelaus's Theorem; with our Fig. 8, "the three points m, n, p be in a straight line." We recognize in Articles 29 and 30 the special cases where adjacent pairs of vertices of the inscribed hexagon merge to produce a triangle or a quadrilateral; the secant on a pair of merging points becomes a tangent to the curve.

6 Brianchon

We now come to Charles-Julien Brianchon (1783–1864). He entered the *École Polytechnique* in 1804 and, as a student in 1806, published Brianchon's Theorem: For a hexagon circumscribed about a conic section, the diagonals on opposite vertices are concurrent. In the proof, he made use of the dual, Pascal's Hexagon Theorem, citing not Pascal but Carnot's (Carnot, 1803, Theorem 45) of 1803. In the proof, he began with a three-dimensional extension of Desargues' Theorem, leading to what we recognize as the *pole-polar* property applied to a hexagon inscribed in a circle. (La Hire, in La Hire (1995), 1685, had this pole-polar property. This proposition, with proof, is also found as Theorem 10 of Carnot's (Carnot, 1806) of 1806, where Carnot credited Brianchon.)

We give special attention to Brianchon's (Brianchon, 1810) of 1810 since it was the one work that Poncelet explicitly noted in his book of 1822 as providing "the first idea of his work" (Poncelet, 1822, p xxiv). Brianchon wrote "I carry out 'la perspective' onto a plane parallel to that determined by the eye and the [given] line of [fixed points] . . . so that the sides, in place of turning on the fixed points, move in a parallel fashion among themselves." In other words, the line containing the fixed points is the vanishing line—the line mapped to infinity—of a plane-to-plane projection. Poncelet applied such projections in several problems, including problems of inscribing and circumscribing polygons to conics, problems that call to mind those in Carnot (1803).

Here is the problem Brianchon was solving: We are given a conic and n fixed collinear points, P_1, P_2, ..., P_n, none on the conic, and we find $n-1$ points on the conic:
(x_1, y_1), (x_2, y_2), ..., (x_{n-1}, y_{n-1}), so the side on the first two such vertices, (x_1, y_1) and (x_2, y_2), lies on P_1, the side on the second and third vertices, (x_2, y_2) and (x_3, y_3), lies on P_2, and so on. We can select (x_1, y_1) as we like on the conic; then (x_2, y_2) is the other point on the conic and on the line on (x_1, y_1) and P_1. We continue this way until we have (x_{n-1}, y_{n-1}) on the conic.

How do we find (x_n, y_n)? It must lie on the line joining (x_1, y_1) to P_n and on the line joining (x_{n-1}, y_{n-1}) to P_{n-1} and that is how we will define (x_n, y_n). The claim is that the point (x_n, y_n) lies on some fixed conic, sliding on that fixed conic as (x_1, y_1) slides on the given conic. Brianchon first makes a plane-to-plane projection, as described above, that maps the line on the fixed points to the line at infinity. As a result, in place of the n fixed points, we have slopes $m_1,\ m_2,\ m_3, ..., m_n$ of the n sides. This provides n linear equations in variables $x_1, y_1, x_2, y_2, ..., x_n, y_n$:

$$(y_{k+1} - y_k) = m_k(x_{k+1} - x_k) \text{ for } k = 1, 2, 3, ..., n-1, \text{ and } (y_1 - y_n) = m_n(x_1 - x_n).$$

Further, we obtain $n-2$ linear equations by substituting pairs (x_k, y_k) into the equation of the conic and subtracting for consecutive values of k. If, for example, $n = 3, m_1 = 1, m_2 = 2$ and the given conic is $x^2 + 4y^2 = 4$, we have ${x_1}^2 + 4{y_1}^2 = 4$ and ${x_2}^2 + 4{y_2}^2 = 4$. Subtracting gives

$$({x_2}^2 - {x_1}^2) + 4({y_2}^2 - {y_1}^2) = 0, \text{ i.e., } 4(y_2 - y_1)(y_2 + y_1) = -(x_2 - x_1)(x_2 + x_1).$$

Now, $(y_2 - y_1) = m_1(x_2 - x_1)$, so

$$4m_1(y_2 + y_1) = -(x_2 + x_1).$$

In this way we get $n-2$ additional linear equations. So we have a system of $2n-2$ linear equations in $2n$ unknowns. In general, there will be a solution expressing

(x_1, y_1) in the two parameters, x_n and y_n. Substituting into the equation of the given conic, we get a second degree equation in x_n and y_n, as we claimed. If we project back to the original figure, we have the n collinear points; note that in a plane-to-plane projection, a second degree curve is mapped to a second degree curve. (Any line that twice meets a second degree curve in one plane twice meets the projected image in the other plane.)

7 Poncelet

Jean-Victor Poncelet (1788–1867) was a student at the *École Polytechnique* from 1807 to 1810. He went with Napoleon's army into Russia in late 1812, and was captured during the winter retreat in November of that year. Poncelet arrived at a prisoner of war camp at Saratov in March 1813, worn out and ill after a trek of several months. As soon as he recovered, he began putting together a series of mathematical *cahiers* or notebooks. Poncelet finally edited those notebooks and published them in 1862 and 1864 (Poncelet, 1813). The first was an ordering of ideas going back to his days at the *École Polytechniques* (Belhoste, 1998). After that he developed new ideas, all leading to his *Traité des Propriétés Projectives des Figures* of 1822 (Poncelet, 1822).

The interest in collinear points is clear from the first *cahier*, about "systems of circles." Any two circles of different sizes, C and C', neither inside the other, have common tangents that meet in a *centre de similitude*, O, which we view as the center for a dilation mapping one circle to the other.

In *Cahier* 1, Poncelet showed the following, in Propositions I., V., and VI., as in Fig. 9, based on Poncelet's $Fig.$ 8 and 10 of 1813/1862:

i. Let t be on circle C'. On circle C draw a radius CT' parallel to $C't$. Then line $T't$ meets line CC' at a point, O, at which common tangents to the two circles meet. (O will later be called the *center*.)

ii. Given any circle tangent to the two given circles, at T' and t', then T' and t' are collinear with the center, O.

iii. Let a secant to the two circles be drawn on O and a point T' on circle C. As in the figure, let this secant meet circle C at T and T'. This secant will meet circle C' at t', described above, and at t. Draw tangents to circle C at T and T', and tangents to circle C' at t and t'. Then the tangents at T' and t' will meet at a point α' so $\alpha't' = \alpha'T'$ and tangents at T and t will meet at a point α so $\alpha t = \alpha T$. Further, let the tangents at T and T' meet at β and let the tangents at t and t' meet at β'. Then β and β' are collinear with O.

iv. Let the line on α and α' be called the *common secant* or *common chord* of circles C and C'. Then any point whose tangents to the two circles are equal lies on the common secant. (The proof shows, by several applications of the Pythagorean Theorem, that D, the point at which line $\alpha\alpha'$ meets line CC', is independent of the choice of the secant TO on O.)

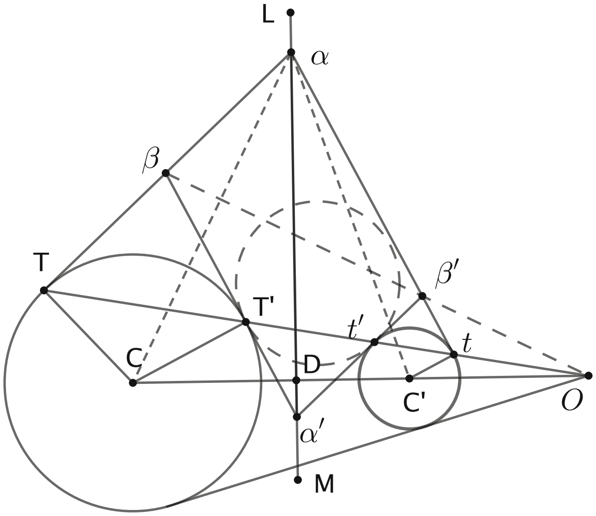


Fig. 9 Poncelet 1813/1862, based on Figs. 8 and 10

In numerous cases, Poncelet followed the approach of Brianchon in his 1810 (Brianchon, 1810), but expanded in that not only is a line, the vanishing line, mapped to the line at infinity but a conic not on the vanishing line is mapped to a circle. Such a mapping appears, justified as his *Quatrième Principe*, in proofs of Pascal's Hexagon Theorem and the dual theorem in *Cahier* 3 of 1813 (Poncelet, 1813, p 135). Here, from *Cahier* 7, is a proof of a pole-polar property: Given a conic and a line LM outside the conic, then there is a point, O, the *pole* of LM, such that when tangents are drawn from any point, m, on LM and the points of tangency joined (in the *polar* of m) the resulting chords all lie on O. When the plane in this proposition is mapped so line LM goes to the line at infinity and the conic to a circle, then the tangents from a point such as m on LM are mapped to parallel lines. In the diagram on the right in Fig. 175, our Fig. 10, we recognize that the chords joining points of tangency have been mapped to diameters of a circle, which are all concurrent at the center of the circle, and the preimage of that center is the point sought.

In the *Traité* itself, we find less emphasis on collinear points. The emphasis has moved elsewhere, but much of his earlier work with collinear points and concurrent lines evolves into concepts of projective geometry. Especially important, the common secant of a system of two circles becomes the *axis* of *homologous* figures. The first of several approaches to homologous figures is in Article 65. When a cone is cut by two different planes that meet in a line m, each point of one resulting conic section has its homologue in the other conic section. When one of the two planes is revolved about m until it lies coplanar with the other, then m is a *common secant*, characterized by the property that homologous lines, such as tangents, meet in points that are collinear, on m. And all the pairs of homologous

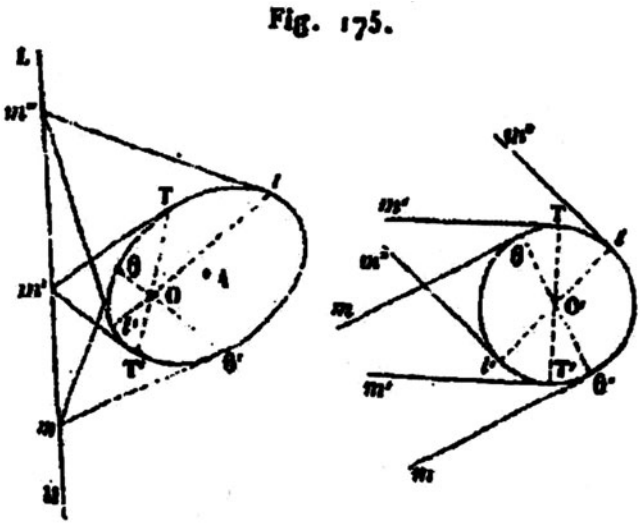


Fig. 10 Poncelet 1813/1862, Figure 175, O is the *pole* of line LM

points are joined by lines concurrent at a particular point, the *center*. m will later, in Article 291, be called the $axis$ of the projection.

References

Archimedes, *The Works of Archimedes with the Method of Archimedes*, ed T. L. Heath, Cambridge University press, 1912, reissued by New York: Dover, 2002.

Bruno Belhoste, De l'École polytechnique à Saratoff, les premiers travaux géométriques de Poncelet, *Société des amis de la bibliothéque de l'École Polytechnique (SABIX)*, www.sabix.org/bulletin/b19/belhoste.html. (original Bulletin No. 19, 1998).

Charles-Julien Brianchon, Sur les surfaces courbes du second degré, *Journal de l'École Polytechnique* Cahier 13, Tome 6, 1806, 297–311.

Charles-Julien Brianchon, Solution de plusieurs problémes de géométrie, *Journal de l'École polytechnique*, Cahier 10, Tome 4, 1810, 1–15.

Charles-Julien Brianchon, Géométrie de la règle, *Correspondance sur l'École Impériale Polytechnique*, No. 5, 2^e vol., 1813, 384–387.

Lazare Carnot, *Géométrie de position*, Paris: Duprat, An XI/1803.

Lazare Carnot, *Essai sur la théorie des transversales*, Paris: Courcier, 1806.

Giovanni Ceva, *De lineis rectis*, Milan, 1678.

Herbert Oettel, Giovanni Ceva, in *New Dictionary of Scientific Biography*, Scribner's Sons, 2008.

Michel Chasles, *Aperçu historique sur l'origine et developpement des méthodes en géométrie*, Brussels, 1837.

Girard Desargues, *Brouillon project d'une atteinte aux événements des rencontres d'un cône avec un plan*, in Field and Gray (1987) (trans Field) and original http://gallica.bnf.fr/ark:/12148/bpt6k105071b/f1.image Paris, 1639.

Girard Desargues, Note: - Extrait de la perspective de Bosse 1648, et faisant suite à la perspective de Desargues de 1636, in David Eugene Smith *A Source Book in Mathematics*, New York: Dover, vol 2, 1959, 307–309.
Euclid, *Euclid's Elements*, edited and translated by David Joyce, https://mathcs.clarku.edu/~djoyce/java/elements/elements.html, 1998.
J. V. Field and J. J. Gray, *The Geometrical Work of Girard Desargues*, New York: Springer, 1987.
Philippe de La Hire, *Sectiones Conicae en novem libros distributae*, Paris 1685; French translation by Jean Peyroux, *Grand Livre des Sections Coniques*, Paris: Blanchard, 1995.
Gaspard Monge, *Géométrie descriptive. Leçons données aux Écoles normals, l'an 3 de la République*, Paris: Baudouin, an VII/1799.
Pappus of Alexandria, *Pappus d' Alexandrie: La collection mathématique*, translated and edited by Paul ver Eecke, Bruges: Desclée de Brouwer, 1933.
Blaise Pascal, *Essay pour les coniques*, 1640, English transl. in Field and Gray (1987), 180–184.
Blaise Pascal, *Oeuvre de Blaise Pascal*, edited by C. Bossut, Paris: La Haye, 1779.
J V Poncelet, *Applications d'analyse et de géométrie qui ont servi, en 1822, de principal fondement au traité des propriétés projectives des figures, etc.*, 2 tomes, Paris: Mallet-Bachelier, 1862–64.
J V Poncelet, 7 cahiers de 1813–14, in Poncelet (1862/4), Tome 1, 1–441.
J. V. Poncelet, *Traité des Propriétés Projectives des Figures*, Paris: Bachelier 1822.
Ptolemy, *Ptolemy's Almagest*, edited by G. J. Toomer, New York: Springer-Verlag, 1984.
René Taton, L' "Essay pour les Coniques " de Pascal, *Revue d'histoire des sciences et de leurs applications*, tome 8, no.1, 1955, 1–18.

Cauchy, Le Verrier et Jacobi sur le problème algébrique des valeurs propres et les inégalités séculaires des mouvements des planètes

Roger Godard

Résumé Dans ce présent travail, on analyse deux approches numériques sur le problème algébrique des valeurs propres, une d'après le polynôme caractéristique par Le Verrier en 1840, et l'autre par Jacobi en 1846.

En 1829, Cauchy introduit la notion du polynôme caractéristique d'une matrice et son théorème sur le spectre des valeurs propres réelles pour des systèmes symétriques.

La méthode de Le Verrier fut créée pour l'étude des variations séculaires des planètes. Elle resta pendant longtemps la méthode pour calculer les valeurs propres. Le processus du calcul revient à déterminer successivement les dérivées d'un système d'équations différentielles linéaires et du premier ordre, à calculer les traces d'un système d'équations linéaires et homogènes, puis à utiliser un théorème de Girard-Newton. La méthode de Le Verrier consiste seulement à trouver les coefficients du polynôme caractéristique. Il faut ensuite trouver par approximations les racines de ce polynôme.

Cauchy and Le Verrier inspirèrent Jacobi, qui publia 'en 1846' une méthode puissante mais complexe pour des matrices symétriques à coefficients réels. Dans ce cas, toutes les valeurs propres sont réelles comme cela avait déjà été prouvé, mais il faut supposer que les valeurs propres sont aussi distinctes. Jacobi fut capable de construire un système orthogonal. Sa méthode est basée sur une suite de matrices orthogonales $\{\mathbf{O_k}\}_{\mathbf{k=1}}^{+\infty}$ telles que $\mathbf{A_{k+1}} = \mathbf{O_k^t A_k O_k} \rightarrow \mathbf{D}$, où $\mathbf{D}$ est une matrice diagonale.

R. Godard (✉)
Department of Mathematics and Computer Science, Royal Military College of Canada, Kingston, ON, Canada

M. Zack, D. Schlimm (eds.), *Research in History and Philosophy of Mathematics*, Annals of the Canadian Society for History and Philosophy of Mathematics/Société canadienne d'histoire et de philosophie des mathématiques,
https://doi.org/10.1007/978-3-030-95201-3_4

1 Introduction

Les acquis mathématiques en algèbre linéaire furent très féconds de 1829 à 1855 environ. Mentionnons un vocabulaire mathématique nouveau introduit par Cauchy avec le polynôme caractéristique. Le Verrier en 1840 travaille sur le calcul des valeurs propres à partir du polynôme caractéristique et la trace d'une matrice. Grassmann propose une algèbre nouvelle en 1844. Hamilton introduit en 1845 le mot vecteur et les composantes scalaires (Moore 1995: 241) ; Jacobi étudie en 1846 les problèmes de symétrie et des systèmes orthogonaux ou des systèmes diagonaux. Il faut aussi inclure les travaux de Sylvester et Cayley en 1850 et 1855 sur les systèmes matriciels. Ici on va s'intéresser surtout aux techniques de preuves de Cauchy, aux calculs de Le Verrier et aux travaux de Jacobi sur les problèmes algébriques de valeurs propres, symétriques par rapport à la diagonale principale. Dans cette période qui nous concerne, on va distinguer trois corpus : un sur l'astronomie, un sur les progrès d'algèbre linéaire, un sur la recherche des racines d'un polynôme.

2 L'inégalité séculaire des planètes

Les inégalités séculaires des mouvements des planètes sont liées au développement de la théorie mathématique des petites oscillations pendant le XVIIIe siècle, à laquelle d'Alembert, Daniel Bernoulli et Lagrange contribuèrent (Hairer et al. 1987: 25–29; Brechenmacher 2007). En 1759, Lagrange publie un système d'équations différentielles du second ordre modélisant la propagation du son en utilisant la loi de Hooke sur l'élasticité. Lagrange observe qu'on obtient un système d'équations similaire au problème de la corde vibrante. Puis en 1762, il modélise un système de masses ponctuelles lié à un ressort. Il obtient un système d'équations différentielles du second ordre avec des coefficients constants. Il choisit alors des solutions particulières du type $y_i = c_i e^{pt}$. En remplaçant ces solutions dans le système différentiel, il obtient un problème algébrique de valeurs propres où les inconnues à déterminer sont les p^2. Lagrange doit alors supposer que les p^2 sont réels, négatifs, et distincts pour la stabilité mécanique d'un système. Il s'appuie pour cela sur des raisonnements intuitifs basés sur la physique. Plus tard, Lagrange commencera à étudier les inégalités séculaires des planètes où il utilisera le même type de raisonnements.

Les inégalités séculaires des mouvements des planètes concernent les perturbations lentes des mouvements des planètes par rapport au temps, à l'échelle d'un siècle ou plus. Déjà, Newton soupçonnait que les lois de Kepler étaient seulement un modèle approximatif et qu'il fallait tenir compte des forces d'attraction des autres planètes ; Euler, Clairaut, d'Alembert, Lagrange, et surtout Laplace contribuèrent à ce projet. Dans son monumental *Traité de Mécanique Céleste*, publié en (1799 1: 327–366), Laplace consacre le chapitre VII aux inégalités séculaires des mouvements célestes. La technique mathématique est la même que celle de Lagrange en 1759 et 1762 : remplacer un système différentiel du premier ordre par un problème algébrique de détermination de valeurs propres. Laplace appelle ses valeurs propres g. Il écrit (Laplace 1799: 346–347) :

> Il suit de ce qui précède que les excentricités des orbites et les positions de leurs grands axes sont assujetties à des variations considérables qui changent à la longue, la nature de ces orbites, et dont les périodes, dépendantes des racines g, g_1, g_2, etc. embrassent, relativement aux planètes, un grand nombre de siècles . . . . Ces variations sont très sensibles dans les satellites de Jupiter, et nous verrons dans la suite qu'elles expliquent les inégalités singulières observées dans le mouvement d'un troisième satellite. Mais les variations des excentricités ont-elles des limites?

Laplace discute alors du problème des valeurs propres, qu'il appelle les racines :

> Mais il n'en serait pas de même si quelques-unes des racines g, g_1, g_2, etc. étaient égales ou imaginaires : les sinus et les cosinus . . . se changeraient en arcs de cercle ou en exponentielles, et comme ces quantités croissent indéfiniment avec le temps, les orbites finiraient à la longue par être fort excentriques ; La stabilité du système planétaire serait alors détruite, et les résultats que nous avons trouvés cesseraient d'avoir lieu. Il est donc très intéressant de s'assurer que toutes les racines g, g_1, g_2, etc. sont toutes réelles et inégales.

Rappelons que Laplace travaillait avec des systèmes d'équations linéaires homogènes symétriques, et que les propriétés des valeurs propres, et leur calcul, dans ce cas, constituent la pierre d'angle de notre travail.

3 Cauchy et le problème des valeurs propres en 1829

Cet article de 1829 s'intitule ≪Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planètes≫ (Cauchy 1829). Cauchy reprend ici les travaux qu'il avait déjà esquissés en 1821 dans *les Exercices de Mathématiques.* L'article de 1829 sera un long article de 21 pages pour prouver que pour un système symétrique, les valeurs propres sont réelles et distinctes. Cet article a déjà été analysé en détails par Hawkins (1975), et aussi par Brechenmacher (2007). Cauchy va seulement dire que ses travaux s'appliquent aussi aux inégalités séculaires : ≪cette équation (le polynôme caractéristique) sera semblable à celle que l'on rencontre dans la théorie des inégalités séculaires des planètes, et dont les racines, toutes réelles, jouissent de propriétés dignes de remarques≫ écrit Cauchy, mais c'est surtout un article d'algèbre linéaire. D'un autre côté, le problème des inégalités séculaires était bien connu à cette époque. Cauchy va partir d'une fonction homogène du second degré avec des propriétés de symétrie :

$$\begin{aligned} &s = A_{xx}x^2 + A_{yy}y^2 + A_{zz}z^2 + \cdots + 2A_{xy}xy + 2A_{xz}xz + 2A_{yz}yz + \ldots, \\ &A_{xy} = A_{yz},\ \ A_{xz} = A_{zx},\ \ A_{yz} = A_{zy}, \ldots \end{aligned} \tag{1}$$

Cette fonction est sujette à une contrainte d'égalité, dite ≪équation de condition≫ :

$$x^2 + y^2 + z^2 + \cdots = 1. \tag{2}$$

À la recherche des maxima et minima de cette fonction, Cauchy va étudier le système d'équations linéaires et homogènes:

$$\begin{cases} (A_{xx} - s)\,x + A_{xy}y + A_{xz}z + \cdots = 0, \\ A_{xy}x + \left(A_{yy} - s\right)y + A_{yz}z + \cdots = 0, \\ A_{xz}x + A_{yz}y + (A_{zz} - s)\,z + \cdots = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases} \tag{3}$$

Dans ce système symétrique, Cauchy va travailler avec le déterminant principal, son ≪polynôme caractéristique≫ et les déterminants mineurs. Le mot polynôme caractéristique n'apparaîtra dans le vocabulaire mathématique de Cauchy que vers 1839–1840. On trouvera le terme ≪valeurs propres≫ dans un mémoire sur ≪l'intégration des équations linéaires≫ (Cauchy 1840), mais sans en souligner l'importance. Bien sûr, par commodité, on sera obligé d'utiliser un vocabulaire plus moderne que celui connu de Cauchy, pour les vecteurs propres, les valeurs propres, et les systèmes matriciels. Cauchy soulignera qu'à partir du déterminant de son système symétrique, on déduit une fonction entière de degré n. Les composantes des vecteurs propres $x_1,\, y_1,\, z_1;\; x_2,\, y_2,\, z_2;\; \ldots;\, x_n,\, y_n,\, z_n$ seront appelées les systèmes de valeurs de $x,\, y,\, z,\; \ldots$ En manipulant les lignes du système (3), Cauchy montre alors que si les racines $s_1, s_2, s_3, \ldots$ sont inégales entre elles, on obtient le tableau suivant :

$$\begin{array}{ll} x_1^2 + y_1^2 + z_1^2 + \cdots = 1; & x_1x_2 + y_1y_2 + z_1z_2 + .\cdots = 0, \\ x_2^2 + y_2^2 + z_2^2 + \cdots = 1.\ldots & \end{array} \tag{4}$$

où on reconnaît implicitement les normes des vecteurs propres et leurs propriétés d'orthogonalité. Puis Cauchy déterminera au signe près les valeurs des composantes $x_1, y_1, z_1, \ldots$ et il prouve que les racines $s_1, s_2, \ldots$ ne peuvent pas être imaginaires. Notons que les résultats (4) avaient déjà été anticipés par Lagrange (Hawkins 1975: 17). Cependant, le vrai théorème spectral pour les matrices symétriques dit que toutes les valeurs propres sont réelles et que si s est une valeur propre de multiplicité k, alors la matrice $A_{n,\,n}$ possède k vecteurs propres linéairement indépendants. Dans sa preuve pour les valeurs propres distinctes, Cauchy ne tient pas compte du fait que certains déterminants mineurs peuvent être nuls et qu'on peut avoir indétermination, ce que soulignera plus tard Kronecker en 1874 (Brechenmacher 2007: 9–10). La symétrie d'un système matriciel ne permet donc pas de conclure que les valeurs propres sont toutes réelles et distinctes, ce dont on a besoin pour la preuve mathématique concernant l'inégalité séculaire du mouvement des planètes. Les propriétés algébriques et la multiplicité des racines d'un système matriciel ne peuvent donc correspondre à la stabilité mécanique concernant les inégalités séculaires des planètes.

4 Le Verrier (1840) : Trouver les coefficients du polynôme caractéristique

En (1840), Le Verrier cherchait à étudier les variations séculaires des éléments elliptiques des sept planètes principales: Mercure, Vénus, la Terre, Mars, Jupiter, Saturne et Uranus. Ce sera un article de 45 pages sans trop de détails et d'explications. Il commence par définir l'excentricité e de Mercure, et la longitude de son périhélie ϖ. Dans une orbite képlérienne, l'excentricité est le rapport de la distance entre les foyers et du grand axe, tandis que le périhélie est le point de l'orbite d'un corps céleste où la distance de ce corps au soleil est minimale. Il définit :

$$\begin{aligned} h &= e \sin \varpi, \\ l &= e \cos \varpi. \end{aligned} \tag{5}$$

Puis pour Vénus, la Terre, Mars, Jupiter, Saturne et Uranus, il donne des indices aux paramètres ,, ,,, ,,,, $^{\mathrm{IV}}$, $^{\mathrm{V}}$, $^{\mathrm{VI}}$. Il suit alors les notations de Laplace dans son *Traité de Mécanique Céleste* (Laplace 1799, livre II: 337). Alors les variations de h, l, $h^{'}$, $l^{'}$, ... dépendront du système différentiel suivant :

$$\begin{aligned}
\frac{dh}{dt} &= [(0,1)+(0,2)+\ldots]\, l - \boxed{0,1}\, l^{,} - \boxed{0,2}\, l^{,,} - \ldots, \\
\frac{dl}{dt} &= -[(0,1)+(0,2)+\ldots]\, h + \boxed{0,1}\, h^{,} + \boxed{0,2}\, h^{,,} + \ldots, \\
\frac{dh^{,}}{dt} &= [(1,0)+(1,2)+\ldots.]\, l^{,} - \boxed{1,0}\, l - \boxed{1,2}\, l^{,,} - \ldots, \\
\frac{dl^{,}}{dt} &= -[(\,1,0+(1,2)+\ldots\Big]\, h^{,} + \boxed{1,0}\, h + \boxed{1,2}\, h^{,} + \ldots, \\
& \qquad etc. \qquad\qquad\qquad etc.
\end{aligned} \tag{6}$$

Soit 14 équations au total. Les coefficients (0,1), (0,2), ..., $\boxed{0,1}$, $\boxed{0,2}$ ... ne dépendent que des masses perturbatrices et des grands axes des orbites. Ils sont liés deux à deux. Pour intégrer ces équations, Le Verrier pose comme solutions particulières :

$$\begin{aligned}
h &= N \sin(gt+\beta), & l &= N \cos(gt+\beta), \\
h^{,} &= N^{,} \sin(gt+\beta), & l^{,} &= N^{,} \cos(gt+\beta), \\
& etc. & & etc.
\end{aligned} \tag{7}$$

où les valeurs propres g sont à déterminer. Le Verrier nommera les valeurs propres, les racines ou arguments par g comme Laplace. Il peut alors dériver à gauche en fonction de t, et après substitution dans le système différentiel, il obtient un système algébrique d'équations linéaires homogènes :

$$\begin{aligned} &(g-(0,1)-(0,2)-\dots)\,N+\boxed{0,1}N^{'}+\boxed{0,2}N^{''}+\cdots=0,\\ &\boxed{1,0}N+(g-(1,0)-(1,2)-\dots)\,N^{'}+\boxed{1,2}N^{''}+\dots\quad=0,\\ &\boxed{2,0}N+\boxed{2,1}N^{'}+\big(g-(2,0-(2,1)-\dots)\,N^{''}+\dots\quad=0,\\ &\qquad\qquad etc. \end{aligned} \tag{8}$$

Comme Jacobi en (1846), on a légèrement modifié l'ordre pour rendre le système homogène plus compréhensible et symétrique. Le Verrier a obtenu ainsi un problème algébrique de valeurs propres. Si on calcule le déterminant du système, on obtient un polynôme caractéristique en g de degré 7 qui est difficile à calculer. Soient $g, g_1, g_2, \dots, g_6$ ces racines. Le Verrier va donc chercher des solutions plus raffinées du type :

$$\begin{aligned} h&=N\sin(gt+\beta)+N_1\sin(g_1t+\beta_1)+etc.,\\ l&=N\cos(gt+\beta)+N_1\cos(g_1t+\beta_1)+etc., \end{aligned} \tag{9}$$

etc.

Le Verrier, comme Lagrange et Laplace avant lui, utilisera une approche intuitive et physique plutôt qu'une approche purement mathématique. Le Verrier le dira explicitement :

> Les conditions nécessaires pour la stabilité de notre système planétaire, relativement aux excentricités, sont de deux sortes : les unes ont rapport à la nature des racines de l'équation en g, les autres à la grandeur absolue des coefficients $N, N_1, \dots, N', N'_1, \dots etc.$
>
> Il est indispensable que les racines de l'équation en g soient toutes réelles et en outre inégales. Autrement les expressions des excentricités contiendraient des termes ayant le temps en facteur ou en exposant, et par là elles croîtraient indéfiniment. Or il est facile de prouver que cet accroissement indéfini des excentricités est impossible . . .

Puis Le Verrier reconnaît que la détermination des racines est la partie la plus délicate de l'analyse, et pour plus de simplicité, il écrira le système homogène comme :

$$\begin{aligned} (g-a)\,N+bN^{'}+cN^{''}+\dots\qquad&=0,\\ a^{'}N+\left(g-b^{'}\right)N^{'}+c^{'}N^{''}+\dots\quad&=0,\\ a^{''}N+b^{''}N^{'}+\left(g-c^{''}\right)N^{''}+\cdots&=0,\\ etc.\qquad\qquad& \end{aligned} \tag{10}$$

Lorsqu'il développe le déterminant, il obtient une équation du septième degré. Puis il trouve que la somme des racines de l'équation est égale à la somme des coefficients $(a+b^{'}+c''+\dots)$, c'est-à-dire la trace du système matriciel (10).

Soit :

$$g + g_1 + g_2 + \cdots = a + b^{'} + c^{''} + \ldots \tag{11}$$

Notons que le mot trace était alors inconnu. En (1847) M. C.-W. Borchard, un élève de Jacobi, écrira un article fort élégant, et moderne sur ce sujet, ainsi que la preuve que les valeurs propres sont réelles dans un système symétrique.

Puis Le Verrier va dériver par rapport au temps son système différentiel (6) pour obtenir un nouveau problème algébrique de valeurs propres en g^2. Il écrit son nouveau système d'équations :

$$\begin{array}{llll} \left(g^2 - A\right) N + & BN^{'} & + \quad CN^{''} & + \cdots = 0, \\ A^{'} N + & \left(g^2 - B^{'}\right) N^{'} & + C^{'} N^{''} & + \cdots = 0, \\ A^{''} \quad + & B^{''} N^{'} & + \left(g^2 - C^{''}\right) N^{''} & + \cdots = 0. \end{array} \tag{12}$$

Le Verrier conclut qu'on peut effectuer le même raisonnement que pour le système (10) et que la somme des carrés des racines (les valeurs propres au carré) est égale à la trace du nouveau système d'équations. Il ajoute :

> Il n'est pas besoin de plus longues explications pour s'apercevoir qu'on obtiendra la somme des cubes, des quatrièmes puissances... des racines, au moyen des dérivées troisièmes, quatrième... de *h*, *h'*, ... Et d'ailleurs, quand on aura les sommes des puissances semblables des racines jusqu'à la septième, il sera facile de calculer les coefficients de l'équation (le polynôme caractéristique) par les fonctions symétriques.

En écriture plus moderne, Le Verrier fait le raisonnement suivant : si on a un système différentiel du premier ordre, linéaire et à coefficients constants $\dot{\mathbf{h}} = \mathbf{Ah}$, on aura en dérivant $\ddot{\mathbf{h}} = \mathbf{A}\dot{\mathbf{h}} = \mathbf{A^2h}$, et par induction mathématique $\mathbf{h^{(k)}} = \mathbf{A}^k\mathbf{h}$. Donc pour les valeurs propres g, on aura $\mathbf{Ah} = g\mathbf{h}$ et $\mathbf{A^2h} = \mathbf{A}g\mathbf{h} = g^2\mathbf{h}$, de même que $\mathbf{A}^k\mathbf{h} = g^k\mathbf{h}$. Donc pour les traces des matrices, on posera :

$$S_1 = Tr\,(\mathbf{A}) = \sum_{i=1}^{n} g_i, \text{et } S_k = \sum_{i=1}^{n} g_i^k = \sum_{i=1}^{n} a_{ii}^{(k)}. \tag{13}$$

Le Verrier ne donne pas plus de détails, mais il dit qu'on pourra consulter son article de (1843) dans *La Connaissance du temps* pour plus de renseignements. Cet article avait été présenté à l'Académie des Science le 16 septembre 1839. Le Verrier va alors utiliser le théorème de Newton-Girard pour relier les coefficients du polynôme caractéristique avec les racines. Il faudra ensuite trouver les racines de ce polynôme. Dans la section 13 de cet article, Le Verrier cherche les variations des racines du polynôme caractéristique, et il prendra l'exemple d'un polynôme de degré 3 :

$$g^3 + a_1 g^2 + a_2 g + a_3 = 0, \tag{14}$$

qu'il compare à :

$$(g - g_0)\,(g - g_1)\,(g - g_2) = 0. \tag{15}$$

Par identification, Le Verrier obtient les relations suivantes :

$$\begin{aligned} a_1 &= g_0 - g_1 - g_2, \\ a_2 &= g_0 g_1 + g_0 g_2 + g_1 g_2, \\ a_3 &= -g_0 g_1 g_2. \end{aligned} \tag{16}$$

Il doit alors généraliser à un polynôme de degré 7, et trouver un algorithme pour relier ensemble les équations du type (16) et les équations du type (13) (Fröberg 1969: 53).

Notons que la revue *Connaissance du temps* était une publication du Bureau des Longitudes à Paris, une académie fondée en 1795 pendant la Révolution Française. À l'époque de Le Verrier, quatre astronomes étaient membres: Bouvard, Arago, Biot et Mathieu. Ce fut Arago qui suggéra à Le Verrier de travailler sur les irrégularités de la planète Uranus. Ceci fut couronné de succès, puisqu'en 1846, Le Verrier, par le calcul [ou: ses calculs], découvrit la position de la planète Neptune.

5 Trouver les racines d'un polynôme

Les racines d'un polynôme ont toujours fasciné les mathématiciens avec le théorème fondamental de l'algèbre (Gilain 1991). C'est-à-dire que tout polynôme de degré $n \geq 1$ peut se décomposer en un produit de n facteurs du premier degré, ou qu'il possède exactement n racines (distinctes ou confondues). Aussi, le théorème de Ruffini-Abel précise que pour tout polynôme de degré égal ou supérieur à cinq, il n'existe pas de formule analytique pour calculer les racines. On sera concerné alors par l'aspect itératif et algorithmique du problème. La méthode consiste en une boucle : une première phase de localisation des racines (Benis-Sinaceur 1992), puis la détermination d'une racine par approximations successives (Chabert et al. 1993 ; Laubenbacher et al. 2001), une troisième phase de division synthétique ou de déflation, puis relocaliser une autre racine et ainsi de suite jusqu'à épuisement.

La méthode de Le Verrier permettait de trouver les coefficients du polynôme caractéristique mais pas la détermination numérique des sept racines. Il va procéder par approximations en localisant d'abord les trois grosses planètes Jupiter, Saturne et Uranus, donc en négligeant les autres planètes. Ceci peut se faire par calcul, puis il va procéder par corrections successives pour l'ensemble des sept planètes.

6 La méthode de Jacobi (1846)

Jacobi a montré un intérêt certain pour les mathématiques appliquées et les méthodes. En 1826, il publie sa version de la méthode inventée par Gauss pour trouver par approximations la valeur d'une intégrale. Puis en 1845, ce sera sa méthode par approximations successives d'un système d'équations linéaires inhomogènes. Ceci se montrera d'une grande utilité pour la solution des grands systèmes d'équations que l'on retrouve dans les traitements par moindres-carrés (Chabert et al. 1993: 337, 490). Nous étudions ici son article sur les variations séculaires et les valeurs propres d'un système symétrique (Jacobi 1846). Ce sera un article majeur où il présente une nouvelle méthode pour trouver les valeurs propres et les vecteurs propres, et où il reprend le problème de Le Verrier de 1840. Jacobi va utiliser pleinement les propriétés de la symétrie pour montrer que les vecteurs propres sont orthonormés, que les valeurs propres sont réelles, et il sera capable de trouver des systèmes orthogonaux et de diagonaliser. Son article, écrit en allemand, s'intitula (en français) ≪Comment résoudre aisément les équations qui apparaissent dans la théorie des perturbations séculaires≫. Il commença ainsi [1]:

> Dans la théorie des perturbations séculaires et des petites oscillations, nous avons à connaitre un système d'équations linéaires où les différentes inconnues sont affectées de paramètres symétriques par rapport à la diagonale, et où on ajoute à tous les coefficients de la diagonale la même valeur (die Gröfse) λ ... Cette méthode permet de surmonter les difficultés relatives à la résolution de l'équation dont les racines sont les valeurs de λ. Elle consiste à transformer le système de telle sorte que l'on obtienne pour tous les λ, des valeurs très rapprochées et que pour chaque λ une procédure d'approximation convergeant rapidement...

Son article est divisé en deux parties. La première partie est plus théorique et s'inspirera d'un autre article de Jacobi publié en latin en 1834. Dans la deuxième partie, il présente sa méthode et reproduit les travaux de Le Verrier.

6.1 La partie théorique

Jacobi commence la partie théorique de son article (1846) en présentant son système d'équations homogènes $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$ où le vecteur propre $\mathbf{x}$ est représenté par ses composantes qui sont écrites en lettres grecques. Bien sûr, Jacobi ne connaissait pas le mot vecteur propre ni la notation matricielle[2]. Jacobi écrivait donc : $\mathbf{x}' = (\alpha', \beta', \gamma', \ldots, \omega')^t$. Il va tout de suite supposer que les vecteurs sont

[1] Ceci est notre traduction de l'allemand vers le français.

[2] Rappelons que le mot vecteur émergea des travaux d'Hamilton sur les quaternions en 1845 (Moore 1995: 265).

normalisés et il pose: $\alpha'\alpha' + \beta'\beta' + \gamma'\gamma' + \ \ldots\ + \omega'\omega' = 1$. Puis il présente son système d'équations :

$$
\begin{array}{l}
\left\{(a,a) - \lambda'\right\}\alpha' + (a,b)\,\beta' + (a,c)\,\gamma' + \cdots + (a,p)\,\omega' = 0,\\
(b,a)\,\alpha' + \left\{(b,b) - \lambda'\right\}\beta' + (b,c)\,\gamma' + \cdots + (b,p)\,\varpi' = 0,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
(p,a)\,\alpha' + (p,b)\,\beta' + (p,c)\,\gamma' + \cdots + \left\{(p,p) - \lambda'\right\}\omega' = 0.
\end{array}
\tag{17}
$$

Notons que pour l'écriture des coefficients, Jacobi est plus moderne que Cauchy. Son système est aussi d'ordre n. Jacobi va insister sur la symétrie du problème par rapport à la diagonale principale avec $(a,b) = (b,a)$, $(a,c) = (c,a)$, $(b,c) = (c,b)$, et ainsi de suite. Les inconnues sont λ', α', β', $\ldots$, ω'. Donc Jacobi augmente la dimensionnalité du problème par rapport à Le Verrier. À cause du manque de connaissance de l'écriture matricielle, l'écriture de Jacobi sera lourde et un peu acrobatique, car il ne pourra travailler qu'avec des systèmes d'équations linéaires. On s'intéresse ici à ses techniques de preuve. Il va multiplier la première ligne du système d'équations inhomogènes par α'', la deuxième ligne par β'', et ainsi de suite, donc il multiplie la dernière ligne par ω''. Ainsi α'', β'', γ'', $\ldots$, ω'' sont les composantes d'un autre vecteur propre $\mathbf{x}''$. Il obtiendra alors :

$$
\begin{array}{l}
(a,a)\,\alpha'\alpha'' + (a,b)\,\beta'\alpha'' + (a,c)\,\gamma'\alpha'' + \cdots + (a,p)\,\omega'\alpha'' = \alpha'\alpha''\lambda',\\
(b,a)\,\alpha'\beta'' + (b,b)\,\beta'\beta'' + (b,c)\,\gamma'\beta'' + \cdots + (b,p)\,\varpi'\beta'' = \beta'\beta''\lambda',\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
(p,a)\,\alpha'\varpi'' + (p,b)\,\beta'\varpi'' + (p,c)\,\gamma'\varpi'' + \cdots + (p,p)\,\omega'\varpi'' = \omega'\varpi''\lambda'.
\end{array}
\tag{18}
$$

Il va faire la même chose pour le système avec la valeur propre λ'' et le vecteur propre $\mathbf{x}''$ en multipliant la première ligne par α', la deuxième ligne par β' et la dernière ligne par ω'.

$$
\begin{array}{l}
(a,a)\,\alpha''\alpha' + (a,b)\,\beta''\alpha' + (a,c)\,\gamma''\alpha' + \cdots + (a,p)\,\omega''\alpha' = \alpha''\alpha'\lambda'',\\
(b,a)\,\alpha''\beta' + (b,b)\,\beta''\beta' + (b,c)\,\gamma''\beta' \cdots + (b,p)\,\varpi''\beta' = \beta''\beta'\lambda'',\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
(p,a)\,\alpha''\varpi' + (p,b)\,\beta''\varpi' + (p,c)\,\gamma''\varpi' + \cdots + (p,p)\,\omega''\varpi' = \omega''\varpi'\lambda''.
\end{array}
\tag{19}
$$

Jacobi additionne alors les équations entre elles des systèmes inhomogènes (18) puis (19) et il observe que les parties à gauche sont identiques. Par exemple, le terme

$(a,b)\beta^{'}\alpha''$ de la première ligne du système (18) est équivalent au terme $(b,a)\alpha''\beta^{'}$ de la deuxième ligne du système (19). Il obtient alors la relation suivante :

$$\left(\alpha^{'}\alpha^{''}+\beta^{'}\beta^{''}+\cdots+\omega^{'}\omega^{''}\right)\lambda^{'}=\left(\alpha^{''}\alpha^{'}+\beta^{''}\beta^{'}+\cdots+\varpi^{''}\varpi^{'}\right)\lambda^{''}. \tag{20}$$

Or, Jacobi sait de Cauchy (1829) que pour son système symétrique, les valeurs propres seront réelles et supposées distinctes, il conclut alors que dans ce cas, les vecteurs propres doivent être mutuellement orthonormés. Donc :

$$\alpha^{'}\alpha^{''}+\beta^{'}\beta^{''}+\cdots+\omega^{'}\omega^{''}=0. \tag{21}$$

Ensuite Jacobi va travailler avec les expressions linéaires suivantes :

$$\begin{aligned}
p_{\mathbf{1}} &= \alpha^{'}q_{\mathbf{1}}+\alpha^{''}q_{\mathbf{2}}+,,,+\alpha^{(n)}q_{\mathbf{n}},\\
p_{\mathbf{2}} &= \beta^{'}q_{\mathbf{1}}+\beta^{''}q_{\mathbf{2}}+,,,+\beta^{(n)}q_{\mathbf{n}},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
p_{n} &= \varpi^{'}q_{\mathbf{1}}+\varpi^{''}q_{\mathbf{2}}+,,,+\varpi^{(n)}q_{\mathbf{n}}.
\end{aligned} \tag{22}$$

Les quantités $q_{\mathbf{1}}$, $q_{\mathbf{2}}$, $\ldots q_{\mathbf{n}}$ ne seront jamais explicitement données. Alors, les $p_{\mathbf{1}}, p_{\mathbf{2}}, \ldots, p_{\mathbf{n}}$ correspondront au système transposé. Jacobi parlera de coefficients horizontaux et de coefficients verticaux. Puis, il multiplie la première ligne de (22) par $\alpha^{'}$, la deuxième par $\beta^{'}$, et ainsi de suite, puis fait l'addition des lignes et simplifie en tenant compte des relations d'orthogonalité. Par symétrie, on va avoir les relations suivantes pour exprimer les quantités q en fonction des p :

$$\begin{aligned}
q_{\mathbf{1}} &= \alpha^{'}p_{\mathbf{1}}+\beta^{'}p_{\mathbf{2}}+\cdots+\varpi^{'}p_{\mathbf{n}},\\
q_{\mathbf{2}} &= \alpha^{''}p_{\mathbf{1}}+\beta^{''}p_{\mathbf{2}}+\cdots+\varpi^{''}p_{\mathbf{n}},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
q_{n} &= \alpha^{(n)}p_{\mathbf{1}}+\beta^{(n)}p_{\mathbf{2}}+\cdots+\varpi^{(n)}p_{\mathbf{n}}.
\end{aligned} \tag{23}$$

En remplaçant dans (22) les $q_{\mathbf{1}}, q_{\mathbf{2}}, \ldots, q_{\mathbf{n}}$ par les relations (23), Jacobi obtient les identités suivantes d'orthogonalité :

$$\begin{aligned}
&\alpha^{'2}+\alpha^{''2}+\cdots+\alpha^{(n)^2}=1,\\
&\alpha^{'}\beta^{'}+\alpha^{''}\beta^{''}+\cdots+\alpha^{(n)}\beta^{(n)}=0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned} \tag{24}$$

Bien sûr, Jacobi ne connaissait pas les outils matriciels, mais en fait, il avait construit une matrice orthogonale du type :

$$\begin{pmatrix} \alpha' & \alpha'' & \ldots & \alpha^{(n)} \\ \beta' & \beta'' & & \beta^{(n)} \\ . & & & . \\ . & & & . \\ . & & & . \\ \omega' & \varpi'' & \ldots & \omega^{(n)} \end{pmatrix} = \left(\mathbf{x}' \; \mathbf{x}'' \; \ldots \; \mathbf{x}^{(\mathbf{n})} \right). \tag{25}$$

Les vecteurs colonnes sont les vecteurs propres, et ils sont mutuellement orthogonaux entre eux. Le fait de découvrir des systèmes orthogonaux remonte à Leonard Euler en 1770 (Dorier 1995: 238). Une matrice orthogonale est définie par le fait que sa matrice inverse est simplement la matrice transposée. Puis, Jacobi va essayer de construire son système diagonal par rapport aux valeurs propres, mais sa méthode sera compliquée. À partir du système représenté par $\mathbf{Ax}' = \lambda'\mathbf{x}'$, il va multiplier la première ligne par la valeur propre λ', la deuxième par λ'', et ainsi de suite jusqu'à la dernière par $\lambda^{(n)}$. Il aurait dû en fait partir des systèmes $\mathbf{Ax}' = \lambda'\mathbf{x}'$, $\mathbf{Ax}'' = \lambda''\mathbf{x}''$, $\ldots$, $\mathbf{Ax}^{(\mathbf{n})} = \lambda^{(\mathbf{n})}\mathbf{x}^{(\mathbf{n})}$, et faire une concaténation des vecteurs propres du type :

$$\mathbf{A}\left[\mathbf{x}' \; \mathbf{x}'' \; \ldots \mathbf{x}^{(\mathbf{n})}\right] = \begin{pmatrix} \lambda' & & & \\ & \lambda'' & & \\ & & . & \\ & & & \lambda_n \end{pmatrix} \left[\mathbf{x}' \; \mathbf{x}'' \; \ldots \mathbf{x}^{(\mathbf{n})}\right]. \tag{26}$$

Ou :

$$\left[\mathbf{x}' \; \mathbf{x}'' \; \ldots \mathbf{x}^{(\mathbf{n})}\right]^t \mathbf{A} \left[\mathbf{x}' \; \mathbf{x}'' \; \ldots \mathbf{x}^{(\mathbf{n})}\right] = \begin{pmatrix} \lambda' & & & \\ & \lambda'' & & \\ & & . & \\ & & & \lambda_n \end{pmatrix}. \tag{27}$$

Contrairement à Le Verrier, Jacobi obtiendra toutes les valeurs propres à la fois. Bien sûr, comme il ne connait pas les $\mathbf{x}'$, $\mathbf{x}''$, $\ldots$, $\mathbf{x}^{(\mathbf{n})}$, il va procéder par approximations successives.

Ainsi on peut dire que sa méthode est basée sur une suite de matrices orthogonales $\{\mathbf{O_k}\}_{k=1}^{+\infty}$ telles que $\mathbf{A_{k+1}} = \mathbf{O_k^t A_k O_k} \to \mathbf{D}$, où $\mathbf{D}$ est une matrice diagonale. Donc, Jacobi, à partir de la matrice $\mathbf{A_k}$, devra faire une double multiplication matricielle avec les matrices $\mathbf{O_k}$ et $\mathbf{O_k^t}$, alors que la théorie des matrices n'était pas développée ! Il va donc travailler avec une fenêtre du système homogène représenté par $(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = 0$. Son secret consistera en des multiplications de lignes par un scalaire et des additions de lignes. On sait que le mot matrice fut proposé par Sylvester en 1850, mais si on lit l'article de Cayley en (1855): ≪Remarques sur la notation des fonctions algébriques≫, publié en français dans le *journal de Crelle* (le journal de Jacobi), on y trouvera toutes des structures mathématiques et les multiplications matricielles qui manquaient à Jacobi en 1846, soit un écart de neuf ans. Jacobi mourut de la variole en 1851, à l'âge de 46 ans.

6.2 La partie pratique

La partie pratique de l'article (1846) de Jacobi est mieux connue et mieux décrite dans la plupart des livres d'analyse numérique que la partie théorique. Jacobi va chercher à reproduire les résultats de Le Verrier de 1840. Nous allons simplifier un peu les équations de Jacobi en un système 2×2 pour alléger l'écriture. Aussi pour nous distinguer de ce que fit Jacobi. Il va choisir la matrice de rotation à deux dimensions. C'est une matrice orthogonale (Fröberg 1969: 121–122) :

$$\mathbf{O} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \tag{28}$$

pour effectuer les opérations suivantes :

$$\mathbf{D} = \mathbf{O}^{\mathbf{t}}\mathbf{A}\mathbf{O} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} a_{ii} & a_{ik} \\ a_{ki} & a_{kk} \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}. \tag{29}$$

En fait, il va partir d'une fenêtre du système d'équations de Le Verrier où il a déjà effectué une multiplication matricielle :

$$\begin{array}{l} \cdots + (\lambda - a)\, M_i \;+ \cdots + cM_k \qquad\qquad + \cdots + h_{i,l} M_l \;+ \cdots = 0 \\ \cdots + \qquad\qquad cM_i + \cdots + (\lambda - b)\, M_k + \cdots + h_{k,l} M_l + \cdots = 0 \end{array}, \tag{30}$$

avec:

$$\begin{aligned} M_i &= (\cos\alpha)\, P_i - (\sin\alpha)\, P_k, \\ M_k &= (\sin\alpha)\, P_1 + (\cos\alpha)\, P_k, \end{aligned} \tag{31}$$

où P_k et P_i sont les composantes des vecteurs propres inconnus. Essayons d'expliquer en notation matricielle moderne ce que fit Jacobi. Partant d'un système 2×2 :

$$\begin{aligned} (\lambda - a)\, P_1 + \quad\quad c P_2 &= 0, \\ c P_1 \quad\quad + (\lambda - b)\, P_2 &= 0. \end{aligned} \tag{32}$$

Il va effectuer l'opération $\mathbf{AOP} = \mathbf{0}$. Soit (équation 31) :

$$\begin{pmatrix} M_1 \\ M_2 \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \end{pmatrix}. \tag{33}$$

Et (équation 30) :

$$\begin{pmatrix} \lambda - a & c \\ c & \lambda - b \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} = \begin{pmatrix} \lambda - a & c \\ c & \lambda - b \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{34}$$

Développons le système (34), ce que fit Jacobi en remplaçant dans (30), les expressions de M_1 et M_2 obtenus de (31). On obtient :

$$((\lambda - a)\cos\alpha + c\sin\alpha)\, P_1 + ((\lambda - a)\,(-\sin\alpha) + c\cos\alpha)\, P_2 = N_1, \tag{35}$$

$$((c\cos\alpha + (\lambda - b)\sin\alpha)\, P_1 + (c\,(-\sin\alpha) + (\lambda - b)\cos\alpha)\, P_2) = N_2. \tag{36}$$

Il faut maintenant effectuer la nouvelle multiplication matricielle $\mathbf{O}^t\mathbf{N} = \mathbf{0}$, avec $\mathbf{N} = \begin{pmatrix} N_1 \\ N_2 \end{pmatrix}$. Jacobi le dit de la façon suivante: on multiplie (35) par cos α et (36) par sin α, et on additionne les deux lignes, puis il effectuera les opérations suivantes: multiplier (35) par $-$ sin α et la deuxième ligne par cosα et additionner. Il aura alors effectué sa double multiplication matricielle ! Jacobi a bien utilisé la matrice transposée qui est aussi orthogonale et le concept de la multiplication d'un vecteur ligne avec un vecteur colonne. Il obtient deux équations homogènes du type (Jacobi 1846: 65) :

$$\begin{aligned} &\left\{\lambda - a\cos^2\alpha - b\sin^2\alpha + 2c\sin\alpha\cos\alpha\right\} P_1 + \{(a - b)\sin\alpha\cos\alpha + c\cos 2\alpha\}\, P_2 = 0 \\ &\{(a - b)\sin\alpha\cos\alpha + c\cos 2\alpha\}\, P_1 + \left\{\lambda - a\sin^2\alpha - b\cos^2\alpha - 2c\sin\alpha\cos\alpha\right\} P_2 = 0 \end{aligned} \tag{37}$$

Jacobi va ensuite forcer le système (37) à être diagonal en forçant les composantes hors diagonale à être égales à zéro. Il pose donc :

$$(a - b)\sin\alpha\cos\alpha + c\cos 2\alpha = 0.\ \text{Soit : } \tan 2\alpha = \frac{2c}{b - a}. \tag{38}$$

On avait pris l'exemple simple d'un système 2×2, mais Jacobi va progressivement étendre sa fenêtre de diagonalisation. La méthode de Jacobi sera itérative comme sa méthode pour résoudre des systèmes d'équations linéaires, c'est-à-dire qu'il procédera par approximations successives pour mettre les éléments hors de la diagonale principale égaux à zéro.

7 Conclusion

L'algorithme de Le Verrier demeurera longtemps la méthode pour trouver les valeurs propres. Elle était encore enseignée dans les années 1960–1970, surtout avec la modification de Faddeev et Faddeeva (1963). Sur 350 livres environ sur les méthodes numériques, couvrant la période de 1960 à l'an 2000, on a trouvé onze références sur la méthode de Le Verrier. La méthode de Jacobi gagna de l'importance avec les ordinateurs numériques. Un programme de la méthode de Jacobi contient 50 instructions environ. Pour connaitre l'histoire de la méthode de Jacobi, citons Golub et van der Vorst (2001: 216–217) :

> Dans les années 1960, la popularité de la méthode déclina, à cause de la popularité grandissante d'abord de la méthode de Givens, et peu après de la méthode de Householder. Ces deux méthodes réduisent d'abord la matrice en une forme tridiagonale et puis utilisent une procédure efficace pour calculer les valeurs propres de la matrice tridiagonale.

Rappelons cependant qu'avec les langages formels comme MAPLE et la possibilité d'un nombre de chiffres significatifs flexibles, des variantes améliorées issues de la technique de Le Verrier sont redevenues intéressantes.

Finalement, le problème numérique des valeurs propres a été peu étudié par les historiens de mathématiques, pourtant Jacobi a bien contribué à la théorie des systèmes orthogonaux, aux vecteurs propres, et aux multiplications matricielles.

Références

Borchardt M C-W (1847) Développements sur l'équation à l'aide de laquelle on détermine les inégalités séculaires du mouvement des planètes. Journal de Math Pures et Appl: 50-67

Brechenmacher F (2007) L'identité algébrique d'une pratique portée par la discussion sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des planètes (1766-1874). Sciences et Techniques en Perspectives, 11e série, fasc 1: 5-85

Chabert J L et al. (1993) Histoire d'algorithmes, du caillou à la puce. Belin, Paris

Cauchy L A (1829) Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planètes. Exer. de Mathématiques 4. Les Œuvres (2)9: 174-195.

Cauchy L A (1840) Mémoire sur l'intégration des équations linéaires. Exercices d'analyse et de physique mathématique. Bachelier imprimeur-libraire, Paris, I: 53-100. Les Œuvres, II, t. XI :75-88

Cayley A (1855) Remarques sur la notation des fonctions algébriques. Crelle's J.: 282-285. The Collected Mathematical Papers, Vol. II, Cambridge University Press, Cambridge (1889): 185-188

Dorier J-L (1995) A General Outline of the genesis of Vector Space Theory. Historia Mathematica, 22: 227-261

Faddeev D K Faddeeva V N (1963) Computational Methods of Linear Algebra. W.H. Freeman editor, San Francisco. First published in Russian in 1960.

Fröberg C-E (1969) Introduction to numerical analysis. Addison-Wesley, Reading

Gilain C (1991) Sur l'histoire du théorème fondamental de l'algèbre: théorie des équations et calcul intégral. Archive for History of Exact Sciences. 42(2): 91-132

Golub G H, van der Vorst H A (2001) Eigenvalue computation in the 20th century. In Historical Developments in the 20th Century, C. Brezinski et L Wuytack eds: 209-237

Hairer E, Norsett S P, Wanner G (1987) Solving ordinary differential equations. Springer-Verlag, Berlin

Hawkins T (1975) Cauchy and the spectral theory of matrices. Historia Mathematica: 1-29

Jacobi C G J (1846) Uber ein Leeichtes Verfahren Die in der Theorie der Sacularstorungen Vorkommendern Gleichungen Numerisch Aufzulosen, Crelle's J., 30, 51-94

Laplace P-S (1799) Traité de Mécanique Céleste, Réédité par Éditions Jacques Gabay, Paris, (2006), tome I

Laubenbacher R McGrath G Pengelley D (2001) Lagrange and the Solution of Numerical Equations. Historia Mathematica, 28: 220-231

Le Verrier U J-J (1840). Sur les variations séculaires des éléments elliptiques des sept planètes principales : Mercure, Vénus, la Terre, Mars, Jupiter, Saturne et Uranus, J. de Liouville, tome 5, 220-254.

Le Verrier U J-J (1843) Mémoire sur les variations séculaires des éléments des orbites, pour les sept planètes principales, Mercure, Vénus, la Terre, Mars, Jupiter, Saturne et Uranus. La Connaissance du temps, Paris: 3-66

Moore G H (1995) The Axiomatization of Linear Algebra: 1875-1940. Historia Mathematica: 262-303

Sinaceur H (1992) Cauchy, Sturm et les racines des équations. Revue d'histoire de sciences, 45: 51-68

Mathematics in Astronomy at Harvard College Before 1839 as a Case Study for Teaching Historical Writing in Mathematics Courses

Amy Ackerberg-Hastings

Abstract Writing a research paper is an important but challenging milestone for secondary and undergraduate students, one that does not need to be limited to disciplines that have been traditionally writing-oriented, such as history or English. Yet, instructors in any discipline may find that teaching the research and writing process—especially in a way that produces, for instance, both sound history and valid mathematics—can be challenging. Meanwhile, the first five men to hold Harvard's Hollis Professorship of Mathematics and Natural Philosophy all had personal lives and professional achievements that catch students' interest, but little scholarly work has been done on topics such as the intersection between their teaching of mathematics and their teaching of astronomy. This paper brings together principles useful for teaching historical research and writing with the specific historical problem of how mathematics was employed in the teaching of astronomy at Harvard before the college established an observatory in 1839. It thus provides students and instructors with an opportunity to experience the application of historical thinking skills as a warm-up for undertaking original research in the history of mathematics.

1 Introduction

How might instructors teach students to delve into the past of mathematics and turn what they learn into the history of mathematics? Perhaps this is a particularly important concern for those who have been pressed into teaching a history of mathematics course without having studied the subject themselves or those whose institutions encourage substantive research and writing in all disciplines. To be sure, numerous excellent research and writing handbooks exist (for instance, on

A. Ackerberg-Hastings (✉)
Independent Scholar, Rockville, MD, USA
e-mail: aackerbe@verizon.net

M. Zack, D. Schlimm (eds.), *Research in History and Philosophy of Mathematics*, Annals of the Canadian Society for History and Philosophy of Mathematics/Société canadienne d'histoire et de philosophie des mathématiques,
https://doi.org/10.1007/978-3-030-95201-3_5

research training in any discipline, see Booth et al. (2008) as well as Turabian (2013); manuals in mathematics and the history of mathematics include Delaware (2019), Higham (2019), and Krantz (2017); and in history some guides designed for undergraduates are Cronon (2008–2009), Rael (2004), and Storey (2009)). There is only so far, though, that *telling* students what to do in a research project can take them. Therefore, this article guides readers through the initial stages of articulating a research plan for a specific large-scale historical project, pausing frequently for "methodological moments" that share insights into what historians may have on their minds at various points along the way. The aim is *showing* students what it means to think critically about the past and *involving* them in the legwork and mental effort required to find and analyze primary sources, so that they may identify what happened and interpret how and why developments unfolded as they did. In other words, the paper can serve an intermediate instructional role between initially covering assignment guidelines and turning students loose to undertake their own projects. Additionally, I encourage instructors to ask their students to return to the article as they think through their research processes, taking the time to revisit each methodological moment and determine how the points raised apply to their topics.

The concrete historical problem provided as the frame for the methodological moments was motivated by this question: How mathematical was the teaching of astronomy by the five men who served Harvard College as the Hollis Professor of Mathematics and Natural Philosophy in the years before the college established an observatory? (Lovering appears in Table 1 to answer readers' questions about who held the chair in 1839, when Harvard built its observatory, but his half-century-long career will not be covered in any depth here.) Even though the chair has merited an entry in encyclopedias such as Lawson (2008), it appears to have received very little historical study as a phenomenon in its own right since "Some of Harvard's" 1926–1927, 1927–1928 was published. Rather, historians of American mathematics, for instance, have tended to focus solely on the mathematical activities of individual Hollis Professors (the range of work over time includes Cajori (1890), Coolidge (1924), Tarwater (1977), and Ackerberg-Hastings (2000)). Yet, as was typical of similarly named academic positions elsewhere in Western Europe and North America in the eighteenth and nineteenth centuries, the professors all not only taught mathematics but also lectured on natural philosophy, astronomy, navigation, and similar subjects. In fact, the professors listed in Table 1 built up a notable and long-lived tradition of astronomy instruction despite lacking a

Table 1 Hollis Professors of Mathematics and Natural Philosophy before 1839

Tenure	Name
1728–1738	Isaac Greenwood
1739–1779	John Winthrop
1780–1788	Samuel Williams
1789–1806	Samuel Webber
1807–1836	John Farrar
1838–1888	Joseph Lovering

formal observatory. Some scholars have looked at this sort of "pre-history" of Harvard astronomy (most notably, Portolano, 2000), while others have written about the professors' contributions to areas such as physics and meteorology (for instance, Varney, 1908, Cohen, 1950, and Elliott and Rossiter, 1992), but few historians of science or mathematics appear to have asked whether teaching and studying mathematics influenced the Hollis Professors' approaches to teaching and studying natural philosophy (or vice versa). The historical material discussed here begins to contemplate what it means to define the teaching of astronomy as "mathematical" as well as to make preliminary suggestions about how scholars might compare instruction in the discipline of mathematical astronomy as it was offered by professors whose main interests lay in mathematics to that given by those who preferred natural philosophy. Readers who turned to this paper hoping to learn about mathematics and astronomy at Harvard in the eighteenth and early nineteenth centuries should find some useful information in its analysis of the textbooks utilized by the college, lectures delivered by professors, and examinations completed by students. However, a full, monograph- or dissertation-length study of this rich body of primary-source material—much of it now generously digitized by Harvard librarians and archivists—has been left to a future researcher. It is possible that, in addition to encountering concerns in historiography and method that arise even in research paper assignments intended to be completed in a few weeks, instructors and students will also find portions of these resources appealing and utilize them as topics for short-term classroom projects.

Methodological Moment: Establishing Topic Parameters—Determining its scope is the first point of a historical research project at which a few words on thinking like a historian are offered. For instance, where should the beginning and ending points be placed? To state the question another way, if history is about identifying patterns of continuity and change, then where were the turning points after which astronomy at Harvard was different than it had been before? The number of preserved records and the level of activity both appear to have begun to increase with the advent of the Hollis Professorship in 1728, nearly 100 years after the college's founding in 1636. The establishment of the Harvard College Observatory in 1839, after decades of advocacy by Harvard faculty and other Cambridge leaders, similarly provides a convenient ending point. Besides the subsequent existence of a formal institution offering expanded opportunities to collect and analyze astronomical data, other transformations had occurred in Harvard mathematics and natural philosophy. In particular, Benjamin Peirce (1809–1880), who was not a Hollis Professor but rather held a new, initially unendowed professorship, was in the process of reshaping the Department of Mathematics, which had been separated from the other subjects (Kent, 2005; Hogan, 2008). Meanwhile, a clear distinction in the teaching of astronomy between research opportunities offered to a select few students and a largely non-mathematical, descriptive, classroom-based approach for the many had recurred throughout the previous era. In the next few decades this trend evolved even more dramatically into the emergence of graduate education and professional research. For more on the theory of periodization in historical interpretation, see Bentley (1996) and Stearns (2009). Rufus (1921) developed one such scheme for American astronomy. As implied by the adjective "American," other components of a project's scope can include geographical area and the people, institutions, or ideas involved.

Methodological Moment: Philosophical Foundations for Historical Research Questions—Setting temporal and geographical boundaries for a research project can also foster reflections on the purpose of researching and writing history. If our aim is only to list developments that apparently led to the observatory, then we risk neglecting parts of the story that do not fit a preconceived narrative of inevitable success. However, besides indicating the complexity of human thoughts and actions, apparent failures are often the most interesting and revelatory tales from the past. Further, pursuing what I think of, in part, as "how we got to today" history can distract us from understanding the people of the past on their own terms, from giving attention to their historical context. Ivor Grattan-Guinness's famous essays contrasting the *history* of mathematics with its *heritage*, especially (2004), are the classic recommended reading on this aspect of historical practice for mathematicians. Instructors may also wish to refer to two brief and useful summaries—from the many that are readily available—of how historians ground their research and writing in this respectful conception of the past (American Historical Association, 2019; Edwards, 2011). Drawing on the model for stating research questions outlined in Booth, Colomb and Williams (2008, pp. 35–48). I might thus summarize the parameters of this proposed project: The purpose of my research project is to better understand how and why astronomy was taught by Harvard's Hollis Professors of Mathematics and Natural Philosophy between 1728 and 1839, assess the extent to which their presentation of the subject may have been especially "mathematical" as a consequence of their job titles and own interests, and explore how those professors attempted to create physical and metaphorical spaces for making, recording, and analyzing observations. As noted above, this research question sets forth an ambitious agenda. While secondary students, undergraduates, and perhaps even beginning graduate students should be guided toward proposals that are more limited in scope, they still can be challenged to aim for at least this level of specificity in stating what they want to learn about the past.

2 John Winthrop

London merchant Thomas Hollis (1659–1731) endowed the Hollis Professorship in 1727, five years after funding a professorship in divinity that was also named after him. One of the people who advised Hollis about encouraging the physical sciences was **Isaac Greenwood (1702–1745, AM 1724)**, who subsequently returned to his hometown of Boston and was installed in 1728 as the first holder of the chair (Quincy, 1840; "Some of Harvard's" 1926–1927, 1927–1928). The college previously had no faculty explicitly charged with natural philosophy and mathematics; tutors who conducted the bulk of the teaching were assigned by class year and thus covered all subjects. The Hollis Professor's chief duty was to lecture once a week on natural philosophy. Under Greenwood's supervision, Harvard tutors apparently continued an existing tradition that required third-year students to recite in Latin from Pierre Gassendi's 1647 astronomy textbook, *Institvtio Astronomica*, in which he discussed the properties of spherical bodies and compared the systems of Ptolemy, Copernicus, and Tycho Brahe (Gassendi, 1647). Greenwood also wrote articles for *Philosophical Transactions of the Royal Society* about weather charts based upon observations of meteors made on land and at sea (1728) and on changes in the aurora borealis (1731), and he published a book on the college's scientific apparatus (Greenwood, 1734; Leonard, 1981; Simons, 1934). His most significant contribution to Harvard astronomy, though, may have been training **John Winthrop (1714–1779, AB 1732, AM 1735)**.

Winthrop succeeded the older man when Greenwood was dismissed for intemperance in 1738. He was already engaged in astronomical observation; by 1743 *Philosophical Transactions* accepted his 1740 letter discussing that year's transit of Mercury and lunar eclipse (Winthrop, 1743; "Sketch" 1891; Brasch, 1939; Turner, 1970). Winthrop also lectured more frequently, on a broader range of topics and at a deeper level, than Greenwood had. He updated the curriculum to incorporate the discoveries of Newton and other late seventeenth-century figures (Cohen, 1995). Students in their final year were required to attend a lecture course in natural philosophy that included astronomy. Early in his career, Winthrop simply added one lecture on the motions of planets and orreries on the last day of the 3-month session (Winthrop, 1746–1747). Over the next two decades, he wrote up thirteen "problems" on topics such as locating the sun on the ecliptic; calculating mean, apparent, and period times; and anticipating the phenomena of lunar and solar eclipses. Solving these problems required consulting tables and subtracting differences, so the mathematics was not challenging. Surviving student notes indicate that at least the descriptions of several of the problems were presented in the classroom (Gannett, 1766–1769; Wadsworth, 1766–1769).

Methodological Moment: Identifying Relevant Primary Sources—The fundamental question for historians is, "How do we know?" In other words, what primary source evidence exists to help us make sense of the past? As I have explained elsewhere (Ackerberg-Hastings, 2014), materials that document the history of education were often intended for temporary use and may not have been preserved. On the other hand, the twenty-first century has been characterized by widespread efforts to digitize print and manuscript sources and make them as available as possible. Thus, in this case, the Harvard Archives hold only a few sets of notes for Winthrop's teaching of astronomy, but the three notebooks described above are posted online in full. (A next step might be crowdsourced transcriptions, such as those underway in various museums within the Smithsonian Institution (2019).) Besides comparing the notebooks to each other, a researcher would want to examine them alongside any textbook required of students or read by instructors. I have not yet, though, found a college "calendar" for this era, the publication that served as what we think of as a course catalog and that often listed the books students were assigned, nor have I come across a secondary source that confirms whether an astronomy textbook was in use during Winthrop's tenure. This sort of brainstorming about types of sources that may document what we want to learn about the past and searching virtual and physical libraries and archives to determine whether such sources exist must start happening early in any historical research project. Students generally expect such legwork, although they may have to be taught not only which databases to try but also how to productively mine those catalogs. In my experience, they are more surprised to discover that they will come up with new ideas for potential sources and thus need to keep digging throughout the course of their research.

While several scholars have characterized Winthrop as a man who loved to explain mathematics and who turned physical demonstrations into performances, his chief passion was for furthering astronomical knowledge. As he continued to contribute papers to *Philosophical Transactions* on the transits of Venus in 1761 and 1769, various eclipses, and sightings of meteors (Winthrop, 1761, 1764, 1769), he constructed an international reputation as an observer and theorist. He seemed to prefer working closely with a small group of able students, most famously

Fig. 1 John Winthrop with telescope manufactured by James Short (ca 1740) and diagram of the transit of Venus, by John Singleton Copley, ca 1773. Harvard University Portrait Collection, Gift to Harvard College by the executors of the estate of John Winthrop and heirs of Mrs. Andrews, 1894

taking two Harvard students to Newfoundland to help observe the 1761 transit of Venus. His telescope and an astronomical diagram from his report on the expedition feature prominently in his official college portrait (Fig. 1). Winthrop also gave public lectures on comets, and he and his personal collections were instrumental in helping Harvard recover from a 1764 fire that devastated the facilities for teaching mathematics, natural philosophy, and astronomy (Brasch, 1916; Brasch, 1939; Cohen, 1950). Note that the sentences in this paragraph serve two functions in the preliminary research unfolding in this paper. First, they list Winthrop's known accomplishments in astronomy, activities that could be described in more depth in an appropriate venue, such as a thesis or extended journal article. In other words, they tell us what Winthrop did. Second, they begin to make claims about the significance of Winthrop's actions: unlike Greenwood, he was mainly an astronomer; other scholars recognized him as such; and he was willing to talk about astronomy to wide audiences but preferred teaching a few highly able students. These claims thus draw conclusions from the evidence about how and why Winthrop acted as he did. Novices in history often assume that historians just make lots of lists of facts, but interpretation is in fact essential to turning the past into history, to helping us better understand both our predecessors and ourselves. The next methodological moment provides a very brief introduction to some of the conceptual models that may influence scholars as they analyze primary-source evidence.

Methodological Moment: Approaches to Historical Interpretation—In particular, we can let the assertions that Winthrop preferred research over teaching and related better to the few most talented students than the entire student body lead us into some examples of how historians interpret evidence. Different scholars might find most illuminating different schools of interpretation, such as Marxism or post-structuralism or the debate between "internalist" (only the mathematical ideas matter) and "externalist" (mathematicians are influenced by their social and cultural contexts) analyses that long shaped mathematics historiography (Howell and Prevenier, 2001; Green and Troup, 1999; Mehrtens et al., 1981). In this instance, some less comprehensive methodological tools, which might also be called explanatory mechanisms or recurring themes, are more relevant. For example, when writing the history of women in science and mathematics, academics often distinguish between whether women were acting as "creators" or "consumers" in the field under study. In other words, were the women doing original research through which they articulated new theories, or were they reading what others had published on the subject? Similarly, Dave Roberts has argued that trends in American mathematics education have historically veered between efforts to make an entire population mathematically literate and emphases on producing mastery in the few most capable students, who would be expected to go on to become research mathematicians (Roberts, 2012). Despite giving lip service to the goal of ensuring all his students were competent enough to solve basic problems in astronomy, Winthrop appears to have preferred teaching the current state of knowledge to the few in order to develop creators of new mathematics. When I gave talks about this paper, I additionally mentioned a concept of "take-off" that was common in mid-twentieth-century scholarship on the Industrial Revolution (Rostow, 1960). In this context, I meant that I see Winthrop as the driving force behind Harvard's rise to prominence among American colleges in astronomy education. Instructors who assign historical research in mathematics courses likely will not have the time to provide their students with much background in approaches to historical interpretation, but if they themselves gain even a slight awareness of the lenses through which historians evaluate evidence, this knowledge can help teachers and professors ask questions about underlying assumptions they might notice in student writing—as well as in any secondary sources they may include among course readings or that their students are consulting.

3 Samuel Williams

One of the students who accompanied Winthrop to Newfoundland in 1761, **Samuel Williams (1743–1817, AB 1761)**, was appointed his successor as Hollis Professor of Mathematics and Natural Philosophy. (Although he was employed as a church minister at the time, Williams also observed the 1769 transits of Venus and Mercury in Newbury and Salem, Massachusetts (Williams, 1786).) Fittingly, one of his first acts was to undertake an expedition to Penobscot Bay, Maine, for the solar eclipse on 21 October 1780. Williams had to get permission to go behind British lines since this was during the Revolutionary War; the location he was allowed to use turned out to be just out of range for totality, although he was able to observe and record the phenomena now known as Baily's Beads (Schechner, 2012; Rothschild, 2009; Bates, 1965). Back in Cambridge, where astronomy had remained mixed in with natural philosophy during Winthrop's tenure, Williams grew the subject into a separate course offered near the end of a student's senior year, designed to serve as the culmination of his studies in mathematics and science (Schechner, 1992; Smith and Ginsburg, 1934). However, students apparently did not read a separate printed textbook as they learned astronomy. Rather, they recited for their tutors from Willem Jacob 'sGravesande *Mathematical Elements of Natural Philosophy* (1720), both published in Latin and translated into English by John Theophilus Desaguliers in 1720.

Methodological Moment: Making Sense of Primary Sources—We know from secondary sources, including those cited above, that Williams taught a separate course on astronomy; what can we find out about the content of those lectures? As a history professor and longtime volunteer with the National History Day program (2019; please encourage your students to judge at a local contest and consider mentoring and judging projects yourself), I believe strongly that students of any age are capable of analyzing primary sources. However, as we mused upon earlier, the content, context, and significance of documents from the past are not necessarily self-evident. Historians must engage with the authors across time and space, and it often turns out that sources can be read for meaning at multiple levels. Further, primary sources must be compared against each other—even an item that may seem to fully document an individual's life, such as a diary, needs to be read alongside other materials from the time period, such as newspaper articles and correspondence, to provide context on the influences that shaped the writer's thoughts and actions. (Some student-friendly primers providing techniques for analyzing primary sources include Connolly-Smith, 2007, Cronon, 2008–2009, American Social History Project et al., 1998–2021, and Walkowitz, 2011. Readers might also look to Raven's overview of the book history approach to historical interpretation (2018), since treatises and textbooks are so important in the history of mathematics.)

In this case, Gordon Joseph Schiff (1981, pp. 15–37) and others have pointed out that Williams prepared an outline of 15 lectures for a strongly mathematical course in astronomy that was approved by Harvard Corporation on 2 May 1785. Additional details are in the Corporation Records (1778–1795). Williams's papers, though, show that he began lecturing on astronomy nearly as soon as he arrived in 1780 (Williams, 1780–ca 1790). Dates jotted on the set of 11 lectures surviving from his tenure suggest he delivered these talks, which were mainly focused on sunspots and fixed stars, at least three times, in 1780–1781, 1784–1785, and 1786–1787. Notations such as dates in lecture notes and textbooks can be especially important in the history of mathematics education to help historians reconstruct what actually happened in classrooms, and not be limited to what instructors hoped would happen when they wrote course materials. A second set of notes, apparently prepared in the first decade of the 19th century after Williams had moved to Vermont, during a period when he hoped to publish a course of 13 lectures, mostly covered comets, although the first lecture consisted of multiple sections on the theory of heat. The content of the notes in this set of two volumes is almost entirely descriptive rather than mathematical, with a few tables of temperatures and the like. A researcher would need to carefully study and compare the contents of the two manuscripts to ascertain, for instance, the extent to which Williams's interests in astronomy changed over time. The condition of the notes also poses challenges to readers, as the second volume of the second set in particular is full of inserted reminders and other characteristics of rough drafts. As Cohen (1995), Rothschild (2009, pp.

240–243), and others have done with famous alumni such as John Adams and John Quincy Adams, historians might additionally seek out lists of Harvard students and check whether they left class notes among their own papers or wrote about their coursework in their daily journals.

During his professorship, Williams also began encouraging juniors and seniors to prepare and defend year-end "mathematical theses" that were lavishly drawn on oversized pieces of paper (Mathematical Theses 1782–1839; Badger, 1888; Jones and Boyd, 1971). While some of the theses offered plans for surveying or architecture and others consisted of solutions for various problems in algebra or fluxions, most of the students depicted eclipses for their projects. The practice persisted long after Williams departed in 1788 amid personal and professional conflicts, so that 404 examples were produced between 1782 and 1839. One of these was submitted in 1803 by the eventual fifth Hollis Professor, **John Farrar (1779–1853, AB 1803, AM 1806)**, who was a student under Williams's successor, **Samuel Webber (1760–1810, AB 1784, AM 1787)** (Palfrey, 1853; Ackerberg-Hastings, 2000, 2010). Farrar presented a diagram and tables for a solar eclipse that was to occur on 21 January 1814, and which was visible from south Sudan to Peru (Fig. 2). Calculations provided on the sheet included the duration of the entire eclipse, the duration of maximum eclipse, and solar and lunar positions. Farrar noted that the event would be invisible at Harvard, but unfortunately the coordinates listed in the

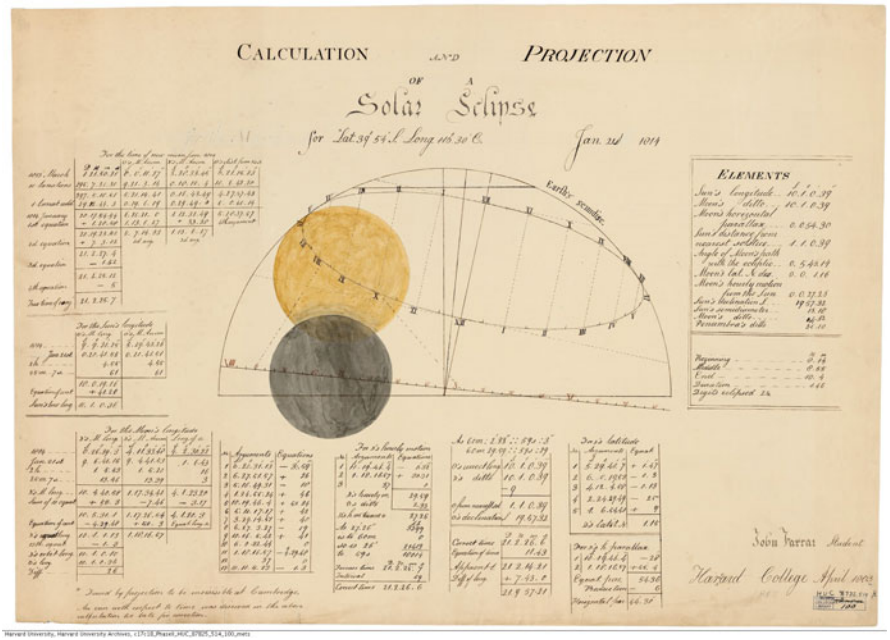


Fig. 2 John Farrar's 1803 thesis, "Calculation and Projection of a Solar Eclipse for Lat. 39°54′ S.; Long. 116°30′ E., Jan. 21st, 1814." Mathematical theses, 1782–1839, HUC 8782.514 (100). Harvard University Archives

thesis's title put the observer in the Indian Ocean, south of western Australia, where the eclipse was also invisible. He probably meant to give the longitude as 116°30′ *West*, which is in the south Pacific Ocean, about midway between west-central Chile and New Zealand, and very close to the southern limit at which the eclipse was visible (Espenak, 2010, 2019). A spot check of the calculations indicated they fell within range for the revised coordinates. Harvard students sometimes hired copyists to draw their thesis papers, so it is unknown whether the error was made by Farrar or by an unknown scribe. Whoever created the document also left the "v" out of "Harvard" (Farrar, 1803).

Methodological Moment: Visual Sources—Images, like texts, require careful and thoughtful historical analysis that makes sense of the content, considers how the original creators and audiences "read" visual sources, and locates materials in historical context (Ackerberg-Hastings, 2019). While I strongly advocate encouraging students to utilize visual sources as primary evidence in their research papers, these materials can also have pedagogical applications in subject-specific courses. For example, the visual nature of images can stimulate students interest via the seemingly passive function of illustration (Swetz, 2015). Additionally, history educators have provided teaching tools for building class lessons around images (Lee, 2018; National Archives, 2018). Either technique might be employed in courses to introduce the mathematical theses described above. The story of mathematical astronomy at Harvard might also be explored through images by searching for Hollis Professors in Harvard's Collection of Historical Scientific Instruments, where the men have been connected to the astronomical instruments in use during their tenures (Collection of Historical Scientific Instruments, 2017).

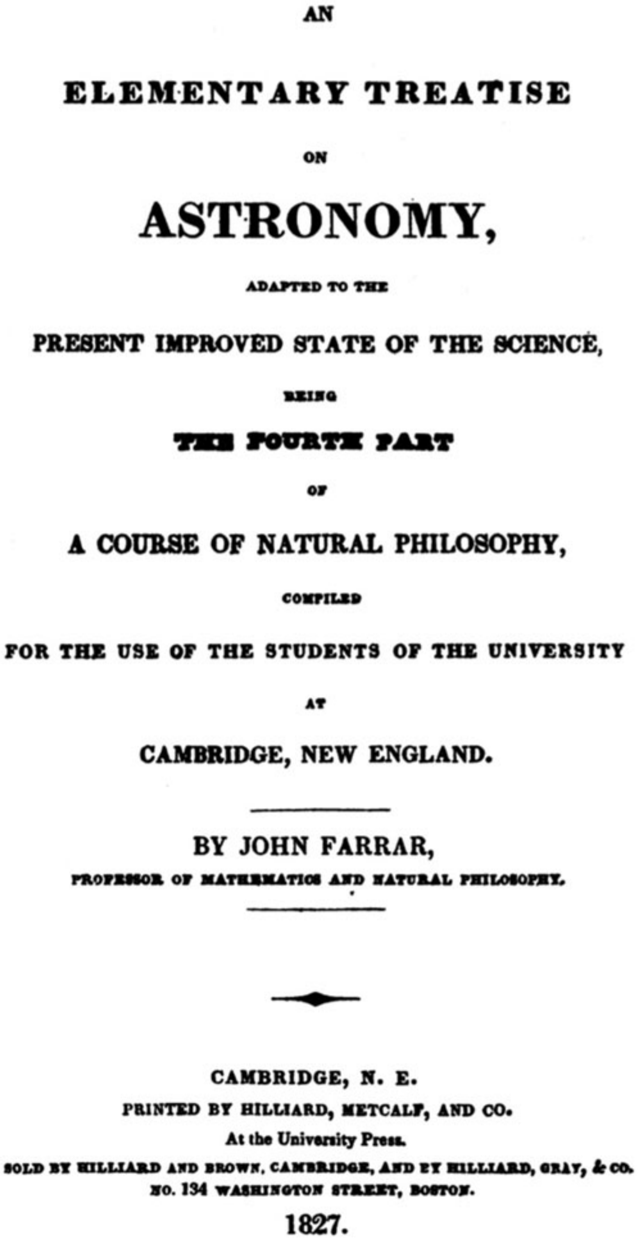
4 John Farrar

Even though students such as Farrar prepared senior theses on future or past eclipses, Webber himself observed and published on a solar eclipse that occurred 3 April 1791, and he made building an observatory a central goal of his later presidency of Harvard (1806–1810), astronomy largely dropped out of the college's course of study during Webber's tenure, which spanned 1789–1806 (Webber, n.d.; Webber, 1793; Cohen, 1950; Greene, 1984). In particular, he did not offer a separate lecture course on the subject. While he did include a chapter on spherical astronomy in his 1801 compendium textbook, *Mathematics*, he treated the topic as part of mixed mathematics (Webber, 1801). After Farrar became Hollis Professor in 1807, he resumed the astronomy lectures for seniors, and he resurrected a practice of both Winthrop and Williams of offering evening observing sessions on the roofs of various campus buildings (Farrar, 1810–1831).

The course and associated recitation sections with tutors also required students to study a textbook. However, for the first two decades of Farrar's tenure, it is not yet clear which manual for natural philosophy or astronomy was used when. Instead, different primary sources name different texts as being in use, some during overlapping time periods (Harvard College Papers, 1797–1825; Farrar, 1810–1831; Greene, 1984). Farrar's reports to college administration suggest he tried out several books before deciding to supply one himself. An additional explanation for the confused situation is that Harvard tutors in the 1810s and 1820s apparently were allowed to choose textbooks, and they did not all select the same ones. For the period between 1807 and 1827, then, textbooks employed for astronomy included: around 1812, Jean Delambre's astronomical tables (1792); by 1821 James Ferguson's *Astronomy* as edited by David Brewster (1811); by 1824 Webber's edition of William Enfield's *Institutes of Natural Philosophy, Theoretical and Experimental* (1802), which covered astronomy in the fifth of its seven books; and by 1825 John Gummere's *Elementary Treatise on Astronomy* (1822).

Farrar is most often known to historians of mathematics education for a series of translations of mathematics and natural philosophy textbooks, which were mainly of French works that were written around the turn of the nineteenth century (Hogan, 1977; Ackerberg-Hastings, 2000, 2010). Unsurprisingly, he ultimately resorted to a translation for astronomy, based upon the second edition of Jean-Baptiste Biot's *Traité élémentaire d'astronomie physique* (1810–1811). Since his name is on the title page, Farrar has traditionally received credit for preparing the translation (1827; Fig. 3). However, manuscript evidence shows that students and colleagues supplied the translations for several of the mathematics textbooks, including the unnamed student who translated Legendre's *Éléments de géométrie* (1819) and George Barrell Emerson, who worked on Étienne Bézout's calculus (1824) and Louis Bourdon's algebra (1831) (Day Papers (n.d.); Cajori, 1890, p. 130; Harvard College Papers, 1826–1838). No similar reference to *An Elementary Treatise on Astronomy* has yet been found, but it seems likely that Farrar received assistance with it as well.

Methodological Moment: Details Matter—Performing historical analysis on mathematics and science textbooks can be an arduous process, made even more challenging when translation between languages is involved. (For general techniques of historical analysis of textbooks, see, for example, Schubring, 1987, Johansson, 2005, and Fan, 2013. On writing histories of translations, see Pym, 1998 and Venuti, 2005.) Both versions can be compared line-for-line, to determine where the translator added or subtracted material. A researcher might also examine the extent to which the translation was faithful. Were the original's underlying philosophical or pedagogical assumptions changed with and for its new locale and audience? Similarly, textbook authors sometimes modified the text before new editions were published. Setting different editions next to each other can reveal changes in knowledge or in intellectual and cultural preferences. In one famous example, Charles Darwin (1859) addressed religious and scientific objections by modifying each of *On the Origin of Species*'s second through sixth editions. This is also a reason why historians who study the publication and reading of books distinguish between *printings*, when additional copies of books were printed to meet market demands, and *editions*, when authors had the opportunity to make revisions (Kidwell, Ackerberg-Hastings and Roberts, 2008, pp. 3–20). We may also use this methodological moment to reiterate that historians must collect a representative, comprehensive body of primary source evidence and compare sources against each other. One reason this process of analyzing evidence is so important is that it permits us to identify patterns or trends that help us make sense of the past's mass of factual information. For instance, one ongoing motif in the history of astronomy at Harvard that has not been explored in this paper is the interest in comets that spanned generations of Hollis Professors. To briefly describe a few examples, Sara Schechner noted that Winthrop secularized discussions of comets (Schechner, 1992), Williams developed a separate lecture course on comets, alluded to above, and Farrar published a paper on the theory of comets near the end of his active career (Farrar, 1836). Thus, it is also essential to sort evidence into


AN

ELEMENTARY TREATISE

ON

ASTRONOMY,

ADAPTED TO THE

PRESENT IMPROVED STATE OF THE SCIENCE,

BEING

THE FOURTH PART

OF

A COURSE OF NATURAL PHILOSOPHY,

COMPILED

FOR THE USE OF THE STUDENTS OF THE UNIVERSITY

AT

CAMBRIDGE, NEW ENGLAND.

BY JOHN FARRAR,

PROFESSOR OF MATHEMATICS AND NATURAL PHILOSOPHY.

CAMBRIDGE, N. E.

PRINTED BY HILLIARD, METCALF, AND CO.

At the University Press.

SOLD BY HILLIARD AND BROWN, CAMBRIDGE, AND BY HILLIARD, GRAY, & CO.

NO. 134 WASHINGTON STREET, BOSTON.

1827.

Fig. 3 Title page of Farrar (1827). Public domain

chronological order, in part so we can test the validity of our suspected causes and effects. In general, of course, the historian's aim is to tell stories that are as accurate and complete as possible.

Preliminary analysis of *Elementary Treatise* suggests that, like Biot's French original, its mathematical content was a small part of the overall presentation of the subject (Biot, 1810–1811; Farrar, 1827). The body of the textbook is almost entirely descriptive, hundreds of pages telling readers how the universe is put together. Readers must turn to the endnotes to find mathematical equations; the proofs there mainly employ trigonometry or show how algebraic formulas described in the text were derived. Presumably few students would have explored this semi-technical mathematics. In addition to endnotes, Biot's text also had footnotes—which occasionally contained equations and mathematical expressions—in the main text, but these were omitted by the translator. *Elementary Treatise* did nearly exactly follow the chapter organization of Biot, whose volumes were grouped into five parts: general phenomena, theory of the sun, theory of the moon, theory of the planets and comets, and fixed stars. Numerous paragraphs and sections were cut from each part for the translation, in order to reduce the number of pages from nearly 2000 across three volumes in French to about 425 pages in one volume in English.

Methodological Moment: Verify Generalizations and Appreciate Complexity—For some amount of time, all seniors took astronomy and read *Elementary Treatise*. However, we know from Harvard records and Farrar's correspondence that, by the end of his tenure in 1836, fourth-year students were allowed to choose between astronomy and the advanced study of modern languages. This information could call into question an existing generalization about the history of mathematics education. Most historians who research the subject as it unfolded in the American colonies and United States are aware that the full story is a complicated one. Nonetheless, it is common in that literature to speak in shorthand about the existence of a "common course" in American colleges before the late 19th century, meaning that every student in a given institution followed the same (usually classics-oriented) course of study. While that was generally true in theory, in practice we occasionally see special topics, independent studies, and other activities that look a lot like elective classes in 18th- and 19th-century curricula. Thus, researchers need to remember that summaries they read or are trying to write can be misleading, since in fact the past is messy and complicated (Wrobel, 2008; Branstiter, 2016). Facts that counter generalizations are also useful to researchers, instructors, and students as reminders that challenging our assumptions is healthy.

5 Conclusion

I transition into ending this paper by presenting one of those necessary but oversimplified generalizations: The subject of astronomy ebbed and flowed over the century between Isaac Greenwood and John Farrar, as the amount of teaching and the activity level of students increased dramatically under John Winthrop, leveled off somewhat with Samuel Williams, although he also contributed research achievements and curriculum, and then largely moved from the formal environment of the classroom to the more ad hoc settings in which individual students prepared their mathematical theses during the tenure of Samuel Webber. Yet, throughout all of these forms of instruction, the mathematical aspects of astronomy appear so far to have been generally downplayed, except for the cases of a few highly able students.

Some of these worked directly with professors, particularly Winthrop, but there is also evidence that full-fledged advanced courses were sometimes available, as during the professorships of Williams and perhaps Farrar. Indeed, it is clear that the uniform college course that American colleges claimed to offer before the middle of the nineteenth century was not always uniform, as students had access to electives and other choices from time to time. Additional work needs to be done, though, to assess the mathematical content of the teaching of astronomy at Harvard and to determine whether any differences that existed between Harvard professors could be attributed to the passion and ability that man had for mathematics over natural philosophy. Additionally, this paper has left other researchers to answer questions about the extent to which a tension existed between teaching at a level where the few mastered the subject and offering a course in which everyone theoretically achieved a general literacy about astronomy. Overall, the narrative has suggested the complex interplay of factors that a historian seeking to complete this story ought to consider.

Further, separate papers might be written on the Hollis Professors' travails with telescopes and other instruments, from the 1764 fire to the use of the replacements for decades after some of them became outdated, and on the interactions between Harvard astronomy and the social, economic, and cultural context of an eastern Massachusetts that relied heavily on the shipping industry and thus needed expertise in navigation, among other skills. Marlana Portolano (2000) has already published an article that details the lobbying conducted by Webber, Farrar, Nathaniel Bowditch, William Cranch Bond, and others before the administration finally authorized an observatory in 1839 and the Great Comet of 1843 helped jump-start fundraising for construction. (See also Greene, 1984.) After the facility was completed, the focus of astronomy at Harvard shifted toward research and, eventually, moved into the establishment of graduate education. Near the turn of the twentieth century came the band of talented women astronomers whose achievements mainly turned away from earlier generations' emphasis on comets to groundbreaking work on the classification of stars.

While the later period might currently be better known (Mack, 1990; Hoffleit, 1993; Sobel, 2016; "Women Computers" 2019), the pre-1839 era also merits the attention of scholars. It provides the background for these stories, of course, as well as the efforts of Peirce to change the direction of the mathematics department (Kent, 2005; Hogan, 2008). This essay has also argued that constructing a narrative of mathematics in astronomy at Harvard opens up what historians have called a "messy past." To make sense when source evidence is neither systematic nor comprehensive—as we saw with the haphazard preservation of lecture notes and inconsistent employment of textbooks—researchers must utilize the principles of sound historical practice to get as close to what really happened as possible. Indeed, although readers may have learned information about the history of astronomy teaching at Harvard before the college established an observatory—and about the digitization of resources relevant to writing such a history—the chief purpose of this article has been to use the planning of that research project as a case study for the process of historical research and writing. Its "methodological moments" have highlighted several fundamental activities: identifying, narrowing, and articulating a productive research question for a topic; understanding what historians do; locating relevant and reliable sources; analyzing and comparing primary sources, whether textual or visual; developing historical interpretations that are supported by the primary-source evidence; paying attention to differences between sources such as translations or editions; and considering challenges and frameworks specific to a given topic, akin to the struggles to identify what instructors and students actually did in classrooms and the perennial tension between educating the many in the basics and the few in the profession of mathematics that are confronted by historians of mathematics education. By attaching these concepts to concrete examples from the story of the interplay between mathematics and astronomy as both subjects were taught and learned under the supervision of Harvard's first five Hollis Professors, I have hoped to help instructors and students experience how historians get going on their research and gain the confidence to undertake their own historical research projects in their disciplines. Thus, while the topic of this paper may in itself contribute to historical scholarship, the essay has also added to historiography, the theory and practice of writing history, and shown instructors of mathematics and other subjects another way to do faithful and thoughtful history with their students.

Acknowledgements The author thanks the anonymous referees as well as audiences at the 2019 CSHPM Annual Meeting, MAA MathFest, and Special Session on Relations between History and Pedagogy of Mathematics of the AMS Fall Central Section Meeting for helpful questions and comments. Conversations with Janet Heine Barnett and Danny Otero also helped her conceptualize the article. Finally, she thanks the staff of Harvard University Archives and Harvard Art Museums for assistance with the illustrations.

References

Ackerberg-Hastings Amy (2000) Mathematics is a Gentleman's Art: Analysis and Synthesis in American College Geometry Teaching, 1790–1840. Ph.D. diss., Iowa State University.

Ackerberg-Hastings Amy (2010) John Farrar and Curricular Transitions in Mathematics Education. International Journal for the History of Mathematics Education 5(2): 15–30.

Ackerberg-Hastings Amy (2014) Mathematics Teaching Practices. In: Karp Alexander, Schubring (Eds) Handbook on History of Mathematics Education, 525–540. Cham: Springer.

Ackerberg-Hastings Amy (2019) Teaching Mathematics with Ephemera: John Playfair's Course Outline for Practical Mathematics. MAA Convergence 16. https://www.maa.org/press/periodicals/convergence/teaching-mathematics-with-ephemera-john-playfairs-course-outline-for-practical-mathematics. Accessed 29 October 2019.

American Historical Association (2019, June) Statement on Standards of Professional Conduct. https://www.historians.org/jobs-and-professional-development/statements-standards-and-guidelines-of-the-discipline/statement-on-standards-of-professional-conduct. Accessed 8 February 2021.

American Social History Project/Center for Media and Learning, Graduate Center, CUNY, and the Roy Rosenzweig Center for History and New Media, George Mason University (1998–2021) Making Sense of Evidence. History Matters: The U.S. Survey Course on the Web. http://historymatters.gmu.edu/browse/makesense/. Accessed 8 February 2021.

Badger Henry C (1888) Mathematical Theses of Junior and Senior Classes, 1782–1839. Cambridge: Library of Harvard University.

Bates Ralph S (1965) Scientific Societies in the United States (3rd ed). Cambridge: The M.I.T. Press.

Bentley Jerry H (1996) Cross-Cultural Interaction and Periodization in World History. The American Historical Review 101(3): 749–770.

Biot Jean-Baptiste (1810–1811) Traité Élémentaire d'Astronomie Physique (2nd ed, 3 vol). Paris: J. Klostermann.

Booth Wayne C, Colomb Gregory G, Williams Joseph M (2008) The Craft of Research (3rd ed). Chicago: The University of Chicago Press.

Branstiter Allan (2016) Madness and a Thousand Reconstructions: Learning to Embrace the Messiness of the Past. AHA Today, 21 June. https://www.historians.org/publications-and-directories/perspectives-on-history/summer-2016/madness-and-a-thousand-reconstructions-learning-to-embrace-the-messiness-of-the-past. Accessed 29 October 2019.

Brasch Frederick E (1916) John Winthrop (1714–1779), America's First Astronomer, and the Science of His Period. Publications of the Astronomical Society of the Pacific 28(165): 153–170. https://adsabs.harvard.edu/full/1916PASP...28..153B. Accessed 3 March 2022.

Brasch Frederick E (1939) The Newtonian Epoch in the American Colonies (1680–1783). Proceedings of the American Antiquarian Society, n.s. 49: 314–332.

Cajori Florian (1890) The Teaching and History of Mathematics in the United States. Washington, DC: Government Printing Office.

Cohen I Bernard (1950) Some Early Tools of American Science: An Account of the Early Scientific Instruments and Mineralogical and Biological Collections in Harvard University. New York: Russell & Russell.

Cohen I Bernard (1995) Science and the Founding Fathers: Science in the Political Thought of Thomas Jefferson, Benjamin Franklin, John Adams & James Madison. New York and London: W. W. Norton & Company.

Collection of Historical Scientific Instruments (2017) Waywiser. Harvard University. http://waywiser.rc.fas.harvard.edu/collections. Accessed 29 October 2019.

Connolly-Smith Peter (2007) Writing on History at Queens College. http://qcpages.qc.cuny.edu/writing/history/considerations/index.html. Accessed 8 February 2021.

Coolidge Julian L (1924) The Story of Mathematics at Harvard. Harvard Alumni Bulletin 26: 372–378.

Cronon William (Ed) (2008–2009) Learning to Do Historical Research: A Primer for Environmental Historians and Others. http://www.williamcronon.net/researching/index.htm. Accessed 29 October 2019.

Darwin Charles (1859) On the Origin of Species by Means of Natural Selection. London: John Murray.

Day Papers: Letters to Jeremiah Day (n.d.) Beinecke Rare Book & Manuscript Library, Yale University, New Haven, CT.

Delambre Jean Baptiste Joseph (1792) Tables astronomiques calculées sur les observations les plus nouvelles, pour servir à la triosème édition de l'astronomie. Paris.

Delaware Richard (2019) More Than Just a Grade: The HOM SIGMAA Student Contest Fosters Writing Excellence at UMKC. MAA Convergence 16. https://www.maa.org/node/1647874. Accessed 8 February 2021.

Edwards, Laura F (2011, January) Writing between the Past and the Present. Perspectives on History 49(1): 31–32. http://www.historians.org/publications-and-directories/perspectives-on-history/january-2011/writing-between-the-past-and-the-present. Accessed 8 February 2021.

Elliott Clark A, Rossiter Margaret W (Eds) (1992) Science at Harvard University: Historical Perspectives. Bethlehem: Lehigh University Press.

Enfield William (1802) Institutes of Natural Philosophy, Theoretical and Practical ... With Some Corrections; Change in the Order of the Branches; and the Addition of an Appendix to the Astronomical Part, selected from Mr. [Alexander] Ewing's Practical Astronomy, Webber Samuel (ed). Boston: Thomas & Andrews.

Espenak Fred (2010, 21 July) Five Millennium Catalog of Solar Eclipses. NASA Eclipse Web Site. https://eclipse.gsfc.nasa.gov/SEcat5/SE1801-1900.html. Accessed 29 October 2019.

Espenak Fred (2019, 22 March) Annual Solar Eclipse of 1814 Jan 21. EclipseWise.com: Predictions for Solar and Lunar Eclipses. http://eclipsewise.com/solar/SEprime/1801-1900/SE1814Jan21Aprime.html. Accessed 29 October 2019.

Fan Lianhuo (2013) Textbook Research as Scientific Research: Towards a Common Ground on Issues and Methods of Research on Mathematics Textbooks. ZDM 45(5): 765–777.

Farrar John (1803, April) Calculation and Projection of a Solar Eclipse for Lat. 39°54′ S.; Long. 116°30′ E., Jan. 21st, 1814. Mathematical Theses, 1782–1839. HUC 8782.514, HUC 8782.514 (100), Harvard University Archives. https://id.lib.harvard.edu/ead/c/hua17004c00118/catalog. Accessed 29 October 2019.

Farrar John (1810–1831) Correspondence and Faculty Reports by John Farrar, Hollis Professor of Mathematics and Natural Philosophy. UAI 15.963, Harvard University Archives. https://id.lib.harvard.edu/ead/hua24011/catalog. Accessed 29 October 2019.

Farrar John (1819) Elements of Geometry, by A. M. Legendre. Cambridge, MA: Cummings & Hilliard.

Farrar John (1824) First Principles of the Differential and Integral Calculus, by Etienne Bézout. Cambridge: Hilliard, Metcalf, and Co.

Farrar John (1827) An Elementary Treatise on Astronomy, Adapted to the Present Improved State of the Science, Being the Fourth Part of a Course of Natural Philosophy, Compiled for the Use of the Students of the University at Cambridge, New England by Jean-Baptiste Biot. Cambridge: Hilliard, Metcalf, and Co.

Farrar John (1831) Elements of Algebra, by Louis Bourdon. Boston: Hilliard, Gray, Little and Wilkins.

Farrar John (1836) Arago on Comets. North American Review 42: 196–216.

Ferguson James (1811) Ferguson's Astronomy: Explained Upon Sir Isaac Newton's Principles, Brewster David (Ed). Edinburgh: J. Ballantyne.

Gannett Caleb (ca 1766–1769) Astronomy notebook, Caleb Gannett Collection. HUM 314, Harvard University Archives. https://id.lib.harvard.edu/ead/c/hua11018c00001/catalog. Accessed 29 October 2019.

Gassendi Pierre (1647) Institvtio astronomica, iuxia hypothesis tam vetervm, qvam Copernici, et Tychonis. Paris.

Grattan-Guinness Ivor (2004) History or Heritage? An Important Distinction in Mathematics and for Mathematics Education. The American Mathematical Monthly 111(1): 1–12.

Green Anna, Troup Kathleen (Eds) (1999) The Houses of History: A Critical Reader in Twentieth-century History and Theory. New York: New York University Press.

Greene John C (1984) American Science in the Age of Jefferson. Ames: Iowa State University Press.

Greenwood Isaac (1728) A New Method for Composing a Natural History of Meteors Communicated in a Letter to Dr. Jurin. Philosophical Transactions of the Royal Society 35(401): 390–402.

Greenwood Isaac (1731) An Account of an *Aurora Borealis* Seen in New-England on the 22d of October, 1730, by Mr. Isaac Greenwood, Professor of Mathematicks at Cambridge in New-England. Philosophical Transactions of the Royal Society 37(418): 55–69.

Greenwood Isaac (1734) Explanatory Lectures on the Orrery, Armillary Sphere, Globes and Other Machines. Boston.

Gummere John (1822) An Elementary Treatise on Astronomy: in two parts: the first containing a clear and compendious view of the theory; the second, a number of practical problems, to which are added solar, lunar, and other astronomical tables. Philadelphia: Kimber & Sharpless.

Harvard College Papers (1797–1825) (vols 4–11) UAI.5.131, Harvard University Archives.

Harvard College Papers (1826–1838) (2nd ser, vols 1–8) UAI.5.131.10, Harvard University Archives.

Harvard University (1778, 5 May–1795, 31 August) Corporation Records: Minutes (vol 3). UAI.5.30, Harvard University Archives. https://hollisarchives.lib.harvard.edu/repositories/4/archival_objects/1180693. Accessed 8 February 2021.

Harvard University (1782–1839) Mathematical Theses, 1782–1839. HUC 8782.514, Harvard University Archives. https://id.lib.harvard.edu/ead/hua17004/catalog. Accessed 29 October 2019.

Higham Nicholas J (2019) Handbook of Writing for the Mathematical Sciences (3rd ed). Philadelphia: Society for Industrial and Applied Mathematics.

Hoffleit Dorrit (1993) Women in the History of Variable Star Astronomy. Cambridge, MA: American Association of Variable Star Observers. http://adsabs.harvard.edu/full/1993whvs.book.....H. Accessed 29 October 2019.

Hogan Edward R (1977) Robert Adrain: American Mathematician. Historia Mathematica 4: 157–172.

Hogan Edward R (2008) Of the Human Heart: A Biography of Benjamin Peirce. Bethlehem: Lehigh University Press.

Howell Martha, Prevenier Walter (2001) From Reliable Sources: An Introduction to Historical Methods. Ithaca: Cornell University Press.

Johansson, Monica. (2005) Math Textbooks—The Link between the Intended and Implemented Curriculum? Paper presented at Reform, Revolution and Paradigm Shifts in Mathematics Education, Johor Bahru, Malaysia, 25 November–1 December 2005.

Jones Bessie Zaban, Boyd Lyle Gifford (1971) The Harvard College Observatory: The First Four Directorships, 1839–1919. Cambridge: Harvard University Press.

Kent Deborah Anne (2005) Benjamin Peirce and the Promotion of Research-Level Mathematics in America: 1830–1880. Ph.D. diss., University of Virginia.

Kidwell Peggy Aldrich, Ackerberg-Hastings Amy, Roberts David Lindsay (2008) Tools of American Mathematics Teaching, 1800–2000. Baltimore: The Johns Hopkins University Press.

Krantz Steven G (2017) A Primer of Mathematical Writing (2nd ed). Providence: American Mathematical Society.

Lawson Russell (Ed) (2008) Research and Discovery: Landmarks and Pioneers in American Science (3 vol). Armonk, NY: M. E. Sharpe, Inc.

Lee John (2018) Using Historical Ephemera in the Classroom. TeachingHistory.org. https://www.teachinghistory.org/teaching-materials/teaching-guides/25028. Accessed 29 October 2019.

Leonard David C (1981) Harvard's First Science Professor: A Sketch of Isaac Greenwood's Life and Work. Harvard Library Bulletin 29: 135–168.

Mack, Pamela E (1990) Strategies and Compromises: Women in Astronomy at Harvard College Observatory, 1870–1920. Journal for the History of Astronomy 21(1): 65–76.

Mehrtens Herbert, Bos Henk, Schneider Ivo (Eds) (1981) Social History of Nineteenth Century Mathematics. Boston: Birkhäuser.

National Archives and Records Administration. (2018, 18 December) Document Analysis Worksheets. Educator Resources. https://www.archives.gov/education/lessons/worksheets. Accessed 29 October 2019.

National History Day (2019) https://www.nhd.org/. Accessed 8 February 2021.

Palfrey John Gorham (1853) Notice of Professor Farrar. Christian Examiner 55: 121–136.

Portolano Marlana (2000) John Quincy Adams's Rhetorical Crusade for Astronomy. Isis 91: 480–503.

Pym Anthony (1998) Method in Translation History. Manchester, England: St. Jerome.

Quincy Josiah (1840) The History of Harvard University (vol 2). Cambridge: John Owen.

Rael Patrick (2004) Reading, Writing, and Researching for History: A Guide for College Students. Brunswick, ME: Bowdoin College. https://courses.bowdoin.edu/writing-guides/. Accessed 8 February 2021.

Raven James (2018) What is the History of the Book? Cambridge, UK: Polity Press.

Roberts David Lindsay (2012) American Mathematicians as Educators, 1893–1923: Historical Roots of the "Math Wars". Boston: Docent Press.

Rostow Walt Whitman (1960) The Stages of Economic Growth: A Non-Communist Manifesto. Cambridge: Cambridge University Press.

Rothschild Robert Friend (2009) Two Bridges for Apollo: The Life of Samuel Williams, 1743–1817. New York: iUniverse.

Rufus W Carl (1921) Proposed Periods in the History of Astronomy in America. Popular Astronomy 29: 393–404.

Schechner Genuth Sara J (1992) From Heaven's Alarm to Public Appeal: Comets and the Rise of Astronomy at Harvard. In: Elliot Clark A, Rossiter Margaret W (Eds) Science at Harvard University: Historical Perspectives, 28–54. Bethlehem, PA: Lehigh University Press.

Schechner Sara J (2012) Astronomy behind Enemy Lines: Colonial American Field Expeditions, 1761–1780. Paper presented at Commission 41, Inter-Union Commission on the History of Astronomy Scientific Meetings, International Astronomical Union XXVIII General Assembly, Beijing, 22–23 August 2012. http://aramis.obspm.fr/~dvy/c41/icha-website/icha/index.html. Accessed 29 October 2019.

Schiff Gordon Joseph (1981) The Efforts and Accomplishments of Samuel Williams. Harvard Undergraduate Thesis, HU92.81.774, Harvard University Archives.

Schubring Gert (1987) On the Methodology of Analysing Historical Textbooks: Lacroix as Textbook Author. For the Learning of Mathematics 7(3): 41–50.

s'Gravesande, Willem Jacob (1720) Physices elementa mathematica, experimentis confirmata, sive introductio ad philosophiam Newtonianam. Leiden.

Simons Lao Genevra (1934) Isaac Greenwood, First Hollis Professor. Scripta Mathematica 2: 117–124.

Sketch of Professor John Winthrop (1891) Popular Science 39: 837–842.

Smith David Eugene, Ginsburg Jekuthiel (1934) A History of Mathematics in America Before 1900. The Carus Mathematical Monographs Number Five, The Mathematical Association of America.

Smithsonian Institution. (2019) Smithsonian Digital Volunteers: Transcription Center. https://transcription.si.edu/. Accessed 29 October 2019.

Sobel Dava (2016) The Glass Universe: How the Ladies of the Harvard Observatory Took the Measure of the Stars. New York: Viking.

Some of Harvard's Endowed Professorships (1926–1927, 1927–1928) Harvard Alumni Bulletin 29: 65–69, 145–150, 387–393, 1026–1028; 30: 138–40.

Stearns Peter N (2009) Long 19th Century? Long 20th? Retooling that Last Chunk of World History Periodization. The History Teacher 42(2): 223–228.

Storey William Kelleher (2009) Writing History: A Guide for Students (3rd ed). Oxford: Oxford University Press.

Swetz Frank J (2015) Pantas' Cabinet of Mathematical Wonders: Images and the History of Mathematics. MAA Convergence 12. https://www.maa.org/press/periodicals/convergence/pantas-cabinet-of-mathematical-wonders-images-and-the-history-of-mathematics-introduction. Accessed 29 October 2019.

Tarwater Dalton (Ed) (1977) The Bicentennial Tribute to American Mathematics, 1776–1976. The Mathematical Association of America.

Turabian Kate L (2013) A Manual for Writers of Research Papers, Theses, and Dissertations (8th ed). Revised by Wayne C. Booth, Gregory G. Colomb, Joseph M. Williams, and the University of Chicago Press Editorial Staff. Chicago and London: The University of Chicago Press.

Turner G. L'E (1970) Winthrop, John. In: Gillispie Charles Coulston (Ed) Dictionary of Scientific Biography (vol 14), 452–453. New York: Charles Scribner's Sons.

Varney B M (1908) Early Meteorology at Harvard College. Monthly Weather Review 36(5): 140–142.

Venuti Lawrence (2005) Translation, History, Narrative. Meta 50(3): 800–816.

Wadsworth Benjamin (1766–1769) An Abridgment of What I Extracted While an Undergraduate at Harvard College. HUC 8766.314, Harvard University Archives. http://id.lib.harvard.edu/aleph/007084739/catalog. Accessed 29 October 2019.

Walkowitz Judith R (2011) On Taking Notes. Perspectives on History 49(6): 38–41. http://www.historians.org/publications-and-directories/perspectives-on-history/january-2009/from-notes-to-narrative-the-art-of-crafting-a-dissertation-or-monograph/on-taking-notes. Accessed 8 February 2021.

Webber Samuel (n.d.) Papers of Samuel Webber. UAI 15.878, Harvard University Archives. https://id.lib.harvard.edu/ead/hua06005/catalog. Accessed 29 October 2019.

Webber Samuel (1793) Observations of an Annular Eclipse of the Sun, at Cambridge, April 3d, 1791. Memoirs of the American Academy of Arts and Sciences 2(1): 20–22.

Webber Samuel (1801) Mathematics, Compiled from the Best Authors (2 vols) Boston: Thomas & Andrews.

Williams Samuel (1780–ca 1790) Lectures. Papers of Samuel Williams. HUM 8, Harvard University Archives. https://id.lib.harvard.edu/ead/c/hua05010c00024/catalog. Accessed 29 October 2019.

Williams Samuel (1786) An Account of the Transit of Venus over the Sun, June 3d, 1769, as Observed at Newbury, in Massachusetts and An Account of the Transit of Mercury over the Sun, November 9th, 1769, as Observed at Salem, in Massachusetts. Transactions of the American Philosophical Society 2: 246–251.

Winthrop John (1743) A Letter from Mr. John Winthrop, Hollisian Professor of Mathematics and Astronomy at Cambridge in New-England, to C. Mortimer, M.D. Sec. R. S. Concerning the Transit of Mercury over the Sun, April 21, 1740, and of an Eclipse of the Moon, Dec. 21, 1740. Philosophical Transactions of the Royal Society 42(471): 572–578.

Winthrop John (1746–1747) The Summary of a Course of Experimental Philosophical Lectures. Papers of John and Hannah Winthrop, 1728–1789. HUM 9, Harvard University Archives. https://id.lib.harvard.edu/ead/hua07010/catalog. Accessed 29 October 2019.

Winthrop John (1761) An Account of a Meteor Seen in New England, and of a Whirlwind Felt in That Country. Philosophical Transactions of the Royal Society 52: 6–16.

Winthrop John (1764) An Account of Several Fiery Meteors Seen in North America and Observation of the Transit of Venus, June 6, 1761, at St. John's Newfoundland. Philosophical Transactions of the Royal Society 54: 185–188, 279–283.

Winthrop John (1769) Observations of the Transit of Venus over the Sun, June 3, 1769. Philosophical Transactions of the Royal Society 59: 351–358.

Women Computers. (2019) Astronomical Photographic Plate Collection. Harvard College Observatory. https://platestacks.cfa.harvard.edu/women-computers. Accessed 29 October 2019.

Wrobel David M (2008) Historiography as Pedagogy: Thoughts about the Messy Past and Why We Shouldn't Clean It Up. Teaching History: A Journal of Methods 33(1).

‘Lectures for Women’ and the Founding of Newnham College, Cambridge

James J. Tattersall and Shawnee L. McMurran

Abstract During the Victorian Era, a national focus on education reform led to marked improvements in the British educational system. In particular, academic opportunities for women in the middle classes grew tremendously. Among the variety of exams instituted to measure and certify academic proficiency was the Cambridge Examination for Women established in 1869 with the intent to certify a candidate’s qualifications for teaching. In order to help women prepare for the exam, and thereby raise the level of teaching, a group led by Henry Sidgwick, Millicent Garett Fawcett, Ann Clough, and John Couch Adams began sponsoring a series of lectures specifically for women to be offered by university faculty in Cambridge, England. Early lecturers in the program included the mathematicians Arthur Cayley, William Kingdon Clifford, and Norman Macleod Ferrers, the economist Alfred Marshall, and the logician John Venn. As a result of this successful venture, the ‘Lectures for Women’ formed a cornerstone for the establishment of Newnham College, Cambridge. In this chapter we highlight the first 10 years of the mathematical lectures.

1 Prologue

Prior to the twentieth century, the quality and scope of education available to a British subject varied widely depending on one’s gender, social class, and often religion. Inspired by powerful and influential activists, social and legislative

James J. Tattersall
Department of Mathematics, Providence College, Providence, RI, USA
e-mail: tat@providence.edu

Shawnee L. McMurran (✉)
Department of Mathematics, California State University San Bernardino, San Bernardino, CA, USA
e-mail: shawn.mcmurran@csusb.edu

M. Zack, D. Schlimm (eds.), *Research in History and Philosophy of Mathematics*, Annals of the Canadian Society for History and Philosophy of Mathematics/Société canadienne d’histoire et de philosophie des mathématiques,
https://doi.org/10.1007/978-3-030-95201-3_6

reforms during the Victorian Era led to marked improvements in the British educational system. Literacy rates in Great Britain rose dramatically during the nineteenth century, and the increase for women was particularly notable. In 1841, approximately 51% of brides and 67% of grooms could write their names. By 1900, this number had risen to approximately 97% for both groups (Altick, 1957, p. 171). It is not surprising that there was a dire need of education reform for the poor and working classes. Perhaps more surprising might be the dismal state of education for women in the middle classes during the first part of the century and the profound effect that education reform had for women from those classes.

Until the late nineteenth century, girls and young women in England were predominantly educated at home.[1] In the middle and upper classes, this education primarily consisted of training in skills that were deemed to make a young woman an attractive candidate for marriage, with little thought given to academic pursuits. Influential women and men from more enlightened families, who recognized that women were not only capable of undertaking the same academic pursuits as their brothers, but that educated English women would benefit society as a whole, began to make their voices heard. Where better to begin reform than with the teacher? In 1841, the London-based Governesses' Benevolent Association pioneered a movement to train women as teachers and grant certificates of proficiency to successful scholars. In 1848, the Christian socialist Frederick Denison Maurice established Queen's College, London, the first English institution established strictly for the education of women. Located on Harley Street, near London's Regent's Park, the college offered day and evening classes for girls over the age of 12. Many of the instructors came from King's College London, which itself had witnessed an unsuccessful attempt to open its lectures to women in 1832. In his inaugural speech, aware of opposition and criticism being published in the press, Maurice endeavoured to allay the doubts of detractors by carefully laying out the curriculum, giving particular attention to mathematics, a subject in which he expected to 'encounter the charge of giving a little learning which is dangerous' (Billings, 2000, p. 39).

Other educational opportunities, geared towards those women who wanted to enter the teaching profession, soon followed. In 1849, Elizabeth Jesser Reid founded Ladies' College (now Bedford College) in London's Bedford Square, bringing in lecturers mainly from University College London. The next year, Frances Mary Buss started the North London Collegiate School on Camden Town Road, and in 1855 Dorothea Beale founded the Cheltenham Ladies' College in Gloucestershire. It is of interest to note that both Buss and Beale were part of the first group of students to attend lectures at Queen's College, London. The schools themselves were not equivalent to colleges in the modern sense, rather, they were more akin to college preparatory schools where a foundational education was provided to women preparing to become teachers or governesses. However,

[1] A notable exception appears in Quaker society where equitable educational opportunities were afforded to both boys and girls in Quaker schools. For the most part, the rest of British society did not share this progressive view.

their establishment provided some of the first stepping stones towards access to higher education for women in England. The availability of such schools spread throughout Britain during the latter half of the nineteenth century and attracted an increasing number of women, yet support from the general population grew slowly. A majority of families remained rooted in tradition and did not encourage their daughters to pursue higher education. The belief that a woman’s mind should or could benefit from such an education struggled to gain a foothold.

It would not be until 1880 that a university degree would be awarded to a woman in Great Britain. A strong move towards this goal occurred in 1868 when the Schools Inquiry Commission issued the Taunton Report,[2] which included an informative and enlightening account of the conditions of middle-class education in Great Britain. In an effort to draw attention to the report’s findings concerning the inadequate education of girls, the social reformer Millicent Garrett Fawcett cited several reports debunking the myth that a woman’s mind was less capable than a man’s and outlined six principal causes for ‘the present defective education of girls’ (Fawcett, 1872). In Fawcett’s words, these causes were:

1. That girls’ schools are not sufficiently places of intellectual training.
2. That in the education of girls undue prominence is given to accomplishments, especially to music.
3. The want of properly trained governesses.[3]
4. The want of an external stimulus and test of the quality of teaching, such as supplied to boys’ schools by the universities and public schools.
5. The indifference of parents to the mental development of their daughters.
6. That owing to all the professions being closed to women, their higher education is unremunerative.

In her summary, Fawcett concluded that such causes were removable via social or legislative reform. From her analysis, it would not be too presumptuous to surmise that she would have supported the view that what girls actually needed was good teachers and access to good books, i.e., the same educational resources afforded to boys.

Meanwhile in Newcastle, Anne Clough[4] and Josephine Butler, both prominent in the Victorian women’s suffragist movement and each with a keen interest in women’s education, founded the North of England Council for Promoting the Higher Education of Women (Fig. 1).[5] The Council, originally representing associ-

[2] The Commission was chaired by Lord Taunton (Henry Labouchere). The report, in twenty volumes, was released from 1858 to 1868.

[3] With respect to mathematics, governesses were expected to have a thorough knowledge of functions, the rule of three, and be able to teach mental arithmetic (Clough, 1868).

[4] Her brother was the poet Arthur Hugh Clough.

[5] The Council was an amalgamation of smaller groups that existed in Liverpool, Manchester, Sheffield, Leeds, and Newcastle.

Fig. 1 Anne Clough (left) and Josephine Butler (right). Image of Butler courtesy of the Women's Library at the London School of Economics and the Josephine Butler Society

ations of school mistresses from several large northern English towns, established and coordinated a program of lectures and courses for women.[6]

The lecture program, inspired by the goal of improving the intellectual standards prevalent in girls' schools by providing a stronger academic background to teachers, led to demand for an examination of a higher calibre than that offered by the current Senior Local Examination.[7] Clough and Butler petitioned establishments of higher education in England to offer women an examination that would be of a sufficiently high standard to validate their qualifications for teaching. In 1869, the Cambridge University Local Examination Syndicate initiated and oversaw the Higher Local Examination for Women[8] aimed at women over the age of 18 who wanted to enter the teaching profession. Women who passed the exam would receive certification and the opportunity to further their education. The application fee of £2 likely ensured that only serious candidates would apply.[9] Each exam venue had a local

[6] In 1867, the Council sponsored a series of lectures on astronomy given in Manchester, Liverpool, Leeds, and Sheffield by James Stuart of Trinity College, Cambridge. It was estimated that altogether about 550 women attended his lectures (Hamilton, 1936, p. 73).

[7] An examination for students under the age of 18, usually taken when leaving secondary school, that covered a wide range of subjects. Girls were first admitted to the examinations in 1867. With respect to mathematics, the Senior Local Examination covered Euclid, algebra, trigonometry, and applied mathematics.

[8] Known at the time as the Cambridge Examination for Women.

[9] The fee of £2 had an average purchasing power of about 250 USD in 2019 (Officer and Williamson, 2021).

committee to oversee accommodations for exam takers, as well as a local secretary to receive the exam papers and supervise the exam.[10] In July 1869, in Cambridge, thirty-six women sat for the first such examination (Hamilton, 1936, p. 77).

The Higher Local Examination was composed of five subject groups:

A. Religious knowledge, arithmetic, English history (including relevant geography), English language and literature, and English composition.
B. Latin, Greek, French, German, and Italian.
C. Mathematics.
D. Political economy and logic.
E. Botany, geology and physical geography, zoology, and chemistry (theoretical and practical).

Satisfactory performance was acknowledged with a certificate that specified the subjects in which the candidate had passed and whether she passed with distinction. According to the guidelines for the exam (Markby, 1870), in order to receive a certification a candidate was required to pass Group A and at least one of the Groups B, C, D, or E. A demonstration of proficiency in all topics was required to pass Group A, with an exception allowed for religious knowledge.[11] After passing Group A, candidates could continue their studies and choose to be examined in any of the other groups in subsequent years.

A pass in Group B, D, or E required proficiency in one subtopic from the group, while a Certificate of Honour was earned by demonstrating proficiency in at least two subtopics. Group C, mathematics, was further divided into the following subtopics, of which the first two were required for a pass, and at least an additional two required for a Certificate of Honour:

- Euclid's *Elements* Books I, II, III, IV, VI, and Book XI to Proposition 21, inclusive.
- The elementary parts of algebra, i.e., the rules for the fundamental operations upon algebraical symbols, with their proofs; the solution of simple and quadratic equations; arithmetical and geometric progressions; the Binomial Theorem; and principles of logarithms.
- The elementary parts of plane trigonometry so far as to include the solution of triangles.
- The simpler properties of conic sections, treated either geometrically or analytically.
- The elementary parts of statics including the equilibrium of forces acting in one plane; properties of the centre of gravity, the law of friction, and the mechanical powers.

[10] Mrs. Annette Peile served as local secretary for the Cambridge committee. She and her husband John, later Master of Christ's College, Cambridge, were staunch supporters of higher education for women.

[11] Any candidate had the right to object to be examined in religious knowledge.

- The elementary parts of astronomy so far as they are necessary for the examination of the more simple phenomena.
- The elementary parts of dynamics including the laws of motion, gravity, and the theory of projectiles.

Recommended readings to prepare for Group C included Isaac Todhunter's *Trigonometry for Beginners*, Harvey Goodwin's *An Elementary Course of Mathematics*, George Airy's *Popular Astronomy: A Series of Lectures delivered at Ipswich*, and George Hale Puckle's *An Elementary Treatise on Conic Sections and Algebraic Geometry*.[12]

2 Lectures for Women

Proponents of the higher local examination recognized that exam candidates would require appropriate academic preparation. The success of the lecture program offered in the North suggested a similar program be offered in Cambridge itself, where talented lecturers were readily available and already teaching the relevant subjects. In December 1869, a meeting to discuss an exam preparation program for women was held in the drawing room of the home of Millicent and Henry Fawcett[13] in Cambridge (Fig. 2).[14] It was agreed that Cambridge, with its large number of trained and practiced teachers who might be persuaded to extend the sphere of their instruction, offered exceptional resources for an experiment of this kind. Suggestions included a plan for the first lectures to be given in private homes or some public room in town (Sidgwick, 1869). To those ends, a Committee of Management of the Lectures for Women in Cambridge was established (Hamilton, 1936, p. 89).[15,16]

The Committee subsequently proposed a series of lectures for women to be offered during term time[17] in subjects considered essential for the basis of a liberal education such as the classics, mathematics, natural science, and moral science (social science and philosophy). The group felt that these inaugural lectures could

[12] Two years later, the list was expanded to include: Hamblin Smith's *Elementary Algebra* and *Elementary Trigonometry*; Todhunter's *Algebra for the Use of Colleges and Schools* and *Mechanics for Beginners*; J. Norman Lockyer's *Lessons in Elementary Astronomy*; and Charles Taylor's *Geometrical Conics*.

[13] Henry Fawcett, who was blind, was Professor of Political Economy at Cambridge. He went on to be an innovative Postmaster General in the second Gladstone administration.

[14] Activists at the meeting included Henry Sidgwick, fellow of Trinity College Cambridge, who would play an instrumental role in the founding of Newnham College.

[15] Originally proposed as 'Lectures for Ladies'.

[16] Prominent women serving on the Committee included Eliza Adams, Eleanor Bonham Carter, Susan Cayley, Julia Kennedy, Annette Peile, and Susanna Venn.

[17] The Cambridge academic year is divided into three 8-week terms: Michaelmas (fall), Lent (winter), and Easter (spring).

Fig. 2 Henry and Millicent Fawcett (1868)

later be expanded to include more topics from the higher local examination beyond those needed for basic certification, as well as more advanced university subjects. Except for modern languages, the lectures would be delivered by resident members of the university who were engaged in academic teaching. The immediate object of the lectures would be to afford a means of higher education to women residing in Cambridge. Anticipating that there would likely be women from outside of Cambridge with interest in attending the lectures, the Committee agreed to take on the responsibility of acquiring and certifying appropriate lodging for those women who had no friends in Cambridge with whom they could reside. In order to move forward with this scheme, the group deemed it necessary to first obtain consent of sufficient members of the University to lecture on the relevant subjects. In addition, it would be necessary to obtain a list of names of ladies and gentlemen residing in Cambridge who approved the scheme and whose names would be a guarantee to parents of the genuineness of the undertaking, as well as its propriety.

The Committee's scheme commenced Lent Term 1870. It was proposed that lectures be given twice a week between the hours of 2 and 5 pm. The primary venue was Mr. Clay's house[18] in Trumpington Street. An ambitious list of sixteen proposed lecture courses was set forth, of which eight ran. Six student registrations were set as the minimum necessary for a course to be offered. In mathematics, the proposed courses included algebra with Arthur Cayley and practical arithmetic with J.F. Moulton, which were offered, and geometry with W.K. Clifford, which was

[18] Charles John Clay, the University Printer (1854–1894), held an M.A. from Trinity College, Cambridge (Venn and Venn, 1944).

not.[19] Compensation for lecturers was set at £1 for each registered student. Special arrangements might be made when the number of registrants fell below six.[20]

The second set of lectures was held during Easter Term 1870, again with an offering of eight courses: English literature by W.W. Skeat (34 registrants), English history by F.D. Maurice, Latin by J.E.B. Mayer, French by L. Boquel (18 attendees each); harmony by G.M. Garrett (11 attendees); practical arithmetic by J.F. Moulton (7 attendees); algebra by Arthur Cayley, and logic by John Venn (4 attendees each) (Lectures for Women, 1870). In the autumn of 1870, there were sufficient registrations to offer lectures in English history, English literature, Latin, French, and harmony, but prior to the start of the term, elementary Latin, German, and practical arithmetic had insufficient enrolment. Initial interest was, understandably, predominantly in those topics required to pass the mandatory Part A of the Higher Local Exam.

The standard fee per course was set at one guinea, and half-rate for those preparing to be teachers or already in the profession.[21] A certain level of academic preparation was expected in order to gain permission to register for a course of study. Passing the Senior Local Examination or a distinguished performance on the Junior Local Examination[22] provided admittance to the lectures on that basis alone. Other women had to be recommended and state on their application their age and qualifications.

An interesting artefact of the lecture series was the 'Stranger's Ticket' (Fig. 3). Any person registered for a course could purchase a Stranger's Ticket to be used for admittance to any single lecture. Such tickets were to be forwarded to the lecturers on the day before the lecture. In addition, to further encourage increased lecture attendance, women enrolled in a course were permitted to invite a guest to a single lecture, provided the lecturer was notified in advance.

Interest in the lectures for women continued to grow, as did resources and support for the women who desired to attend them. A lending library was established and the Committee procured annual scholarships for lecture attendees sponsored by John Stuart Mill, Helen Taylor, and Eliza Adams. In the early years, while some lecture topics were clearly quite popular, other topics struggled to garner sufficient interest.

The number of courses increased as various subjects experienced rising popularity. By Lent Term 1871, interest had risen sufficiently for the Committee to schedule lectures in English literature, English history, Latin, Greek, French, German, logic, political economy, geology, geometry, practical arithmetic, and

---

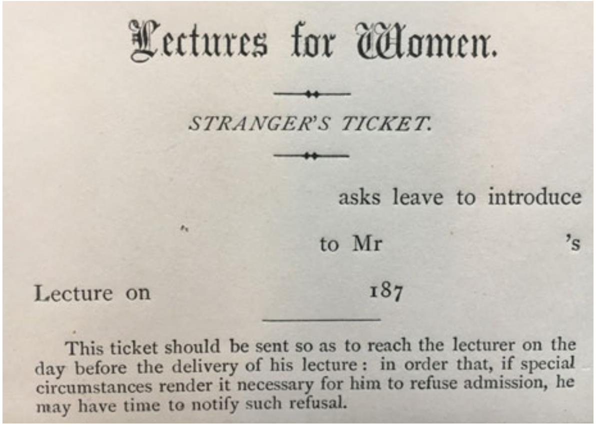
[19] The other six courses that ran were English Language and Literature, English History, Latin, French, Logic, and Harmony. Those that did not were botany, chemistry, geology and physical geography, political economy, Greek, German, and the theory of sound in its application to music.

[20] Expense records from 1873 to 1880 indicate that some lecturers received £6 even though the registration for their lectures was less than six (Fees for Lecturers, n.d.).

[21] A guinea was a pound sterling and a shilling, equivalent to about £1.05 (having the purchasing power of about 130 USD in 2019).

[22] An exam for students under the age of sixteen.

Fig. 3 A Stranger's Ticket


Lectures for Women.

STRANGER'S TICKET.

asks leave to introduce

to Mr 's

Lecture on 187

This ticket should be sent so as to reach the lecturer on the day before the delivery of his lecture: in order that, if special circumstances render it necessary for him to refuse admission, he may have time to notify such refusal.

harmony. The only lectures originally advertised that remained under-subscribed were P.T. Main's chemistry lectures, scheduled for the laboratories at St John's, and Arthur Cayley's algebra lectures, to be offered at his house on the River Cam.

The increasing popularity of the lectures brought with it interest from women located outside of Cambridge. A small market town at the time, Cambridge offered limited options for lodging. In the academic year 1871–1872, the glimmerings of what would become Newnham College emerged when, under the guidance of Henry Sidgwick, the Committee of Management secured lodgings at 57 Regent Street[23] for five women coming from outside of Cambridge to attend the lectures. Anne Clough agreed to take on supervision responsibilities for the attendees (Fig. 4).[24]

Demand for housing for those who wanted to attend the lectures continued to escalate. In the fall of 1872, the Committee of Management leased a larger accommodation at Merton Hall on Queen's Road near St John's College that could house twenty-four women.[25] Rules for the Merton Hall lodgers included the following (Merton House and the Cambridge Lectures for Women, n.d.):

- Students were expected to inform the principal[26] what places of worship they chose for regular attendance.
- Students were expected to consult with the principal on receiving invitations from friends, and also if they wished to make excursions in the neighbourhood.

[23] The building was renumbered 41 in the early nineteenth century. Built in 1830, it served as a residence until 1910 when the Bird Bolt Family and Commercial Temperance Hotel relocated there from the corner of Downing and St Andrew's Streets. From 1922 to 1985 it was the Glengarry Hotel and since then the privately owned Regent Hotel (Regent Hotel, n.d.).

[24] The lodgings fee was set at £20 per term, with a discount of £5 for those preparing for the educational profession.

[25] Use was made of Merton Hall from October 1872 to September 1874 when the lease expired.

[26] Anne Clough served as principal from 1871 until her death in 1892.

Fig. 4 Henry Sidgwick (left) and Anne Jemima Clough (right)

- Students were expected to be home by 6:30 pm during Michaelmas and Lent terms, and at 8:30 pm during Easter Term and all Sundays throughout the academic year.

That same fall, to alleviate the pressure to find adequate housing in Cambridge and expand access to the academic preparation provided by the lecture courses, the Committee initiated a series of correspondence classes under the direction of Annette Peile[27] and supported by university faculty. The correspondence courses provided an invaluable service, providing access to higher learning for many women whose various circumstances prevented them from coming to Cambridge, even had housing been available. The lectures by correspondence continued to be offered until 1894.

In October 1873, in order to unify efforts and gain more support for their venture, the Committee of Management reestablished itself as the Association for the Promotion of Higher Education for Women in Cambridge (Committee of Management Note, 1873). They felt the move would help place their system of lectures on a more formal, broadly based, and democratic footing. The move aided further development of the system of lectures for women on the more advanced subjects on the Higher Local Examination and other branches of academic study. The Association solicited donations, stressing several needs including supplements to the fees paid for more advanced courses, the establishment of scholarships,

[27] Annette Peile supervised the correspondence courses for the next 12 years.

and assistance for women with insufficient means, especially those preparing to be teachers, that would enable them come to Cambridge to attend the lectures. The Clothworkers' Company offered five annual scholarships (Association for the Promotion of Higher Education of Women Annual Report for 1874–1875, 1875). The Goldsmiths' Company and Drapers' Company soon followed with their own scholarships (Association for the Promotion of Higher Education Annual Report for 1875–76, 1876). Fifteen courses ran that fall, including courses in algebra, arithmetic, and geometry, taught, respectively, by Richard T. Wright (with twelve students), W.H.H. Hudson (ten students), and Norman Macleod Ferrers (six students).

Although the primary object of the lecture courses was to provide a thorough preparation for the Higher Local Examination, more advanced instruction began to be offered for those students who, having passed the examination, desired to proceed further in a line of study of their choice. It was the intention of the promoters of the scheme to extend to all women, who were willing and qualified to profit by it, as complete an education as the university could afford. By the academic year 1873–1874, the number of women taking advantage of the lectures averaged almost ninety per term (Higher Education for Women, n.d.). Women could attend the public lectures of twenty-two professors of the university, a course in elementary mechanics was added, and lectures on botany were opened to women. The number of tickets sold each term were: Michaelmas (193), Lent (170), and Easter (212). On account of the larger enrolments, the lectures were moved from Mr. Clay's house to the Christian Men's Association at 1 Alexandra Street.[28]

One measure of the effectiveness of the lecture series may be the success of the participants who chose to be examined. For example, of the eighteen lecture course participants who took the June 1874 Higher Local Examination, only one failed to pass.[29] During the academic year 1874–1875, three women felt prepared enough to test their mettle against that of the men by sitting, albeit informally, for the formidable tripos exams.[30] Had they been men, their performances would have entitled them to honours degrees: Edith Creak attained a second class standard in the Classical Tripos and third class in the Mathematical Tripos exam, Mary Paley[31] and Amy Bulley in Moral Sciences, the former with an Honour Standard and the latter with a second class standard. In subsequent years, participants in the lecture series would not only continue to show success on the Higher Local Exams, many would move on to more advanced studies, achieve honours, and demonstrate distinction in a variety of academic areas, including various tripos exams.

[28] Now the location of the Lion Yard shopping arcade.

[29] In 1874, men were first permitted to take the examination and the name changed from the Examination for Women to the Higher Local Examination. However, no men took part that year.

[30] Success on a Cambridge tripos examination was necessary for a university man to obtain a Cambridge honours degree. Tripos subjects included mathematics, classics, moral sciences, history, and law.

[31] Paley was the daughter of William Paley, author of *Natural Theology*. In 1877, she married the economist Alfred Marshall.

Another measure of effectiveness might be seen in the inclusion of lecture series alumnae in the lecturer ranks. The process to include women lecturers began in Michaelmas Term 1875, when Mary Paley became the first female resident lecturer, teaching political economy to nine students of whom six attained distinction on the Higher Local Examination. Later, Amy H. Ogle, resident lecturer in natural science, lectured on botany and zoology. Of the six students in each class, three received first-class certificates in Group E of the Higher Local Examination. In Michaelmas Term 1880, Charlotte Scott of Girton College,[32] who was bracketed with the Eighth Wrangler on the 1880 Mathematical Tripos, and Sarah Jane Dugdale Harland[33] gave mathematical lectures in algebra, trigonometry, and conics. In 1883, Agnes Bell Collier, a Clothworkers' Scholar who had earned a second class on the 1883 Mathematical Tripos, was hired as a college lecturer and director of studies in mathematics. Collier offered lectures in algebra, geometry, and mechanics until the 1920s.[34] By 1894, over 60% of the lectures for women were taught by women, with that number rising to more than 70% by 1900.

Interest in mathematical subjects continued to grow, and with this growth came wider and more advanced offerings. Elementary mechanics first ran in Lent Term 1873, trigonometry in 1874, advanced algebra, conic sections, and mechanics in 1875, and statics and dynamics in 1877. In Easter Term 1878, the class with the largest enrolment (33) was Hudson's lectures on arithmetic (Fees for Lecturers, n.d.).[35]

As a conclusion to this section, we provide a summary of the lecturers for the mathematical subjects offered from 1870 to 1883 in the tables 'Basic mathematics course lecturers' and 'Advanced mathematics course lecturers'.

3 Newnham College

By 1874, when the lease on Merton Hall expired, the escalating demand for lodgings became critical. From October 1874 to June 1875 supplementary housing was provided at 7 Trumpington Street and at two adjacent residences on Bateman

[32] At the time, Girton and Newnham Colleges were the only two establishments in Cambridge offering higher education to women. The foundation under which Girton was established held a different philosophy from that of Newnham. For example, examinations were compulsory for women coming to Girton, but optional for women at Newnham. For more information see (McMurran and Tattersall, 2017).

[33] Harland went on to do research in the Cavendish Laboratory and married Sir William Napier Shaw, a meteorologist who studied air pollution. She continued to lecture on arithmetic until 1889.

[34] In the 1890s small classes in mathematics were formed when necessary.

[35] William Henry Hoar Hudson, Third Wrangler on the 1861 Cambridge Mathematical Tripos, was a fellow of St John's College. Hudson offered arithmetic lectures by correspondence until 1882, when he was appointed professor of mathematics at King's College London and Queen's College on Harley Street. He also served on the Council of the London Mathematical Society where he was engaged in the improvement of mathematical education in schools and colleges.

Basic mathematics course lecturers[a]				
Year	Arithmetic	Algebra	Geometry	Trigonometry
1870	Moulton	Cayley	Clifford	
1870–1871	Moulton	Cayley	Clifford	
1871–1872	Hudson	Cayley	Ferrers	
1872–1873	Hudson	Cayley, Wright	Ferrers	
1873–1874	Hudson, Smith	Cayley, Wright	Ferrers	
1874–1875	Hudson, Wright	Cayley, Smith	Ferrers	Smith
1875–1876	Hudson, Torry	Smith	Ferrers	Wright
1876–1877	Hudson	Torry	Ferrers	Wright
1877–1878	Hudson	Torry	Ferrers	Cox
1878–1879	Hudson		Ferrers	Torry
1879–1880	Hudson	Torry	Ferrers	Torry
1880–1881	Hudson	Torry, Harland	Ferrers	Harland
1881–1882	Hudson, Harland	Harland	Ferrers, Harland	Harland
1882–1883	Harland	Harland	Ferrers, Harland	Harland

[a] John Cox (Trinity, 9W, 1874), John Fletcher Moulton (St. John's, SW, 1868), Charles Smith (Sidney, 3W, 1868), Alfred Freer Torry (St. John's, 4W, 1862), and Richard Thomas Wright (Christ's, 5W, 1869). The notation indicates the Order of Merit on the Mathematical Tripos, an examination for an honours degree at Cambridge. The top ranked student was designated the Senior Wrangler (SW). He was followed by the Second Wrangler (2W), and so on. The last man in the ranking was called the Wooden Spoon.

Advanced mathematics course lecturers[a]						
Year	Conics	Mechanics	Dynamics	Statics	Adv. Algebra	Calculus
1874	Smith	Stuart				
1874–1875	Smith	Garnett	Garnett		Cayley	
1875–1876	Wright	Garnett		Cayley		
1876–1877	Smith		Garnett	Garnett	Cayley	Freeman
1877–1878		Garnett		Garnett	Cayley	
1878–1879	Smith		Garnett	Garnett		Hudson
1879–1880	Smith, Scott		Garnett	Garnett		Hudson
1880–1881	Scott		Garnett	Garnett	Harland	Hudson
1881–1882	Scott		Garnett	Garnett	Torry	Scott
1882–1883	Scott		Johnson			

[a] Alexander Freeman (St. John's, 5W, 1861), William Garnett (St. John's, SW, 1873), William Ernest Johnson (King's, 11W, 1882), and James Stuart (Trinity, 3W, 1866)

Street. A search throughout the Cambridge area convinced the Association of the impossibility of finding a suitable extant house. It was therefore decided to build a large house for students lodging in Cambridge or its vicinity. With the object of defraying the cost of this undertaking, the Newnham Hall Company[36] was formed

[36] The name likely refers to the intent to build the house within, or in proximity to, the council ward of Newnham, Cambridgeshire. The ward is home to several Cambridge University colleges.

Fig. 5 Newnham Hall circa 1875

in the spring of 1874.[37] The main objectives of the Company were to establish and maintain a house or houses in or near Cambridge in which female students may reside and study while attending the lectures, as well as provide for the delivery of lectures for women. A call for donations brought in £4650 and two and a half acres of land in Newnham Parish, just outside the town of Cambridge, was purchased from St John's College.[38] Newnham Hall was completed in the summer of 1875 and opened for students in the fall. The building consisted of a dining hall, a kitchen, a principal's room, two lecture rooms, two student rooms, offices on the ground floor, one lecture room and twelve student rooms on the first floor, thirteen student rooms on the second floor, and on the top floor: a music room, four student rooms, four servant rooms, and two rooms for storage. Each student room served as a bedroom and a study. The modern reader may see a somewhat desolate setting when looking at early images of Newnham Hall, but to the pioneering women who resided there, the hall provided a home in which a true learning community was given the opportunity to thrive (Fig. 5).[39]

[37] In 1876 and 1877, the Company paid a dividend of 4% for any shares bought before June 19, 1876.

[38] The estimated cost of the planned building, capable of housing 25 students and the principal, was £6000. The expense included necessary furnishings, as well as the cost of laying out a roughly one-acre garden and a tennis court (Higher Education for Women, n.d.).

[39] We refer the reader to Lamberton (2014) for a descriptive perspective of academic collaboration among early Cambridge women.

This opportunity, however, did not come without a certain set of ground rules. Women who resided at Newnham Hall were subject to the following terms of admission (Newnham Hall Rules and Terms of Admission, 1875):

- No students could come into residence without the approval of the principal.
- Students wishing to leave were expected to give 3 months' notice.
- The principal may require any student to withdraw if in her opinion the student is not profiting by the course of study.

The cost of room and board was £20 per 8-week term, with a discount of £5 for those intending to become teachers.[40] If a student wished to have a fire in her room, the cost was an additional four guineas per year, or two guineas for those intending to become teachers. Payment for instruction varied according to the line of study, but rarely exceeded four guineas for ordinary students, with a reduction of one-half for those women intending to become teachers.

In Michaelmas Term 1875, Newnham Hall had twenty-seven residents. Five additional women, referred to as out-students, resided at 14 Trumpington Street with house rules similar to those employed earlier for Merton Hall. Other out-students lived with a parent or guardian in Cambridge, or, if over 30 years of age, lived in a residence approved by the principal. The remaining lecture attendees consisted of students who came in for the day from neighbouring villages.

The popularity of the lecture series continued to grow. According to Clough, 'Students were encouraged to begin with familiar subjects, such as arithmetic, history, English language and literature, and then take up whatever special department may interest them the most (Clough, 1873).' French or logic were popular choices for the latter. Once they had passed the basic requirements of the Higher Local required to earn a University Certificate, students were encouraged to pursue a course of study in a more advanced subject. In Michaelmas Term of 1877, forty women enrolled in George Prothero's lectures on English history, thirty-eight in John Robert Seeley's modern history lectures, thirty in Walter William Skeat's English Literature lectures, twenty-seven in W.H.H. Hudson's arithmetic lectures, fifteen in J.N. Keynes' logic lectures, and fifteen in L. Boquel's French lectures. The continued growth in popularity of the lecture series was reflected in the ever-increasing demand for lodgings. Of the eighty-two resident students in fall 1877, thirty-three were housed at Newnham Hall. In order to accommodate the demand for additional housing, Norwich House on Panton Street,[41] which could hold sixteen women residents, was leased for 3 years from Cavendish College (Association for the Promotion of Higher Education Annual Report for 1876–77, 1877). Women who chose to be examined continued to distinguish themselves. Of the eighteen

[40] To assist with lecture fees and book purchases, students of scanty means could apply for an interest-free loan to Anna Bateson who was in charge of the loan fund. She was mother of the journalist Margaret Heitland and historian Mary Bateson. Her husband, William Henry Bateson, was master of St John's College, Cambridge, and founder of the Cambridge University School of Genetics.

[41] Near the present site of the University Botanical Gardens.

Newnham women who entered for the Higher Local in 1877, fifteen were examined in advanced subjects, eight obtained first class honours, and one obtained a double first class. One student (informally) passed the Mathematical Tripos (Newnham College Report, 1877).

As the number of applicants for residence continued to rise, the Association proposed building a second hall on the grounds of Newnham Hall[42] and combining the two groups to better carry out the work of both. As a consequence, in 1880, Newnham College Association for Advancing Education and Learning Among Women in Cambridge was formed from the amalgamation of the Newnham Hall Company and the Association for the Promotion of the Education of Women in Cambridge, uniting in itself the whole work of the two bodies under the new name 'Newnham College' whose objectives were (Newnham College Objectives, n.d.):

- To establish and maintain at or near Cambridge a house or residence, or houses or residences, in which women students may reside and study.
- To provide the instruction of women students, and for the delivery of lectures to such students, or to other women at or near Cambridge.
- For the above purposes to receive and apply donations from persons desiring to promote the objects of the college.
- The doing of such other things as are incidental or conducive to advancing education and learning among women in Cambridge and elsewhere.

The direction of the affairs of the college was put in the hands of a Council, elected annually from its first promoters, benefactors, and lecturers. At its first meeting, the mathematician Arthur Cayley was elected president of the Council, a role he filled until his death in 1887.[43]

In the 10 years since the seed had been planted with five resident students coming from outside of Cambridge for the purpose of attending the Lectures for Women, Newnham College had firmly taken root. The influence of Newnham graduates in educational fields was growing and would continue to do so throughout the remainder of the century. By 1878, of the 140 resident students who had completed their studies, fifty-two had become teachers, seven were head mistresses, twenty-four were assistant mistresses in secondary schools, and seven were lecturers (Association for the Promotion of Higher Education Annual Report for 1877–78, 1878). By 1897, those numbers would more than triple. Out of 490 resident students having completed studies, 220 were teachers and at least 62 women were employed in other professions (Occupations of Former Students, 1897).

Between October 1871 and June 1880, approximately 215 students had come to Cambridge to attend the lectures and 184 passed examinations (Newnham College Report, 1880). In 1879, the average number of women attracted to Cambridge by the lectures was 82 (Association for the Promotion of Higher Education Annual

[42] At an estimated cost of £11,000.

[43] John Peile, husband of Annette Peile, served as president of the Council from 1890 to 1909.

Report for 1878–79, 1879). According to the Treasurer's Report of November 1880 (Association for the Promotion of Higher Education for Women: Treasurer's Final Report, 1880), Newnham College had reached sixty-eight women in residence: thirty-two in Newnham Hall, sixteen in Norwich House, thirteen in other college lodgings, and seven boarding in local residences.

Improvements in facilities developed in tandem with the growing student body. By the spring of 1879, Newnham Hall[44] had added a laboratory,[45] a gym, and a small hospital at the back that could be isolated in case of infectious diseases. North Hall[46] was opened in October 1880 with a library, dining hall, and two lecture rooms. Each student had a room that served as a bedroom and study.[47]

After 1880 the focus at Newnham College shifted from passing the Higher Local Examination and obtaining a certificate to preparing for a tripos examination.[48] The latter generally necessitated residence for 9–12 terms. According to university regulations at the time, admission to a tripos examination required either that a candidate had passed the university's Previous Examination[49] or had passed one of the following exams: Senior Local Examination, Higher Local Examination, or a test given by the Oxford and Cambridge Schools Examining Board, which included problems of elementary and additional maths, algebra, plane trigonometry, statics, and dynamics. Students were allowed to take the Previous Exam if they preferred to do so, but the college did not, except in special cases, provide instruction in preparation for it. Women were encouraged to get such exams out of the way before coming to Newnham.

Beginning Easter Term 1881, mainly on account of Charlotte Scott's first class performance on the formidable Mathematical Tripos the previous year, the University Senate approved the following recommendations:

- Female students at Girton and Newnham Colleges who have fulfilled the residence requirement of the university be admitted to the previous and tripos exams.
- Names of female students who are successful on the tripos will appear on the list of Order of Merit in the place they would have appeared.

[44] Then South Hall, now referred to as Old Hall.

[45] Up to then, women sitting for the natural science lectures had access three times a week to the natural science museum and laboratories at St John's College.

[46] Now Sidgwick Hall.

[47] Clough Hall opened in 1888. Pfeifer Hall was added to Old Hall in 1893. A block of buildings called Kennedy Buildings opened in the spring of 1900.

[48] For women, the exam was still informal and given by courtesy of the examiners. In May of 1877, Leonard Courtney had brought forward a motion in the House of Commons that would have enabled the Universities of Oxford and Cambridge to examine female students concurrently with males. The motion failed to pass.

[49] A preliminary examination to demonstrate that a student had a basic command of mathematics and the classics.

- Each successful female student will receive a certificate.[50]

In subsequent years, availability of lectures continued to grow. During the academic year 1881–1882, thirty public lectures, given by university professors, were open to women. By 1885–1886 the list of Cambridge lectures open to women had grown to ninety. As noted in the previous section, the Lectures for Women courses consistently included a good number of mathematics lectures where the ranks of lecturers included both university professors and women associated with Newnham. After Easter Term 1883, specific mathematics classes under the auspices of Newnham College were offered by Sarah Jane Harland and Agnes Collier. By 1905, lectures and individual instruction in mathematics were being given to Newnham College students by Arthur Beery (King's), Agnes Collier (Newnham), John Forbes Cameron (Caius), William Hewson Gunston (St John's), G.H. Hardy (Trinity), Andrew Munro (Queen's), Arthur Stanley Ramsey (Magdalene), Lillian Mary Reynolds (Newnham), and Mary Ellen Rickett[51] (Newnham).

4 Conclusion

Among the many plans devised for improving the education of women during the nineteenth century, an important place may be claimed for the scheme of lectures established in Cambridge in 1870. These lectures gave women an opportunity to strengthen and deepen their knowledge of subjects taught in schools. For many, this academic background led to careers in education and fostered their capacity to become effective and knowledgeable educators, who might in turn better prepare their students to pursue higher education. Many women also took advantage of the opportunity to pursue advanced university studies in their fields of interest. We have drawn the reader's attention to the remarkable fact that these lectures were a free-will offering of higher culture made to women by members of the university. Through the 'Lectures for Women' program, support and instruction from highly respected academics prepared women to pass, and often excel, on the Examination for Women over 18, later the Cambridge Higher Examination. By 1874, women who participated in the program felt prepared to tackle, with success, the Cambridge tripos (honours) exams. The success of the lecture program and its students was instrumental in opening university lectures to women, founding Newnham College, providing an opportunity for women to pursue a course of university studies on a par with that of the men and earn prestigious certifications of their qualifications, and eventually, albeit not until 1948, be granted a bona fide Cambridge university degree. In conclusion, we might also reflect on the pressure that early Newnham

[50] It would not be until 1948 that the right to earn a Cambridge undergraduate degree was granted to women. Women could receive titular graduate degrees at Cambridge beginning in the 1920s.

[51] Rickett was Newnham's first Wrangler when she earned a first class on the 1885 Mathematical Tripos.

students may have felt to ensure that the experiment worked, the dedication of many to extend their learning to other women, and, perhaps, the satisfaction they may have gained from building close relationships with other like-minded women.

Acknowledgements The authors thank Anne Thomson, archivist at Newnham College, for her generous advice and assistance with archival materials. Except for the picture of Josephine Butler, all pictures appear with the permission of the Principal and Fellows of Newnham College, Cambridge. We are also grateful to the referees for their thoughtful comments and suggestions.

References

Association for the Promotion of Higher Education of Women Annual Report for 1874–1875, October 1875. Newnham College Archives EC/2/2/6.

Association for the Promotion of Higher Education Annual Report for 1875–76, November, 1876. Newnham College Archives EC/2/2/6.

Association for the Promotion of Higher Education Annual Report for 1876–77, November, 1877. Newnham College Archives EC/2/2/6.

Association for the Promotion of Higher Education Annual Report for 1877–78, November, 1878. Newnham College Archives EC/2/2/6.

Association for the Promotion of Higher Education Annual Report for 1878–79, November, 1879. Newnham College Archives EC/2/2/6.

Association for the Promotion of Higher Education for Women: Treasurer's Final Report, November 1880. Newnham College Archives EC2/2/6.

Committee of Management Note, November 1873. Newnham College Archives EC2/2/2.

Fees for Lecturers. Newnham College Archives EC/2/2/6.

Higher Education for Women. Records of Newnham College, 1871–1881. Newnham College Archives AD/5/4/1.

Lectures for Women, Easter Term 1870. Newnham College Archives EC/2/2/2.

Merton House and the Cambridge Lectures for Women. Records of Newnham College, 1871–1881. Newnham College Archives AD/5/4/1.

Newnham College Objectives. Records of Newnham College, 1882–1890. Newnham College Archives AD/5/4/1.

Newnham College Report, 1877. Records of Newnham College, 1871–1881. Newnham College Archives AD/5/4/1.

Newnham College Report, 1880. Records of Newnham College, 1871–1881. Newnham College Archives AD/5/4/1.

Newnham Hall Rules and Terms of Admission, 1875. Records of Newnham College, 1871–1881. Newnham College Archives AD/5/4/1.

Occupations of Former Students, November 1897. Newnham College Archives EC/2/4/3.

Regent Hotel: History. Online at https://www.regenthotel.co.uk/history.

Altick, Richard D., *The English Common Reader: A Social History of the Mass Reading Public, 1800–1900*. Ohio State University Press, 1957.

Billings, Malcolm, *Queen's College: 150 Years and a New Century*. James and James, London, 2000.

Clough, Anne J., *Suggestions for the Training and Examination of Governesses*. London, Robson and Son, 1868, pp. 2–3. Newnham College Archives EC/2/1/6.

Clough, Anne J., *Merton Hall and the Cambridge Lectures for Women (1873)*. Records of Newnham College, 1871–1881. Newnham College Archives AD/5/4/1.

Fawcett, Millicent Garrett, Education of Girls, *Essays and Lectures on Social and Political Subjects*. Fawcett, H. and Fawcett, M.G. Macmillan, London, 1872, p. 196.

Hamilton, Mary Agnes, *Newnham: An Informal Biography*. Faber and Faber, London, 1936.

Lamberton, L.J., "A Revelation and a Delight": Nineteenth-Century Cambridge Women, Academic Collaboration, and the Cultural Work of Extracurricular Writing. College Composition and Communication, **65**, No. 4 (June 2014).

Markby, Thomas, Regulations for the Examination for Women given July 4–7, 1870. Newnham College Archives EC/2/4/1.

McMurran, S.L. & Tattersall, J.J., Fostering Academic and Mathematical Excellence at Girton College, 1879–1940, *Women in Mathematics*, ed., Berry. J.L., et al, Springer, New York, 2017, 1–35.

Officer, Lawrence H., & Williamson, Samuel H., Computing 'Real Value' Over Time with a Conversion between U.K. Pounds and U.S. Dollars, 1791 to Present, *MeasuringWorth*, 2021.

Sidgwick, Henry, Prospectus for Lectures for Women, 18 December 1869. Newnham College Archives EC/2/2/2.

Venn, John & Venn, J.A., *Alumni Cantabrigienses, Vol. 2: From 1752 to 1900.* Cambridge, 1944.

Are Euclid's Diagrams 'Representations'? On an Argument by Ken Manders

David Waszek

Abstract In his well-known paper on Euclid's geometry, Ken Manders sketches an argument against conceiving the diagrams of the *Elements* in 'semantic' terms, that is, against treating them as representations—resting his case on Euclid's striking use of 'impossible' diagrams in some proofs by contradiction. This paper spells out, clarifies and assesses Manders's argument, showing that it only succeeds against a particular semantic view of diagrams and can be evaded by adopting others, but arguing that Manders nevertheless makes a compelling case that semantic analyses ought to be relegated to a secondary role for the study of mathematical practices.

1 Introduction

Should the diagrams of Euclid's geometry be conceived of in *semantic* terms, that is, as *representations* of some kind? On the basis of Euclid's pervasive use of seemingly 'impossible' diagrams in proofs by contradiction, Manders (2008) argues for a negative answer: a semantic approach to mathematical diagrams is not fruitful, he claims, neither for Euclid nor in general. This paper aims at clarifying and evaluating Manders's argument, a task that requires distinguishing different senses in which Euclid's diagrams may be called 'representations'.

I first show that, despite seemingly being about semantic accounts of diagrams *in general*, Manders's argument actually targets a *very specific* thesis about the relation between the text of Euclid's propositions and proofs on the one hand, and the corresponding diagram on the other (Sect. 2): roughly, the thesis that the diagram is (or depicts) what the text *is about*—that the text *talks* about lines and circles while the diagram directly (though perhaps approximately) *displays* them. Against this thesis, which I dub the 'classical view' of diagrams, Manders's observation that

D. Waszek (✉)
McGill University, Department of Philosophy, Montréal, QC, Canada
e-mail: david.waszek@posteo.net

M. Zack, D. Schlimm (eds.), *Research in History and Philosophy of Mathematics*, Annals of the Canadian Society for History and Philosophy of Mathematics/Société canadienne d'histoire et de philosophie des mathématiques,
https://doi.org/10.1007/978-3-030-95201-3_7

reductio proofs often rely on apparently incorrect diagrams indeed seems cogent: for instance, in some reductio proofs, the text talks about properties of circles but the diagram shows (indeed, *has to* show) something else, namely ellipses or other irregular closed curves, because circles satisfying the conditions imposed by the text cannot exist. Section 3 then explores an objection (hinted at but not really discussed by Manders) that a fairly straightforward refinement of the classical view, according to which the diagram only depicts *some* of the properties ascribed by the text to the objects it discusses, may allow evading the argument. I show that the case of reductio proofs, while no longer sufficient to provide a downright refutation of such a refinement, nevertheless diminishes the classical view's appeal and plausibility.

I then introduce another approach to Euclid's diagrams—inspired by Barwise and Etchemendy's work on diagrammatic reasoning—which certainly deserves to be called 'representational' or 'semantic', but in a different sense than the classical view (Sect. 4): in it, diagrams and text are placed on an equal footing, and both are given a semantics in a common domain. I argue that this approach easily accommodates reductio proofs, so that Manders's argument, while successful against a limited target, cannot warrant sweeping conclusions about semantic accounts of diagrams in general. Finally, Sect. 5 returns to the bigger picture: I conclude that, although semantic analyses of Euclid's diagrams remain possible despite reductio proofs, they might not bring much to the study of a mathematical practice such as Euclid's. In the end, Manders's broader point—that semantic analyses, even if possible, should be relegated to a secondary role—remains compelling.

2 Reductio Proofs and the 'Classical View' of Diagrams

While Manders starts his discussion of 'semantic' approaches to Euclid's diagrams in very general terms,[1] the target of his argument is actually a quite specific thesis. The goal of this first section is to spell out this target, which I shall dub the 'classical view' of diagrams, and to explain why Manders's discussion of reductio proofs is a cogent objection to it.

Here is how Manders introduces the thesis he sets out to refute:

> Long-standing philosophical difficulties, on the nature of geometric objects and our knowledge of them, arise from *the assumption that the geometrical text is in an ordinary sense true of the diagram or a 'perfect counterpart'*.[2]

Roughly, this 'assumption' may be pictured as in Fig. 1; I shall refer to it as the 'classical view' of diagrams, although it is probably best thought of not as a single well-defined view, but as a family of related ones—depending on

[1] 'Artifacts in a practice that gives us a grip on life are sometimes thought of in semantic terms—say, as representing something in life. There is, of course, an age-old debate on how geometrical diagrams are to be treated in this regard.' (Manders, 2008, 84).

[2] Manders (2008, 84); my emphasis.

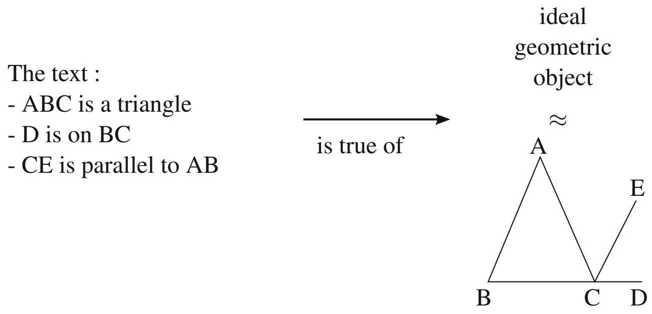


Fig. 1 The 'classical view' of diagrams

whether concrete diagrams or ideal objects take centre stage, and on how the relation of approximation that may obtain between them is understood. The word 'representation' is notoriously slippery and there is little to be gained, for our purposes, from a terminological discussion about whether—on such an account—diagrams should properly be called representations, and if so representations of what; but notice that the crucial semantic relation here holds between the text and the diagram and in fact goes *from* the text *to* the diagram. The diagram itself may secondarily be taken to 'represent' the ideal version of itself that it approximates (assuming one decides to defend such a version of the classical view), but this second relation is of no import for Manders's argument.

On the face of it, something like the classical view has considerable plausibility. Elementary geometry would be about 'perfect' or 'ideal' geometrical configurations, to which the diagrams we actually draw give us some kind of approximate access; the fact that in practice, our lines have thickness, tend not to be perfectly straight, and so on, does not for the most part threaten the reliability of our inferences.

Against this classical view, Manders brings up the puzzling case of proofs by contradiction:

> [A] genuinely semantic relationship between the geometrical diagram and text is incompatible with the successful use of diagrams in proof by contradiction: reductio contexts serve precisely to assemble a body of assertions which patently could not together be true; hence no genuine geometrical situation could in a serious sense be pictured in which they were. This simple-minded objection has nothing to do specifically with geometry: proofs by contradiction never admit of semantics in which each entry in the proof sequence is true (in any sense which entails their joint compatibility).[3]

To clarify what is at stake here, let us look at an example, namely Proposition III.5 (i.e., Proposition 5 of Book III of Euclid's *Elements*), which asserts that 'if two circles cut one another, they will not have the same centre.'[4] In his proof, Euclid

[3] Manders (2008, 84).

[4] See Heath (1908, II: 12) = Vitrac (1990, I: 399–400) = Heiberg (1883, I: 176–177).

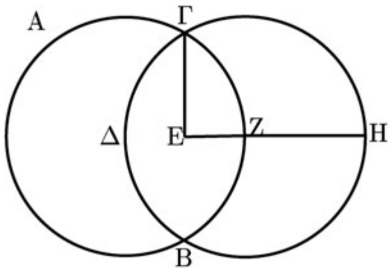


Fig. 2 Diagram for Proposition III.5 of Euclid's *Elements* found in Codex B, as reproduced by Saito (2011, 54)

introduces two circles ΑΒΓ and ΓΔΗ that cut one another in two points Β and Γ, assumes that they have the same centre Ε, and goes on to show that these hypotheses are contradictory. On the face of it, the conclusion means that it is impossible to come up with a correct diagram of two concentric circles that cut one another in two points. Nevertheless, as is always the case in Euclid, the *Elements* do provide a diagram (Fig. 2)—but this diagram cannot and does not conform to the hypotheses: while it does show two circles cutting one another and also displays a point Ε, this point does not look remotely like the centre of either circle. (Admittedly, the vagaries of textual transmission make it impossible to know for sure what kind of diagrams Euclid himself may have used; taking advantage of recent historical scholarship, I have chosen to reproduce sketches of the diagrams of 'Codex B', one of the oldest extant manuscripts of the *Elements*.[5]) Thus it is clear—here as in every reductio proofs of the *Elements*—that the assumptions in the text cannot all be true of the diagram.

The reason this is, according to Manders, a serious problem is that it undermines the classical view's main virtue, to wit, the fact that it nicely explains why Euclid-style geometry is able to unproblematically rely on diagrams in the course of its proofs:

> If diagram imperfections only were in play, one might well hold that the function of diagrams could fruitfully be approached by first elaborating a notion of perfect geometricals of which the text is literally true, then treating diagrams actually drawn in geometrical demonstrations as approximations to perfect ones; finally deriving from all this an

[5] A few elements of context about this choice are in order. As recently shown by Ken Saito, Byzantine manuscripts of Euclid—which, save for isolated fragments, are the oldest we have—typically display diagrams that differ significantly from those of modern editions, including critical ones. (For an introduction, see Saito and Sidoli, 2012, Saito, 2009, 817–825 or Saito, 2012—the latter discussing the very Proposition III.5 taken as an example here. For a fuller overview of diagrams in manuscript sources of the first books of Euclid's *Elements*, see Saito, 2006, 2011.) The difference is of particular interest in the case of reductio proofs, as the manuscript diagrams are often much more blatantly 'wrong' or 'impossible' than those of modern editions. This is why I have chosen, in this paper, to reproduce the diagrams from 'Codex B', a 888 C.E. manuscript, which, though removed from Euclid himself by almost 1200 years, is one of the oldest still extant. (The letters standardly used to refer to manuscripts of the *Elements* go back to Heiberg's authoritative nineteenth-century critical edition of the Greek text; for a list, see Heiberg, 1883, I: V–X or Saito, 2006, 95–96.)

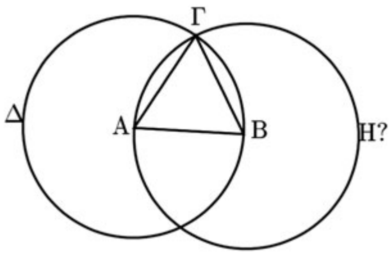


Fig. 3 Diagram for Proposition I.1 of Euclid's *Elements* in Codex B (Saito, 2006, 97)

> understanding of the bearing of the imperfect diagram on inferences in the text. But no detour through ontology and semantics, which treats of truth in a diagram in a sense which entails joint compatibility of all claims in force in the reductio context can speak to the difficulty with the role of diagrams in reductio arguments, which are pervasive in Euclid.[6]

Let us spell out the argument here. All the geometrical diagrams we draw are (at least a little) off. Accordingly, it would be hopeless, say, to try and determine whether the angles of a triangle are exactly equal by way of measurements performed on some concrete drawing of it. On the other hand, it is plausible to think that *some* properties of our diagrams are not impacted by such imperfections and that they adequately reflect the properties of the ideal geometrical objects we are aiming at. (Such, for instance, is the position defended by Panza (2012).) In fact, it is well-known that Euclid's proofs often rely on their diagrams;[7] the most famous example of this is Proposition I.1 (Fig. 3), which shows how to construct an equilateral triangle on a given line.[8] Modern criticisms notwithstanding, authors like Panza and Manders argue that—in the framework of Euclid's plane geometry if not by later standards—it is safe and legitimate to conclude on the basis of the diagram that the two circles intersect, and in general that Euclid's geometry licenses reading off from diagrams *some* properties that are unaffected by standard drawing imperfections.[9]

Once the inferential reliance of Euclid's geometry on its diagrams is recognized, Manders's challenge to what I called the 'classical view' arises from the observation that *reductio proofs too* rely on diagrams—despite their diagrams being blatantly wrong. Again, Proposition III.5 can help us pinpoint the issue. After introducing the circles ΑΒΓ and ΓΔΗ, their intersections Β and Γ, and their common centre Ε (see Fig. 2 above), Euclid reasons as follows:

---

[6] Manders (2008, 85–86).

[7] For a survey of the role of Euclid's diagrams in his proofs, see for instance Netz (1999, 175–182).

[8] Heath (1908, I: 241–242) = Vitrac (1990, I: 194–195) = Heiberg (1883, I: 10–13).

[9] In general, Manders calls 'co-exact' those properties that can be read off from diagrams in Euclid; as for Panza, the fact that diagrams of Euclidean geometry allow attributing some of their properties to the corresponding geometrical objects is what he calls their 'local role', and those properties that geometrical objects are taken to inherit from their diagrammatic representations are what he calls 'diagrammatic attributes' (see Panza 2012, in part. 72–82).

> [L]et EΓ be joined, and let EZH be drawn through at random. Then, since the point E is the centre of the circle ABΓ, EΓ is equal to EZ. Again, since the point E is the centre of the circle ΓΔH, EΓ is equal to EH. But EΓ was proved equal to EZ also; therefore EZ is also equal to EH, the less to the greater: which is impossible.[10]

This proof relies on the line EZH, which is 'drawn through at random' and is taken to meet ABΓ in Z and ΓΔH in H (in this order). Importantly, in Euclid's geometry *nothing but the diagram* can provide us with the existence of a line satisfying these conditions—a modern diagram-free version of this proof would require an additional axiom, just as in the case of Proposition I.1.

Manders's argument should by now be clear. Reductio proofs in the *Elements*—just like direct proofs—essentially rely on their diagrams. But whereas, in the case of direct proofs, this reliance can be justified by taking the diagram to adequately reflect some of the properties of a corresponding perfect geometrical object (to wit, those properties that are not impacted by the coarseness of our drawing), this answer appears unavailable for reductio proofs, since what such proofs show is precisely that geometrical objects obeying their hypotheses cannot exist. Hence, unless one is prepared to claim that diagrams play a fundamentally different role in direct proofs and in reductio proofs—and that the classical view is adequate to the former but not to the latter—the classical view seems refuted.

3 A Partial Version of the Classical View

The argument just presented against the classical view gives rise to a simple objection. It is clear that in propositions proved by contradiction, like III.5, the hypotheses of the reductio cannot *all* be true of the diagram. But *some* of them will be, and that may be enough to account for whatever inferences are drawn from the diagram—in much the same fashion as for usual proofs. Manders concedes as much:

> It does not follow that there could not be a picturing-like relationship between the diagram and *some* claims in force within a reductio context.[11]

This section discusses the prospects of such a 'partial' classical view and puts it to the test of Proposition III.5 as well as of thornier cases, Propositions III.2 and III.10. We shall see that while, strictly speaking, the argument of the previous section is not

[10] Trans. from Heath (1908, II: 12), where (for consistency with the Codex B diagram) I have replaced Heath's Roman letters with Heiberg's Greek letters. See also Vitrac (1990, I: 400) = Heiberg (1883, I: 177).

[11] Manders (2008, 85); KM's emphasis. The claims 'in force' within a reductio context refer to the hypotheses under which one arrives at a contradiction; the terminology here comes from an analogy with natural deduction, in which inferences are relative to a context defined by the undischarged assumptions under which it is made.

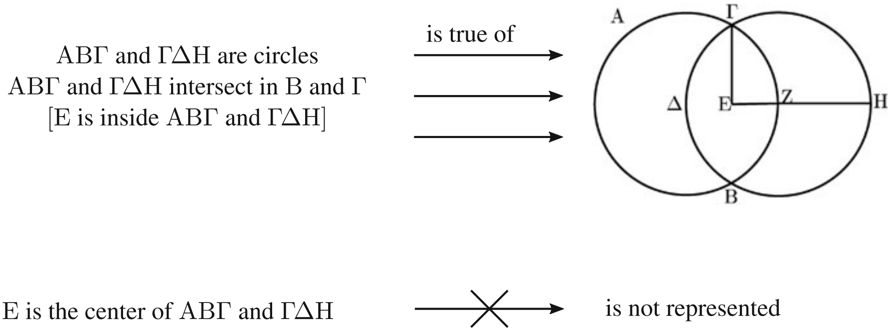


Fig. 4 A partial version of the classical view of diagrams, on the example of Proposition III.5

enough to knock down the partial classical view, it still puts significant pressure on it.

Let us first return to Proposition III.5. To rescue the classical view of diagrams from the reductio objection, the idea is to turn it into a 'partial' variant—pictured on Fig. 4—according to which only *some* of the claims in the text would be true of the diagram. In this case, while the diagram does display two circles cutting one another in two points—this part of the hypotheses is accurately shown—point E is not their centre—this hypothesis is simply not reflected by the diagram. This makes the inferential reliance on the diagram unproblematic. Indeed, the diagram adequately represents two circles as well as a point *that is inside both of them*; on this unassailable basis, one can construct Z and H. One can then derive a contradiction by reference to the *additional hypothesis* (not reflected by the diagram) that E is the centre of both circles—but crucially, no further reference to the diagram is required for this.

However, the reductio problem raised by Manders retains some force, for two different reasons. First, there are reductio proofs that are harder to analyse than III.5. Consider for instance III.2—'if on the circumference of a circle two points be taken at random, the straight line joining the points will fall within the circle'[12] (Fig. 5)—or III.10—'A circle does not cut a circle at more points than two'[13] (Fig. 6). In the first case, the straight line AB cannot be diagrammed as straight; in the second, at least one of the circles cannot be a circle. Thus, in order for the same strategy as above to go through, the textual claims that one would have to assume are not represented concern the very identity of some of the objects in the diagram (e.g., that AB is a straight line, or that ABΓ and ΔEZ are circles). The problem is that, by contrast with III.5, what remains when such fundamental claims are removed is not a run-of-the-mill Euclidean diagram made up of straight lines and circles but involves objects that are not normally discussed by Euclid, namely arbitrary (non-straight)

[12] Heath (1908, II: 8–9) = Vitrac (1990, I: 394–395) = Heiberg (1883, I: 168–171).

[13] Heath (1908, II: 23–24) = Vitrac (1990, I: 412–413) = Heiberg (1883, I: 192–195).

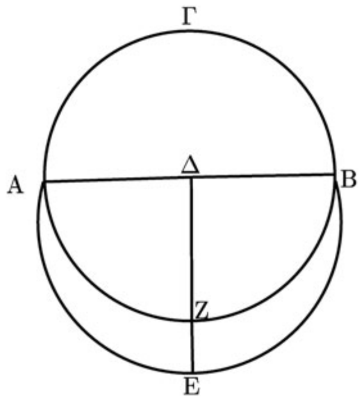


Fig. 5 Diagram of Proposition III.2 in Codex B (Saito, 2011, 52)

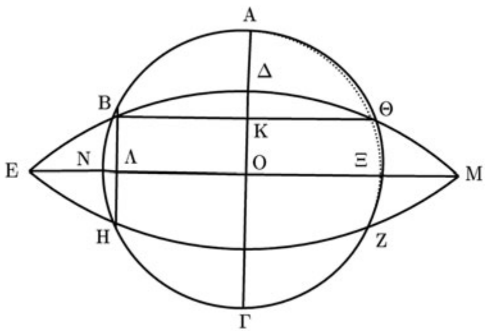


Fig. 6 Diagram of Proposition III.10 in Codex B (Saito, 2011, 59)

lines and arbitrary (non-circular) closed curves. Admittedly, this is not a knock-down argument against the (partialized) classical view: if the point of the classical view, and of its partial variant, was to defend the reliance on diagrams in proofs—by arguing that diagrams adequately reflect some properties of corresponding ideal geometrical objects—then examples such as Propositions III.2 and III.10 do not make this strategy impossible. However, they do make it more costly, because they force one to admit that, in the course of a proof, diagrams can allow discerning properties of geometrical objects—like arbitrary curved lines—that go beyond the usual circles and straight lines that Euclid's geometry is ostensibly about.

A second difficulty for the partial classical view concerns the norms governing partial representation. It is well and good to suggest, as done above, that the diagram of III.5 does not depict E as the centre of both circles but merely as a point inside them, or that the diagram of III.2 does not depict the line AB as straight, but merely as a line. Euclid's text, however, does not say anything about what the diagram should or should not depict. Tellingly, in order to make the 'partial' interpretation of III.5 fully explicit on Fig. 4, I have been led to introduce a phantom claim—that E is inside ABΓ and ΓΔH—which is stated nowhere in the text. In other words, while the classical view of diagrams introduced in the previous section was simple and clear, its partial variant seems to obey subtle and implicit norms: what claims is one allowed not to represent on the diagram? What kind of distortions on circles and lines are allowable?[14]

This second difficulty is presumably why, in Manders's view, conceding that the diagram is not a straightforward (though perhaps approximate) model of the text is the crucial step. Once this is granted, one is forced to accept that the way diagrams operate in reductio proofs is not self-evident and requires a deeper analysis:

> Thus one is forced back to a direct attack on the way diagrams are used in reductio argument; the problem of the relationship between diagram and geometric inference here turns out to be one of standards of inference not reducible in a straightforward way to an interplay of ontology, truth, and approximate representation. But once this is admitted, there seems to be no reason why direct inferential analysis of diagram-based geometrical reasoning should not be the approach of choice to characterizing geometrical reasoning overall, with or without reductio.[15]

The 'direct inferential analysis' Manders advocates focuses on the way diagrams are used rather than on what they might be taken to represent. Explaining in detail how he does this is beyond the scope of this paper, but roughly, his main strategy is to explain the norms of the practice in terms of the requirement—seemingly crucial to Euclid's kind of mathematical practice—that geometers should be able to reach agreement on controversial cases and adjudicate disputes conclusively. In a nutshell, he argues that the reason the norms of Euclidean practice do not license observing, say, equality of lengths on a diagram is that this cannot be done reliably: different geometers, constructing two diagrams according to the same rules, might reach different conclusions. On the other hand, properties that can be reliably reproduced by different practitioners—for instance that two circles constructed as in Proposition I.1 intersect—can unproblematically be read off from diagrams.

4 A Semantics for Diagrams

As discussed in the previous sections, Manders's argument targets a very specific account of the semantic role of Euclid's diagrams, to wit, the 'classical view' according to which the diagrams are what the text is about (or perhaps approximations of the objects that the text is about). But—regardless of whether, as discussed in the previous section, his refutation of the classical view can ultimately be evaded—other 'semantic' accounts of diagrams are possible, which his discussion does not address. This section explores one such account, in which the diagrams, rather than being—so to speak—the semantics of the text (as was the case in the above), are

[14] As Rabouin (2015, 115–118, 126–131) shows, the kinds of distortions that reductio proofs require easily produce incorrect results in other situations: some form of *selective* control over diagram distortions is clearly going on; see Manders (2008, 109–118) for further discussion.

[15] Manders (2008, 86).

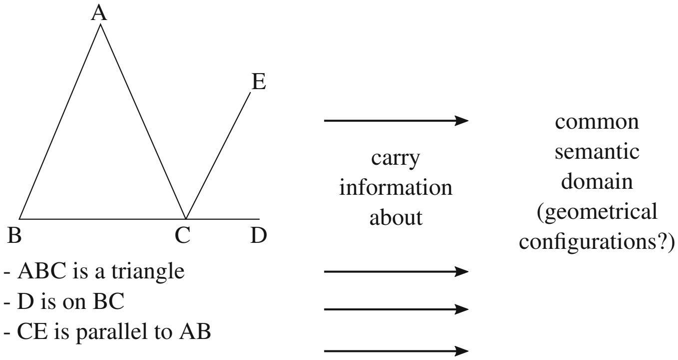


Fig. 7 A different semantic view of geometrical diagrams

given a semantics themselves. I then argue that the case of reductio proofs does not raise particular problems for such an account.

As a first observation, note that the straightforward conclusion that follows from Manders discussion of reductio proofs is, in his own words, that diagrams should be treated as '*textual* components of a traditional geometrical text or argument, rather than semantic counterparts'[16]—but that textual components of a proof *are* usually given a semantics. Indeed, Manders's conclusion is perfectly compatible with a view according to which both text *and* diagrams are treated as ways to represent a common subject-matter, that is, as admitting a semantics in a common domain. Roughly, the resulting view could be pictured as in Fig. 7. Notice that here, the relevant semantic relation does not relate the text with the diagram, as it does according to the classical view discussed earlier, but relates both text and diagram with a common subject-matter. Such a picture of diagrams as information bearers underlies much of the diagrammatic reasoning program initiated by Barwise and Etchemendy;[17] this is, for instance, how Shin (1994) approaches Venn diagrams. In a similar spirit, two formal diagrammatic systems for Euclid's geometry have been produced in recent decades that approach diagrams as text-like proof symbols and define for them a formal semantics of the kind meant here: the systems of Miller (2007) and Mumma (2006, 2019).[18]

To get a better grasp on what such a semantics of diagrams would look like—and importantly, on how it could help make sense of reductio proofs—let us make a detour through a reductio proof in which no diagram is required. Consider

[16] Manders (1996, 391). (This quotes comes, not from his most famous 2008 paper, but from a previous publication on the topic; his view on this did not change, however.)

[17] See, in particular, the collective volume Allwein and Barwise (1996).

[18] The original version of Mumma's system did not define a formal semantics for its diagrams, but this is possible and is done in Mumma (2019).

Proposition VII.22, from Euclid's books on number theory:[19] 'The least numbers of those that have the same ratio as them are prime to one another.'[20] Euclid proves this by reductio. He starts from two numbers A and B and makes two distinct hypotheses about them, from which he then derives a contradiction: (1) that they are the smallest among those that have the same ratio as them (call this hypothesis 'p') and (2) that they are not prime to one another (call this 'q'). It follows from q that A and B have a common factor, hence that there are smaller numbers C and D that have the same ratio, which contradicts p.

Now, both p and q taken on their own can unproblematically be given a semantics, and so can their conjunction $p \wedge q$. For instance, in a suitable model-theoretic framework, such a semantics might roughly be determined by: (1) for p, the set of pairs of numbers (A, B) such that they are the smallest among those having the same ratio; (2) for q, the set of pairs of numbers (A, B) that are not relatively prime; (3) for $p \wedge q$, the empty set—precisely because the two hypotheses are, in fact, incompatible. Of course, the fact that the semantics of $p \wedge q$ reflects the incompatibility of p and q does not in and of itself preclude using p, q, or $p \wedge q$ in the course of a (reductio) proof, since there are no syntactic principles that would allow deducing this incompatibility immediately.

With this example in mind, let us go back to a reductio proof involving one of the 'impossible' diagrams discussed earlier—say, Proposition III.10. My claim is that it can be treated in a similar fashion. In very broad terms, and without entering into a precise discussion of how Euclidean diagrams may be formalized, there would be at least two different strategies to give a semantics to the diagram of III.10 (Fig. 6). First, in the spirit of our discussion in Sect. 3, one could take this diagram to represent two closed curves ΑΒΓ and ΔΕΖ intersecting in four points, and nothing more. The diagram would then play the role of, say, hypothesis p from the preceding example, while q would correspond to the further (and incompatible) claim that ΑΒΓ and ΔΕΖ are circles. A second, perhaps less convoluted reading would take the diagram at face value and accept that it does represent two *circles* intersecting in four points; on this reading, it would be comparable to the conjunction $p \wedge q$ rather than to p (or q)—i.e., it would have an empty set of models as semantics, but would not self-destruct, because nothing in the practice of Euclidean geometry licenses immediately reading off the contradictoriness of such a diagram from its appearance (indeed, one could hardly practice Euclidean geometry at all if one were allowed to reject any diagram as contradictory if some circle *appeared* imperfect). The 'semantic' contradictoriness of such a diagram—that is, the fact that it admits no models because it involves incompatible assumptions—would no more preclude

[19] As a matter of fact, the propositions from Euclid's arithmetical books also contain diagrams of sorts, which represent by way of lines the numbers discussed in the text; but, in contrast to the geometrical case, these diagrams do not play much of a role in proofs. See, e.g., Mueller (1981, 67).

[20] Heath (1908, II: 323–324) = Vitrac (1990, II: 328) = Heiberg (1883, II: 234–237). In modern terms, if two integers A and B are such that there are no *smaller* integers C and D such that $\frac{C}{D} = \frac{A}{B}$, then A and B are relatively prime.

its use in proofs than the incompatibility of p and q precludes assuming $p \wedge q$ and reasoning from there. This may still strike us as strange, but, as the comparison with our number-theoretic example shows, the strangeness in question is about the semantic analysis of reductio proofs *in general* and has nothing specific to do with Euclid's geometry: reductios involving seemingly impossible diagrams then become no more (but also no less) semantically problematic than any other reductio proofs. They certainly license no conclusions unique to the semantics of Euclid's diagrams.

Thus, approaching diagrams in 'semantic terms' is possible, and in fact fully compatible with Manders's diagrams-as-text account of Euclidean practice; but, instead of establishing a semantic relation between text and diagram, this involves endowing with a semantics both diagram and text, now placed on an equal footing.

5 By Way of Conclusion: What Is a Semantics Good for?

In sum, Manders's reductio argument, though framed as an attack against any kind of 'semantic' approach to Euclid's diagrams, is in fact only cogent against a particular account, the one I called the classical view (in passing, we have seen that even this view may be defended, more or less convincingly, with appropriate contortions). The idea of treating diagrams—just like the sentences of Euclid's text—as representations that admit of a semantics is unaffected, in principle, by the case of reductio proofs.

There is a sense, however, in which this discussion misses the deeper lesson of Manders's study. It may well be that giving a semantics to Euclidean diagrams is possible; his analysis, however, suggests that it would be rather futile. As he puts it:

> If this order of analysis [a direct inferential analysis of diagram-based geometrical reasoning] proves fruitful, ontological and semantic considerations will seem decidedly less central to the philosophical project of appreciating geometry as a means of understanding. For in their then remaining role of making the standards of geometrical reasoning seem appropriate, ontological-and-semantic pictures will have to compete with other types of considerations which we will find have potential to shape a reasoning practice.[21]

According to Manders, what accounts for the way diagrams are used in Euclid is not at bottom what they *represent*, but rather pressures and constraints of a different kind—centrally, the requirement that practitioners be able to straightforwardly reach agreement. This, of course, does not preclude coming up with some kind of semantics (perhaps along the lines suggested in Sect. 4) that would account for the norms of the practice as one finds them—but it would serve no purpose save '[making] the standards of geometrical reasoning seem appropriate' *after the fact*, that is, legitimizing norms whose real source is elsewhere.

[21] Manders (2008, 86).

Manders's conclusion, then, is that the historian of a mathematical practice should focus on its inferential norms; one may add a semantic analysis later on, but trying to start from it is liable to lead astray rather than bring any benefit. What matters is not so much to argue for the impossibility of a 'representational' approach to mathematical diagrams as to *marginalize* such an approach in favour of an inferential analysis.

Acknowledgements Without David Rabouin, who initially drew my attention to this issue and discussed it with me on numerous occasions, this paper would not exist. I would also like to thank Marco Panza, Jeremy Avigad, Ken Manders, and Dirk Schlimm for useful discussions.

References

Allwein G, Barwise J (ed) (1996) Logical Reasoning with Diagrams. OUP

Vitrac B (ed) (1990–2001) Euclide d'Alexandrie, Les Éléments, 3 vol. PUF, Paris

Heath T L (ed) (1908) The Thirteen Books of Euclid's Elements, 3 vol. CUP

Heiberg J L (ed) (1883–1888) Euclidis Elementa, 5 vol. Teubner, Leipzig

Manders K (1996) Diagram contents and representational granularity. In: Seligman J, Westerståhl D (ed) Logic, Language and Computation, 389–404. CSLI Publications, Stanford

Manders K (2008) The Euclidean Diagram (1995). In: Mancosu P (ed) The Philosophy of Mathematical Practice, 80–133. OUP

Miller N (2007) Euclid and His Twentieth Century Rivals. CSLI Publications, Stanford

Mueller I (1981) Philosophy of Mathematics and Deductive Structure in Euclid's *Elements*. MIT Press, Cambridge

Mumma, J (2006) Intuition Formalized. PhD thesis, Carnegie Mellon University. http://johnmumma.org/Writings_files/Thesis.pdf

Mumma, J (2019) The Eu approach to formalizing Euclid. Notre-Dame J Form Log 60(3)

Netz, R (1999) The Shaping of Deduction in Greek Mathematics. CUP

Panza, M (2012) The twofold role of diagrams in Euclid's plane geometry. Synthese 186(1):55–102

Rabouin, D (2015) Proclus' Conception of Geometric Space and Its Actuality. In: De Risi V (ed) Mathematizing Space, 105–142. Springer

Saito K (2006) A preliminary study in the critical assessment of diagrams in Greek mathematical works. SCIAMVS 7:81–144

Saito K (2009) Reading Ancient Greek Mathematics. In: Robson E, Stedall J (ed) The Oxford Handbook of the History of Mathematics, 801–826. OUP

Saito K (2011) The diagrams of book II and III of the *Elements* in Greek manuscripts. In: Saito K (ed) Diagrams in Greek Mathematical Texts, 39–80. http://greekmath.org/diagrams/Diagrams_in_Greek_Mathematical_Texts_Report_Ver_2_03_20110403.pdf

Saito K (2012) Traditions of the diagram, tradition of the text. Synthese 186(1):7–20

Saito K and Sidoli N (2012) Diagrams and arguments in ancient Greek mathematics. In Chemla K (ed) The History of Mathematical Proof in Ancient Traditions, 135–162. CUP

Shin S-J (1994) The Logical Status of Diagrams. CUP

Abstraction by Embedding and Constraint-Based Design

Bernd Buldt

Abstract The traditional approach to concept formation via abstraction is tantamount to removing properties and making the corresponding concept less rich. In other words, the more abstract a concept is, the less content it has. This approach to abstraction does not, however, provide an adequate model for concept formation in mathematics. We need to replace the traditional account with one that is, on one hand, true to mathematical practice and the mathematical experience and is, on the other hand, compatible with insights from cognitive science. This chapter adds to the existing literature by homing in on a question that lies at the intersection of the two criteria just mentioned; it is the question of why abstract concepts are perceived not only as more powerful but also richer, not poorer, in content.

1 Introduction

Most people will agree that modern mathematics is abstract. And also that it is difficult. And some will add that it is beautiful. But all these statements do not mean much unless we make them more precise. So what do we mean when we say modern mathematics is abstract?

Common parlance knows at least two meanings of the word "abstract." On one hand, it may mean that something is not concrete, not tangible, not an object of the senses but an object of our thinking only. Accordingly, the oak tree in my backyard is concrete, but a metric space is abstract. On the other hand, something may be

I wish to thank two anonymous referees for their careful reading and many helpful observations and remarks.

B. Buldt (✉)
Department of Mathematical Sciences, Purdue University Fort Wayne, Fort Wayne, IN, USA
e-mail: buldtb@pfw.edu

M. Zack, D. Schlimm (eds.), *Research in History and Philosophy of Mathematics*, Annals of the Canadian Society for History and Philosophy of Mathematics/Société canadienne d'histoire et de philosophie des mathématiques,
https://doi.org/10.1007/978-3-030-95201-3_8

abstract because we obtained it by a process of abstraction. The general concept "tree," for instance, is abstract in this sense because we stripped it from all aspects particular to one or more concrete, individual trees. Other uses of "abstract" seem derivative from those two; for example, when we speak of "abstract art" because a painting lacks a concrete object it depicts or music is without a catchy melody.

Mathematicians, too, use the word "abstract" and, like common parlance, they use it without an agreed-on or technical meaning. In the eighteenth century, abstract mathematics was used as the opposite of mixed, that is, applied mathematics, and it was synonymous with pure mathematics (Anonymous, 1797, vol. 1, p. 29). In the nineteenth century, with the emergence of non-Euclidean geometry and the arithmetization of analysis, "abstract" referred to the perceived loss of guidance from geometric or spatial intuition (see, e.g., Volkert, 1986). In the twentieth century, it was interchangeably used, as we all know, to refer to (i) the axiomatic method, (ii) the growing dominance of abstract algebra, (iii) the adoption of a set theoretic framework, or any mix of those three. We believe all these uses have their justification, and in this chapter, we will not try to standardize language or identify the "true meaning" of abstract. Actually, for our present purposes, most everything that falls within the fuzzy boundaries of abstract as delimited by any of these three meanings is fine. For we will be more concerned with what abstract is not.

Our main target in this paper is the mismatch between the perceived poverty and the power of abstract mathematics. The perceived poverty of modern mathematics results from the everyday understanding of abstraction according to which an abstract concept is less rich since it has been stripped from many or most of its specifics. It is the difference between a stick figure and a real human being. But abstract mathematics in any form we mentioned above has been advanced precisely because it is more powerful. So why, then, is modern, abstract mathematics more powerful?

We shall proceed in three steps. First, we briefly look at the idea that a concept is abstract if and when it was obtained by a process of abstraction; it is this idea that undergirds the most common understanding of abstraction. We need to clarify this notion because we argue that it does not apply to our case. We then home in on our topic and look, first, at the process of embedding one theory in another. This process is frequently being used to achieve results that cannot be established otherwise and is often characterized as an increase in abstraction. We argue that calling this process or its outcomes "abstract" is misleading. We then look, second, at how mathematicians arrive at concepts that are powerful enough to make embeddings into their theories rewarding in the first place. While the resulting concepts are likewise characterized as abstract, we make a similar argument and claim that calling them "abstract" is equally misleading. Rather, we argue, these concepts are obtained by a process we call "constraint-based design."

Things will become clearer as we go, we hope, but not in regard to how we use the words "concept" and "power." We use the word "concept" the way mathematicians do, without digging any deeper, and presume concepts are given by their definitions. In regard to ideas about one theory being more powerful than another, we likewise follow informal custom among mathematicians. This comprises two aspects, namely, first, that the language is expressively more powerful (can express more and more subtle ideas) and, second, that the resulting theories are deductively more powerful (can prove more and more subtle results). Topology, for example, is more powerful than classical analysis in that it allowed Banach to replace Weierstrass' proof for the existence of continuous nowhere differentiable functions—a result Weierstrass obtained by constructing a specific example—with the simple but "abstract" observation that F, the set of all functions in $C[0, 1]$ that have a finite derivative, is a first-category set.

2 Stick-Figure Abstraction

We understand the traditional or everyday account of abstraction, which is tied to the idea that we can slice up the world we experience into objects and their properties, to be as follows. A person we meet (i.e., the concrete object) has certain properties: height, weight, age; skin and hair color; dressed to signify a certain gender; personal names and relatives; and so on. The concept of that person, say, Julia, is then the sum total of all these properties. Properties are in turn conceived as concepts again that, in that role, are variously called subconcepts, attributes, notions, or marks (we use notion). The contents of a concept, its intension, are then the set of all the notions it contains. Accordingly, an individual concept (i.e., the concept of an individual object), such as "Julia," which contains potentially infinitely many notions (each space–time location turned into a property), is therefore as rich as possible in contents, while more abstract concepts, such as "animal," are poorer in content but have greater scope, that is, they apply to an entire class of objects, not just one or a few individual objects. The subconcept relation gives rise to a preorder among concepts, traditionally depicted as a conceptual hierarchy with the most abstract concept (e.g., "being" or "something") on top and proper names, such as "Julia" (as concepts of concrete individuals), at the bottom. The inverse relation that obtains between the intension (or the contents of a concept) and the extension (or the scope of a concept) corresponds to different movements along that hierarchy: moving up (abstraction) means losing contents but gaining scope, while moving down (instantiation) means losing scope but gaining in contents.

This complex of ideas we just sketched has been vastly influential since the Ancient Greek. It offered philosophers throughout the centuries a template for how to account for abstract ideas, and it could be used by philosophers, theologians, and early scientists as a metaphysical blueprint for understanding the world (the so-called *scala naturae* or the great chain of being). These ideas had a second career during our lifetimes—sometimes under different names (e.g., classification tasks, proto- or stereotype reasoning, etc.)—in psychology and later in cognitive science.

We put so much emphasis on these non-mathematical observations to remind us how deeply rooted in our basic cultural fabric these ideas are.[1]

What will be helpful to recall when we speak about embeddings in the next section is how Frege modified the traditional understanding of conceptual contents. The traditional, intensional definition of conceptual content (according to which the contents consist of all notions a concept includes) with an extensional definition that equates conceptual content with everything a concept logically entails.

> Any attempt to [...construe] a concept as the sum of its notions was far from my mind. [...] The conclusions that can be drawn from one judgement when combined with certain others [...is what] I call [...] *conceptual content.*[2]

Thus, according to Frege, a concept is poorer in contents if it entails less, or richer in content if it entails more.

I do not say these ideas about concept formation via abstraction are wrong per se. What I rather claim is that abstraction thus understood causes trouble when it gets applied to modern mathematics. Define a dilemma to be a pair of sentences that are both believed to be true but cannot be true simultaneously. Then the dilemma posed by abstraction is obvious:

(D1) Modern mathematics is more abstract.
(D2) Modern mathematics is more powerful.

For if (D1) is true and modern mathematics is abstract, then its concepts have fewer contents and are therefore less powerful since they entail less; this contradicts (D2). Likewise, if (D2) is true and modern mathematics is more powerful, then its concepts must entail more and hence cannot be more abstract; this contradicts (D1).

The same observations apply if we talk about abstraction not in terms of conceptual content as before but conceive of it as loss of intuitive content, here understood very broadly as something we can see, touch, feel, produce, or manipulate in a tangible way. This contrast has been expressed in various ways; for example, as the contrast between:

- Calculus and analysis (e.g., the metaphorical language of motion we employ when we work with limits such as $\lim_{h\to 0} f(x+h) - f(x)$ versus the motion-free and hence more abstract language of ϵ-δ definitions)
- Euclidean and modern geometry (e.g., using spatial intuition versus proving the plane separation from Pasch's axiom)

[1] The ideas go back mostly to Aristotle who used them to show how we arrive at mathematical ideas (*aphairesis*) and general concepts (*epagoge*); see Bäck (2014) for a recent systematic exposition. Boethius merged the Aristotelian ideas about *aphairesis* and *epagoge* and rendered them into a single Latin term: *abstractio*. It has been a staple ever since; see Weinberg (1973) for an overview. It were Arnauld and Nicole who introduced the language of intension (*la compréhension*) and extension (*l'étendue*) in their *Logic of Port Royal*; see (Arnauld and Pierre, 1662, pt. 1, ch. vi). See Lovejoy (1936) for the great chain of being.

[2] See (Frege, 1879, pp. 49, 53); emphasis in the original.

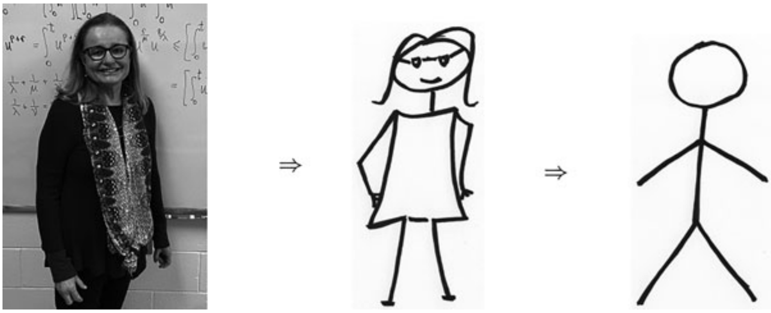


Fig. 1 Stick-figure abstraction

- Constructing and proving (e.g., solving a system of equations versus giving a non-constructive proof that a solution must exist) Or, framing the issue more generally,[3] the contrast between:
- Operative deployment and conceptual extension

What is mixed into all these instances of alleged abstractness is the perceived loss of content, whether it is construed as conceptual or intuitive or both. And because there is less content to go with, the dilemma remains. To have a common name for this mixed bag of instances, whether the emphasis is on the loss of conceptual or the loss of intuitive contents, I refer to them collectively as stick-figure abstraction (see Fig. 1).

Mathematicians working in a subject area X are trained to adhere to the definitions of the technical terms used in X and not supply them with their own, informal meaning. If you work in topology, then you go with the definitions of open and closed sets and not their informal meaning so that clopen sets will not confuse you. Non-technical terms, however, words from outside the specialized vernacular, must often be used in their informal meaning, and we must supply those. Since no textbook in mathematics defines "abstract," we have to supply our own, informal meaning of abstract. And when we do, we invariably will supply stick-figure abstraction; for, as we have shown above, it has been the default conception for more than 2000 years and the only game in town that offers itself for free. I will argue that this is wrong; and because it is wrong, the dilemma goes away.

[3] We owe the last pair of contrasts to Boniface (2007).

3 Abstraction by Embedding

Trisection of Angles Mathematicians[4] working in the mould of traditional, synthetic geometry have been unable to conclusively answer the question of whether or not it is possible to trisect an arbitrary angle with ruler (i.e., straightegde) and compass. This changed only in the nineteenth century when mathematicians learned how to tackle this problem with algebraic tools.[5] Here is a reminder how.

We translate points, straight lines, and circles into points and point sets (i.e., subsets) of $\mathbb{R}^2$, the latter two defined by polynomial equations:

$$\begin{array}{lll} \text{a.} & \text{point:} & \langle x, y\rangle \in \mathbb{R}^2\ , \\ \text{b.} & \text{straight line:} & ax + by + c = 0, \\ \text{c.} & \text{circle:} & (x-a)^2 + (y-b)^2 + c^2 = 0. \end{array} \tag{1}$$

Construction by ruler and compass (i.e., constructions of intersecting straight lines and circles) gets thus translated into solutions of systems of polynomial equations. Algebra then shows that the intersection of two straight lines amounts to the solution:

$$x = \frac{b_1c_2 - b_2c_1}{a_1b_2 - a_2b_1}\ , \text{and } y = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1}, \tag{2}$$

while the intersection of a straight line with a circle amounts to

$$\sqrt{B^2 - 4AC}, \tag{3}$$

where the equation to be solved: $Ax^2 + Bx + C = 0$ has coefficients A, B, C that are rational functions in a_i, b_i, c_i, for $i = 1, 2$. (Note that the intersection of two circles can be reduced to the intersection of one circle with their common chord, which is a straight line.) Equations (2) and (3) then entail that the constructible numbers, viz., numbers for which we can construct a corresponding line segment with ruler and compass, are numbers that result from performing the arithmetical operations of adding, subtracting, multiplying, or dividing numbers and taking the square root. (And indeed, given any two straight line segments e and f, there are known ruler-and-compass constructions for performing these operations on e and

[4] What got me started to think about embeddings in the present context was an inspiring talk by Irina Starikova on Gromov's work (on hyperbolic groups and Cayley graphs, see Starikova (2012)).

[5] The historical record does not seem to be entirely clear; see (Smoryński, 2008, ch. 4). While Wantzel (1837) announced the results first, Petersen first proved it in the modern vein using the fact that for polynomials over $\mathbb{Q}$ to have roots their degree must be a power of 2; see (Peterson, 1878, pp. 159, 166f.). (Note that we use letters such as "$\mathbb{Q}$" indiscriminately for both the set and the set equipped with some algebra.)

f.) Finally, one shows that trisecting an angle amounts to solving a cubic equation that lacks the x^2-term:

$$ax^3 - cx - d = 0, \tag{4}$$

and which has no constructible number as its root. The last step, proving the non-constructibility of the root, can be demonstrated either by algebraic observations specific to Eq. (4) under consideration (see Yates, 1942, ch. 1.6) or by proving that the minimal degree of the polynomial in question must be a power of 2 (which Eq. (4) clearly is not); that latter fact in turn can be established directly (see Klein, 1895, pt. 1, ch. 1) or, as it normally done, as a corollary to results about field extensions of $\mathbb{Q}$ or Galois theory proper (in which case we start with polynomials over $\mathbb{Q}$, of course, not over $\mathbb{R}$); see (Beachy and Blair, 2006, ch. 6.3) and (Cox, 2012, ch. 10), respectively.

The example just given seems to speak to the point we are interested in: polynomials seem more abstract than straight lines and circles, while employing ideas from the algebra of fields extensions is more powerful than geometric constructions; the former solved a millennia-old problem that was beyond the reach of the latter.

I take the increased power to be uncontroversial, but not the claim about abstractness as I hope to establish in the next section.

What Defines a Line We first remark on the importance that any mathematical concept must be operationalized by a suitable definition. If we say, for instance, an even number is a number that ends on $0, 2, 4, 6,$ or 8, then "even number" is not yet a mathematical concept proper for a mere listing does not support nor enable mathematical work. We could call such an intuitive, not yet operationalized version a proto-mathematical concept. It is a step in the right direction when we define "even number" as "divisible by two without remainder"; but this is still not yet good enough because we cannot deploy it in a proof without reformulating it first. A better operationalization of "n is even" is "there is an integer k such that $n = 2k$." And it is better because we can plug it straight into proofs (see (Alcock, 2013, § 3.5) for more on this specific example). The great importance of good definitions seems uncontroversial in light of observations from mathematics education, from the history of newer mathematical disciplines, and from anecdotal evidence of eminent mathematicians who stressed the importance of a well-chosen terminology. The reason to belabor the obvious is that we need to come back to it repeatedly.

Returning to our example, it seems obvious that we do not arrive at an equation such as "$ax + bx + c = 0$" by looking at a drawn line and then jettisoning conceptual or intuitive excess contents until what remains is a polynomial. So if there is abstraction involved, it is not of the stick-figure variety. Moreover, a polynomial is but one among different formulations of what a straight line, or a segment thereof, is; other familiar options include

synthetic geometry:	________________	(a drawn line segment)
pre-calculus:	$y - b_1 = \frac{b_2 - b_1}{a_2 - a_1}(x - a_1)$	
geometry in $\mathbb{R}^2$:	$[A : B : C]$	(equiv. class of coefficient vectors)
vector space:	$\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}),\quad 0 \le \lambda \le 1$	
topology:	$f : [0, 1] \to X,\quad f$ continuous	

None of the other four options listed above is obtained by stick-figure abstraction from the first. The only candidate that was the result of such an ill-fated abstraction process was the one Euclid recorded and which is arguably among the worst operationalization attempts we know: a straight line is an extension without breadth that lies evenly on itself with what has no parts (i.e., Euclidean points forming a Euclidean line). And the reason why it is a failed attempt is the fact that we do not get any mileage out of this definition for proving theorems.[6]

We thus see that a drawn line segment as such, that is, non-operationalized, is not a mathematical object; it is an object in my field of vision. Consequently, if I compare a drawn line segment with any of the other options listed above, then I do not compare two mathematical concepts but compare a mathematical object with a perceptual one—apples with oranges, if you will. Then, almost by definition, a concept will always be considered more abstract than an object of my senses; there is nothing specific to mathematics here.

Context and Custom What criteria do we use when we talk about definitions and judge one formulation to be more abstract than the other? My contention is that it is mostly a matter of context and, consequently, of custom.

Context means that any of the operationalizations listed above was developed to do work within the context of the indicated mathematical discipline. And it means that within that disciplinary context the definition lends itself very naturally and highly efficiently to the tasks at hand. This is because it was operationalized to do that, tuned for performance. (More on this what we call "constraint-based design" below in Sect. 4.4.1.)

Custom means that those who work in the discipline do not suffer feelings of abstractness; rather on the contrary, definitions and their meaning are part of the vernacular they feel very much at home speaking. If, however, someone leaves behind their mathematical mother tongue and learns the language of another mathematical discipline, or, when a student takes their first more advanced classes, then their first reaction might very well be a feeling of alienation caused by the new concepts. But the root cause is not that the new language is more abstract but that it is new. And to master a new language that comes with a densely woven conceptual fabric—think abstract algebra or topology—is a daunting task at first. What this suggests is that the word "abstract" is a misnomer for the alienating experience of learning a new language. And we mislabel "alienating" as "abstract" because

[6] See Fitzpatrick (2007) for the text of the *Elements* and Heath (1908) for discussion.

the new language is not just a foreign language in which we learn to speak about well-known daily routines such as greetings and asking for directions but is a new language in which we learn to speak about yet totally unknown mathematical objects that are, as concepts, eo ipso abstract.

If our analysis is correct, then it helps to explain why such a substantial effort goes toward the re-proving of known results that makes up a sizeable portion of published research. Mathematicians like to translate results back into their first language because it boosts their confidence in the result. For example, if you were raised a graph theorist, you might find Whitney's proof that a graph is planar if and only if it has a dual, the most convincing; if you were raised an algebraist, much clearer may be Mac Lane's approach to restate the question in the vernacular of edge and cycle spaces and their basis; if your first training was in topology, you find full confidence in the result only if you start with the Jordan Curve Theorem for path-connected sets and go from there.[7]

As humans we need coherent narratives and as mathematicians we are no different; one wants to understand a result, and this goes beyond the following steps in a proof and requires putting the result into the broader context of what else one knows. Or, as Rota (2008, p. xxii) once quipped:

> We often hear that mathematics consists mainly of 'proving theorems.' Is a writer's job mainly that of 'writing sentences?'

I take the fact that mathematicians feel the urge to reprove and translate results from one vernacular into another as evidence that our analysis is correct.[8]

Synthetic Geometry Does our proposal that "abstract" is a misnomer for the alienating experience of being confronted with a new language, which we introduced in terms of translating among various technical vernaculars, include the case of synthetic geometry? That is, does it apply to the four Euclidean operations (E1) through (E4)?

(E1) Given two points, draw a line connecting them.
(E2) Given two non-parallel lines, find their point of intersection.
(E3) Given a point p and a length r, draw a circle with radius r centered at p.
(E4) Given a circle, find its points of intersection with another circle or line.

It seems as though our proposal does not apply. For there is nothing in the reasoning about and the manipulation of points, lines, and circles that would strike us as abstract; rather, it seems a fairly concrete business. Suppose, however, that we had grown up in a nomadic tribe where all naturally occurring or artificially constructed shapes were crooked, never straight or smoothly curved, and that we

[7] See Whitney (1932), (Knauer, 2011, ch. 6), and Thomae (1992), respectively. Our point here is informed by recent discussions on why mathematicians redo proofs; see, e.g., Rav (1999) and Dawson (2015).

[8] Clearly, the last paragraph appeals to and makes sense only if a certain view of mathematics is adopted; for example, a view (we unfortunately cannot argue here) that emphasizes the importance of explanation.

had never encountered before in our lives the word points, lines, circles, triangle, etc.; how "abstract" would it be to learn Euclidean geometry? While it is difficult for us to put ourselves in that situation, we can get a ballpark by having a stab at the thorough assimilation of the Huzita–Hatori axioms for paper folding (i.e., the Origami version of Euclidean geometry); I suspect it will not be easy.[9]

What I here propose is the following: If we acknowledge the extent to which we have imbibed the basic vocabulary of Euclidean geometry with our cultural mother's milk, then we realize that learning Euclidean geometry can be as "abstract" for the uninitiated as modern algebra is. We just fail to notice because those concepts form part of our original cultural furnishing. Admittedly, counterfactual stories do not prove a point; they can make it plausible, though. And I think I made it sufficiently plausible that the case of synthetic geometry is included in the general case of mistaking "new and alienating" for "abstract."[10]

Emdeddings We conceive of an embedding to include at least a translation τ from one language into another. To give a precise recursive definition of τ is a standard procedure in mathematical logic and need not be repeated here;[11] the hint we gave (i.e., (1.a)–(1.c) above) suffices for our purpose. Let T and T' be two mathematical theories, each formulated in their own special language $\mathscr{L}$ and $\mathscr{L}'$, respectively, and let $\tau : \mathscr{L} \to \mathscr{L}'$ be a translation. We use square brackets to denote the deductive closure of a theory T, viz., the set of all propositions logically entailed by that theory: $[T] = \{p \in \mathscr{L}_T : T \models p\}$. One way to express the fact that a theory T' is equally or more powerful than a theory T is by introducing a partial order "$\prec$" among theories:

$$T \preceq T' \text{ iff } \tau[T] \subseteq [T'], \tag{5a}$$

$$T \prec T' \text{ iff } \tau[T] \subsetneq [T']. \tag{5b}$$

There might be more parameters involved in an embedding than just a translation τ and comparative deductive strength; for example, a commonly heard goal is that of obtaining "more information" about an object or that of "unifying" two or more theories. So in order not to prejudge the situation, we write $T \hookrightarrow T'$ for an embedding that satisfies (5a), possibly along with other criteria, and write $T \hookrightarrow^+ T'$ for a non-conservative embedding that satisfies (5b) possibly plus some other criteria. In some instances, we employ an embedding "$\hookrightarrow$" that does not

[9] For an overview of the Huzita–Hatori axioms, their history, and algebra, see Lee (2017).

[10] There is more to the plausibility of our claim than just one counterfactual story. I see it in the context of the growing awareness that culture and natural environment might lead to marked differences in our basic mathematical competencies. The here relevant buzz words include ethnomathematics, WEIRD people, neuroanthropology, and the encultured brain; see, e.g., Rosa et al. (2017), Henrich et al. (2010), Northoff (2010), or Lende and Downey (2012). Some of the challenges this open empirical question poses for mathematical cognition were recently formulated in Beller et al. (2018). I expect that the outcomes will support our claim.

[11] See, e.g., Hinman (2005, ch. 2.6).

necessarily enlarge the pool of known theorems because what we want are shorter or more transparent proofs or a different conceptual framing than what is available in T itself. In other instances, however, we look for an embedding "$\hookrightarrow^{+}$" precisely because we want to prove results in T' we were unable to obtain in T.

The existence of so many mathematical disciplines cannibalizing one another for the power grab that comes with embeddings is a characteristic feature of modern mathematics and is reflected not only at all three levels of the Mathematics Subject Classification (MSC)[12] but also in its length; in the nineteenth century there was no need for a classification scheme roughly one hundred pages long. (We use the expression "cannibalize" (which we borrow from an MSC-editor) because what gets embedded need not be an entire theory when just parts of it suffice (as tools to get a job done).)

What exactly it is that causes the power boost in each individual instance of an embedding "$\hookrightarrow^{+}$" is not always easy to tell, though; well, beyond the tautological response that T' or its axioms are more powerful. For instance, in regard to our example of the embedding of geometry into algebra: $G \hookrightarrow^{+} A$, Felix Klein proposed that it yields stronger results because Euclidean geometry lacks the notion of an algorithm that algebra has.[13] What he probably had in mind was that we have an algorithm for solving systems of equations but no comparable algorithm for sequencing geometric constructions toward a desired outcome. But I would rather locate the reason elsewhere, namely, in the circumstance that talking about fields allows us to talk about closure under operations.[14] This is how we know that no matter how complicated the constructible numbers become, they will never be the root for certain cubic equations.

Summary In this section, we made three claims about embeddings $T \hookrightarrow T'$ of theory T into theory T':

(C1) The embedding may yield solutions in T' for problems unsolvable in T.
(C2) Embeddings are ubiquitous in mathematics.
(C3) Embeddings may cause (unjustified) feelings of "abstractness."

We see the ubiquity of embeddings, i.e., (C2), as a consequence of their usefulness, i.e., (C1). I take both claims to state obvious facts that do not require argument. We nonetheless looked at one concrete example, the embedding of geometry into algebra: $G \hookrightarrow^{+} A$, and referred to what the MCS reflects about the entire discipline in order to illustrate these issues. My critical claim was the third,

[12] See https://mathscinet.ams.org/msc/msc2010.html. While not every mention of "algebraic," "topological," or "category/-ical" in the MSC (to name some of the usual suspects) indicates a case of embedding, their ubiquity still corroborates our point.

[13] See (Klein, 1895, p. 3): "This new method of attack is rendered necessary because elementary geometry possesses no general method, no *algorithm*, as do the last two sciences [sc. algebra and higher analysis]" (emphasis in the original).

[14] The closure property was striking enough for Dedekind at the time to make it—"so closed and complete in itself" (*in sich so abgeschlossen und vollständig*)—the feature that defines a field; see (Dedekind, 1871, p. 424).

i.e., (C3): embeddings may produce an initial feeling of alienation that often gets misattributed to the other theory as being either abstract or more abstract. We argued that this is a misattribution based on what we summarized as context and custom, namely, the observation that when we embed a theory T, formulated in language $\mathscr{L}$, into a theory T', formulated in different language $\mathscr{L}'$, then this may indeed come with a feeling of "abstractness" but only for those fluent in the language of T but not (yet) proficient in the language of T'. Practice makes this feeling go away as soon as sufficient fluency in $\mathscr{L}'$ has been achieved. And while abstraction may be at work during the embedding, it is not of the stick-figure variety if it is. Thus, our section title should be amended from "abstraction by embedding" to "misattributed feelings of abstractness by embedding."

We now turn to the question of where the more powerful target theories T' and their resourcefulness come from. We could continue our example from above and study the origins of the various representations of what a straight line (segment) is in various contexts. But we better put our arguments on a broader historical basis.

4 Abstraction by Constraint-Based Design

Analysis is arguably one of the pillars on which modern mathematics rests. To wit, Dieudonné (1960, p. ix) remarked:

> All branches of modern mathematics involv[e...] "analysis" (which in fact means [it is] everywhere, with the possible exception of logic and pure algebra).

Asked what has informed and shaped the historical development of modern analysis the most, two candidates are most likely to be singled out for their towering influence: Fourier Analysis and complex analysis.[15] In this section, we conduct a case study that is connected to the development of Fourier Analysis.

4.1 Fourier Analysis and a Challenge It Posed

As is known, Joseph Fourier (1768–1830) claimed that on a closed interval an arbitrary function $f(x)$ can be represented by an infinite sum of sine and cosine functions with appropriate coefficients a_n and b_n:

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx). \tag{6}$$

[15] See, e.g., Van Vleck (1914), Gonzalez-Velasco (1992), or (Hawkins, 1979, chs 1–2) for Fourier Analysis and Bottazini and Gray (2013) for complex analysis.

His main argument was that Eq. (6) must be correct since its coefficients can be calculated:

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\, dx, \text{ for } n = 0, 1, 2, 3, \ldots \tag{7a}$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\, dx, \text{ for } n = 1, 2, 3, \ldots \tag{7b}$$

provided we interpret the integrals in (7a) and (7b) as areas under the graph. (We here make the usual simplification that $f : [-\pi, \pi] \rightarrow \mathbb{R}$.) Fourier, Carothers noted, thereby "transformed the question of existence of the series representation into the geometrically 'obvious' fact that the area under a curve can always be computed" (Carothers, 2000, p. 141). It turned out that Fourier's claim raised many challenging questions about concepts that lie at the very foundation of analysis, such as convergence, continuity, differentiability, and integrability as well as the very notions of function and integral themselves.

One such question was whether the Fourier representation is unique in the following sense, first formulated by Heine in 1870 (Heine, 1870). Given a function $f(x)$ and a corresponding Fourier sum:

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx) = 0,$$

which converges for (almost) all values of x, will all its coefficients a_n, b_n be zero? Heine could show that this is true provided we assume uniform convergence for all x except for finitely many points where the series does not converge or converges to a value $\neq 0$, so-called critical points. Subsequently, in papers published 1870–1872, Cantor improved Heine's result by showing, first, that the assumption of uniform convergence can be dispensed with and, second, that the result still holds when we assume that there are countably many critical points. In order to prove the second result, Cantor had to establish certain facts about infinite sets (e.g., about critical points in the interval $[-\pi, \pi]$), and this raised a question about one infinite set in particular: the set of all real numbers and how to define it.[16]

The attention that was given to critical points was related to a discussion that took place in integration theory about points of discontinuity and whether they form dense or nowhere dense sets. This latter discussion had been initiated by Dirichlet when he investigated under what conditions the Fourier series of a function converges, and it eventually led to insights about continuity, integrability, and differentiability that were not even conceivable to earlier generations. To clarify,

[16] See Cantor (1932, chs II.1–5) for his work on Fourier series and Cantor (1872, pp. 92–97) for his theory of the real numbers. We should mention that it were these questions that led him to the development of set theory.

the most rigorous formulations at Fourier's time had been given by Augustin-Louis Cauchy (1889–1857; see Grabiner, 1981). He championed d'Alembert's language of limits and defined continuity similar to what we still do it in our calculus classes (i.e., in terms of the limit $\lim_{h\to 0} f(x+h) - f(x)$ (Cauchy, 1821, p. 43)) and by making this more precise by employing arbitrarily small numbers ϵ and δ ([*d*]*ésignons par* δ, ϵ *deux nombres très petits* (Cauchy, 1823, p. 44)). He still assumed, though, that continuity and differentiability would go hand-in-hand and did not distinguish between point-wise and uniform convergence. It is therefore instructive to see, in his proof of the intermediate-value theorem—which can be said to capture what it means for a function f to be continuous on the closed interval $[a, b]$—how he relied on geometric intuitions for both the existence and the identity of limits. His argument for the special case $c = 0$ (sometimes called Bolzano's Theorem) was as follows (see Cauchy, 1821, pp. 378ff. (= Note III)):

> Let $f : \mathbb{R} \to \mathbb{R}$ and $[a, b]$ be an interval such that a sign change occurs between $f(a)$ and $f(b)$. Divide the interval $[a, b]$ into m equal parts and choose $i < j < m$ such that the sign change occurs in $[a_1, b_1]$, for $a_1 = \frac{a}{i \cdot m}$ and $b_1 = \frac{b}{j \cdot m}$. Lather, rinse, repeat to construct two sequences (a_n) and (b_n) such that the sign change occurs in $[a_k, b_k]$, for all $k \leq n$. Since (a_n) is bounded and increasing while (b_n) is bounded and decreasing there must be a limit x they both converge to: $\lim_n a_n = x = \lim_n b_n$. Continuity of f on $[a, b]$ and construction of the sequences (a_n), (b_n) then entail: $f(x) = 0$.

As traditional ideas about continuity, integrability, and differentiability started to fall apart in the wake of Fourier, more robust ideas were needed (see Hawkins, 1979). This led, among many others, to the question of whether the existence and uniqueness of certain limits can be secured by a more rigorous definition of the real numbers (i.e., one where we avoid undue reliance on geometric intuitions).

We thus find that the problem of how to define the real numbers came up twice in the context of Fourier Analysis. It is the question we turn to now. And while we need to produce a historical narrative, this is not our primary goal. We plan to engage with the topic not as historians but rather to work as lab technicians: we take a familiar sample (in our case from the history of analysis) but stain it differently. And the stain we apply is meant to bring out new or hitherto ignored features. In our case, these are features of the epistemic practices mathematicians engage in when they introduce new concepts.[17]

4.2 How to Define the Real Numbers

We can formulate the problem as follows: Can we put the point set of a geometric line in a one–one correspondence with a number set yet to be determined (i.e., the set

[17] One referee wondered whether the case study is too long. I do not believe it is, we simply exercise due diligence. Making a historical argument requires that we look at all the evidence available to make it less likely that our claims are the result of biased readings of selected sources.

of numbers we now call the real numbers and denote with $\mathbb{R}$)? If so, then this would secure that the limits Cauchy had assumed to exist based on geometric evidence do in fact exist as numbers. To frame the question this way, as a correspondence between a point set and a number set, is how mathematicians at the time conceived of the problem. It was explicitly mentioned in the first textbook written in the new spirit of increased rigor, Dini's *Fondamenti* (1878), and Du Bois-Reymond perceived it as so problematic that he made it a recurring topic throughout the entire *General Function Theory* (1882). Years later, in 1907, when Hobson caught up the English mathematical community on what has happened on the Continent, he still found it worthwhile to devote a couple of sections to the topic.[18]

Today we normally define the completeness of the real numbers by requiring that every non-empty subset bounded above has a supremum. Historical accounts, in contrast, secured completeness by providing a definition of the irrational numbers, thereby "filling the gaps" among the rational numbers. We look at such three such proposals made by Heine, Cantor, and Dedekind, all in 1872, and with an eye at how they arrived at their notion of irrational number and whether stick-figure abstraction played a role in it.[19]

---

[18] See (Dini, 1878, §İ0), (Du Bois-Reymond, 1882, *passim*), and (Hobson, 1907, §§ 40–44), respectively.

[19] For the record we have to mention three more proposals that were made at the time. We mostly ignore them, however, and do so for two reasons. First, they offer little else than their definition; second, they had no noticeable impact at the time. This renders them mostly useless for our goal of learning about epistemic practices.

1. Charles Méray (1835–1911), a French mathematician in Dijon, was the first to publish what later became known as Cantor's definition of the reals (see Méray, 1870). So what we call Cantor's definition should really be labeled Méray-Cantor. Although he subsequently used his definition in two textbooks (Méray, 1872; Méray, 1894), it went unnoticed at the time. It seems it was only three decades later that Jules Molk drew attention to it when he translated and heavily revised Pringsheim's article (Pringsheim, 1898) on irrational numbers and convergence for the French version of Klein's *Enzyklopädie der mathematischen Wissenschaften*; see (Pringsheim and Molk, 1904, § 6). In the same year, 1904, Tannery published a revised edition of his textbook and also acknowledged the priority of Méray (Tannery, 1904, p. vii); but he did so only after Méray had called him out for the oversight (Méray, 1894, p. xxiii). While Méray waxed poetic when it came to criticizing the sad state of analysis in France of the time, he never said much about his definition beyond stating it. When and how Méray's priority became known more widely, I have no tried to determine; but see Dugac (1970) for some general background.

2. Weierstrass never published the theory of irrational numbers he presented in his lecture course on analytic function theory that he started to offer on a fairly regular basis in the fall of 1861 (see Weierstrass, 1903, pp. 355–360). Thus, lacking *ipsissima verba*, scholarly caution prevents us from including his theory to our discussion below. For later references, however, we need to know what transpired at the time. We have a somewhat incoherent account by Kossak (1872) based on lectures given in the winter term 1865–1866 (incidentally, this is a lecture that Cantor, too, might have attended) and a thorough reconstruction by Pincherle (1880) based on lectures he attended in the academic year 1877–1878. While both are mentioned by Biermann (1887), his reconstruction of Weierstrass' theory seems to owe Pincherle everything. We should mention that, in private correspondence, Weierstrass criticized Kossak for having "butchered" (*verhunzt*) his theory (Mittag-Leffler, 1910, p. 12), while his student Schwarz called it "defaced" (*entstellt*; Dugac, 1973, p. 144). Biermann's account met with even more anger; Weierstrass would have

All authors (including Méray and Weierstrass) employed the same general strategy, so we will not make specific mention of it below. They showed (1) that the law of trichotomy: "either $p > q$, or $p = q$, or $p < q$," holds for their newly defined irrational numbers and that (2) the new numbers are closed under the usual laws of addition, multiplication, and their inverses. This fact, viz., that the new numbers behave in all relevant aspects like the more familiar numbers (i.e., natural numbers, integers, rational numbers) and are therefore indistinguishable in the language of arithmetic, was the main ("structuralist") argument they made for calling them "numbers" and declaring them rightful denizens of the mathematical world.[20]

4.2.1 Heine

Eduard Heine (1821–1881) of Heine–Borel fame was a student of Dirichlet and a senior colleague of Cantor at the University of Halle. In 1872, he published the paper "The elements of the theory of functions" (*Die Elemente der Functionenlehre*) to provide the analysis community of the time with a systematic exposition (*im Zusammenhange entwickelt*) of results on continuous functions that Weierstrass had obtained but which were circulating by word of mouth or in handwritten notes without any official stamp of approval. This systematic exposition forms Part B

intervened with the publisher had he known what was brewing (Dugac, 1973, pp. 142f.). It seems that Dantscher (1908), who published his version as late as 1908—dating back to a lecture Weierstrass gave in 1872!—was given almost no attention at the time. Due in particular to Dugac (1973), and later supplemented by Ullrich (1989), we now have more reliable information, which is sufficient to vet as accurate the little that we need on Weierstrass below.

Weierstrass was the first to use sequences; or, to be more precise, sequences of convergent partial sums (i.e., series). He defined an irrational number as an infinite series of positive and negative rational numbers (q_n) such that the partial sum of every finite subseries (q_{n_k}) is bounded by a number b: $\sum_{i=1}^{k} q_{n_i} < b$. Each q_n was a unit fraction, i.e., the fraction of a chosen base unit, say, a: $q_n = 1/a$, and b a multiple of that base unit: $b = m \cdot a$; to show convergence of the sums, Weierstrass employed the notion of absolute convergence.

3. Giulio Ascoli (1843–1896) proposed to use what we now call the Nested Interval Theorem. Let $\{I_n\}_{n\in\omega}$, with $I_n = [a_n, b_n]$, be a sequence of closed intervals $I_1 \supseteq I_2 \supseteq \ldots \supseteq I_n \supseteq \ldots$, and assume $\lim_n I_n = 0$. Then there is exactly one point p such that: $p \in \bigcap_{n\in\omega} I_n$ and thereby defined by $\{I_n\}_{n\in\omega}$. (See Ascoli, 1895, p. 1065; we here interpret Ascoli's "∞" to mean "ω.") While his approach is now one of the usual suspects for defining the completeness of $\mathbb{R}$, it does not seem to have garnered much interest at the time.

[20] The last step was even more pronounced in Méray (in his book from 1894), Weierstrass (lectures notes taken by Hurwitz in 1878, see (Dugac, 1973, App. I)), and later in Pringsheim (1916): they first established that the respective properties and laws hold for sequences that define rational numbers; then, everything else being equal, what other than prejudice could be the reason for discarding the remaining sequences that define irrational numbers?

of his paper, while in Part A he puts it all on a firm basis by defining what the irrational numbers are.[21]

Heine started out his paper with the observation that, on one hand, the definition of irrational numbers has not yet been fixed (*der nicht völlig feststehenden Definition der irrationalen Zahlen*), while, on the other hand, geometric intuitions such as lines being generated by movement (*die Erzeugung einer Linie durch Bewegung*) have muddied the waters (*Vorstellungen der Geometrie [die...] oft verwirrend eingewirkt haben*). His goal therefore was to avoid all intuitive hunches and to achieve completeness of the real numbers by, first, introducing new number signs for which, second, the usual arithmetical operations are well-defined in such a way that, third, a function $f(x)$ in one variable is well-defined for any choice of x, whether x is rational or irrational (*so dass eine einwerthige Function für jeden einzelnen Werth der Veränderlichen, sei er rational oder irrational, gleichfalls einen bestimmten Werth besitzt*).[22] His new number signs denote sequences of rational numbers (q_n) whose differences become eventually arbitrarily small (*ibid.*, p. 174):

$$(q_1, q_2, q_3, \ldots, q_n, \ldots) \text{ s.t. } \forall \eta > 0\, \exists n_0 \in \mathbb{N}\colon q_{n_0} - q_{n_0+\nu} < \eta, \text{ for all } \nu \in \mathbb{N}.$$

Heine is not willing, however, to make much of an ontological commitment. The new entities can be called numbers since order, and hence equality can be defined for them and since they allow a seamless extension of the four basic arithmetical operations. That is as far he willing to go. And even though he defines what the limit is (*ibid.*, p. 178):

$$\mathfrak{A} \text{ is the limit of } (q_n) \text{ iff } \lim_{n\to\infty} \mathfrak{A} - q_n = 0,$$

he is not willing to say it does exist.

> I do not answer the question what a number is by giving a conceptual definition, or even by introducing the irrational numbers as limits—whose existence was an assumption. I take the purely formal stance by calling certain concrete signs numbers such that the existence of these numbers is never in question.[23]

---

[21] It is not entirely clear from the text itself how much exactly Heine owes to Weierstrass or to Cantor, respectively. Heine himself stated that his only contribution was to bring oral communications of Weierstrass into a publishable, organized form and that discussions with Cantor had left their mark on how he presented the part on irrational numbers (*bedeutenden Einfluss auf die Gestaltung*), a part, he noted, that had been finished some time ago (*seit längerer Zeit vollendet*). Furthermore, he remarked that Cantor's definition was a felicitous improvement of Weierstrassian ideas he had used earlier (*glückliche Fortbildung der ursprünglichen Einführungsart*; p. 173). So it might seem that Heine had replaced Weierstrassian series with Cantorian sequences (modulo some adjustments) in an otherwise completed part of the manuscript. This is not, however, how Cantor remembered it after Heine's death; he claimed (Cantor, 1887, p. 385) that Heine's entire approach was due to him.

[22] All quotations are from (Heine, 1872, p. 172).

[23] "Die Frage, was eine Zahl sei, beantworte ich [...] nicht dadurch[,] dass ich die Zahl begrifflich definire, die irrationalen etwa gar als Grenze einführe, deren Existenz eine Voraussetzung wäre. Ich

The caution Heine displayed clearly shows how unfamiliar the whole enterprise still was; he was not willing to take any backlash because he had exposed himself too much.

4.2.2 Cantor

When Georg Cantor (1845–1918), a student of Kronecker and Weierstrass, came to Halle, Heine drew his attention to Fourier analysis, and Cantor soon proved results that had eluded all his precursors (see Sect. 4.1 above). But in order to do so, he had to establish facts about infinite sets and among them what the real numbers are.

He defined real numbers in terms of what he called "fundamental sequences" of rational numbers (a_n) that have a limit, and a sequence has a limit if eventually all its members are in an ϵ-neighborhood:

$$\forall \epsilon \in \mathbb{Q}^+ \, \exists n_0 \in \mathbb{N} \colon |a_{n+m} - a_n| < \epsilon, \text{for all } n, m \in \mathbb{N} \text{ and } n \geq n_0$$

(*ibid.*, p. 93; here $\mathbb{Q}^+ = \{q \in \mathbb{Q} : q > 0\}$). For the sake of conceptual clarity, he distinguished (like Heine did) between sequences (a_n) of first order, which only have rational numbers as members, and sequences of n-th order, which may have members of order $n - 1$. Both Heine and Cantor proved that a number defined by a sequence of arbitrary order (in common parlance: a Cauchy sequence) can already be defined by a first-order sequence.

The interesting point is how Cantor thought about these fundamental sequences (see *ibid.*, pp. 96f.). He first makes the observation that for every point p on the number line, identified by measurement or construction, we can find a corresponding fundamental sequence (a_n) such that:

$$d(0, p) = \lim_{n\to\infty} a_n$$

where "$d(x, y)$" denotes the chosen distance function. The converse claim, however, viz., that for every fundamental sequence (a_n) there is a point p on the number line such that: $\lim_n a_n = d(0, p)$, he called an axiom "because it is in its nature not to be generally provable" (*weil es in seiner Natur liegt, nicht allgemein beweisbar zu sein*). A little later he elaborated: "According to this conception [... point sets on the number line] form a genus among all conceivable point sets" (*denkbare Punktmengen*; *ibid.*, p. 98). Cantor did not name the two-way correspondence between points and numbers; for later reference we do.

Correspondence Principle *For each point p on the number line ℓ, there is a fundamental sequence (a_n), such that $d(0, p) = \lim_n a_n$.*

---

stelle mich [...] auf den rein formalen Standpunkt, indem ich gewisse greifbare Zeichen Zahlen nenne, so dass die Existenz dieser Zahlen also nicht in Frage steht." ((Heine, 1872, p. 173))

Correspondence Axiom *For each fundamental sequence* (a_n), *there is a point* p *on the number line* ℓ, *such that* $\lim_n a_n = d(0, p)$.[24]

4.2.3 Dedekind

Richard Dedekind (1831–1916) published his account of the real numbers in 1872 in direct response to Heine's publication but had developed his ideas already in 1858 when he had to offer calculus but felt that as a theory it was lacking a truly scientific foundation (*wirklich wissenschaftliche Begründung;* Dedekind, 1872, p. 9).

Dedekind first remarked that he had observed that the existence of suprema is sufficient to replace any appeal to geometric intuitions in proofs.[25] Thus, his starting

[24] We should not keep quiet about a certain ambiguity in the original, German text; it concerns the word *Reihe* that I translated as "sequence" although today it means "series." This, and the observation that Cantor was a student of Weierstrass (who had employed series (see footnote 20), prompted some to understand Cantor's definition in terms of convergent series (see, e.g., Spalt, 1991, p. 357). And indeed, either reading was adopted by mathematicians during Cantor's lifetime and with reference to him: Stolz used series (Stolz, 1885, pp. iv, 97ff.), Bachmann sequences (Bachmann 1892, pp. iv, 4–9). To the present author, however, it is clear that Cantor meant sequences. This is how he used the term *Reihe* in his work on Fourier analysis when he distinguished between a sequence (*Reihe*) and its sum (*Summe*; e.g., (Cantor, 1932, p. 73); later he used the compound word *Reihensumme* for series (*ibid.*, pp. 105, 131, 180, 190) and reserved the word *Folge* (the present-day word for sequence) for transfinite well-orderings (*ibid.*, p. 147, 444). Moreover, Cantor not only portrayed Weierstrass' definition as very different from his own, but also said this was identical to what Lipschitz had done later (Cantor, 1883, pp. 184ff.); and Lipschitz used sequences, not series (Lipschitz, 1877, § 15).

That Cantor would point to Lipschitz is not without irony, though. We know from the correspondence between Lipschitz and Dedekind that Lipschitz doubted that introducing cuts or proving completeness were even called for. He believed we can find it all in Euclid and the "fundamental property of a line without which no one can conceive a line" (*Grundeigenschaft einer Linie [. . .], ohne die kein Mensch sich eine Linie vorstellen kann*; *ibid.*, p. 475). Accordingly, while Lipschitz freely mentioned predecessors in his book, he makes no references at all to Cantor or Dedekind when he defines the irrational numbers but rather maintains that "we owe the [Ancient] *Greeks* the study of such sequences of fractions that converge towards a limit. They elevated the limit associated with such sequences [. . .] to a concept in its own right and realized how to extend the [basic] operations [of arithmetic] to these concepts" (*Die Betrachtung solcher Folgen von Brüchen die sich einer Grenze nähern, verdanken wir den* Griechen. *Sie haben die in einer solchen Folge von Brüchen zugehörige Grenze zu einem selbstständigen Begriff erhoben, und erkannt, wie die Operationen [. . .] auf diese Begriffe auszudehnen seien*, Lipschitz, 1877, pp. 36f.). Thus, we must assume that all arguments by Dedekind to the contrary were wasted on him and that Lipschitz believed he can do without Dedekind or Cantor, while at the same time Cantor saw him as executing his own program.

[25] It is not entirely clear from the text itself what exactly Dedekind had in mind (not that it would matter, both formulations are equivalent, see (Propp, 2013, pp. 395f.), but we should still address it). The most straightforward reading suggests: every bounded monotone increasing sequence converges toward a limit. But his proof in § 7 suggests reading it as the existence of suprema. Here is an interpolated version of what he wrote: "Every magnitude which grows continually but not beyond all limits [= every set of numbers bounded above], must certainly approach a limiting value [= converges towards a limit or has a least upper bound]." (*Jede Größe, welche beständig, aber*

point was extensive experience with proofs. Based on this experience he was able to pinpoint a requirement that had promise to serve as a fruitful axiom. He did not choose the existence of suprema as his axiom, though, for he believed this to be a fairly abstract idea and rather sought something that was closer to a basic arithmetical truth and which, at the same time, reflects how the continuity of the line goes beyond the density of the rational numbers on it. He found this truth in the following insight which he believed to be evident.

Cut Principle *If p is a given point on a line ℓ, then p divides ℓ into two disjoint classes A and B such that all points of A are to one side of all points of B.*

The converse proposition, however, which is equivalent to the existence of suprema, he did not consider to be a given but something that must be axiomatically required.

Cut Axiom *If all points of a line ℓ fall into two disjoint classes A and B, such that all points of A are to one side of all points of B, then there exists a unique point p that produces this division.*

The principles and axioms given by Cantor and Dedekind are clearly identical twins. And just like Cantor before him, Dedekind states that the converse direction has to be introduced as an axiom "for I am utterly unable to adduce any proof of its correctness, nor has any one the power."[26]

4.3 Discussion

Dedekind presented his thoughts in a 30-page pamphlet, while Heine and Cantor had to limit themselves to brief remarks in journal articles. We therefore follow mostly Dedekind and supplement his observations by pronouncements made by Heine and Cantor.

A goal that was common to all three authors was to establish the existence of irrational numbers (e.g., as limits or roots) without undue appeal to geometric ideas. Heine dismissed, as we saw above, geometrical ideas lock, stock, and barrel since they "muddy the waters" (*verwirrend eingewirkt haben*). Dedekind concurred:

> Obviously, nothing is gained by vague remarks upon the unbroken connection in the smallest parts; the problem is to indicate a precise characteristic of continuity that can serve as the basis for valid deductions.[27]

nicht über alle Grenzen wächst, [muss] sich gewiß einem Grenzwerth nähern. (Dedekind, 1872, pp. 9, 29))

[26] "[D]enn ich bin außer Stande, irgend einen Beweis für seine Richtigkeit beizubringen, und Niemand ist dazu im Stande. Die Annahme dieser Eigenschaft der Linie ist nichts anderes als ein Axiom ..." (*ibid.*, p. 18)

[27] "Mit vagen Reden über den ununterbrochenen Zusammenhang in den kleinsten Theilen ist natürlich Nichts erreicht; es kommt darauf an, ein präcises Merkmal der Stetigkeit anzugeben, welches als Basis für wirkliche Deductionen gebraucht werden kann." (*ibid.*)

But Dedekind and Cantor, as we have seen, were willing to take ideas about continuity of the line into consideration. And they arrived at almost identical formulations, namely, what we called the Correspondence Principle and the Cut Principle, respectively. Both principles are not abstracted, stick figure or otherwise, from spatial intuition but express facts about measuring or constructing line segments. Their respective converse statements, however, the Correspondence Axiom and the Cut Axion, do not express intuitive facts. There is nothing in intuition that would allow us to distinguish the apparent continuity of the rational numbers between zero and one from actual continuity (i.e., the completeness). For the only intuitive "fact" that could be summoned is infinite divisibility; but in this respect $\mathbb{Q}$ and $\mathbb{R}$ are the same: for any two numbers a and b, with $a < b$, there is a third number c that is in between: $a < c < b$. This is what I think Dedekind meant when he wrote:

> If space has real existence at all, it is not necessary for it to be continuous; innumerable of its properties would remain the same even if it were discontinuous.[28]

But if the continuity of the reals, their completeness, respectively, is not a fact of intuition but a requirement made by the needs of analysis, then this means neither of the two axioms that define it can be obtained by abstraction; that is, not under any definition of abstraction that maintains that to obtain some X by abstraction requires X to be already present somehow. This is why Cantor and Dedekind claimed that continuity cannot be proved but must be mandated, introduced as an axiom; to repeat the previous quotations:

> It is in its nature not to be generally provable. (Cantor)
> For I am utterly unable to adduce any proof of its correctness, nor has any one the power. The assumption of this property of the line is nothing else but an axiom. (Dedekind)

Heine, as we saw above, dodged the issue altogether by avoiding existence axioms and retreating to a noncommittal formal point of view instead.[29]

Dedekind is very clear where the continuity or completeness of the reals comes from if not from intuition. The completeness of the real numbers, the continuity of the geometric or the number line, respectively, is our own creation; it is us who impose it on the number line, we force it to be true whether actual space happens to be continuous or not.

> [It is] an axiom by which we attribute to the line its continuity in the first place, it is us who add the idea that the line has this property.

---

[28] "Hat überhaupt der Raum eine reale Existenz, $_{18}|^{19}$ so braucht er doch nicht nothwendig stetig zu sein; unzähliger seiner Eigenschaften würden dieselben bleiben, wenn er auch unstetig wäre." (*ibid.*, pp. 18f.)

[29] Méray displayed similar reservations when he called sequences with rational limits "effective" and those with irrational limits "ideal." (Méray, 1894, § 50) For the sake of a convenience, however, he wanted to keep the language simple and assigned those ideal limits, although they do not really qualify, number signs that were previously reserved for actually existing quantities (*un nombre ou une quantité* incommensurable, *et qu'on représente par le même signe que si elle existait réellement*).

> And if we knew for certain that space was discontinuous, there would be nothing to prevent us, in case we so desired, from filling up its gaps and thus making it continuous; this filling up, however, would consists in a creation of new point-individuals and were to be executed according to the above principle [i.e., the cut axiom].[30]

Or, rephrased in the language of free creations of the mind which became Dedekind's hallmark:

> It becomes absolutely necessary that the instrument R [i.e., $\mathbb{Q}$] constructed by the creation of the rational numbers be essentially improved by the creation of new numbers such that the domain of numbers shall gain the same completeness, or as we may say at once, the same *continuity*, as the straight line.[31]

Let me stress this point. Did Dedekind obtain his cut principle by abstraction? He is quite clear that it is the opposite. If any line we ever encountered or conceived had the order type of the rationals instead of the order type of reals, we would not have noticed. For there is no resolution high enough, not for the eye nor for the mind's eye, to tell the set $[0, 1] \subset \mathbb{Q}$ apart from its counterpart $[0, 1] \subset \mathbb{R}$. So the cut principle cannot be obtained by abstraction. For it is not a property of the line that makes the principle true. It is us who impose that finer-grained continuity on it:

> The assumption of this property of the line is nothing else but an axiom by which we attribute to the line its continuity in the first place, it is us who add the idea that the line has this property.[32]

Heinrich Weber (1843–1913) came to the same conclusion:

> Continuity as well as density are properties whose very nature makes it inaccessible to our sense perception [. . .] as much as it seems rooted in the nature of our intuition. It is possible, however, to construct systems of pure concepts . . . that are accorded density and continuity.[33]

[30] "ein Axiom, durch welches wir erst der Linie ihre Stetigkeit zuerkennen, durch welches wir die Stetigkeit in die Linie hineindenken. [. . . $_{18}|^{19}$ . . .] Und wüßten wir gewiß, daß der Raum unstetig wäre, so könnnte uns doch wieder Nichts hindern, falls es uns beliebte, ihn durch Ausfüllung seiner Lücken in Gedanken zu einem stetigen zu machen; diese Ausfüllung würde aber in einer Schöpfung von neuen Punct-Individuen bestehen und dem obigen Principe gemäß auszuführen sein" (Dedekind, 1872, pp. 18f.)

[31] "es wird daher unumgänglich nothwendig, das Instrument R, welches durch die Schöpfung der rationalen Zahlen construiert war, wesentlich zu verfeinern durch eine Schöpfung von neuen Zahlen der Art, daß das Gebiet der Zahlen dieselbe *Stetigkeit* gewinnt, wie die gerade Linie." (*ibid.*, p. 16; emphasis in the original)

[32] "Die Annahme dieser Eigenschaft der Linie ist nichts als ein Axiom, durch welches wir erst der Linie ihre Stetigkeit zuerkennen, durch welches wir die Stetigkeit in die Linie hineindenken." (*ibid.*, p. 18)

[33] "Stetigkeit sowohl als Dichtigkeit sind Eigenschaften, die der Natur der Sache nach unserer Sinneswahrnehmung unzugänglich sind [. . .] wie sehr sie uns auch im Wesen unserer Anschauung zu liegen scheinen. Es lassen sich aber wohl reine Begriffssysteme konstruieren, denen . . . Dichtigkeit und Stetigkeit zukommen." (Weber, 1895, p. 5) See also *ibid.*, p. 9, where Weber states that measurability (and, consequently, the Archimedean Axiom) is not a given fact but is imposed by the thinking observer (*durch den denkenden Beobachter hineingelegt*). Some may want to object that Weber, as a close friend of Dedekind, cannot count as an independent source.

Some may want to object that we have geometric constructions (e.g., $\sqrt{2}$ as the diagonal of the unit square) from which we infer the existence of irrational numbers and therefore could attempt to abstract a general concept. Clearly, as we have seen above, these constructions informed how mathematicians of the time thought about the subject matter. To wit, Dedekind wrote:

> It may generally be conceded that the proximate reason for extending the concept of number were such tie-ins with non-arithmetical [sc. geometrical] ideas; but this is certainly no reason to incorporate these alien considerations into arithmetic itself.[34]

And working from geometry cannot be successful; for at least three reasons. First, it is not clear how geometric constructions can account for transcendental numbers that at the time were known to exist (in 1768 Lambert conjectured and Liouville proved their existence in 1844, 1851 resp.). Second, we have no purely geometric notion of closure under construction that we could use to demonstrate completeness. Third, the irrationality of $\sqrt{2}$ is not a fact of or given in geometric intuition. Rather, it is an artifact of the mathematician's attempt to use numbers in geometry by introducing a unit length.[35] In other words, the original sin was to impose numbers on the line in the first place; and it can only be redeemed by imposing more numbers on the line.

4.4 *Epistemic Practices*

In the previous section, we argued that and why stick-figure abstraction cannot be assigned any conceivable role in the formation of the new concept of irrational number and, consequently, the concept of real number. What we have rather seen is an example of what I call constraint-based design.

4.4.1 Constraint-Based Design

Mathematicians had devised a mathematical theory, calculus, to model the behavior of bodies and forces (i.e., Newtonian mechanics). Upon closer inspection, the

---

I would disagree; Weber was an eminent mathematician and therefore had to have an independent mind.

[34] "Daß solche Anknüpfungen an nicht arithmetische Vorstellungen die nächste Veranlassung zur Erweiterung des Zahlbegriffs gegeben haben, mag im Allgemeinen zugegeben werden [. . .]; aber hierin liegt gewiß kein Grund, diese fremdartigen Betrachtungen selbst in die Arithmetik [. . .] aufzunehmen" (Dedekind, 1872, p. 17).

[35] Russell made a similar remark: "In the past, the definition of irrationals was commonly effected by geometrical considerations. This procedure was, however, highly illogical; for if the application of numbers to space is to yield anything but tautologies, the numbers applied must be independently defined." (Russell, 1903, p. 278)

language and the techniques of that theory were found to be insufficient; in this case, insufficiently rigorous both for itself and for answering follow-up questions it had triggered (e.g., Fourier). After exploring the issue for a while from various angles, mathematicians had a good understanding of what a good fix would look like. For example, any proposal should:

(1) Be rigorous (e.g., by employing algebraic methods)
(2) Be consistent with the existing body of mathematical knowledge (e.g., Hankel's principle of permanence; Hankel, 1867, § 3)
(3) Be well-operationalized to be fit for deployment in proofs (see Sect. 3 above)
(4) Deliver what is needed (e.g., existence of limits or roots)
. . . etc.

I call requirements such as (1) through (4) constraints on the design space for a solution. Finding a solution is similar to finding a missing piece while completing a puzzle; or like furnishing a room with limited means given its blueprint; or engineering a device based on certain performance specifications. Especially the latter analogy emphasizes something we will dwell on later: solutions may be custom-tailored for or reverse-engineered from a problem.[36]

We now show that the historical record proves that what we proposed is not a fictitious analysis but that it captures exactly how the problem was conceived by those involved. For a start, we see constraints (1), (2), and (4) being acknowledged by the fact that they all—Méray, Weierstrass, Heine, Cantor, Dedekind—proposed a strictly arithmetical account and that they all employed the same basic three-step strategy. Step 1: show that equality and order can be defined as well as closure under the four basic operations; Step 2: infer that if the new entities behave like numbers, they must be numbers; Step 3: prove completeness (often skipped or only hinted at).

Operationalization Dedekind was very clear about (3), the proper operationalization of concepts (see footnote 27):

> Obviously, nothing is gained by vague remarks upon the unbroken connection in the smallest parts; the problem is to indicate a precise characteristic of continuity that can serve as the basis for valid deductions.

Cantor was more explicit on this (he had the most horses in the race, after all). To answer early critics of set theory, he devoted an entire section of his paper "Foundations of a general theory of aggregates" (1883) to his views on concept formation in mathematics. It is in this section that he coined his famous phrase that "the nature of mathematics is its freedom" (*denn das Wesen der Mathematik liegt gerade in ihrer Freiheit*; Cantor, 1883, p. 182). Mathematically speaking, he

[36] We refrain from calling it conceptual engineering for that would run counter to the deeply felt conviction among mathematicians that finding the right concept is not an engineering task but more like artistic production. If we agree that what is peculiar to solving a mathematical constraint-based design problem is the fact that it happens entirely in the realm of concepts, then "conceptual artist" (a phrase no longer at our disposal to use) would be more fitting than "conceptual engineer."

said, there are only concepts. Whether or not they denote objects—i.e., whether or not they have also "transcendental meaning" (*transiente Bedeutung*) beyond their mathematical meaning—was a question he left to metaphysicians to answer (*ibid.*, p. 207, footnote 7/8). But any mathematical concept, he continued, albeit freely created, comes with its own corrective force (*Korrektiv*): it will be dropped (sc. by the research community) if found "impractical" (*unbrauchbar*, i.e., unfit for the job), not "fruitful or inexpedient" (*unfruchtbar oder unzweckmässig*; *ibid.*). Applying those criteria, he arrived at the following conclusions.

> Dedekind cuts are practical, even superior to the use of sequences due to their one-one correspondence with the real numbers, but not expedient because they cannot be used 'as is' (they must first, "with great artistry and intricacy" (*mit großer Kunst and Umständlichkeit*), be brought into a form suitable for the purposes of analysis; *ibid.*, p. 185). Sequences are practical, although they require the detour via equivalence classes. The additional complication of forming and checking sums makes Weierstrass sequences less practical than Cantor sequences. The latter have the additional benefit of being applicable to transfinite numbers (*ibid.*, p. 190); but they really shine when it comes to expediency. Cantor sequences have the "simplest and most natural definition" which has the upside that it "most directly aligns itself with the analytic calculus" (*einfachste und natürlichste von allen, und man hat an ihr den Vorteil, dass sie sich dem analytischen Kalkül am unmittelbarsten anpasst*; *ibid.*, p. 184). Moreover, they feature "a remarkably flexible and at the same time intelligible language which yields a degree of clarity and transparency not to be underestimated" (*außerordentlich leichtflüffaßliche Sprache [...] nicht zu unterschätzender Gewinn an Klarheit und Durchsichtigkeit erzielt*; *ibid.*, p. 188)

Dedekind begged to differ; he believed his account to be "simpler" and to "bring out more precisely the quintessential point" (*meine Darstellung ... einfacher zu sein und den eigentlichen Kernpunkt präciser hervorzuheben*; Dedekind, 1872, p. 11). He admitted, though, that proving closure under arithmetical operations leads to atrocious ponderosity (*abschreckende Schwerfälligkeit*; *ibid.*, p. 28).

Gregorio Ricci (1853–1925) concurred: working with cuts provides the necessary clarity and simplicity (*chiarezza e semplicitá*). It is clearer since the cuts uniquely define a number; it is simpler since it does not require the machinery of sequences and limits (which is taking the second step before the first), and without such debts to other ideas, it is also more unified (*unitá*). All three are virtues of a good definition (Ricci, 1893, p. 235). In the revised paper, Ricci added the increase in rigor (*rigore*) when one wants to justify the rules for exponents, logarithms, and roots of polynomials (Ricci, 1897, p. 22).

Jules Tannery (1848–1910), comparing the definitions of Weierstrass, Mérat, and Dedekind, likewise saw the most merit in Dedekind's approach (Tannery, 1908, p. 102). He listed a total of five points that all deal with how "straightforwardly"—he employs *immédiatement* four times, while the fifth time a question does not even arise (*la question ne se pose pas*)—certain properties follow. Dedekind cuts immediately yield a unique characterization (*d'une façon univoque*) of (i) the new numbers as well as of (ii) equality and of (iii) order among them; moreover, cuts (iv) illuminate (*éclairer*) the numbers' nature and (v) make the distinction rational–irrational jump into the eye (*la distinction apparaît encore immédiatement*). Tannery also added a sociological observation. He explained the fact that most French mathematicians were taken with (*ait séduit*) Méray's definition by the observation that once your thinking has become rigid by habit (*la pensée cristallisée par l'habitude*, *ibid.*, p. 103), it takes efforts to assimilate something new (i.e., cuts) instead of running with a familiar idea (i.e., sequences and limits).

Observations about convenience in use are also echoed in Giuseppe Peano (1858–1932). Having discussed four completeness axioms, he find himself almost tied between the existence of suprema (Dedekind) and Cauchy sequences (Cantor). And while he makes constant use (*uso quasi costante*) of sequences, his preference is nevertheless for suprema: the fact that they get constantly used in analysis makes them the more convenient concept (*la forma più conveniente del principio che applica costantemente in analisi*; Peano, 1899, p. 139f.).

Free Creations Dedekind and Cantor agreed on the active design aspect. By Dedekind's lights, as we saw earlier, all numbers are "free creations." And, more specifically, he called the natural numbers "an extremely useful tool for the human mind" (*ein überaus nützliches Hülfsmittel für den menschlichen Geist*; *ibid.*, p. 17) and the rational numbers "a tool of infinitely greater perfection" (*ein Instrument von unendlich viel größerer Vollkommenheit*; *ibid.*, p. 12). *Homo faber* is what we are according to Dedekind.[37]

Cantor's account in contrast reads almost like an entry from a cookbook for concepts:

> In my opinion, the process of the correct formation of concepts is the same everywhere: one posits an object without properties; initially it is nothing but a name or a sign A, and in due order assigns it different even infinitely many intelligible predicates that must not contradict each other. Once one is completely done with it, the finished concept has been brought into being.[38]

Gottlob Frege (1848–1925) and Bertrand Russell (1872–1970) had strong reservations precisely about this point. Frege ridiculed the appeal to creative powers without specifying in what they consist and how they are limited; for limited they must be, otherwise I could go ahead and create solutions by *fiat* (see Frege, 1903, §§ 139f.). Russell assented and called out the practice of postulating existence (as done by the correspondence axioms and the cut axiom) for having "the advantages of theft over honest toil" (Russell, 1919, p. 71).

Additional Constraints There were more candidates for what constraints should be imposed on the design space. The most radical criticism came from (Frege, 1903, §§ 68–155) who contended that all proposals were untenable, and actually so deeply flawed as to put them beyond rescue, since they all violated logic: failure

[37] Nothing of this is as straightforward as it seems, though; see McCarty (1995).

[38] "Der Vorgang bei der korrekten Bildung von Begriffen ist m. E. überall derselbe: man setzt ein eigenschaftsloses Ding, das zuerst nichts anderes ist als ein Name oder ein Zeichen A, und gibt demselben ordnungsmäßig verschiedene, selbst unendlich viele verständliche Prädikate [...] die einander nicht widersprechen dürfen. [...] ist man hiermit vollständig zu Ende, so [...] tritt [der Begriff] fertig ins Dasein" (Cantor, 1883, p. 207, footnotes 7+8).

to distinguish between concept, meaning, and sign; failure to adhere to scientific standards of defining; failure to properly delimit the power of free creations, etc.[39] Equally radical was the critique of Leopold Kronecker (1823–1891) who denunciated all infinitary methods and consequently all contemporaneous attempts (he singled out Heine) to define what is irrational (*das "Irrationale" ganz allgemein zu fassen und zu begründen*).[40]

Paul du Bois-Reymond (1831–1889; Du Bois-Reymond, 1882, ch. 1) built on the traditional requirement (for which see Pringsheim, 1898, §§ 3, 7) that any concept of number must at least in principle provide a link to magnitude (or quantity) and its measurement, a link that was broken in the new definitions. Frege argued the same point.[41] That these two were not two lone voices in the wilderness is evinced by the fact that the editors of a journal as reputable as the *Mathematische Annalen* provided a forum for an author without mathematical training to launch an attack on the Weierstrass–Cantor definition of irrational numbers.[42]

What was at stake here is a preconceived notion of what a number is, and which therefore must be conserved through extensions of the number system.[43]

Russell made a similar objection (see Russell, 1903, chs xxxiii–xxxiv). But he vetoed the new proposals not based on tradition but based on what he deemed to be the only logically sound construction of irrational numbers, one that was due to

---

[39] See Tait (1996) for a measured account of Frege's criticism.

[40] See Kronecker (1886, p. 156, fn.) for the quote and Edwards (1995) for a rehabilitation of Kronecker.

[41] "What is the meaning of a number sign just by itself? Obviously, a ratio among quantities." (*Was bedeutet nun dabei ein Zahlzeichen allein [sc. without unit of measurement]? Offenbar ein Grössenverhältnis*; Frege, 1903, § 73).

[42] See Illigens (1889), Illigens (1890). Knowing they were walking on thin ice, the editors added a footnote saying that they wanted to stimulate the discussion of an important question from as many different viewpoints as possible without, however, endorsing the particular views expressed by the author. (*"[Wir] wünschen dazu beizutragen, dass die wichtige Frage der Irrationalzahlen unter möglichst vielseitigen Gesichtspunkten zur wissenschaftlichen Diskussion gelangt [... aber ohne] Verwantwortung für den besonderen Inhalt des betr. Aufsatzes [zu übernehmen]."* (*Math. Ann.* 33 (1889), p. 155, footnote)) For completeness' sake, we have to mention that Klein, as editor of the journal, got Cantor involved beforehand and asked for the latter's opinion (Cantor, 1991, p. 262); for more information on Cantor and Illigens, see (Tapp, 2005, chs 2.4, 11.5).

Since no one else is providing the information (cf. *ibid.*, footnotes 41 and 383 on pp. 73, 353, resp.), we do: Eberhard Heinrich Illigens (October 8, 1858–June 23, 1931) spent most of his life in the small town of Beckum, Westphalia, serving as the manager of the local savings bank his father had founded. Initially, he had studied Catholic theology, which had brought him into contact with a correspondent of Cantor, the Catholic philosopher Gutberlet. But Illigens could not finish his studies; his father's poor health required him to take over as manager of the bank. (The information was kindly provided by Illigens' grandson. Despite numerous attempts I was unable to confirm this information from independent historical sources except for his date of birth and death, graciously confirmed by Andrea Langner, Kreisarchiv Warendorf.)

[43] We should mention that this were not vain complaints of some backward-looking narrow-minded stick-in-the-muds; Du Bois-Reymond, e.g., helped to lay the foundation for progressive non-Archimedean ideas in algebra and analysis; see Ehrlich (2006), Ehrlich (2018).

Peano (who had replaced Dedekind cuts with a similar device he called segments (*segmento*); see Peano (1899)).

In particular the responses just summarized pose the challenge of what to make of constraints that are raised on philosophical grounds alone (such as the "true meaning" of number or "permissible" mathematical means, etc.). Shall one even try to draw a line when the boundary between what is mathematical and what philosophical is fuzzy?

Group Dynamics Looking back at the issues we broached in this section, we see ourselves confronted with the new problem of understanding group dynamics: which considerations tip the balance, and for what reason, when different proposals compete in the same design space?[44]

In regard to our present case study (the completeness of the reals), I would argue that the discussion did not run its natural course but was cut short. It is therefore not representative for how group dynamics normally play out; here is why. The discussion was not only muted by political events (two world wars), but, more importantly, it was sidelined by developments within mathematics itself: Hilbert's axiomatic program and the advent of new disciplines enabled by set theory were both game changers; and so was Bourbaki later. As a result, we postulate the real number system axiomatically as an ordered field and assign those older constructions by Dedekind or Méray–Cantor only a subsidiary role, if any, namely, the role of showing that the axioms have a model.[45]

[44] I plan to address this question more fully in another paper after additional case studies have provided more data.

[45] Our reason for saying the discussion was cut short is that it took about three generations to decide a similar ballgame in topology, namely, the question of what serves better as the most basic concept: open sets or neighborhoods. But Phase 1 (to be explained momentarily) did not last that long.

For the sake of clarity, here is how I see the development during the 80 years from 1870 to 1950; I distinguish three phases. During Phase 1, the late nineteenth century, not all textbook authors included a treatment of the irrational or real numbers (e.g., Bertrand, Duhamel, and Hermite did not), but those who did adopted one of the three approaches: Biermann (1887) followed Weierstrass (who suspected, however, that his name was used solely to promote sales (see Dugac, 1973, p. 143)); Capelli (1895), Capelli (1897), Dini (1878), Jordan (1893), Pasch (1882), Peano (1884), Ricci (1893), Ricci (1897), and Weber (1895) built on Dedekind; Tannery switched from Méray to Dedekind with the second edition (Tannery, 1904); (Thomae, 1880, p. iii) explicitly adopted Heine's "formal stance," while Bachmann (1892) was sympathetic to it; Cantor got employed, without a detour via Heine, by Lipschitz (1877) (but see footnote 24), Stolz (1885), Dini (1892) as well as by Pringsheim (1916) (who gets included to Phase 1 since his lectures still breathe the spirit of the late nineteenth century). It is this phase that I believe was cut short since the tie between Dedekind and Méray–Cantor was never broken, say, by a decisive discussion of their respective theoretical advantages or by exhibiting compelling benefits in the practice of teaching and research. Rather, the question was superseded and made obsolete by the rapid development that followed. The next phase, Phase 2, the early twentieth century, was heralded by Hilbert's proposal (Hilbert, 1900) to postulate the real numbers by axioms, and not by "filling the gaps." And while his ideas took some time to sink in, those who authored textbooks in analysis at the time were anxious to present new results in the framework of set theory and abstract spaces (e.g., Lebesgue measure and integral); they had bigger fish to fry than irrational numbers. Borel (1898)

4.4.2 Abstraction

We have not investigated what cognitive or logical processes were involved in the definitions of the irrational numbers by sequences or cuts. From a methodological point of view, it is not even clear whether the textual basis we have would support such a remote analysis across space and time. But we did argue that whatever it is, it cannot be stick-figure abstraction. The argument was simple. First, we cannot stick-figure abstract what is not present in what is given; second, our main witnesses agreed that the completeness of the reals is not present in what is intuitively given as continuous; our claim then follows.

Abstraction I This raises the follow-up question of what causes the air of abstractness surrounding either approach to defining the irrational numbers if it is not our whipping boy stick-figure abstraction. The answer I propose is similar to what I suggested above: we say "abstract," but what we mean is a feeling of alienation by unfamiliar concepts. My reasoning is as follows.[46] Mastering a new concept C means learning to navigate its conceptual space. In other words, we learn about the place the new concept C occupies in our shared semantic web, viz., we learn what other concepts entail C and what other concepts C itself, usually in conjunction with others, entail. Without the knowledge yet how to connect a new concept C to other concepts new and old, C lacks meaning for us. This initial poverty in meaning we then mistake for another case of poverty in meaning: abstractness. But C does not have little contents due to abstraction, but due to us not

did not define irrational numbers (instead, he discussed Liouville's approximation method), while Carathéodory (1918, § 18) and Hahn (1921, pp. 30f.) took as their starting point the existence of suprema. (Carathéodory, to be fair, prefixed this with a short discussion of the Peano axioms for the natural numbers.) Ten years later, Hahn (1932) did not bother at all about number systems; the real numbers got introduced as a metric space (*ibid.*, § 25.3). According to Landau, all this was the prevailing approach of the day and caused him to write and publish his well-known book (Landau, 1927). Phase 3, the 1950s and beyond, when mathematicians resumed their regular jobs after prolonged war efforts, Bourbaki ruled the day and in their wake the axiomatic method. Consequently, the first textbook in the new spirit, Dieudonné (1960), introduced the reals as an Archimedean ordered field. The reviewer at the time could not agree more: "These axioms can, of course, be proved [...] through the Dedekind or Cantor procedures. Although such proofs have great logical interest, they have no bearing whatsoever on Analysis and teachers should not burden students with them [...]. This is the right attitude shared by this text." (Nachbin, 1961, p. 247) And with translations into 13 languages the arguably most influential textbook in analysis, the *Principles of Mathematical Analysis* by Walter Rudin (1953) (soon to be dubbed "Baby Rudin"), followed suit in its second edition.

The reason that, in respect to phases 1 and 2, we almost exclusively quoted French, German, and Italian authors is that, as far as we know, this specific development took place within that tri-country community. And we can confirm this for England in regard to which Hardy wrote that before the publication of Young (1906) and Hobson (1907), "real function theory was practically unknown [in Great Britain]" (Hardy, 1934, p. 245).

[46] The explanation that follows presupposes that the inferentialist picture of meaning (see, e.g., Brandom (1994, esp. chs 2–3) or the condensed version (Brandom, 2000)), is basically correct.

yet being sufficiently familiar with it; once this has changed and we have mastered C's contents sufficiently well, the feeling of abstractness wanes.

We just made an empirical claim; how do we know whether it is true? As far as I know there are no empirical studies to corroborate the hypothesis that people mistake unfamiliar for abstract. In the absence of such studies, I claim that our point is substantiated by the experience common to all students of mathematics: feelings of abstractness go away as we master new concepts. By the time we graduate, the concept "open set" may very well be more concrete, more rich, and especially much more clear than the concept "oak tree."

Abstraction II Assume we were correct when we suggested the analogy between mathematical concepts (as the outcome of a constraint-based design process) and puzzle pieces or parts machined according to certain specs; then we can continue the analogy as follows. A puzzle piece looked at in insolation will appear to have a totally haphazard shape; once it is put in place, it makes perfect sense and no other puzzle piece would do. A pull-off tool in an automotive shop will totally befuddle the non-mechanic; once it is being used, it makes perfect sense and it is clear that no other tool would do. Accordingly, a new mathematical concept will look totally haphazard to us until we have seen how it fits into the bigger picture. A concept custom-tailored in response to specific challenges will befuddle us until we know from applications how to use it. I take this to be the message John von Neumann was trying to convey when he allegedly said to a physicist:

> Young man, in mathematics you don't understand things. You just get used to them.[47]

The previous analogy with custom-tailored tools might help to understand how demands for proper operationalization may aggravate feelings of alienation (i.e., abstractness). We illustrated the idea of proper operationalization by comparing "n is divisible by two," which is clear and intuitive, with "there is an integer k such that $2k = n$," which is less straightforward but better operationalized for deployment in proofs. We can now make a similar observation in the context of our present case study. Out of the twenty or so potential axioms we could use for postulating the completeness of the real numbers (see Propp, 2013), the one that we usually pick is the one that posits the existence of suprema: every non-empty subset bounded above has a least upper bound. To the beginner, this is less clear and less intuitive (more "abstract") than completing the order à la Dedekind (real closed field) or completing the topology à la Méray-Cantor (complete metric space), but it is better operationalized to go with proofs in analysis. The lesson is this: the learning curve for new concepts may be even steeper and hence their force to alienate (i.e., their perceived abstractness) even greater than necessary because the concepts are introduced not in an intuitive, easily recognizable way but groomed for operational success.

[47] The quote and the story that goes with it are reported in (Zukav, 1979, p. 208, footnote)

Abstraction III Finally, the analogy about custom-tailored concepts machined according to narrow specifications to solve certain problems gives away our answer to the question of where the increased expressive or deductive power of abstract theories comes from: it is baked right into their concepts, and they were specifically designed to have that power. In our case, Dedekind's cut axiom and Cantor's correspondence axiom were custom-tailored to yield exactly what was missing: the completeness of the reals either as a real closed field or as a complete metric space. We thus see that the power boost that comes with abstract theories is no coincidence: their concepts were explicitly designed to provide for that.

4.4.3 Beyond Abstraction

As a collateral, the historical record also revealed aspects of epistemic practices other than abstraction. (We here take "epistemic practices" as an umbrella term for what people do, or avoid doing, individually and socially, when they create, share, or teach knowledge.) Among the recurring themes a few stand out: fit into existing theories (Cantor, Ricci, Tannery); alignment with existing techniques (Cantor, Peano); overall ease of use (Cantor, Dedekind); conceptual simplicity Dedekind, Ricci, Tannery); conceptual clarity (Dedekind, Ricci, Tannery); computational complexity (Cantor, Dedekind, Ricci); support role in proofs (Dedekind, Peano). We saw that the practices we find centered around such themes do not have to be uniform or embraced by all. Cantor and Dedekind, for instance, emphasized the freedom to create whatever is useful as a tool (Dedekind) or what is fruitful and expedient as a concept (Cantor), while Du Bois-Reymond and Frege insisted that the traditional nexus between number and measurement cannot be severed. This tension between traditional and progressive ideas became even more pronounced in Kronecker's rejection of infinitary methods and Russell's characterization of the axiomatic methods as theft. All these considerations worked, although with unequal force, as constraints on the design space.

4.4.4 Summary

I believe that the case study above has established two claims:

(C1) There are "abstract" mathematical concepts that are not the result of stick-figure abstraction (but the outcome of a top-down design process).
(C2) There are instances where a new concept, its design, and operationalization, will alienate its first-time user and thus cause an impression of "abstractness."

Do these claims generalize? Do the following more general claims hold?

(C3) The introduction of new mathematical concepts occurs in a design space that is defined by specific constraints that limit the range of viable options.
(C4) The competition within the design space may be decided for reasons not entirely mathematical.

(C5) The power of abstract theories is a direct consequence of the design process that created them.

While I believe that our case study makes C3–C5 plausible candidates, we need a more representative sample, consisting of several case studies, to corroborate such more general claims.[48]

5 Concluding Remarks

What We Did Not Do We did not amend the title. Instead of "Abstraction by embedding and constraint-based design," it should read "Feelings of perceived abstraction caused by embedding and constraint-based design." This would have been more accurate. But it is a mouthful; and ugly. So we did not.

Next, we did not say that stick-figure abstraction is absent in mathematics. There may actually be natural candidates for it; pattern recognition, for example, when we rewrite sums:

$$\begin{aligned}&\tfrac{1}{3}+\tfrac{2}{9}+\tfrac{4}{27}+\tfrac{8}{81}+\ldots=\tfrac{1}{3}+\left(\tfrac{1}{3}\cdot\tfrac{2}{3}\right)+\left(\tfrac{1}{3}\cdot\tfrac{4}{9}\right)+\left(\tfrac{1}{3}\cdot\tfrac{8}{27}\right)+\ldots=\\&\quad\left(\tfrac{1}{3}\cdot(\tfrac{2}{3})^0\right)+\left(\tfrac{1}{3}\cdot(\tfrac{2}{3})^1\right)+\left(\tfrac{1}{3}\cdot(\tfrac{2}{3})^2\right)+\left(\tfrac{1}{3}\cdot(\tfrac{2}{3})^3\right)+\ldots=\sum_{n=1}^{\infty}\frac{1}{3}\cdot\left(\frac{2}{3}\right)^{n-1}\end{aligned}$$

or when we generalize $\mathbb{R}^3$ to $\mathbb{R}^n$, or when we recognize certain law-like regularities:

from instances such as: $n+0=n$, and: $m\cdot 1=m$, abstract the law: $a * e = a$.

It seems varieties of stick-figure abstraction, or some close cousin, may be involved in these cases. What we did say, however, is that stick-figure abstraction cannot be the whole story. Likewise, we did not say what abstraction in mathematics is; able hands have done so already (see, e.g., Marquis (2013), Marquis (2016)—not

[48] My plan was to exactly do that in the present paper. But I ran out of space. So it has to be left to a follow-up paper. But not just the sample size, also the methodology raises concerns. I described what I take to be aspects of the mathematical experience and defended certain claims. In so doing, I did not base my claims on empirical studies. Rather, I continued the tradition of modern epistemology that, to a considerable extent, has been philosophical psychology that, in the absence of relevant empirical studies, appeals to experiences believed to be familiar to everyone (in our case: everyone who received a college-level education in mathematics). While this is shaky methodology, it seems legitimate as a first step; in the end, however, we cannot shirk empirical studies to (dis)confirm what we found in the first step. But I also used historical case studies to gather data about epistemic practices among mathematicians. This was based on the unproven assumption that these practices have not undergone a fundamental change in the last 150 years and therefore still tell us something about mathematics today.

to mention the cottage industry around Frege, see Cook (2007), Ebert and Rossberg (2017)).

What We Did Do Our starting point was the tension between two statements:

(1) Modern mathematics is more abstract.
(2) Modern mathematics is more powerful.

We noted that the two statements cannot be simultaneously true provided we understand "more abstract" to mean "poorer in content," which then in turn would imply "less powerful." We identified stick-figure abstraction as the culprit. In the absence of a suitable mathematical concept of abstraction, call it abstraction$_M$, the informal, stick-figure version of Aristotelian abstraction (to which we have been exposed as the default account of abstraction for about two millennia) is the imposter that takes the rightful place of abstraction$_M$; inept, however, to properly model concept formation in mathematics, stick-figure abstraction generates the contradiction.

We argued that in our case studies abstract does not and cannot mean stick-figure abstract, i.e., produced by a process of abstraction that trims conceptual or intuitive fat and thereby delivers leaner (i.e., abstract) concepts. It was rather the opposite: instead of bottom-up abstraction, we saw top-down design.

We argued that Statement 1 (provided we replace "more abstract" with "alienating") and Statement 2 are true. We reconciled the two statements by revealing a certain ambiguity in how we use the word "abstract." We reminded ourselves that a concept that is stick-figure abstract has less content and therefore provides less guidance for its use; an experience often described as "breathing the thin air of abstraction." But we encounter situations where words provide less guidance than usual also when we learn a new language (first case study) or master new concepts (second case study). We said that it is the latter experience (i.e., feeling alienated while learning a new language or mastering new concepts) that we live through when we learn modern mathematics; an experience that may be aggravated by constraints on the design that favors performance over ease of learning. So, while the experience may be similar ("it feels the same"), the root cause is very different. We thus obliged Frege by "breaking the power of the word 'abstract' over the human mind, by uncovering illusions that through its use often almost unavoidably arise."[49]

Furthermore, the historical sources we consulted shed light not only on the topic of abstraction but also on other epistemic practices that act as constraints on the design process of new concepts and on how the winner of the design contest is determined.

[49] Cf. Frege (1879, p. vif.): "[i]t is a task of philosophy to break the power of words over the human mind, by uncovering illusions that through the use of language often almost unavoidably arise [...]"

References

Anonymous (1797) *Encyclopædia Britannica, or, A Dictionary of Arts, Sciences, and Miscellaneous Literature; ...*, 18 vols, 3rd ed. Edinburgh: Bell & Macfarquhar (^{3}1797).

Alcock, Lara (2013). *How to Study as a Mathematics Major*, Oxford: Oxford UP (2013).

Arnauld, Antoine & Pierre Nicole (1662). *La Logique, ou l'art de penser*, in: *Œuvres de messire Antonine Arnauld, I–XLI*, ed. G. D. Bellegrade & J. Hautefage, Paris: Darnay (1775–1783; repr. Bruxelles: Culture & Civilisation (1964–1967)), vol. xli, pp. 99–416; first Paris: Desprez (1662).

Ascoli, Giulio (1895). "I fondamenti dell'algebra," in: *Rendiconti di Reale Istituto Lombardo di scienze e lettere*, Serie II, XXVIII (1895), pp. 1060–1071.

Bachmann, Paul (1892). *Vorlesungen über die Natur der Irrationlazahlen*, Leipzig: Teubner (1892).

Bäck, Allan (2014). *Aristotle's Theory of Abstraction* (= New Synthese Historical Library; 73), Cham: Springer (2014).

Beachy, John A. & Blair, William D. (2006) *Abstract Algebra*, Long Grove: Waveland P (^{3}2006).

Beller, Sieghard et al. (2018). "The Cultural Challenge in Mathematical Cognition," in: *Journal of Numerical Cognition* 4:2 (2018), pp. 448–463.

Biermann, Otto (1887). *Theorie der analytischen Funktionen*, Leipzig: Teubner (1887).

Boniface, Jacqueline (2007). "The Concept of Number from Gauss to Kronecker," in: Goldstein C., Schappacher N., Schwermer J. (eds), *The Shaping of Arithmetic after C. F. Gauss's Disquisitiones Arithmeticae*, Berlin: Springer (2007), pp. 314–342.

Borel, Émile (1898). *Leçons sur la théorie des fonctions*, Paris: Gauthiers-Villars (1898).

Bottazini, Umberto & Gray, Jeremy (2013). *Hidden Harmony—Geometric Fantasies. The Rise of Complex Function Theory* (= Sources and Studies in the History of Mathematics and Physical Sciences), New York: Springer (2013).

Brandom, Robert B. (1994). *Making It Explicit: Reasoning, Representing, and Discursive Commitment*, Cambridge, MA: Harvard UP (1994).

——— (2000). *Articulating Reasons: An Introduction to Inferentialism*, Cambridge, MA; Harvard UP (2000).

Cantor, Georg (1932). *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, ed. Ernst Zermelo and a biography by Adolf Fraenkel, Berlin: Springer (1932).

——— (1872). "Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen," in Cantor (1932), pp. 92–102; first in: *Mathematische Annalen* 5 (1872), pp. 123–132.

——— (1883). "Grundlagen einer allgemeinen Mannigfaltigkeitslehre, Nr. 5," in Cantor (1932), pp. 165–209; first as "Ueber unendliche, lineare Punktmannichfaltigkeiten," in: *Mathematische Annalen*, 21 (1883), pp. 545–591.

——— (1887/8). "Mitteilungen zur Lehre vom Transfiniten," in: Cantor (1932), pp. 378–439; first in: *Zeitschrift für Philosophie und philosophische Kritik*, NF 91 (1887), pp. 81–125; 92 (1888), pp. 240–265.

——— (1991). *Georg Cantor. Briefe*, ed. Herbert Meschkowski, Winfried Nilson, Berlin: Springer (1991).

Capelli, Alfredo (1895). *Lezioni di algebra complentare, ad uso degli aspiranti alla licenza universitaria in scienze fisiche e matematiche*, Naples: Pellerano (1895).

——— (1897). "Sulla introduzione dei numeri irrazionali col metodo delle classi contigue," in: *Giornale di Matematiche*, XXXV (1897), pp. 209–234.

Carathéodory, Constantin (1918). *Vorlesungen über reelle Funktionen*, Leipzig: Teubner (1918).

Carothers, Neal L. (2000). *Real Analysis*, Cambridge: Cambridge UP (2000).

Cauchy, Augustin-Louis (1882f.). *Œuvres complètes*, vols I:1–12, II:1–15, Paris: Gauthier-Villar (1882–1974).

——— (1821). *Course d'analyse de l'Ecole Royale Polytechnique. I: Analyse algébrique*, in: Cauchy (1882f.), II:3; first Paris: Debure (1821).

——— (1823). *Le Résumé des Leçons données à l'Ecole royale Polytechnique sur le Calcul infinitésimal*, in: Cauchy (1882f.), II:4, pp. 5–261; first Paris: Debure (1823).

Cook, Roy T. (ed) (2007). *The Arché Papers on the Mathematics of Abstraction* (= Western Ontario Series in Philosophy of Science; 71), Dordrecht: Springer (2007).

Cox, David A. (2012) *Galois Theory* (= Pure and Applied Mathematics; 106), Hoboken, NJ: Wiley (2012).

Dantscher, Victor von (1908). *Vorlesungen über die Weierstrasssche Theorie der Irrationalzahlen*, Leipzig: Teubner (1908).

Dawson, Jr., John W. (2015). *Why Prove it Again? Alternative Proofs in Mathematical Practice*, Basel: Birkhäuser (2015).

Dedekind, Richard (1871). "Ueber die Composition der binären quadratischen Formen" (= Supplement X), in: Peter G. Lejeune Dirichlet, *Vorlesungen über Zahlentheorie*, posthum ed. R. Dedekind, Braunschweig: Vieweg (^{2}1871), pp. 380–497.

——— (1872). *Stetigkeit und irrationale Zahlen*, Braunschweig: Vieweg (1872); Engl. tr by Wooster W. Beman, as "Continuity and irrational numbers," in: *Essays on the Theory of Numbers*, Chicago: Open Court (1901), pp. 1–27.

Dieudonné, Jean (1960). *Foundations of Modern Analysis, Volume 1* (= Pure and Applied Mathematics; 10), New York: Academic P (1960, ^{2}1969, rev. 1972, 1974, 1976, 1977, 1978).

Dini, Ulisse (1878). *Fondamenti per la teorica delle funzioni di variabili reali*, Pisa: Nistri (1878).

——— (1892). *Grundlagen für eine Theorie der Functionen einer veränderlichen reellen Grösse*, rev. tr. of Dini (1878) by Jacob Lüroth & Adolf Schepp, Leipzig: Teubner (1892).

Du Bois-Reymond, Paul (1882). *Die Allgemeine Functionentheorie. Erster Theil: Metaphysik und Theorie der mathematischen Grundbegriffe: Grösse, Grenze, Argument und Function*, Tübingen: Laupp (1882).

Dugac, Pierre (1970). "Charles Méray (1835–1911) et la notion de limite," in: *Revue d'histoire des sciences et de leurs applications*, 23:4 (1970), pp. 333–350.

——— (1973). "Eléments d'analyse de Karl Weierstrass," in: *Archive for History of Exact Sciences*, 10:1/2 (April 1973), pp. 41–176.

Ebert, Philip A. & Rossberg, Marcus (eds) (2017). *Abstractionism. Essays in Philosophy of Mathematics*, Oxford: Oxford UP (2017).

Edwards, Harold E. (1995). "Kronecker on the foundations of mathematics," in: Hintikka (1995), pp. 45–52.

Ehrlich, Philip (2006). "The Rise of non-Archimedean Mathematics and the Roots of a Misconception I: The Emergence of non-Archimedean Systems of Magnitudes," in: *Archive for History of Exact Sciences*, 60:1 (January 2006), pp. 1–121.

——— (2018). *Contemporary infinitesimalist theories of continua and their late 19th- and early 20th-century forerunners*. `arXiv:1808.03345v3` (12/27/18); 73 pp.

Fitzpatrick, Richard (ed.) (2007) *Euclid's Elements of Geometry*, the Greek text of J. L. Heiberg [. . .] ed. & tr. by R. Fitzpatrick, (2007, rev. 2008); URL: http://farside.ph.utexas.edu/Books/Euclid/Elements.pdf

Frege, Gottlob (1879). *Begriffsschrift: eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*, Halle: Nebert (1879); Engl. tr. of Part 1 (§§ 1–12) in: *The Frege Reader*, ed. Michael Beaney, Oxford, Blackwell (1997), pp. 48–78.

——— (1903). *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet*, Bd. 2, Jena: Pohle (1903); Engl. tr. P. Ebert and M. Rossberg, *Basic Laws of Arithmetic: Derived using concept-script*, Oxford: Oxford UP (2013).

Gonzalez-Velasco, Enrique A. (1992). "Connections in Mathematical Analysis: The Case of Fourier Series," in: *American Mathematical Monthly*, 99:5 (May 1992), pp. 427–441.

Grabiner, Judith V. (1981). *The Origins of Cauchy's Rigorous Calculus*, Cambridge, MA: MIT P (1981).

Hahn, Hans (1921). *Theorie der reellen Funktionen. Erster Band*, Berlin: Springer (1921).

——— (1932). *Reelle Funktionen. Bd. 1: Punktfunktionen* (= Mathematik und ihre Anwendungen; 13), Leipzig: Akademische Verlagsgesellschaft (1932).

Hankel, Hermann (1867). *Vorlesungen über die complexen Zahlen und ihre Functionen. I. Theil: Theorie der complexen Zahlensysteme, insbesondere der gemeinen imaginären Zahlen und der Hamiltonischen Quaternionen nebst ihrer geometrischen Darstellung*, Leipzig: Voss (1867).

Hardy, Godfrey H. (1934). "Ernest William Hobson, 1856–1933," In: *Obituary Notices of Fellows of the Royal Society*, 1:3 (1934), pp. 236–249.

Hawkins, Thomas (1979). *Lebesgue's Theory of Integration. Its Origins and Development*, Providence, RI: AMS (2002); first Madison: University of Wisconsin Press (1970); New York: Chelsea (^{2}1979).

Heath, Thomas L. (ed.) (1908). *The Thirteen Books of Euclid's Elements*, tr. from the text of Heiberg, with introduction and commentary by T. L. Heath, Cambridge; Cambridge UP (1908, ^{2}1926); repr. New York: Dover (1956).

Heine, Eduard (1870). "Ueber trigonometrische Reihen," in: *Journal für die reine und angewandte Mathematik* (= Crelle's/Borchardt's Journal), 71 (1872), pp. 353–365.

Heine, Eduard (1872). "Die Elemente der Functionenlehre," in: *Journal für die reine und angewandte Mathematik* (= Crelle's/Borchardt's Journal), 74 (1872), pp. 172–188.

Henrich, Joseph & Heine, Steven J. & Norenzayan, Ara (2010). "The weirdest people in the world?," in: *Behavioral and brain sciences*, 33:2–3 (June 2010), pp. 61–135.

Hinman, Peter G. (2005). *Fundamentals of Mathematical Logic*, Wellesley, MA: Peters (2005).

Hilbert, David (1900). "Über den Zahlbegriff," in: *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 8 (1900), pp. 180–183.

Hintikka, Jaakko (ed.) (1995). *From Dedekind to Gödel. Essays on the Development of the Foundations of Mathematics* (= Synthese Library; 251), Dordrecht: Kluwer (1995).

Hobson, Ernest W. (1907). *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, Cambridge: Cambridge UP (1907).

Illigens, Eberhard H. (1889). "Zur Weierstrass'-Cantor'schen Theorie der Irrationalzahlen," in: *Mathematische Annalen*, XXXIII (1889), pp. 155–160.

——— (1890). "Zur Definition der Irrationalzahlen," in: *Mathematische Annalen*, XXXV (1890), pp. 151–455.

Jordan, Camille (1893). *Cours d'analyse de l'École Polytechnique*, t. 1: *Calcul différentiel*, Paris: Gauthioer-Villars (1882, ^{2}1893).

Klein, Felix (1895). *Vorträge über ausgewählte Fragen der Elementargeometrie*, red. by F. Tägert (= Festschrift zu der Pfingsten 1895 in Göttingen stattfinden dritten Versammlung des Vereins zur Förderung des mathematischen und naturwissenschaftlichen Unterrichts), Leipzig: Teubner (1895); cited according to Engl. tr. by W. W. Beman & D. E. Smith, *Famous problems of elementary geometry*, Boston: Ginn (1897).

Knauer, Ullrich (2011). *Algebraic Graph Theory: Morphisms, Monoids and Matrices* (= De Gruyter Studies in Mathematics; 41), Berlin: de Gruyter (2011).

Kossak, Ernst (1872). *Die Elemente der Arithmetik* (= Programmabhandlung des Werderschen Gymnasiums, Berlin), Berlin: Nicolai (1872).

Kronecker, Leopold (1886). "Ueber einige Anwendungen der Modulsysteme auf elementare algebraische Fragen," in: *Werke*, ed. Kurt Hensel, Bd. 3.1, Leipzig: Teubner (1899), pp. 145–208; first in: *Journal für die reine und angewandte Mathematik* (= Crelle's/Borchardt's Journal), 99 (1886), pp. 329–371.

Landau, Edmund (1927). *Grundlagen der Analysis. (Das Rechnen mit den ganzen, rationalen, irrationalen, komplexen Zahlen) Ergänzung zu den Lehrbüchern der Differential- und Integralrechnung*, Leipzig: Akademische Verlagsgesellschaft (1927); numerous reprints.

Lee, Haw Y. (2017). *Origami-constructible numbers*, Athens, GA: U of Georgia (2017); url: https://getd.libs.uga.edu/pdfs/lee_hwa-young_201712_ma.pdf.

Lende, Daniel H. & Downey, Greg (eds) (2012). *The Encultured Brain: An Introduction to Neuroanthropology*, Cambridge, MA: MIT P (2012).

Lipschitz, Rudolf (1877). *Lehrbuch der Analysis. Bd. 1: Grundlagen der Analysis*, Bonn: Cohen (1877).

Lovejoy, Arthur O. (1936). *The Great Chain of Being: A Study of the History of an Idea*, Cambridge: Harvard UP; many reprints.

Marquis, Jean-Pierre (2013). "Mathematical Abstraction, Conceptual Variation and Identity," in: Schroeder-Heister, Peter *et al.*, *Logic, Methodology and Philosophy of Science. Proceedings of the 14th International Congress (Nancy). Logic and Science Facing the New Technologies*, London: College Publications (2014), pp. 299–322.

——— (2016). "Stairway to Heaven: The Abstract Method and Levels of Abstraction in Mathematics," in: *The Mathematical Intelligencer* 38:3 (2016),pp. 41–51.

McCarty, D. Charles (1995). "The mysteries of Richard Dedekind," in: Hintikka (1995), pp. 53–96.

Méray, Charles (1870). "Remarques sur la nature des quantités définies par la condition de servir de limites à des variable données," in: *Revue sociétés des savantes. Sciences mathématiques, physique er naturelle*, Second Series IV (Janvier–Juin 1869; 1870), pp. 280–289.

——— (1872). *Nouveau précis d'analyse infinitésimale*, Paris: Savy (1872).

——— (1894). *Leçons nouvelles sur l'analyse infinitésimale et ses applications géométriques. Première partie: Principes généraux*, Paris: Gauthier-Villars (1894).

Mittag-Leffler, Gösta (1910). "Sur les fondements arithmétiques de la théorie des fonctions d'après Weierstrass," in: *Compte rendu du congrès des mathematiciens, tenu à Stockholm, 22–25 Septembre 1909*, ed. Gösta Mitag-Leffler & Ivar Fredholm, Leipzig: Teubner (1910), pp. 10–31.

Nachbin, Leopoldo (1961). "Review of Dieudonné (1960)," in: *Bulletin of the American Mathematical Society*, 67:3 (1961), pp. 246–250.

Northoff, Georg (2010). "Humans, brains, and their environment: marriage between neuroscience and anthropology?," in: *Neuron*, 65 (March 2010), pp. 748–751.

Pasch, Moritz (1882). *Einführung in die Differential- und Integralrechnung*, Leipzig: Teubner (1882).

Peano, Guiseppe (1884). *Calcolo differenziale e principii di calcolo integrale, pubblicato con aggiunte*, Torino: Bocca (1884).

——— (1899). "Sui numeri irrazionali," in: *Revue de mathématiques (Rivista di matematica)*, VI:4 (1899), pp. 126–140.

Peterson, Julius (1878). *Theorie der algebraischen Gleichungen*, Kopenhagen: Höst (1878).

Pincherle, Salvatore (1880). "Saggio di una introduzione alla Teoria delle funzioni analitiche secondo i principii del Prof. C. Weierstrass," in: *Giornale di Matematiche* XVIII (1880), pp. 178–254, 317–357.

Pringsheim, Alfred (1898). "Irrationalzahlen und konvergente unendliche Prozesse," in: *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. 1: *Arithmetik und Algebra*, ed. Wilhelm F. Meyer, Leipzig: Teubner (1898–1904), pp. 47–146.

——— (1916). *Vorlesungen über Zahlen- und Funktionenlehre. Bd. 1: Zahlenlehre* (= Teubner's Sammlung von Lehrbüchern auf dem Gebiete der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen; XL:I.1), Leipzig: Teubner (1916, ^{2}1923).

—— & Molk, Jules (1904). "Nombres irrationnels et notion de limite," tr. and rev. of Pringsheim (1898) by J. Molk, in: *Encyclopédie des sciences mathématiqus pures et appliquées, publiée* . . . , ed. J. Molk, t. 1: *Arithmétique*, Paris: Gauthier-Villars (1904), pp. 133–208.

Propp, James (2013). "Real analysis in reverse," in: *American Mathematical Monthly*, 120:5 (May 2013), pp. 392–408.

Rav, Yehuda (1999). "Why do we prove theorems?" in: *Philosophia Mathematica*, 7 (1999), pp. 5–41.

Ricci, Gregorico (1893). "Saggio di una teoria dei numeri reali. Secondo il concetto di Dedekind," in: *Atti del R[eale] Instituto Veneto di Scienze, Lettere ed Arti*, Ser. 7:4, LI (1893), pp. 233–281.

——— (1897). "Della teoria dei numeri reali. Secondo il concetto di Dedekind," in: *Giornale di Matematiche*, XXXV (1897), pp. 22–74.

Rosa, Milton & Shirley, Lawrence & Gavarrete, Maria Elena & Alangui, Wilfredo V. (eds) (2017). *Ethnomathematics and its Diverse Approaches for Mathematics Education* (= ICME 13 Monographs), Cham: Springer (2017).

Rota, Gian-Carlo (2008). "Introduction," in Davis, Philip J. & Hersh, Reuben (1981). *The Mathematical Experience*, Boston: Brikhäuser (1981); *Study Edition*, with Elena A. Marchisotto, Boston: Brikhäuser (1995), pp. xxi–xxiii.

Rudin, Walter (1953). *Principles of Mathematical Analysis* (= International Series in Pure and Applied Mathematics), New York: McGraw-Hill (1953, ^{2}1964, ^{3}1976).
Russell, Bertrand (1903). *The Principles of Mathematics, Vol. 1*, Cambridge: Cambridge UP (1903).
——— (1919). *Introduction to Mathematical Philosophy*, London: Allen & Unwin (1919).
Smoryński, Craig (2008). *History of Mathematics. A Supplement*, New York: Springer (2008).
Spalt, Detlef D. (1991). "Die mathematischen und philosophischen Grundlagen des Weierstraßschen Zahlbegriffs zwischen Bolzano und Cantor," in: *Archive for History of Exact Sciences*, 41:4 (1991), pp. 311–362.
Starikova, Irina (2012). "From Practice to New Concepts: Geometric Properties of Groups," in: *Philosophia Scientiæ*, 16:1 (2012), pp. 129–151.
Stolz, Otto (1885). *Vorlesungen über allgemeine Arithmetik. Nach den neueren Ansichten, Bd. 1: Allgemeines und Arithmetik der reellen Zahlen*, Leipzig: Teubner (1885).
Tait, William W. (1996) "Frege versus Cantor and Dedekind: On the concept of number," in: *Frege, Russell, Wittgenstein: Essays in Early Analytic Philosophy (in honor of Leonard Linsky)*, ed. W. W. Tait, Lasalle: Open Court (1996), pp. 213–248; repr. in: *Frege: Importance and Legacy*, ed. M. Schirn, Berlin: de Gruyter (1996), pp. 70–113.
Tannery, Jules (1904). *Introduction à la théorie des fonctions d'une variable*, t. 1, Paris: Hermann (1886, ^{2}1904).
——— (1908). "Review of Dantscher (1908)," in: *Bulletin des sciences mathématiques*, Ser. II, XXXII (1908), pp. 101–105.
Tapp, Christian (2005). *Kardinalität und Kardinäle. Wissenschaftshistorische Aufarbeitung der Korrespondenz zwischen Georg Cantor und katholischen Theologen seiner Zeit* (= Boethius: Texte und Abhandlungen zur Geschichte der Mathematik und der Naturwissenschaften; 53), Wiesbaden: Steiner (2005).
Thomae, C. Johannes (1880). *Elementare Theorie der analytischen Functionen einer complexen Veränderlichen*, Halle: Nebert (1880).
Thomassen, Carsten (1992). "The Jordan-Schonflies theorem and the classification of surfaces," in: *American Mathematical Monthly*, 99:2 (Feb., 1992), pp. 116–131.
Ullrich, Peter (1989). "Weierstraß' Vorlesung zur 'Einleitung in die Theorie der analytischen Funktionen'," in: *Archive for History of Exact Sciences*, 40:2 (1989), pp. 143–172.
Van Vleck, Edward B. (1914). "The influence of Fourier's series upon the development of mathematics," in: *Science*, NS, 39:995 (Jan. 23, 1914), pp. 113–124.
Volkert, Klaus Th. (1986). *Die Krise der Anschauung. Eine Studie zu formalen und heuristischen Verfahren in der Mathematik seit 1850* (= Studien zur Wissenschafts-, Sozial- und Bildungsgeschichte der Mathematik; 3), Göttingen: Vandenhoeck & Ruprecht (1986).
Wantzel, Pierre L. (1837). "Recherches sur les moyens de reconnaître si un Probléme de Géométrie peut se résoudre avec la règle et le compas," in: *Journal de mathématiques pures et appliquées* 10 (1837), pp. 366–372.
Weinberg, Julius (1973). "Abstraction in the formation of concepts," in: *Dictionary of the History of Ideas. Studies of Selected Pivotal Ideas, I–IV*, ed. by Philip P. Wiener, New York: Scribners (1973), vol. 1, pp. 1–9.
Weber, Heinrich (1895). *Lehrbuch der Algebra, Bd. 1*, Braunschweig: Vieweg (1895, ^{2}1912).
Weierstrass, Karl (1903). *Mathematische Werke. Bd. 3: Abhandlungen III*, Berlin: Mayer & Müller (1903).
Whitney, Hassler (1932). "Non-Separable and Planar Graphs," in: *Transactions of the American Mathematical Society*, 34:2 (April 1932), pp. 339–362.
Yates, Robert C. (1942). *The Trisection Problem*, Baton Rouge, LA: Franklin P (1942).
Young, William H. & Chisholm Young, Grace (1906). *The Theory of Sets of Points*, Cambridge: Cambridge UP (1906).
Zukav, Gary (1979). *The Dancing Wu Li Masters. An Overview of the New Physics*, New York: Bantam (1980); first ed. New York: Morrow (1979).

The Birth of Undergraduate Modern Algebra in the United States

Walter Meyer

Abstract The first objective of this paper is to confirm a conjecture of Garrett Birkhoff about the slow adoption of undergraduate modern algebra courses in the United States in the twentieth century. We will chart the chronological development of modern algebra courses in a sample of American colleges and universities. It will be seen that the pace was quite gradual, perhaps puzzlingly so. Our second objective is to describe some of the factors in the mathematical environment of the early decades of the twentieth century which may have served to slow down the development of these courses.

1 Introduction

George Birkhoff, a giant of twentieth century American mathematics, commented in a survey of research in algebra that "...it was not until after World War II that modern algebra became popular at the college level in our country..." (Birkhoff 1973). He does not provide evidence of this timing, which may have been more of a conjecture than assertion, but we will show evidence that he was right. What makes this historically interesting, and was probably puzzling to Birkhoff, is the fact that modern algebra, especially group theory, was already an important area of research

The Cajori Two Group consists of: Dr. Walter Meyer, Emeritus Professor, Adelphi University; Dr. Worku T. Bitew, Farmingdale State College; Dr. Larry D'Antonio, Ramapo State College; Dr. Heather Huntington, Nassau Community College; Dr. Joseph Malkevitch, Emeritus Professor at York College of CUNY & Adjunct Professor at Columbia Teacher's College; Dean Nataro, Nassau Community College; JoAnne Thacker, Nassau Community College; Dr. Jack Winn, (Deceased) Emeritus Distinguished Teaching Professor, Farmingdale State College of SUNY.

W. Meyer (✉)
Adelphi University, Garden City, NY, USA
e-mail: Meyer1@adelphi.edu

M. Zack, D. Schlimm (eds.), *Research in History and Philosophy of Mathematics*, Annals of the Canadian Society for History and Philosophy of Mathematics/Société canadienne d'histoire et de philosophie des mathématiques,
https://doi.org/10.1007/978-3-030-95201-3_9

in mathematics before the twentieth century, at least a half century earlier than its widespread appearance in undergraduate curricula in America.

Birkhoff's single remark about teaching might seem like an odd tangent in his article otherwise all about research. Perhaps, it was Birkhoff's way of encouraging some investigation of what might be thought to be the slow birth of undergraduate modern algebra. We shall attempt this here, but we prefer to say the birth was cautious rather than slow. To call it slow, it would be best to have something to compare it to. Is there comparable data for some quicker adoption of a revolution in mathematics which became a standard part of the curriculum for mathematics majors?[1]

It is not part of our purpose to say anything new about the history of research in algebra. Readers more deeply interested in that might start with the Birkhoff article just noted, or any of the following more focused accounts which we have found excellent: The Evolution of Group Theory (Kleiner 1986), The Genesis of the Abstract Ring Concept (Kleiner 1996), and Field Theory: From Equations to Axiomatization (Kleiner 1999, pp. 677–684).

2 What Is Modern Algebra

By "modern algebra" we will mean the study of topics such as groups, rings, fields, and set theory pursued more or less abstractly. Searching in *Worldcat* and in Google's *Ngram Viewer* suggests that this term grew in usage around the time of van der Waerden's *Moderne Algebra* in 1930.[2]

Such searches will also show that "abstract algebra" was often used to describe the same material. Note, for example, Ore's review article (Ore 1931). In the present work, we consider the terms interchangeable. Starting in the late 1960s, "abstract algebra" started to gain popularity. In the decade 1970–1979, *Worldcat* lists 497 books with "abstract algebra" in the title and about the same number (452) including "modern algebra." However, by 2010–2019, "abstract algebra" figured in 391 book titles, while "modern algebra" had 122.

Earlier works whose titles lack both the words "modern" or "abstract" designation may certainly be regarded as modern algebra or, at least, forerunners thereof. Examples include Weber's *Lehrbuch der Algebra* (1895, F. Vieweg and Son) and B. Peirce's *Linear Associative Algebra* (*1882*, van Nostrand, also published in 1870 in lithograph form).

[1] Probability and statistics bear investigation.

[2] There is, however, a course entitled *Introduction to Modern Geometry and Modern Algebra* in Harvard's 1905 catalog. However, there is no course description so it is unclear what it contained.

3 Algebra in the Nineteenth Century: Examples of a Successful Start[3]

With only a bit of reliance on forerunners, Galois, in 1829, discussed the structure of certain permutations relevant to solving polynomial equations in terms that would today be described as belonging to groups and fields (Galois 1846). It took a while for this to be recognized in the way we would describe it today, but in 1854, Cayley took a big step by framing an abstract definition of a group of functions. In 1870 (Jordan 1870), Camille Jordan demonstrated the pervasiveness of groups in mathematics. As Felix Klein put it (Klein 1979): "Jordan wandered through all of algebraic geometry, number theory, and function theory in search of interesting permutation groups." In 1871, Dedekind defined the concept of an ideal and used it to show that any algebraic number field (any subfield of complex numbers that has finite linear dimension over the rationals) has unique factorization. In 1872, Klein proposed the Erlangen Program in which different geometries (hyperbolic, Euclidean, etc.) should be distinguished from one another via the groups of transformations of the space and their invariants.

Americans were also involved in abstract algebra in the late nineteenth century. Already in 1870, we have important work on linear associative algebras by Benjamin Peirce (Peirce 1881). In a 1893 paper by E. H. Moore, on doubly infinite sets of simple groups, he proves that every finite field is isomorphic to a Galois field of prime power order (Parshall and Rowe 1991). In 1895, Moore also introduced the idea of an automorphism (Moore 1895), and in 1898, American G. A. Miller introduced the idea of the commutator subgroup (Miller 1898).

One good indication of the importance in 1900 of what we today call modern algebra is the fact that in Hilbert's famous 1900 address at the International Congress of Mathematicians in Paris, the fifth of his 23 problems concerns Lie groups (which had been defined and studied starting in 1872–1873.)

4 Modern Algebra in the American Undergraduate Curriculum

4.1 About Cajori Two

Our data comes partly from the Cajori Two Project, a survey of the curricula of 20 well-known colleges[4] at 10-year intervals throughout the twentieth century (Meyer

[3] Anything in this section with no specific citation can be found in the aforementioned works.

[4] The institutions surveyed in Cajori Two are those in Table 1 (with the understanding that JHopkinsAll means the union of both applied and pure mathematics departments at Johns Hopkins University, and likewise for Stanford and the University of Texas at Austin; Aust stands for Austin, Cal stands for California, and Mad for Madison).

Table 1 Data from Cajori Two Schools

Abstract/Modern Algebra with or without Linear Algebra at Cajori Two Schools

Shaded cell interior indicates one or more instances of the course.

	1905	1915	1925	1935	1945	1955	1965	1975	1985	1995	2005
Bowdoin						■	■	■	■	■	■
BrynMawr					■	■	■	■	■	■	■
CityCollege NY				■	■	■	■	■	■	■	■
Colorado Coll						■	■	■	■	■	■
Georgetown						■	■	■	■	■	■
JHopkinsAll					■	■	■	■	■	■	■
MIT			■	■	■	■	■	■	■	■	■
Morgan						■	■	■	■	■	■
Reed					■	■	■	■	■	■	■
Samford							■	■	■	■	■
SanJoseState							■	■	■	■	■
StanfordAll			■		■	■	■	■	■	■	■
Tuskegee							■	■	■	■	■
UCalBerkeley				■	■	■	■	■	■	■	■
USMA(West Point)							■	■	■	■	■
UTexasAustAll				■	■	■	■	■	■	■	■
UWisconsinMad							■	■	■	■	■
Vassar					■	■	■	■	■	■	■
Williams						■	■	■	■	■	■
Yale						■	■	■	■	■	■

In the early part of the twentieth century, "modern algebra" was understood to include linear algebra, and courses in the former were typically year courses. As time went on, the course separated

2013). The institutions surveyed in Cajori Two are those in Table 1. This survey was conducted and made accessible online by the dedicated work of the Cajori Two Group.

With the help of archivists from these 20 institutions, mathematics sections of catalogs covering each of the years 1905, 1915, ..., 2005 were obtained. The catalog excerpts, and tables produced from them, are online at http://matcmp.ncc.edu/taormij/cajori_two/.

The Cajori Two institutions were chosen to exhibit diversity with regard to geography, mission, size, religious affiliation, selectivity of admissions, and focus on particular populations such as black students and women. But in addition, we wanted schools with enough visibility that they would attract the interest of the reader. Within each category, a list of member schools was assembled and random choices were made. This is not a random sample, but we believe that this is the biggest, most diverse, and most representative sample of well-known institutions that has ever been reported on for general curricular issues spanning the twentieth century.

Table 2 Data from Selective Liberal Arts Colleges

Abstract/Modern Algebra with or without Linear Algebra at Selective Liberal Arts Colleges

Shaded cell interior indicates one or more instances of the course.

	1905	1910	1915	1920	1925	1930	1935	1940	1945	1950	1955	1960
Amherst						■	■	■	■	■	■	■
Carleton										■	■	■
Davidson												■
Grinnell										■	■	■
Occidental											■	■
Pomona											■	■
Richmond												■
Swarthmore										■	■	■
Whitman									■	■	■	■
Yale										■	■	■
Washington & Lee											■	■
Williams										■	■	■

Yale and Williams appear in both tables, but this is not a true redundancy because this table has better resolution

4.2 *About Table 1*

In Table 1, shaded boxes show years ending in 5 when there was a course with the content of modern algebra (but possibly under another name provided the course description made clear that the dominant content was in the union of abstract groups, rings and fields). It is interesting to note that even eminent institutions delayed introducing undergraduate modern algebra. We doubt that this was due to inertia. These institutions did not become or stay eminent without a staff full of vigor and initiative.

4.3 *About Table 2*

The Cajori Two data in Table 1 has only 10 year resolution. Thus, if the first shaded cell for an institution is 1955, the course could actually have been instituted in 1946. In addition, the Cajori Two set of institutions contain many which had graduate departments and could have delayed instituting an undergraduate course because advanced students could be allowed to take a graduate course in algebra. For these reasons, we sought a separate confirming sample with 5-year resolution that contained mostly high-ranking liberal arts colleges without graduate departments of mathematics. In this sample, shown in Table 2, the midcentury phenomenon is also clear.

5 The Story in a Nutshell and a Caution

Our next objective is to identify some aspects of the national mathematical environment of the first half of the twentieth century in America that could have acted to create delay in creating modern algebra courses. We do so below, but we must warn the reader that, while these environmental factors clearly existed and it is plausible that altogether they had some effect on the creation of modern algebra courses, documentary evidence for their effect is lacking.

We have not been able to find famous figures or groups who left their exhortations on undergraduate curricula in modern algebra in the written record in the early twentieth century. But even that would be inadequate. Individual institutions often ignore well-meant advice of famous figures or groups in favor of close-to-home priorities and judgment. What would be needed would be testimonies from an adequate national sample of rank and file mathematics faculty. However, rank and file mathematics instructors have rarely been on the record about what went on in their department meetings and why. However, we will report facts and opinions of researchers about modern algebra as these are part of history whether their effects are clearly known or not.

An interesting fact about the factors we discuss is that few of them had any great staying power. Large or small to start with, as the twentieth century proceeded, these potential retarding influences on adoptions of modern algebra mostly all became smaller.

Finally, there is the question of whether we have uncovered every possible environmental influence on curricular decisions in modern algebra. Surely we have not. In every realm where history is practiced, be it military, political, or mathematical, new data and new ideas have a way of deepening our understanding. However, one needs to start somewhere. This is the first attempt, and we will be grateful for suggestions of what could be added to the discussion.

6 The Environment Concerning Modern Algebra Courses

6.1 Doubts About Abstraction and Applications

6.1.1 The Double Mindset: Abstraction Versus Concrete Mathematics

One thing that could have given pause to undergraduate instructors was that there seemed to be two mindsets among researchers about the role of abstraction: As we will describe, in one view it was an occasionally useful tool for classifications or proofs but not really central to the subject matter of mathematics; in the other view, abstract structures were themselves of strong independent interest and were a compelling new way to think of most of mathematics, even if they did not resolve concrete questions.

Galois had no idea of abstract groups, no idea even that there could be structures similar to the permutations he studied. The abstract idea of a group had to wait till Cayley's work in 1854. Unfortunately, no one paid much attention to Cayley's work (Kleiner, The Evolution of Group Theory, 1986, pp. 195–215). Instead, what we see is abstraction making slow headway against the concrete approach which dealt with particular polynomials, particular geometric structures, etc. For those favoring the concrete approach, which included many mathematicians until well into the twentieth century, abstraction was regarded mainly as a tool to simplify proofs in some concrete cases like finding the zeroes of polynomial equations or dealing with number theory or some set of transformations of a geometric space.

One good place to examine these two mindsets is in the work of Heinrich Weber, Hilbert's predecessor at Göttingen, whose 1895 textbook *Lehrbuch der Algebra* (Weber, Lehrbuch der Algebra, 1895) was the dominant algebra text of his era. Two years before the appearance of this book, Weber wrote an article (Weber, Die Allgemeinen Grundlagen der Galois'schen Gleichungstheorie, 1893) in which, according to Corry (1996, pp. 35–38), he took many steps toward abstraction. He was the first to give abstract definitions of a group and a field in the same breath so to speak: A field was defined as a commutative group with an additional operation. Finite and infinite groups were both considered from the abstract point of view in this article, as were finite and infinite fields. The article contains the "first truly modern published presentation of Galois theory" in which the interplay of groups and fields was primary (Gray and Parshall 2007, p. 227).

Despite this emphasis on abstraction in the 1893 article, Weber shows some unease with the abstract point of view, especially in his book which follows his article by 2 years. As Corry (1996, pp. 38–45) points out, in Weber's book he implicitly assumes that his abstract structures are subsets of the real or complex numbers. As a result, his algebraic structures are only halfway abstract. Furthermore, the book emphasizes algebra as having one main aim: the resolution of equations. Weber's thinking about abstraction as being merely a tool is reflected in his remark: "The effect of this viewpoint [the abstract] is that the theory appears as a pure formalism that acquires content and life only after the individual elements are assigned numerical values" (Wussing 1984, p. 247). Weber's abstract definition of a group occurs astonishingly late in the book—only on p. 511 (Corry 1996, p. 41).

Felix Klein also had his reservations about the abstract point of view. Klein acknowledges group theory as an important tool for the unification of mathematics. However, he also writes, concerning abstraction: "Thus there is a complete loss of appeal to the imagination" and "The abstract formulation is excellent for working out of proofs but it does not help one find new ideas and methods" (Wussing 1984, p. 228).

In America too one could hear demurrals about abstraction. In 1902, E. H. Moore, then the leading American research mathematician, used his AMS Presidential Address to wonder: "whether the abstract mathematicians ... are not losing sight of the evolutionary character of all life processes" (Parshall and Rowe 1991, p. 417).

Finally, historian Rüdiger Thiele reports remarks of Carl Ludwig Siegel, Gian Carlo Rota, and Jean Dieudonné which Thiele interprets as reservations about abstraction, at least as it is expressed by van der Waerden's work (Thiele 2007, p. 160). For example, Siegel's remarks, contained in a letter to André Weil, include: "It is completely clear to me which circumstances have caused the gradual sinking of mathematics to its present hopeless level from its high level about 100 years ago

. . . the influence of textbooks in the style of Hasse, Schreier and van der Waerden did perceptible harm to young researchers . . . ”

These remarks of Siegel, Rota, and Dieudonné, respectively in 1959, 1988, and 1970, were too late to influence curricular decisions in the era in which modern algebra made its way into the undergraduate curriculum, but they do suggest that reservations about modern algebra were probably in the air, even among some eminent mathematicians, during that earlier era.

6.1.2 Applications of Modern Algebra Did Not Seem Compelling

Most of the powerful applications of modern algebra that we know of today, in coding theory, for example, were not known in the early twentieth century. The ones which did exist, such as the study of symmetries of crystals and the use of quaternions in mechanics were not durably exciting to mathematicians. These applications were too briefly on the stage to enhance the status of modern algebra enough so as to give birth to an undergraduate course.

6.2 Lack of a Suitable Modern Algebra Textbook

No modern algebra text proved its widespread suitability for undergraduates on the American scene till the appearance of Birkhoff and Mac Lane’s *A Survey of Modern Algebra* in 1941 (Birkhoff and Mac Lane, A Survey of Modern Algebra, 1941). Books by Dickson, Albert, and MacDuffee failed to achieve the undergraduate popularity of Birkhoff and Mac Lane. But why did it take until 1941 for a text like Birkhoff and Mac Lane to appear? Is it possible that the shape of this subject might have seemed uncertain to prospective authors when they had undergraduates in mind? The subject was still young and developing. Indeed, when Birkhoff and Mac Lane got together to compare their separate conceptions of what should go into an undergraduate course, they were “somewhat differing” as they, with wonderfully polite restraint, write in A Survey of Modern Algebra: the Fiftieth Anniversary of its Publication (Birkhoff and Mac Lane 1992, p. 26).

6.3 The “Marginal Interest” and Interrupted Pace of Research

For a course as novel as modern algebra, a department would probably not create it unless there was strong interest, especially in the form of one or more professors steadily at work in the field. However, Birkhoff has remarked that interest was not strong even halfway through the era under consideration here: “Even in 1929 its concepts and methods were still considered to have marginal interest as compared with those of analysis in most universities, including Harvard" (Birkhoff 1973, pp. 760–782).

Furthermore, such interest as there was may have seemed erratic around 1929. In a 1938 survey by E. T. Bell, he points to a dip in total numbers of papers in algebra by Americans, comparing the decade from 1908 to 1917 to the next decade when there about 30% fewer (Bell 1938).[5] Bell convincingly argues that this is not due to World War I or the Great Depression.[6] Algebraists of the time may have noticed this dip by noticing that there were fewer speakers in the sessions where they gave research talks. But were there enough such observant algebraists for this to lessen the pressure for an undergraduate course? The extent of this was probably small.

6.4 Polarizing Anxieties Related to Research

Around the turn of the century, more and more young American mathematicians were successfully pursuing research, often traveling to Germany and other European countries for Ph.D. training. The American Mathematical Society (AMS) was dedicated to furthering the increasing interest in research, but there was no such organization dedicated especially to undergraduate teaching. A movement arose to create such an organization, and it was successful in creating the Mathematical Association of America in 1915. This was not exactly a split of the mathematical community, since mathematicians could, and did, belong to both groups, but the birth of this new organization was not without stress either. Indeed, Frederick Rickey calls it "contentious" in his colorful essay on Benjamin Finkel (http://sections.maa.org/ohio/ohio_masters/finkel.pdf). Novel course proposals such as modern algebra probably tapped into the negative emotions stimulated, in the minds of some, by the debate over the creation of a new mathematical organization. Keeping advanced courses such as modern algebra out of the undergraduate domain might help keep such anxiety at bay. There would have been no need to be explicit about such underlying anxieties in the (unrecorded) debate over modern algebra. A perfectly good argument could have been made that such a course would be quite hard for undergraduates and have minimal practical applications or connections to

[5] E. T. Bell is mostly known today as the author of popular histories of mathematics, histories which are said by some historians to be flawed, especially by exaggeration. In addition, his dissertation was said by a few to be murky (Reid 1993). But the vetting of his research work must also include the approval of his dissertation advisers, which included the distinguished Frank Nelson Cole and Cassius Jackson Keyser, the referees of his 250 papers, and the jury which awarded him the Bôcher Prize for research in 1924. Finally, he was president of the Mathematics Association of America from 1931 to 1933, a distinction which would have been difficult to achieve if there had been a serious grapevine of doubt about him. Anyhow, it seems unlikely that the work for which we quote him—entirely a matter of counting and telling the difference between algebra and other subjects such as analysis—would suffer from murk or exaggeration.

[6] Declines and increases were quite different in different branches of algebra whereas the war and the Depression affected everyone equally.

workhorse courses such as analysis. Therefore, the argument would proceed, such a course would be more appropriate as a graduate course.

7 The Fading of Reasons to Hesitate

7.1 Advances in Research in Algebra

7.1.1 The Arrival of Emmy Noether and Abstract Ring Theory

Dedekind introduced the idea of an ideal already in 1871, but his ideals were concrete in that they were subsets of the real or complex numbers. A modern and fully abstract definition of a ring and its associated ideals finally appeared in Abraham Fraenkel's 1914 paper (Corry 1996, p. 208). The abstract approach was impressively elaborated in the 1920s by Emmy Noether when she studied numbers and polynomials in a unified way, namely as instances of an abstract axiomatization of a ring. Her work turned abstract algebra into a central subject within research mathematics. But it took some doing: The abstract concept of a ring was still so alien to mathematicians in 1921 that Noether, in an article, felt it necessary to prove that there is only one multiplicative identity element in a commutative ring![7] (Corry 1996, p. 227).

7.1.2 The Development of Abstract Field Theory

An abstract definition of a field can be found already in Lehrbuch der Algebra (Weber 1895), but Weber does not exploit the abstraction. The practical initiation of abstract field theory is usually dated by historians to (Steinitz 1910). In this 1910 paper, Steinitz introduces prime fields, separable elements, and the degree of transcendence of an extension field, and he proves that every field has an algebraically closed extension field.

7.1.3 The Use of Group Theory in Quantum Mechanics

In 1927, we have the onset of what some physicists called "group theory disease" (Seitz et al. 1999). That referred to a sudden appearance of numerous fundamental papers and books linking group theory to quantum theory. It started with Eugene Wigner's 1927 paper (Wigner 1927). This paper was the first to observe that group representations offered an essential key to the complex order of the atomic spectra. Wigner declared that this work allowed him to solve problems in atomic

[7] At least her proof was only in a footnote.

spectroscopy "almost without calculation." Between Dec. 1927 and June 1928, Wigner and his childhood friend John von Neumann submitted three additional papers on using group theory to understand spectra and quantum mechanics (*Zeitschrift für Physik* 47(1928):203; and 49(1928):73; and 51(1928):844) (Weyl 1931). See also https://www.perimeterinstitute.ca/videos/effectiveness-group-theory-quantum-mechanics.

Working independently, Hermann Weyl got similar results to those of Wigner, but pushed further and published the book *Gruppentheorie und Quantenmechnaik*. See also the English version (Weyl 1931) translated by H. P. Robertson, and E. P. Dutton, 1931.

We do not argue that anyone wanted to teach these quantum applications to undergraduates or that many mathematics instructors followed physics even well enough to notice them. However, in 1936, the distinguished American mathematician Marshal Stone wrote a glowing review of the work of Wigner, von Neumann, and Weyl in the *Bulletin of the AMS*, a widely read mathematics journal. Most specialists in group theory would have seen it because of the prestige of Stone and the *Bulletin* and thus have been well armed to join any departmental debate as to whether group theory had any important applications.

In his review, Stone states that this work revealed group theory to be not merely a "very useful mathematical device" for understanding the foundations of quantum theory, but something "far more profound" that yields insights not otherwise available. In particular, the quantum nature of atoms and molecules "would be only partially intelligible without the representations of an abstract group by means of linear transformations" (Stone 1936).

7.2 The Appearance of Textbooks

7.2.1 The Appearance of Bartel van der Waerden's *Moderne Algebra*

A key step toward the teaching of the modern conception of algebra to graduate students was taken by Bartel van der Waerden in his book *Moderne Algebra* (van der Waerden, Moderne Algebra, 1931). Saunders Mac Lane describes its importance succinctly: "This beautiful and eloquent text served to transform the graduate teaching of algebra, not only in Germany, but elsewhere in Europe and the United States" (Mac Lane 1997). This book presents a thoroughly abstract and structural view of algebra as pioneered by E. Noether, E. Artin, and others. (Some of those others included van der Waerden himself.)

As a book for undergraduates, van der Waerden leaves something to be desired. For one thing, connections of algebra to other areas of mathematics or to science are not emphasized. For example, after van der Waerden gives the definition of a vector space over a (possibly skew) field of scalars, his first examples are not the ones which would be most useful for beginning students or those with applied interests, namely, the Euclidean spaces. Instead he gives an entirely algebraic example (one of the particular interest in number theory): "Examples of vector spaces are all extension fields of a field K and more generally, of all rings R containing a skew field K as long as the unit element of K is also a unit element of R" (van der Waerden, Modern Algebra, 1953). It is hard to imagine that the appearance of this book had much direct promotional effect on modern algebra for undergraduates: its level and single mindedness were more appropriate for fairly mathematically sophisticated students. However, it exercised a galvanizing and unifying effect on algebraists which probably suggested that an undergraduate presentation would be next on the agenda.

7.2.2 The Appearance of Birkhoff and Mac Lane's *A Survey of Modern Algebra*

The book which provided that undergraduate view of modern algebra was *A Survey of Modern Algebra* (Birkhoff and Mac Lane, A Survey of Modern Algebra, 1941) by Garrett Birkhoff and Saunders Mac Lane. This book expresses much of the abstract and structural point of view of van der Waerden. For example, the version of Galois Theory presented in Chap. 15 is Artin's version based on automorphisms of field extensions.

However, Birkhoff and Mac Lane's *A Survey of Modern Algebra*, a remarkable work that was in wide use for at least a half-century, is hardly just an easier version of *Moderne Algebra*. Noetherian rings, prominently featured in *Moderne Algebra*, are not mentioned by Birkhoff and Mac Lane. In addition, they did not banish all notions of applications from their book and included some ideas from physics. And their primary examples of vector spaces are the Euclidean spaces which numerical analysts and other applied workers eventually came to appreciate once the computer made numerical analysis such a thriving area. Birkhoff and Mac Lane themselves have written a revealing description of the events and influences that led to their first manuscript (Birkhoff and Mac Lane, A Survey of Modern Algebra: the Fiftieth Anniversary of its Publication, 1992). They each started from their own drafts which were different enough that they had "lively discussions" in the process of compromising toward a joint draft. How lively? According to their account, their senior colleague J. L. Coolidge warned them in advance about the possibility of an "uncongenial collaboration".[8] Altogether, *A Survey of Modern Algebra* was a new creation, painstakingly wrung from separate visions and showing a good grasp of what might be taught to strong undergraduates, a grasp based on trying the material at Harvard.

A final example that distinguishes Birkhoff and MacLane's book from that of van der Waerden can be found in their preface: an agenda we would characterize as courageously disputatious. Their preface asserts that they wish to further a "broader interpretation of the significance of modern algebra" than merely setting

[8] A possibility the authors say never came about.

forth “this formal or ‘abstract’ approach.” In our opinion, the broader goals they refer to can be thought of as attempting to refute some reservations expressed by Klein’s and others about abstraction. Their preface speaks of their desire to display the “imaginative appeal” of the subject, to constantly use “familiar examples,” to encourage “the student’s power to think for himself,” and to reveal “applications to other fields: higher analysis, geometry, physics, and philosophy.” These goals make an interesting refutation to our earlier quoted reservations of Klein’s.

7.3 *“Marginal and Interrupted” Give Way to a Plenitude of Research in Four Important Journals*

In our earlier remarks about the slow and interrupted advance of modern algebra research in America, it should be noted that our characterization rests partly upon work and remarks by Bell and Birkhoff that pertain to research in the era up to around 1929. What about the next few decades, decades not addressed by Bell’s work or Birkhoff’s remarks? We now describe some data that suggest an upsurge in research in algebra in those years.

We examined the four leading American research journals of the early twentieth century (*American Journal of Mathematics*, *Annals of Mathematics*, *Transactions of the American Mathematical Society*, *Bulletin of the American Mathematical Society*) to compare 1900 to 1951 with regard to how many American authors wrote or collaborated on at least one article in modern algebra or in a subject so closely related to modern algebra that it could scarcely exist in the same form without the existence of modern algebra (e.g., algebraic topology, parts of algebraic geometry, etc.). Our presumption is that such authors would have non-trivial knowledge of modern algebra and would be likely to support the teaching of the subject to undergraduates in their department. In 1900, 20% of all American authors in these journals wrote at least one article which was algebraic or closely related to algebra. In 1951, the figure was 43%, a near majority of authors.

7.4 *The Fading of the Polarizing Anxieties Related to Research*

The split we earlier alluded to between mathematicians who did research and those who did not must surely have been ameliorated in the aftermath of the founding of the Mathematical Association of America (MAA), an organization firmly dedicated to the undergraduate domain. This amelioration would have been greatly assisted by the cooperative attitude of the AMS. As the years went by and it became clear that the research elite and the AMS were no threat to those who did not do research, it would have been easier to evaluate modern algebra courses on their merits rather than as a shift in the power relations of the two branches of the mathematical community. Even hard feelings fade over time if the reasons for those feelings are not constantly evoked anew.

As for modern algebra being difficult, that was probably a fair assessment, and it may still be the case, but perhaps it came to matter less. It is quite plausible that, starting perhaps as early as during World War II and surely by the immediate postwar era, a greater sense of ambition took hold of the American mathematical community and made things seem possible that had seemed daunting before. As is sometimes asserted, it was a time in America of "leaping tall buildings in a single bound." In our opinion, American mathematicians would have been quite happy to be caught up in this feverish but joyful spirit of enterprise.

7.5 Leadership from the Top

However, beyond this mere generality about the joyful spirit of enterprise, there were specific circumstances that would have motivated the mathematical community to take on harder, even risky, endeavors around midcentury. This was the challenge to which the mathematical community, and indeed all the sciences, was "invited"[9] namely that it should include, in its fundamental purpose, being a keystone of the national interest in matters of both war and peace. This was not a new idea. One sees it already in 1863 with the formation of the National Academy of Sciences (NAS). However, a more concrete and urgent version was propelled into the political arena for debate in 1942 when Senator Harley Kilgore introduced legislation to mobilize American scientists to work in the national interest. This kicked off the debates leading to the formation of the National Science Foundation (NSF) in 1950 (The National Science Foundation 1994). Those discussions clearly suggested that financial support would be forthcoming for ambitious curricular proposals.

But would American mathematicians naturally think of modern algebra in connection with these funding programs? The leadership of the profession certainly did. June 1953 saw the very first instance of NSF funding for college-level mathematics when NSF established a summer institute at the U. of Colorado in which Emil Artin gave an 8-week course on "Modern Developments in Algebra" (Krieghbaum 1968, pp. 6–8).

We should also consider the work of the Committee for the Undergraduate Program in Mathematics (CUPM), a structure in the mathematical community created in 1959. In 1965, this committee recommended, for the first time, all undergraduate mathematics study groups, rings, and fields and also some linear algebra (Committee on the Undergraduate Program in Mathematics 1965).

[9] Those familiar with this era will surely agree that this was an invitation one could not refuse, considering the cold war anxieties inflaming the news and the influx of federal money flowing to academia.

Of course, the forces boosting modern algebra we have just described, involving the NSF and CUPM, occur at a point when most of the institutions in our sample already had adopted modern algebra (see Tables 1 and 2). However, our institutions are not a random sample—they are fairly well-known institutions, many of them accustomed to leadership—and we believe it is possible that many less well-known institutions might have waited longer to institute modern algebra courses. If so, would they have found that the new role for mathematics as part of the national interest, and the CUPM recommendations, were positive factors in the climate of opinion in which they deliberated about modern algebra courses? This would be well-worth research based on a sample of less well-known institutions.

What about the arrival of emigré European mathematicians in the 1930s, often fleeing Nazi persecution, prior to the Second World War? Many of these had done distinguished work in algebra including Emil Artin, Reinhold Baer, Richard Brauer, John Von Neumann, Emmy Noether, and Hermann Weyl. One may ask whether these researchers may have nudged the prospects for some undergraduate courses in America as well. We think the argument for this is weak. Giants of research and their accomplishments are ruefully but routinely ignored by undergraduate instructors anxious that their students could not deal with sophisticated matters coming from the research world. But could the students of these giants of research have made a difference? Of the 69 American PhD students the aforementioned emigres had, only a few came onto the teaching scene as early as the 1940s (according to the Mathematics Genealogy Project)—too late to affect Tables 1 and 2. We regard the role of the emigres as being unlikely for the matter at hand.

7.6 Conclusion

In conclusion, we have shown that there was a considerable delay at many colleges in adopting modern algebra courses at the undergraduate level, but that there were factors urging caution that could have seemed persuasive at the time. The fact that these factors dissipated over the time interval during which courses in modern algebra became more widespread makes it quite plausible that these factors did influence the slow adoption rate. However, we have no documentary evidence to confirm this. Our discussions of these factors must remain footholds for research.

8 Related Wider Issues

The issues raised in this article touch on many wider questions which range too extensively to be satisfactorily addressed in the limited scope of this article. We mention only three—the reader can doubtless think of others.

1. How, and on what time scale, have other research innovations made their way (or not) into undergraduate curricula? Is there a "normal" process here or is the particular nature of the innovation and the particular nature of the curricular environment at the time decisive?
2. Our work relies strongly on course catalogs. They are surely quite reliable about what was and was not offered, but they say nothing about the thinking behind the inclusion or exclusion of a particular course. Can other materials shed light on what mathematics instructors of the time were thinking and why they had those opinions?
3. In many areas of American life, World War II was a watershed. Did it operate that way in the thinking of American mathematicians about undergraduate matters? For example, was there a multiplicity of new curricular ideas that modern algebra was part of?

References

Bell, E. T. (1938). Fifty Years of Algebra in America, 1888-1938. In R. C. Archibald (Ed.), *Semicentennial Addresses of the American Mathematical Society* (Vol. 2, pp. 1-34).

Birkhoff, G. (1973). Current Trends in Algebra. *The Amerian Mathematical Monthly, 80*, 760-782.

Birkhoff, G., & Mac Lane, S. (1941). *A Survey of Modern Algebra.* MacMillan, New York.

Birkhoff, G., & Mac Lane, S. (1992). A Survey of Modern Algebra: the Fiftieth Anniversary of its Publication. (K. V. Parshall, Ed.) *The Mathematical Intelligencer, 14*, 26-31.

Committee on the Undergraduate Program in Mathematics. (1965). *A General Curriculum in Mathematics for Colleges.* Washington DC: Mathematics Association of America.

Corry, L. (1996). *Modern Algebra and the Rise of Mathematical Structures.* Boston: Birkhäuser.

Galois, E. (1846). Ouevres mathématique. *Journal de mathématiques pures et appliquées, T*(XI).

Gray, J. J., & Parshall, K. H. (Eds.). (2007). *Episodes in the History of Modern Algebra (1800-1950).* American Mathematical Society.

Jordan, C. (1870). *Traité des substitutions et des équations algébriques.* Gauthier-Villars, Paris.

Klein, F. (1979). Development of Mathematics in the 19th century. In R. Hermann (Ed.), *Lie Groups: History, Frontiers and Applications* (M. Ackerman, Trans., Vol. IX, pp. 1-361). Math. Sci. Press. Brookline, MA.

Kleiner, I. (1986). The Evolution of Group Theory. *Mathematics Magazine, Vol 59*(4 (Oct.)), p. 195.

Kleiner, I. (1996). The Genesis of the Abstract Ring Concept. *American Mathematical Monthly, 103*, pp. 417-424.

Kleiner, I. (1999, Aug-Sept). Field Theory: From Equations to Axiomatization. *The American Mathematical Monthly, 106*(7), 677-684.

Krieghbaum, H. R. (1968). *To Improve Secondary School Science and Mathematics Teaching.* Retrieved July 1, 2015, from http://files.eric.ed.gov/fulltext/ED032226.pdf

Mac Lane, S. (1997, March). van der Waerden's Modern Algebra. *Notices of the AMS*, 321.

Meyer, W. J. (2013). *The Cajori Two Project*. Retrieved May 19, 2015, from http://matcmp.ncc.edu/taormij/cajori_two/

Miller, G. A. (1898). On the Commutator Groups. *Bulletin of the American Mathematical Society, 4*, pp. 135-139.

Moore, E. H. (1895). The Group of Holoedric Transformations Into Itself of a Given Group. *Bulletin of the American Mathematical Society, 1*, pp. 61-66.

Ore, O. (1931). Some recent developments in abstract algebra. *Bulletin of the AMS, vol. 37, pp 537-548*, 537-548.

Parshall, K. H., & Rowe, D. E. (1991). *The Emergence of the International Mathematical Research Community 1876-1900: J. J. Sylvester, Felix Klein and E. H. Moore.* American Mathematical Society.

Peirce, B. (1881). *Linear Associative Algebra, Amer. J. of Math. 4:1 (1881)* (Vol. 4).

Reid, C. (1993). *The Search for E. T. Bell.* Washington, D. C: Mathematical Association of America.

Seitz, F. (1999), Vogt E, Weinberg AM. Wigner, E.P. *Biographical Memoirs of Fellows of the Royal Society*. 46: 577–592.

Steinitz, E. (1910). Algebrasche Theorie der Körper. *Journal für die Reine und Angewandte Mathematik, 137*, pp. 167-309.

Stone, M. (1936). Four Books on Group Theory and Quantum Mechanics. *Bulletin of the AMS, 42*(3), 165-170.

The National Science Foundation. (1994, July 15). *The National Science Foundation: A Brief History*. Retrieved May 21, 2015, from https://www.nsf.gov/about/history/nsf50/nsf8816.jsp

Thiele, R. (2007). Van der Waerden und Artin - Der Weg Zur Modernen Algebra. In K. a. Reich, *Emil Artin (1898-1962) Beiträge zu Leben, Werk und Persönlichkeit* (pp. 137-167). Augsburg, Germany: Erwin Rauner Verlag.

van der Waerden, B. L. (1931). *Moderne Algebra.* Berlin: Springer.

van der Waerden, B. L. (1953). *Modern Algebra.* Ungar.

Weber, H. (1893). Die Allgemeinen Grundlagen der Galois'schen Gleichungstheorie. *Mathematische Annalen, 43*, pp. 521-549.

Weber, H. (1895). *Lehrbuch der Algebra.* Braunschweig: F. Vieweg und Sohn.

Weyl, H. (1931). *The Theory of Groups and Quantum Mechancs.*

Wigner, E. (1927). Einige Folgerungen aus der Schrödingersche Theorie für die Termstrukturen. *Zeitschrift fur Physik, 43*, 624.

Wussing, H. (1984). *The Genesis of the Abstract Group Concept.* (A. Shenitzer, Trans.) MIT Press. Cambridge, MA.

History as a Source of Mathematical Narrative in Developing Students' Interpretations of Mathematics

Po-Hung Liu

Abstract The stereotypical images for mathematics and narrative are: mathematics is logical, certain, and objective, whereas narrative is often perceived as emotional, indefinite, and subjective. Nonetheless, mathematics and narrative can actually be deeply intertwined. Mathematical narrative is a form of narrative used to communicate or construct mathematical meaning or understanding. The role of history of mathematics in the teaching and learning of mathematics has been discussed for several decades, and the relationship between mathematics and narrative has also drawn much attention among scholars. In this paper, I will address in what way and to what extent history of mathematics can trigger college students' mathematical narrative in the courses of "History of Mathematics" and "Mathematics in Ancient Civilizations." The study reports on how these students redefined mathematics after receiving history-based courses of mathematics. Based on students' interpretations of mathematical knowledge, this chapter addresses the role of history of mathematics in developing students' capacity for mathematical narrative.

1 Introduction

In recent years, there has been a growing awareness of the role of the history of mathematics in the teaching and learning of mathematics (Clark et al., 2016; Clark and Thoo, 2014; Fauvel and van Maanen, 2000). Though scholars have several reasons to call for integrating history into the school mathematics curriculum, the role of history of mathematics can be categorized into three mutually complementary and supplementary aspects: a replacement role, a reorientation role, and a cultural role (Clark et al., 2016). The replacement role refers to replacing the finalized

P.-H. Liu (✉)
National Chin-Yi University of Technology, Taichung, Taiwan ROC
e-mail: liuph@ncut.edu.tw

M. Zack, D. Schlimm (eds.), *Research in History and Philosophy of Mathematics*, Annals of the Canadian Society for History and Philosophy of Mathematics/Société canadienne d'histoire et de philosophie des mathématiques,
https://doi.org/10.1007/978-3-030-95201-3_10

and polished mathematical products by emphasizing the mental processes leading to them. The aim of the reorientation role is to challenge the learner's and the teacher's conventional perception of mathematical knowledge through changing what is familiar to something unfamiliar. The cultural role treats mathematical knowledge as an integral part of human intellectual history. One of the common doctrines of all three roles is that there is a need to demonstrate the dynamic, potentially fallible, and socio-cultural nature of mathematics.

2 A Brief Account of What History Is

Regarding the well-known and frequently debated question "What is history?", there are different ways of interpreting history such as "history as literature," "history as science," "history as social science," and "history as a narrative." A brief review of these arguments is given below to clarify the position of the present study.

2.1 History as Literature

Traditionally, the study of history was considered a form of literature in which historical events were not only connected chronologically but were also intertwined with legends about significant figures to form a scenario played out on the historical stage. In this manner, the writing of history is very much like a storytelling activity based on historical documents. Several classical historians such as Herodotus and Titus Livius were great literary writers. A more recent example is Edward Gibbon's prose style and ironic wit in *History of the Decline and Fall of the Roman Empire* (the first volume was published in 1776), which made history more appealing. Though Gibbon's interpretive observations and conclusions came to no longer be acknowledged by the academic community as new evidence emerged, his work remains a comprehensible historical literary entry for understanding the period. A radical case may be seen in Spence's *Emperor of China; Self-Portrait of K'ang-hsi* (Spence, 1974). Spence wrote in the first person through the eyes of K'ang-hsi, making history more vivid. However, such a subjective approach to interpreting history received criticism for its freedom of flexibility. As Wesseling (1998) indicated, history is not simply literature, and a historian should only write statements for which he/she has evidence. A call for the introduction of scientific methods in history then appeared in a reaction to this situation.

2.2 History as Science/Social Science

The aim of a historian's work is to truthfully, if not correctly, understand the past. With the rise and success of natural science in interpreting the world in the seventeenth and eighteenth centuries, a positivist belief pervaded all branches of knowledge including historiography. In the early nineteenth century, historians tried to promote their activities as a science in its own right through frankly interpreting and reconstructing historical texts and documents. Leopold von Ranke, generally recognized as the father of the scientific historical school, set a model for training historians in systematic, critical research methods. He claimed that:

> To history has been given the function of judging the past, of instructing men for the profit of future years. The present attempt does not aspire to such a lofty undertaking. It merely wants to show how it essentially was (*wie es eigentlich gewesen*). (Ranke, 2011, p. 86)

For achieving such an objective end, in Ranke's mind, history should not be judged according to determined contemporary values or ideas. Rather, it has to be understood on its own terms by empirically establishing how things really were (Boldt, 2014). The school of scientific history (e.g., the Marxist economic model, the French ecological/demographic model, and the American "cliometric" methodology) was based not on new data, but on new models or new methodology. This group of scientific historians was confident that they would, given time, succeed in solving major problems of historical explanation and in identifying general laws of historical changes, as can be seen in Le Roy Ladurie's claim that "history that is not quantifiable cannot claim to be scientific" (Le Roy Ladurie, 1979, p. 15). However, this scientific or social scientific approach was also challenged because it was contradicted by the very powerful notion of the unity of science (Wesseling, 1998). Jenkins (2009) pointed out that history cannot actually provide any objective idea of the past and can never be literally true, fair, and non-positioned because 99% of what happened in the past was never recorded. Munslow (2010) also noted that history cannot claim to be straightforwardly scientific because "historians can know the past for what 'it most likely was and means' through the process of empirical justification and argument to the best explanation" (Munslow, 2010, p.35). Being a completely objective historian is, therefore, unattainable because the historian is also a product of history and of society (Carr, 1961).

2.3 History as a Narrative

Literature and science can be seen as the two ends of the knowledge spectrum. But which is appropriate for history? The answer might lie somewhere in between. Burke (2007) claimed that if the past is a foreign country, then historians might be regarded as translators between the past and the present. History rests not only on storytelling and the statistical analysis of facts, but also on a meaningful interpretation of the past. A new trend categorized as the narrative school and concerned with the feelings, emotions, behavior patterns, and states of mind of the mass population rather than of the elite has become fashionable in recent years. The use of narrative is inevitable when historians try to describe what was going on in people's heads in the past, and what it was like to live in the past (Stone, 1979).

Narrative history "is taken to mean the organization of material in a chronologically sequential order and the focusing of the content into a single coherent story, albeit with sub-plots" (Stone, 1979, p. 3). Therefore, there is no doubt that, no matter how objective the historians were, their interpretation relies on an a priori decision. When historians reconstruct the history of the time, the past facts are transformed into a discourse. It contains metaphors, plot settings, ideologies, and elements that were probably not a part of the past, so history also takes on the characteristics of fiction. In this manner, the narrative historian's task is like that of a storyteller, albeit one for which only those events considered relevant could have a place in their work (Mandelbaum, 1967); hence, it might be methodologically unsound (White, 1984).

Stone (1979) acknowledged that the adoption of the narrative approach is not without its problems, but its challenge is not new. Namely, arguments based on selective incidents are unpersuasive and a rhetorical device is not a scientific proof. Nonetheless, the fundamental reason for the shift from the analytical to the descriptive mode in historiography is a philosophical belief about the way humans choose to interact with the past. It may be useful to convert the definition of history from "a discovery of the events that have happened in the past" into "an exploratory way of understanding the past." Considering the reality that history is an unending dialogue between the present and the past (Carr, 1961) carried on by means of a lost-and-found translation of historical facts (Burke, 2007), it is hard to cast the many-faceted nature of history into any single word. For the time being, "narrative" might serve as a shorthand code word for all that is going on (Stone, 1979).

3 Mathematics and Narrative

Unlike the ambiguity and controversy regarding the true nature of history, mathematics is no doubt a science. But, is it a kind of narrative? A first glance would lead us to reject such a seemingly bizarre claim, yet a second thought might find their similarity. There is a trend of paying attention to the interrelationship between mathematics and narrative (Corry, 2009; Doxiadis and Mazur, 2012). Considering the pessimistic notion that higher mathematics has become practically irrelevant to the mainstream cultural discourse, the British Society for the History of Science (BSHS) organized a summer course, "Mathematics and Narrative," in 2009 to question this opposition between mathematics and culture by exposing the deep and complex interconnections between the mathematical and narrative modes of thought. It is hoped that this line of thought "will not only open up novel perspectives on the history and philosophy of mathematics, but, it is hoped, will help bridge the chasm that now separates mathematics from mainstream thought" (BSHS, 2009). Mathematical narrative is a way of using narrative to communicate or construct mathematical meaning or understanding (Su et al., 2016). A mathematical narrator usually has to, by means of certain metaphors, induce or promote learners', listeners', and readers' mathematical imaginations and interpretations to communicate or construct mathematical meaning or understanding. Chen and Liu

(2018) proposed three main relational forms between mathematics and narrative: (a) mathematics as a narrative, (b) mathematics in narrative, and (c) narrative in mathematics teaching. Note that there is no clear-cut line drawn between the three relational forms. A well-known ancient model of mathematical narrative is the plot in Plato's *Meno* in which Socrates intellectually guides an illiterate slave boy to double the area of a square. To demonstrate his idea of reincarnation to Meno, Socrates uses a geometry problem as a medium (which is a use of mathematics in narrative). Then through a series of careful questioning, counterexamples are arranged to remind the slave boy to acknowledge his misconception about the dimension of length and area—mathematics as a narrative (presenting mathematics as a narrative). The slave boy eventually understands that the required double area is the square of the diagonal of that given square. The whole dialogue can be seen as an exemplification of "narrative in mathematics teaching," though teaching geometry was not Socrates' original intension.

3.1 History as a Source of Mathematical Narrative

Taking the aforementioned arguments that emphasize the narrative features of history and mathematics into account, history may enrich mathematical narrative in terms of the two relational forms: "mathematics as a narrative" and "mathematics in narrative." History of mathematics is a typical example of involving history in the process of "mathematics as a narrative" because its aim is at conveying mathematical conceptions through investigating and revealing the development of mathematical concepts and propositions. In the history of mathematics, each mathematical proposition comes with a background story and each story involves the developmental processes of mathematical facts. However, the story alone does not constitute a narrative. A meaningful engagement is required. While reading or learning the history of mathematics, readers or learners would be attracted to go on only by using both sides of their brains, logical and mathematical thinking on the one side and emotional and imaginative thought on the other. This meaningful engagement is the source of the pleasure of mathematical narrative (Thomas, 2002).

On the other hand, history may play a role in the form of "mathematics in narrative," mathematics in fiction in particular, either by helping the narrator to establish the structure of the whole story or by providing ideas for creating the protagonists and plots. Mathematics in fiction includes at least two kinds of genre, mathematical fiction, and popular mathematics texts, both of which have become more plentiful in recent years. Corry (2009) discussed the triangular relationship among the discourses of mathematics, history of mathematics, and mathematics in fiction. The targeted readers of mathematical texts are usually within a specialized and professional community, whereas mathematics in fiction is for the general public. The readership of the history of mathematics involves both groups. The distinction of readership differentiates the language used in the texts and the expected attitudes of readers (Table 1).

Table 1 Comparison of language, readership, and attitude in different texts

Features	Texts		
	Mathematics	History of Mathematics	Mathematics in Fiction
Language used	Formal/Formalized	Discursive/Natural	Discursive/Natural
Expected readers	Specialized	Specialized/General	General
Expected attitudes	Critical	Critical	Suspension of Disbelief

The texts of mathematics and mathematics in fiction occupy the two ends of the narrative spectrum, the language of mathematical texts being formal or formalized, and that of mathematics in fiction being natural or discursive. Texts on the history of mathematics may lie in between because they not only use discursive and natural language to introduce the timeline, historical context, and people involved in the development of mathematical knowledge, but also employ formal and formalized language alternatively to demonstrate the original or adaptive forms of mathematical propositions. With the truth-finding nature of history, a critical attitude is required of readers while reading the texts of history of mathematics. Nonetheless, in some sense, the interpretation of history inevitably goes with a storytelling approach. Corry (2009) reminds us that an over-dramatized style would appear if the undocumented legends were used too often (e.g., Bell's (1937) *Men of Mathematicians*).

The following sections will report on how Taiwanese college students defined mathematics before and after a history of mathematics course. The purpose is to investigate in what way the history of mathematics may develop their narrative ability in interpreting mathematics.

3.2 The Context of Two History of Mathematics Courses

In line with the reorientation role and cultural role indicated by Clark et al. (2016), two history-based courses of mathematics were offered for Taiwanese college students. One was "Mathematics in Ancient Civilizations" (MiAC), focusing on mathematics in several ancient civilizations including Egypt, Babylon, Greece, Arabia, India, and China, a course for non-mathematics majors. The other was "History of Mathematics" (HoM), introducing the development of mathematics from ancient to modern times, a course for mathematics majors.

MiAC is a general education course offered for all students except mathematics majors. It was not as mathematics intensive as HoM because about one-half of the participating students were from the college of business and liberal arts. Rather, the objective of the course was to compare the similarities and differences in mathematical knowledge of various ancient civilizations, in such aspects as their numerical systems, forms of problems, strategies and tools for solving equations, and mathematical thought hidden within cultures. For instance, students learned

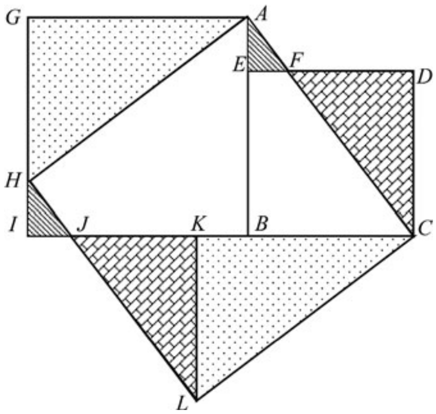


Fig. 1 Liu Hui's Out–In Mutual Patching Technique

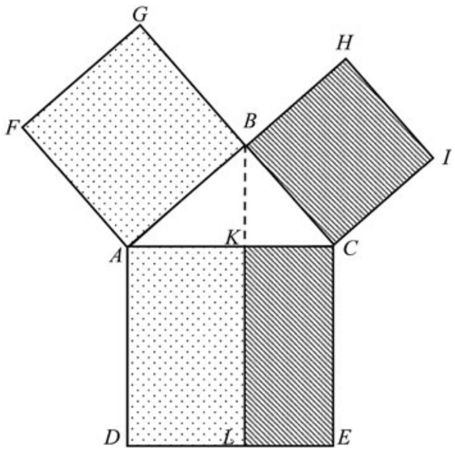


Fig. 2 Euclid's proof of the Pythagorean Theorem

different proofs of the Pythagorean Theorem in ancient China and Greece. Figure 1 illustrates Liu Hui's "Out-In Mutual Patching Technique." Given a right-angle triangle *ABC*, while extending the side of *BC* to get a square BCDE and extending the side of *AB* to get a square *AGIB*, by shifting the triangles *AGH*, *FDC*, *HIJ* to their seemingly congruent counterparts *CKL*, *JKL*, *AEF* and leaving other parts fixed, we can assert that $\overline{AC}^2 = \overline{AB}^2 + \overline{BC}^2$. It appears that this approach is totally empirical. On the contrary, as shown in Fig. 2, for asserting $\overline{AC}^2 = \overline{AB}^2 + \overline{BC}^2$, Euclid proves that the area of *ABGF* is equal to that of *ADLK* and the area of *BCIH* is equal to that of *CKLE* by means of showing the congruence of various triangles, a totally deductive approach.

HoM, on the other hand, delved wider and deeper by adding two additional parts to the course. About one-third of the course introduced ancient mathematics of Egypt, Babylon, Greece, and China. In particular, the inductive fashion of ancient Chinese mathematics was compared to the deductive style of ancient Greece. Another one-third focused on mathematical achievements during Song and Yuan period in China, and European mathematics in the Middle Ages and the Renaissance. It was stressed in the course that this point in time could be seen as the watershed of the development of mathematics between the East and the West. Chinese mathematics reached its peak around 1300 and then gradually declined. European mathematics and science in the Renaissance paved the road for the scientific revolution in the seventeenth century and opened the door to modern mathematics. The last part of the course focused on the rigorization of mathematical analysis, algebra, and geometry. Especially impressed upon the students were the facts that the invention of group theory was caused by the effort of looking for the general solution of quintic equations, and questioning the necessity of the fifth Euclidean postulate resulted in the creation of non-Euclidean geometry.

3.3 Changes in Students' Views

To gain further understanding of the effectiveness of MiAC and HoM, a qualitative questionnaire was used to reveal the changes in students' views about mathematics. Within this study, which aims to demonstrate that history can be an appropriate source of mathematical narrative for interpreting mathematics, this section mainly reports on what ways the two history-based courses helped students redefine mathematics.

MiAC Students' Post-instruction Views There were 52 students in MiAC. In the beginning of the course, they mostly held that mathematics is a quantitative subject involving numbers and computations for solving problems in daily life and workplaces, a typical instrumentalist view. Near the end of the course, they did not totally abandon the instrumentalist views, but were likely to interpret the essence of mathematics in different ways. The student Bou-Guo initially considered mathematics a subject representing things by numbers or symbols, such as the measurement of a piece of land. In the post-instruction interview, he viewed mathematics as an on-going academic knowledge:

> Interviewer: Now you are saying that mathematics is an academic knowledge.
> Interviewee: Yes!
> Interviewer: What do you mean by that?
> Interviewee: Academic knowledge is always evolving. It is not fixed. It keeps going forward and developing.
> Interviewer: What made you change?
> Interviewee: The things that I had learned in the course, especially the origin of mathematical concepts.

Another student, Zan-Zer, also regarded mathematics as a science of computation at the outset but in the post-instruction interview stresses the significance of logical reasoning:

> Interviewer: You said that mathematics is about computation and logic. Could you give me an example?
> Interviewee: For instance... Plato's story in ancient Greece. Someone proposed a solution and he was led back to the origin... He then got a contradiction and realized his reasoning was wrong.

The story mentioned by Zan-Zer was the plot addressed in Plato's *Meno* in which Socrates guided an illiterate slave boy to realize his mistakes in doubling the area of a square. Zan-Zer also mentioned the three crises in the development of mathematics. It seems that many students were impressed by these crises in the history. Ya-Tin, a female student, emphasized logical reasoning in the post-instruction interview. Her response also raised the mathematical crises:

> The first crisis led to the discovery of irrationals. If it hadn't happened, we would not have thought of inventing square root. Like 1.414... It is not so convenient for us to use infinite decimal point. And that paradox... could the barber shave himself? That is too complicated. How could they come up with this problem? This kind of problem could only be thought of by mathematicians. Me? No way!

Though Ya-Tin questioned the value of mathematical crises in reality, she confessed that the debate over the crises might have helped the development of mathematics.

MiAC addressed various mathematical cultures between the East and the West. Knowing the way students interpreted the cultural similarities and differences may shed more light on the effectiveness of the course. In addition to the ancient Egyptian and Babylonian numerical systems, students were attracted by the dialectical culture of mathematical debate in ancient Greece. Student Ming-Fon expressed an extremely conservative attitude toward mathematics. He saw mathematics merely as a useless school subject, the only utility of mathematics being for daily shopping. However, in the post-instruction interview, he was particularly impressed by the ancient Greek academic tradition:

> Interviewer: Among the different approaches to mathematics in ancient civilizations introduced in the course, which one impressed you most?
> Interviewee: Ancient Greece! There were many schools of thought. They competed with each other, even competed with their own people.... Then they discovered some interesting things and promoted the development of mathematics.
> Interviewer: About the same time, there were also several schools of thought in China, such as Confucianism, Taoism, and Moism.
> Interviewee: Not much mathematics in their thoughts.

Ming-Fon further proposed Zeno's paradoxes to endorse his view. It appears that, in minds of our students, compared to ancient Greek philosophy, traditional Chinese philosophy paid little attention to mathematics, which could be the reason that none of them selected ancient Chinese as the topic of their term papers.

HoM Students' Post-instruction Views The 122 students in the HoM were all mathematics majors demonstrating more sophisticated views of mathematics than the MiAC students. At the outset, their definition of mathematics could be summarized by the following five categories:

- A scientific tool: an instrument solving problems through quantification
- A symbolic system: a system consisting of numbers, computations, and symbols
- A science of pattern: a science of looking for numerical and geometrical patterns
- A rigorous science: a knowledge involving definition, axioms, logic, and proofs
- A language of nature: a language revealing the secret of nature

The last three categories reflect HoM students' experience in their professional training and were rarely mentioned by MiAC students. After the course, HoM students expressed a wide range of views difficult to summarize into categories. However, their post-instruction responses exhibited a common feature that their focus had shifted from the knowledge structure of mathematics to the evolving processes of mathematics in the making. Student Zen-Shuan, who at first regarded mathematics as the mother of science, which is logical, rigorous, and flawless, changed his view:

> Interviewee: Instead of saying mathematics is the mother of science, I would say it is the seed of science.
> Interviewer: Why do you say that?
> Interviewee: I found the development of mathematics was not as perfect as I had thought. Some theories even contain flaws, such as the fifth postulate of Euclidean geometry. It might be arguable. Besides, the debate about the foundation of mathematics among the three schools might contradict each other.
> Interviewer: Then, why is it the seed of science?
> Interviewee: Number theory was unexpectedly employed in cryptography. We have no idea when the works of pure mathematicians would be applied to the real world. The seed of science is a more appropriate term than the mother of science because we can never know when seeds will sprout.

Student Jia-Ying originally considered the purpose of mathematics is to encapsulate all kinds of phenomena into a formula, but she said at the end of the course:

> Now I realize that mathematics is not all about boring formulas. It could be very intuitive, abstract, logical, and even beyond your imagination. These are all parts of mathematics. Mathematics has no limitations and cannot be framed as long as it can justify itself…… Mathematicians have seen

> something that we have never seen. Our thought should not be restricted by our eyes.

Asked about the reason of changing her views, she said:

> There are some problems that I used to think could only be solved by using formulas, but they had been solved by intuition in the past. The methods employed in ancient Greece and India were more intuitively understandable... Abstract mathematics in the 19^{th} century, like non-Euclidean geometry and group theory, laid the foundation of physics of 20^{th} century. Number theory, the purest mathematics in Hardy's mind, became the tool of encoding and decoding in World War II. I will not let mathematics be restricted by my mind. This is what I have learned in this course.

It appears students' post-instruction views, to a great extent, were influenced by the historical figures and events introduced in the course. In particular, the story of Fermat's Last Theorem moved many students and made them think that mathematics is a human endeavor. In the beginning, student Zi-Shuan saw mathematics as a scientific tool. He still held that mathematics is a tool at the end of the course, yet that it involves sequential efforts of mathematicians:

> Interviewer: For now, how would you define mathematics?
> Interviewee: Actually, I previously felt that mathematics was nothing but those theorems and formulas that have been proved and could be applied to physics and chemistry. I never paid serious attention to thinking about this until the last week
> Interviewer: Last week?
> Interviewee: In the last week of the course, we watched the documentary of Fermat's Last Theorem. Regardless of the usefulness of this theorem, Fermat claimed he had proof already. Why did those mathematicians bother to take the challenge to look for proof? In particular, Andrew Wiles spent seven years proving it. I finally realized that mathematical achievement is evolutionary. The unsolved problem is a gap, statically standing there, waiting for dreamers or challengers. So long as this gap gets filled, mathematics could move forward. Each great mathematician is one of the road workers.

In Zi-Shuan's mind, mathematics is more like a human endeavor than a scientific instrument. This humanistic view can also be seen in student Yu-Chen's post-instruction interview in which she claimed:

> Interviewee: Prior to taking this history course, I had been majoring mathematics for three and half years. I used to think mathematics is constituted by rigorous definitions, theorems and proofs. [I] appreciated its rigor and astonishing mathematical proofs, pursuing the so-called absoluteness. But in this history of mathematics course, what I learned and appreciated was the context of its evolution, either the chronological development of mathematics in general, or algebra and geometry in particular.

> In this manner, mathematics demonstrates its humanity and warmth.
> Interviewer: What do you mean by humanity and warmth?
> Interviewee: We get to know how the people thought of mathematics at that time. Their works were not only for solving problems, but also for pursuing the truth, for the sake of philosophy and religion as well as for practical reasons.

Yu-Chen then took the instance of the philosophical debate about the foundation of the infinitesimal between Newton and Berkeley to endorse her view. The aforementioned excerpts of students' statements may suggest that, in some sense, the HoM course did trigger these mathematics majors to rethink, reflect, and reevaluate what mathematics really is.

4 Conclusion

Mathematics is often seen as logical, certain, and objective, whereas narrative is emotional, indefinite, and subjective. This chapter demonstrates that history of mathematics may develop college students' abilities in mathematical narrative in interpreting mathematical knowledge. The pre- and post-instruction comparison reveals that the HoM and MiAC courses did have an influence on students' interpretations of mathematics, but the effectiveness may have varied depending on the students' mathematical maturity. Though the MiAC students still held instrumentalist perspectives, their accounts had shifted from mathematics as knowledge ready to be used to mathematics as a logically evolving system. Probably attributed to their limited experiences in doing and thinking about mathematics, a sophisticated understanding was lacking in their views. Nonetheless, the contrast in mathematical culture between the East and the West did impress upon those students. On the other hand, the HoM students' post-instruction claims demonstrated a deep reflection on the nature of mathematics. These mathematics majors initially considered mathematics as a rigorous, logical, and matured science, whereas they later tended to stress the organic development of mathematics by addressing its evolutionary, debatable, and humanistic aspects. As compared to the MiAC students, the HoM students were more likely to endorse their views by recalling the episodes they had learned in the class.

This study investigated whether history of mathematics would trigger Taiwanese college students' reflection on mathematics and imagination to redefine mathematics. Results suggest a positive effect for some students. Another meaningful question worth asking is that once having had this experience of reinterpreting mathematics, would these students thereafter be more likely to develop an imaginative and productive state of mind about mathematical knowledge? It is hoped that further research sheds more light on this issue.

References

Bell, E. T. (1937). Men of Mathematics. New York: Simon & Schuster.

Boldt, A. (2014). Ranke: Objectivity and history. Rethinking History, 18(4), 457–474.

British Society for History of Science (2009). Mathematics and narrative: Bring Mathematics back to the cultural mainstream. A summer course at the Central European University in Budapest, July 20–24, 2009. Retrieved from https://www.bshs.org.uk/mathematics-and-narrative-bringing-mathematics-back-to-the-cultural-mainstream Cited 23 July 2019.

Burke, P. (2007). Lost (and found) in translation: A cultural history of translators and translating in early modern Europe. European Review, 15(1), 83–94.

Carr, E. H. (1961). What is History? London: Palgrave Macmillan.

Clark, K., Kjeldsen, T. H., Tzanakis, C., & Wang, X. (2016). History of mathematics in mathematics education: Recent developments. In L. Radford, F. Furinghetti, and T. Hausberger (Eds.). Proceedings of the 2016 ICME Satellite Meeting of the International Study Group on the Relations Between the History and Pedagogy of Mathematics (pp. 135–179). Montpellier, France: IREM de Montpellier Editors.

Clark, K., & Thoo, J. B. (2014). Introduction to the special issue on the use of history of mathematics to enhance undergraduate mathematics instruction. Problems, Resources, and Issues in Mathematics Undergraduate Studies, 24(8), 663–668.

Chen, T.-S. & Liu, P.-H. (2018). Analysis of the metaphorical role of mathematics in students' writing. In T. Sibbald (Ed.), Teaching interdisciplinary mathematics (pp. 9–27). Champaign, IL: Common Ground Research Network.

Corry, L. (2009). Calculating the limits of poetic license: Fictional narrative and the history of mathematics. Configurations, 15(3), 195–226.

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

Fauvel, J., & van Maanen, J. (Eds.) (2000). History in mathematics education: The ICMI study. Dordrecht, The Netherlands: Kluwer Academic.

Jenkins, K. (2009). At the limits of history. London: Routledge.

Le Roy Ladurie, E. (1979). The territory of the historian. (Ben Reynolds and Sian Reynolds, Trans.) Chicago: University of Chicago Press.

Mandelbaum, M. (1967). A note on history as narrative. History and Theory, 6(3), 413–419.

Munslow, A. (2010). The future of history. New York: Palgrave Macmillan.

Ranke, L. (2011). The theory and practice of history. New York: Routledge.

Spence, J. D. (1974). Emperor of China: Self-portrait of K'ang-Hsi. New York: Alfred A. Knopf.

Stone, L. (1979). The revival of narrative: Reflections on a new old history. Past and Present, 85, 3–24.

Su, Y.-W., Horng, W.-S., Huang, J.-W., Chen, Y.-F. (2016). Mathematical narrative: From history to literature. 2016 ICME Satellite Meeting of the International Study Group on the Relations Between the History and Pedagogy of Mathematics, July 2016, Montpellier, France.

Thomas, R. (2002). Mathematics and narrative. Mathematical Intelligencer, 24 (3), 43–46.

Wesseling, H. (1998). History: Science or art? European Review, 6(3), 265–267. White, H. (1984). The question of narrative in contemporary historical theory. History and Theory, 23(1), 1–33.

White, H. (1984). The question of narrative in contemporary historical theory. History and Theory, 23(1), 1–33.

Thoughts on Using the History of Mathematics to Teach the Foundations of Mathematical Analysis

Fairouz Kamareddine and Jonathan P. Seldin

Abstract This paper discusses ideas for a different approach to teaching the foundations of mathematical analysis. The main idea is to avoid the use of what Keith Devlin in 2005 called "formal definitions," which are definitions that nobody can understand without working with them. For students without mathematical maturity, these definitions can be difficult to understand and use. This paper discusses an approach that uses the history of mathematics to first develop fundamental concepts and only introduces formal definitions after the concepts are understood. The audience for this approach is third-year undergraduate students. Several examples are provided in the paper.

1 The Problem

Our long experience of teaching mathematicians and computer scientists indicates to us that many students of moderate ability have trouble understanding mathematics. A significant part of this problem appears to be that many students think that mathematics is a game of manipulating symbols without worrying about their meaning. They then try to learn mathematics by memorizing the manipulations. Students who try to learn this way will be lost if they forget one detail. This problem seems to start with beginning algebra and becomes more serious after the transition from computational courses like calculus and elementary linear algebra to more advanced courses, which are more theoretical.

F. Kamareddine
School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK
e-mail: f.d.kamareddine@hw.ac.uk

J. P. Seldin (✉)
Department of Mathematics and Computer Science, University of Lethbridge, Lethbridge, AB, Canada
e-mail: jonathan.seldin@uleth.ca

M. Zack, D. Schlimm (eds.), *Research in History and Philosophy of Mathematics*, Annals of the Canadian Society for History and Philosophy of Mathematics/Société canadienne d'histoire et de philosophie des mathématiques,
https://doi.org/10.1007/978-3-030-95201-3_11

At a meeting of the Canadian Mathematical Society in 2005, Keith Devlin gave the plenary address (Devlin 2005) in which he discussed what he called *formal definitions*, which are definitions that nobody can completely understand without first working with them. This contrasts with dictionary definitions of words which are understood easily by anyone who looks up the definition. These formal definitions are common in advanced mathematics courses, including analysis, but for students without mathematical maturity, they can be difficult to learn and understand.

This is the motivation for the authors of this paper writing a book, tentatively entitled *A Non-formal Introduction to Mathematical Analysis*, which is intended for a third-year undergraduate course to introduce students to analysis while avoiding the use of formal definitions as much as possible. Some reactions to a preliminary draft of this book suggest that it might be useful to precede this book with content that provides a bridge between computational courses and the more theoretical courses which follow in the normal mathematical curriculum.

Many textbooks for this type of bridge course begin with formal logic, apparently assuming that this should indicate to students what proofs are and why they are needed. Starting this way appears to place too much of a burden on the students. It would be better for the formal logic to occur later in the book when examples from earlier parts of the course can be used to illustrate how formal logic can serve to analyze them. We believe that this is the best approach for students to recognize the *meaning* of the logical symbols.

For example, analysis textbooks often start with the axioms for a complete ordered field, give the $\varepsilon - \delta$ definition of the limit of a function and the $\varepsilon - N$ definition of the limit of a sequence, and then start deriving theorems. This seems to be the wrong order for many students, since it does not help the students learn why these axioms and theorems are important and what analysis is really about. For both the bridge and analysis courses, it seems that approaching the material in a historical order rather than a logical order may help student learning. Even if we do not have an account of the history that is beyond controversy, we think taking a historical approach to the material will make sense to students and will enable students to visualize what they will be studying.[1]

This paper contains some examples of this approach to learning analysis via a historical ordering of the concepts.

[1] For the importance of visualization in the learning of mathematics, see Nardi (2014) as well as the publications of eric.ed.gov on the Role of Visualization in the Teaching and Learning of Mathematics and especially Mathematical Analysis as for example in the article of Miguel de Guzman in https://files.eric.ed.gov/fulltext/ED472047.pdf.

2 Mathematics as *Reasoning*

Mathematics beyond elementary arithmetic is about *reasoning*. This can be demonstrated by showing that simple algebraic equations can be solved in words without using modern algebraic notation. For example, suppose that the problem is to find a number such that five more than three times the number is twenty. We can solve this as follows:

$$\begin{aligned}\text{Suppose}:&\ \text{Five more than three times the number is twenty.}\\ \text{Then}:&\ \text{Three times the number is fifteen.}\\ \text{Hence}:&\ \text{The number is five.}\end{aligned}$$

This shows that modern algebraic notation is not necessary to do algebra. The key point is that whereas problems in arithmetic tell the student what operations to perform on what numbers, problems in algebra require that reasoning is used to determine those operations and numbers. Furthermore, before the European Renaissance, this was the only way to do algebra, since the modern notation used today began as a kind of shorthand during the Renaissance.

This solution is actually a short proof of the proposition: *If five more than three times the number is twenty, then the number is five.* When the solution is checked, the converse: *if the number is five, then five more than three times the number is twenty* is proved.

To pass from this solution to one using the modern notation involves two steps: first, use the standard notations for arithmetic operations in the sentences in the solution:

$$\begin{aligned}\text{Suppose}:&\ 5 + 3 \times (\text{the number}) = 20.\\ \text{Then}:&\ 3 \times (\text{the number}) = 15.\\ \text{Hence}:&\ \text{The number} = 5.\end{aligned}$$

For the final step, introduce a letter to represent the number: let n be the number. Then

$$\begin{aligned}\text{Suppose}:&\ 5 + 3n = 20.\\ \text{Then}:&\ 3n = 15.\\ \text{Hence}:&\ n = 5.\end{aligned}$$

Note that the role of the notation is to help students notice the patterns needed to complete the reasoning successfully.

It is customary to leave out the logical words out of these solutions, so that the solution might be written as:

$$5 + 3n = 20.$$

$$3n = 15.$$

$$n = 5.$$

This may be part of the reason that there are students who do not realize that there is reasoning involved here.

Leaving out the logical words may also lead students to confusion if the equation has no solution.

Consider, for example, the equation $2\sqrt{x+1} = \sqrt{4x+5}$. If the steps of the solution are written without the logical words, the computation appears as:

$$2\sqrt{x+1} = \sqrt{4x+5}$$

$$4\,(x+1) = 4x+5.$$

$$4x+4 = 4x+5.$$

$$4 = 5.$$

Students who do not understand that this is supposed to be reasoning may feel completely lost at this point. But now inserting the logical words gives:

$$\text{Suppose}: 2\sqrt{x+1} = \sqrt{4x+5}$$

$$\text{Then}: \quad 4\,(x+1) = 4x+5.$$

$$\text{Thus}: \quad 4x+4 = 4x+5.$$

$$\text{Hence}: \quad 4 = 5.$$

This may help students realize that what has been shown is *if* $2\sqrt{x+1} = \sqrt{4x+5}$ *then* $4 = 5$. Since $4 \neq 5$ this shows that $2\sqrt{x+1} \neq \sqrt{4x+5}$ for every x.[2]

The conclusion of this example is the result of a *proof by contradiction*: when a contradiction is deduced from an assumption, the conclusion is that the assumption

[2] This example comes from the Preface of Middlemiss (1952).

is false. The idea is closely related to the debating tactic of *reductio ad absurdum*, or *reduction to an absurdity*.

3 Why Is There a Problem About Calculus?

Consider the function $y = f(x) = x^2$ and the problem of finding its derivative $f'(2)$ at $x = 2$. The process for finding this derivative is to consider first the difference quotient

$$\frac{\Delta y}{\Delta x} = \frac{f(x) - f(2)}{x - 2} = \frac{x^2 - 2^2}{x - 2}$$

This quotient is evaluated as follows:

$$\frac{x^2 - 2^2}{x - 2} = \frac{(x + 2)\,(x - 2)}{x - 2} = x + 2$$

where the last step is valid only if $x \neq 2$, but when "the limit is taken" at $x = 2$ by substituting 2 for x yields 4. Since the conclusion that the difference quotient is equal to x + 2 is based on the assumption that $x = 2$, it appears that an illegal step was taken.

Calculus textbooks justify this last step by saying that the difference quotient is not being *evaluated* at x = 2, but instead, the computation is a limit as $x \to 2$, thus

$$\frac{dy}{dx} = \lim_{x \to 2} \frac{\Delta y}{\Delta x}$$

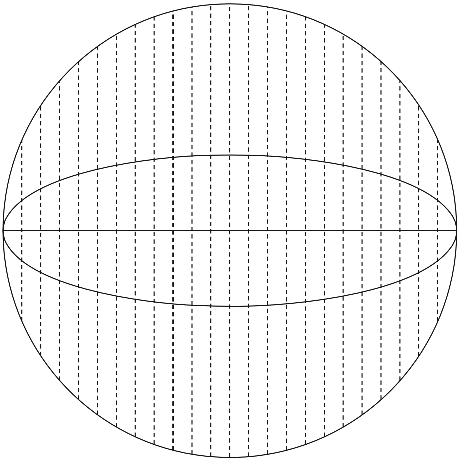
and that all this really means is that if x is *near* 2, then the difference quotient is *near* 4, and further, that is possible to get the difference quotient as close to 4 as desired, simply by choosing x sufficiently close to 2. These standard textbooks typically give a number of rules for evaluating limits with very little justification for these rules.

In the seventeenth century, although the Greek's geometric methods were admired and applied to calculate quadratures, fear of the infinite was abandoned. For example, Johannes Kepler used infinitesimals to determine the area of an ellipse and viewed the circumference of a circle as an infinite sided regular polygon.

Kepler noted that for a circle of radius a and an ellipse of radiuses a and b: (see Fig. 1)

- The ratio of each vertical line within the circle to the vertical line within the ellipse is b/a.
- The area of each of the circle/ellipse is the infinite sum of vertical lines contained in the circle/ellipse.

Fig. 1 Circle and Ellipse



Hence, the ratio of the area of the circle to that of the ellipse is also b/a. This means that

$$\text{the area of the ellipse} = \left(\pi a^2\right) \times (b/a) = \pi ab.$$

Further uses of the infinitesimal as in Kepler's method were developed to calculate more advanced tangents and quadratures and to relate tangents with quadratures. However, there was still a lack of a formal development of the concepts and rules in question. Part of this was filled independently by the work of Newton and Leibniz on the calculus. Infinite series/sequences were crucial for their calculus methods, and they in this way legitimized the use of infinite processes. Although Newton's and Leibniz's work had introduced differentiation and its reverse (integration), they were still short of formulating the definition of limit. Although both Newton and Leibniz invented the calculus independently in the late seventeenth century, they did not have a good explanation of what they were doing. Newton, for example, described what we call

$$\lim_{x \to 2} \frac{\Delta y}{\Delta x}$$

as its "ultimate value," or its value at "the instant of its disappearance." This explanation does not satisfy the normal requirements of a mathematical definition.

In fact, in the 1730s, the inadequacy of this definition was brought home by George Berkeley in Berkeley (1951). Berkeley was an Anglican bishop, and he had become disturbed for the soul of Edmond Halley (the astronomer for whom Halley's Comet is named) who was proclaiming himself an atheist on the basis of Newton's physics. Berkeley happened to be an excellent satirist, and he jumped on Newton's explanation of the limit of the difference quotient as its value at "the instant of its disappearance" by asking why it should not be called "the ghost of a departed quantity." His idea was to suggest that anybody who was prepared to accept Newton's calculus should also have no trouble accepting theology. Despite all this criticism, infinite processes and calculus as developed by Newton and Leibniz continued to be very much in use.

New developments by Euler on the generalization of function were followed by attempts at explaining the notion of limit. This was followed in the nineteenth century by Cauchy's work that successfully combined the new ideas of function and limit in order to give a rigorous formulation of the calculus explaining convergence, divergence, and continuous functions. More rigor was put into explaining the calculus, its notion of limit and continuous function, and even the core on which it is based (the real numbers). This theoretical work is called *analysis*, or the *arithmetization of analysis*. In fact, all analysis can be derived from a set of axioms about the real numbers.

The arithmetization of analysis is one of the greatest intellectual achievements of human history. However, it does not seem possible to appreciate its greatness just by looking at the latest version of the theory itself. It is necessary to consider the entire process by which this theory developed. This process involved major changes in the way mathematicians looked at their subject, the sorts of things they studied, and even at how mathematics could be justified. It really began over 2000 years ago in ancient Greece. In order to fully understand and appreciate analysis, it is important to begin with the ideas of the ancient Greeks.

4 Proofs as Sequences of Statements

In all the solutions of the equations of Sect. 2 of this paper, the arguments consist of sequences of statements. However, proofs were not always sequences of statements. In ancient Greek mathematics, there was a time when a proof consisted of a diagram; understanding the proof meant seeing the diagram the right way.

For example, knowing that the area of a rectangle is the product of its 2 adjacent sides, consider the following proof that the area of a parallelogram is the base times the altitude:

Fig. 2 Area of the parallelogram

This diagram (Fig. 2) shows that the area of the parallelogram is the same as the area of the rectangle with the same base and altitude.

Next, consider the following proof of the formula for the area of a triangle:

This diagram (Fig. 3) shows that the area of the triangle is half of the area of the parallelogram.

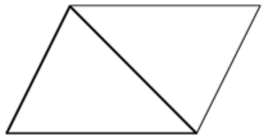


Fig. 3 Area of a triangle

Consider also the following proof (Fig. 4) that the three angles of a triangle (in a Euclidean plane) add up to two right angles:

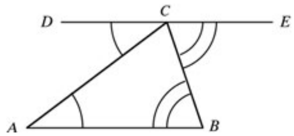


Fig. 4 Angles of a triangle

Similarly, the following diagram (Fig. 5) shows the Pythagorean Theorem:

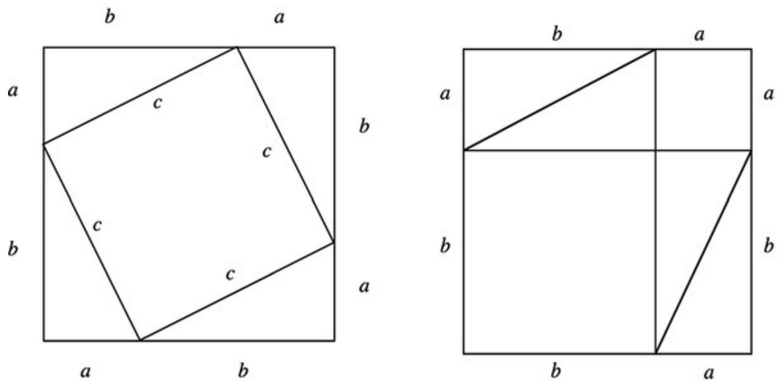


Fig. 5 The Pythagorean Theorem

For this last one, if A is the area of the right triangle with legs a and b and hypotenuse c, then the diagrams show that

$$4A + c^2 = 4A + a^2 + b^2,$$

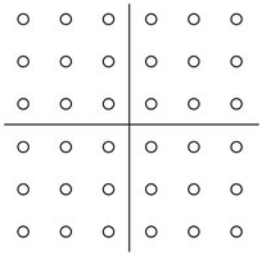
from which $c^2 = a^2 + b^2$ follows easily.

Proofs by diagrams were prevalent in early mathematics. So how and when did proofs become sequences of statements? One possible way this might have happened is given by Wilbur Knorr (1975, pp. 179–180). It is based on Knorr's reconstruction of how the ancient Greeks might have first proved that the side and diagonal of a square are incommensurable; i.e., there is no length that evenly divides both. This is equivalent to the irrationality of $\sqrt{2}$. It is known that the Ancient

Greeks did arithmetic using diagrams of pebbles arranged in certain ways, so these proofs of arithmetic properties can be illustrated by *pebble diagrams*.

Figure 6 is a pebble diagram illustrating the result that *every even square number is a multiple of 4*:

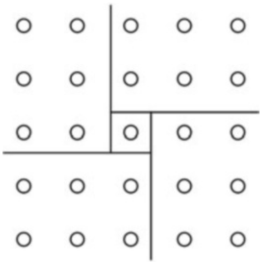
Fig. 6 Pebble diagram for even squares



This only shows one case ($6^2 = 4 \times 3^2$), but it is clear that, at least in principle, a diagram like this one can be created for every even square number. In modern algebraic symbols, this can be shown by $(2n)^2 = 4n^2$.

Figure 7 shows in a similar way that *every odd square number is one more than a multiple of eight*:

Fig. 7 Pebble diagram for odd squares



That every odd square is one more than a multiple of four is immediate from this diagram; that it is one more than a multiple of 8 follows because each of the four rectangular areas is one number times the next one, and of those two numbers one is always even. In algebraic symbols, we have

$$(2n+1)^2 = 4n^2 + 4n + 1 = 4\left(n^2 + n\right) + 1 = 4n\,(n+1) + 1.$$

Now, by the last of these two numerical results, it follows that the sum of two odd squares is never a perfect square, since it must have the form $(8n + 1) + (8m + 1) = 8(m + n) + 2$, which cannot be a perfect square since it is divisible by 2 but not divisible by 4.

Now, consider an isosceles right triangle (i.e., a right triangle whose two legs are equal). Such a triangle can be seen as part of a square with a diagonal. We know by the Pythagorean Theorem that if the legs are both l and the hypotenuse is h, then

$$h = \sqrt{2}l.$$

If h and l are commensurable, then there is a unit for which both h and l are positive integers. Assume that h and l are positive integers. Incommensurability can be proven by showing that this assumption leads to a contradiction. So the assumption implies that h is either even or odd.

- If h is even, then so is l, and this means that the unit can be doubled and still have a length which evenly divides h and l. This is equivalent to cutting each of h and l in half. Obviously, this cannot be continued indefinitely, so eventually there must be a length which divides h an odd number of times.
- But if h is odd, then one leg must be even and the other must be odd, so the triangle is not isosceles.

This is a contradiction, since an even number cannot equal an odd number. So the hypothesis that there is a length that exactly divides both h and l an integral number of times must be false.

Note that there is no way to illustrate this result with one or more diagrams. The argument is a sequence of statements. This is Knorr's explanation of how proofs became sequences of statements. Further details are given in Seldin (1990).

5 Bridge Course: Mathematical Theories

Many students are confused by abstract mathematical theories when they are introduced to them for the first time. They may be able to follow some of the proofs, but they seem not to understand what is really going on and why mathematicians find proofs interesting.

For this reason, we believe that a bridge course should begin with an example. One which students at this stage in their mathematical development are supposed to know but often do not: the mathematics of fractions as formal quotients of positive integers. Students are supposed to know that cross-multiplication is the method to determine whether or not two fractions are equal:

$$\frac{m}{n} = \frac{p}{q} \text{ if and only if } mq = np.$$

The idea is to take the relation determined by cross-multiplication not as equality, which means identity of fractions, but as a relation of "having the same value." This gives

$$\frac{m}{n} = \frac{p}{q} \text{ if and only if } m = p \text{ and } n = q.$$

Now define the relation $\sim$ between fractions by

$$\frac{m}{n} \sim \frac{p}{q} \text{ if and only if } mq = np.$$

It is easy to show that ~ is an equivalence relation. A (positive) rational number as an equivalence class can be defined from the equivalence relation ~, and it is easy to prove that the usual properties follow from this definition, including the fact that unlike in the positive integers, each (positive) rational number has an inverse with respect to multiplication.

Double-entry bookkeeping can be used in the form of accounts, each of which has a credit and a debit to look at formal differences of positive integers. Thus, an account with credit c and debit d will represent the formal difference $c \ominus d$. Two accounts $c \ominus d$ and $e \ominus f$ will be equal if and only if they are identical:

$$c \ominus d = e \ominus f \text{ if and only if } c = e \text{ and } d = f,$$

while the relation $\cong$ for "having the same value" will be defined by

$$c \ominus d \cong e \ominus f \text{ if and only if } c + f = d + e.$$

This theory with respect to addition is very similar to fractions (formal quotients) with respect to multiplication. The main difference between this theory and the theory of fractions discussed earlier is that the positive integers have an identity element with respect to multiplication but not with respect to addition. This means that in this case, an identity and an inverse are being added with this theory.

While these theories are not isomorphic, they are extremely similar. It is instructive for the students to compare them. The desire is for the students to see that most of the words in the results (in fractions with respect to multiplication and in accounts with respect to addition) are identical.

From here it is possible to define a commutative cancellation semigroup to consist of a set S together with a binary operation which satisfies closure, commutativity, associativity, and the cancellation law. We then define a binary operation on ordered pairs of elements of S by

$$(x, y) \approx (z, u) \text{ if and only if } x \circ u = y \circ z.$$

This definition corresponds to the definition of "having the same value" for fractions with multiplication and accounts with addition, and the resulting theory is so close to both of those theories that it can be "instantiated" to either of them. This shows that this formal theory allows us to add an identity (if there is not one already present in S) and an inverse for every element.

The hope is that this collection of examples will help students understand why mathematicians study formal theories and find them useful.

6 Bridge Course: Set Theory and Logic

For the rest of the bridge course, we plan to begin by using equivalence relations and equivalence classes to explain cardinal numbers of sets,[3] including infinite sets, and derive some of the elementary properties of transfinite cardinal numbers. We would, of course, take the equivalence relation on sets to be the relation of the existence of a bijection between the sets. The properties of transfinite cardinal numbers include sets of real numbers which are countable and some which are uncountable. The specific countable sets that should be included are the integers, rational numbers, and algebraic numbers. The important uncountable sets to be studied include the real numbers between 0 and 1 and the larger set of single-valued real-valued functions of real numbers.

From here it is possible to use the language of set theory to define functions. We believe that putting the results on transfinite cardinal numbers before these definitions is, in fact, the historical order in which these ideas were introduced and these definitions entered mathematics.

The next step is to use formal logic to evaluate proofs and informal arguments that have occurred previously in the bridge course.

7 Analysis: Limits

Analysis is based on limits, so any course on analysis must begin by considering them.

First, students should be reminded of some of what they (are supposed to) know about limits of functions and sequences. It would be possible to create the first theories of limits of functions and sequences by taking some of the standard limit theorems as axioms and from those deriving the rest. However, doing this formally may be confusing to some students.

Another approach is to begin with looking at two examples of proofs of limit results from ancient Greek geometry. The results are about circles. The theorems of interest are as follows:

1. A theorem of Euclid which says that the areas of circles are to each other as the squares of their radii
2. Archimedes' theorem giving the area of a circle

According to Knorr (1975, pp. 311–312), both theorems were proved by Eudoxus, using a method that came to be known as the *method of exhaustion*.

---

[3] Given the lack of a proper universal set in axiomatic set theory, we cannot formally define cardinal numbers as equivalence classes, but we can indicate how the idea behind this is approximately that of such equivalent classes to be covered.

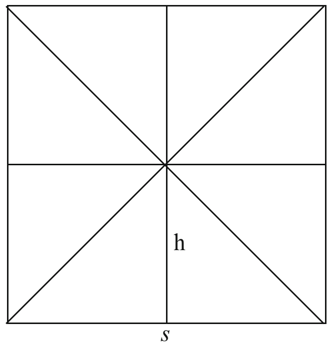


Fig. 8 Square for the proof of the area

For both theorems, a general formula for the area of a regular polygon is needed. To see how this formula is developed, let us look at a square in a new way (Fig. 8). Instead of finding the area simply by taking s^2, start with the bottom of the four triangles obtained by taking the diagonals. The altitude of each triangle is $h = \frac{1}{2}s$, so the area of the triangle is given by $\frac{1}{2}hs$. If we add the areas of all four triangles, we get

$$A = \frac{1}{2}h(4s) = \frac{1}{2}hp$$

where $p = 4s$ is the perimeter. Substituting the values of h and p in terms of s into this formula yields s^2, which is what was expected.

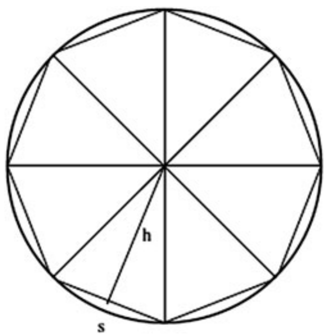


Fig. 9 Octagon for the proof of the area of a circle

Now consider a regular octagon (Fig. 9). If the octagon is divided into eight triangles, each of whose area is $\frac{1}{2}hs$. If all eight triangles are combined noting that $p = 8s$, the result is

$$A = \frac{1}{2}h(8s) = \frac{1}{2}hp.$$

This should be enough to establish the result that the area of any regular polygon is one-half the altitude to a side times the perimeter, or $\frac{1}{2}hp$.

Now what about a circle? If the number of sides of a regular polygon is repeatedly increased, the perimeter will approach the circumference of a circle, and the altitude will approach the radius of the circle. This suggests that the formula for the area of a circle should be

$$A = \frac{1}{2}rC$$

And since π is defined to be the ratio of the circumference of a circle to its diameter, or what amounts to the same thing, the ratio of the circumference to twice its radius, we have

$$\pi = \frac{C}{2r}$$

from which our familiar formula $C = 2\pi r$ follows. If this is substituted into the formula above for the area of a circle, the formula becomes the familiar

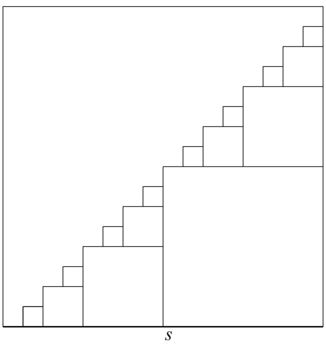
$$A = \frac{1}{2}r\,(2\pi r) = \pi r^2.$$

This must have seemed obvious to the ancient Greeks from an early period in the history of their geometry. But how could they prove it?

At one time, some of them argued that a circle is a regular polygon with infinitely many sides, but they eventually decided that this kind of reasoning is inadequate for mathematical proofs. For just because regular polygons with an increasing number of sides seem to be approaching a circle, deducing the formula for the area is not automatically justified. Evidence like this can be misleading.

Consider the following example: The length of the stepped line (Fig. 10) is clearly $2s$ no matter how many steps there are. But as the number of steps increases, the stepped line seems to approach the diagonal, and the length of the diagonal is $\sqrt{2}s \neq 2s$.

Fig. 10 The stepped line



Thus, although it must have been obvious to the ancient Greeks that the area of a circle is, in our terms, given by the formula

$$A = \frac{1}{2} r C$$

where r is the radius and C is the circumference, it was a long time before a proof was given of this fact. And before that proof was given, the following result, attributed to Eudoxus by Knorr (1975), appeared in Euclid's *Elements* (Heath 1926) as Proposition 2 of Book XII:

Proposition *(The areas of) circles are to one another as the squares on their diameters.*

The original proof of this result, along with a translation into a modern proof, is given in Seldin (1991).

Getting back to the area of a circle, the result on its area was finally published by Archimedes in a book called (in English) "Measurement of a Circle," which can be found in Heath (1912). The statement of this result, which in this book is Proposition 1, is as follows:

Proposition *The area of any circle is equal to a right-angled triangle in which one of the sides about the right angle is equal to the radius and the other to the circumference of the circle.*

This original proof, along with its transformation into a modern proof, is given in Seldin (1991).

At the meeting of the Canadian Society for History and Philosophy of Mathematics (CSHPM) in 1990, Judith Grabiner presented as an invited address (Grabiner 1997) on the work of the Scottish mathematician Colin Maclaurin in which she showed a page from a treatise on analysis in which he proved a limit result using exactly this approach.

He began by saying that if the limit were not the value given in his theorem, it must either be greater or less, and then deriving a contradiction from both of the latter two assumptions.

As Seldin showed (Seldin 1990), if these proofs are rewritten in algebraic notation and the absolute value function is used, the proofs are transformed into $\varepsilon - N$ proofs about the limits of sequences.

Once the $\varepsilon - N$ definition is available for the limit of a sequence, it is fairly easy to justify the $\varepsilon - \delta$ definition of the limit of a function. For saying that $|x - c| < \delta$ for some, δ is roughly like saying that $x > N$ for some N; i.e., that x is close to infinity.

Additional work with proofs using the $\varepsilon - \delta$ and $\varepsilon - N$ definitions for limits provides a foundation for considering the completeness axiom for the real numbers.

8 The Real Numbers

When Newton and Leibniz first published on calculus, they talked about "quantities," which were some-thing of an amalgam of the numbers and magnitudes of the ancient Greeks. These quantities included the positive integers and fractions. To get to the modern approach, the rational numbers need to be extended to the real numbers. This requires adding the Axiom of Completeness to the rational numbers.

To explain the Axiom of Completeness, it is useful to use the $\varepsilon - N$ definition of the limit of a sequence and to consider a strictly monotonically increasing infinite sequence with an upper bound. Intuitively, such a sequence must have a limit.

Now if a candidate for the limit is not an upper bound of all the terms of this sequence, it is not the limit of the sequence; for let ε be the difference between this candidate and a term of the sequence that is greater, and then all terms after this latter term will be further from the number than ε.

On the other hand, if the candidate is an upper bound and there is a lower upper bound, then let ε be the difference between these two upper bounds, and then all the terms of the sequence are further than ε away from the given number.

It follows that the limit of this sequence *must* be the least upper bound of the terms of the sequence.

But then, if we are limited to the rational numbers, there will be sequences like this that do not have limits. Dedikind (1965, p. 13) gives a strictly monotonically increasing sequence of rational numbers converging to $\sqrt{2}$. In the rational numbers, this sequence has no limit.

This makes it clear that to guarantee that every sequence of this kind has a limit, it is sufficient to have the Axiom of Completeness, which says that *every nonempty set of numbers that has an upper bound has a least upper bound.*

From this point, students can study some material on uniform convergence and then consider the Riemann integral (which involves a different kind of limiting process from that for functions and sequences).

9 Conclusion

The goal of our research project is to create materials (a book or pair of books) that exposes students to analysis while avoiding introducing definitions before they are used. This is part of our effort to avoid what Devlin called *formal definitions*.

To make the book as useful as possible, we have considered adding material on the topology of the real line. We also plan to add appendices on two methods of constructing the real numbers from the rational numbers: the first method using Dedekind cuts and the second method using Cauchy sequences. These examples are intended to motivate students to study more complex topics in a real analysis course.

We believe that this way of approaching the introductory analysis topics material has a good chance of helping some students understand this material who would not previously have been able to master it.

References

Berkeley G (1951) The Analyst: or A Discourse Addressed to an Infidel Mathematician. Wherein it is examined whether the object, principles, and inferences of the modern Analysis are more distinctly conceived, or more evidently deduced, than religious Mysteries and points of Faith. First printed in 1734. In Luce A, Jessop T, editors, *The Works of George Berkeley Bishop of Cloyne*, Vol. 4, pp. 53–102. Nelson, London Full text, edited by David R. Wilkins, available on line at https://www.maths.tcd.ie/pub/HistMath/People/Berkeley/Analyst/Analyst.pdf.

Dedekind R (1965) *Stetigkeit und Irrationale Zahlen* Friedr. Vieweg & Sohn, Braunschweig, 7th edition. 1st edition, 1872.

Devlin K (2005) Plenary address Delivered at a meeting on June 3, 2005 of the Canadian Mathematical Society.

Grabiner JV (1997) Was Newton's Calculus a Dead End? the Continental Influence of Maclaurin's Treatice of Fluxions. *American Mathematical Monthly* 104:393–410. Preliminary version presented as an invited address under the title "Was Newton's Calculus a Dead End? A New Look at the Calculus of Colin Maclaurin" to the Sixteenth Annual Meeting of the Canadian Society for the History and Philosophy of Mathematics, University of Victoria (British Columbia), May 31–June 1, 1990.

Heath TL (1912) *The Works of Archimedes* Cambridge University Press Reprinted by Dover (no date given).

Heath TL (1926) *The Thirteen Books of Euclid's Elements* Cambridge University Press, second edition Three volumes. Reprinted by Dover, 1956. The text, along with a Java applet to manipulate the diagrams can be found on line at http://aleph0.clarku.edu/~djoyce/java/elements/elements.html.

Knorr W (1975) *The Evolution of the Euclidean Elements: A Study of the Theory of Incommensurable Magnitudes and Its Significance for Early Greek Geometry* Reidel, Dordrecht and Boston and London.

Middlemiss RR (1952) *College Algebra* McGraw-Hill.

Nardi E (2014) Reflections on visualization in mathematics and in mathematics education In Fried M, Dreyfus T, editors, *Mathematics and Mathematics Education: Searching for Common Ground*, Advances in Mathematics Education. Springer.

Seldin J (1990) Reasoning in elementary mathematics In Berggren T, editor, *Canadian Society for History and Philosophy of Mathematics. Proceedings of the Fifteenth Annual Meeting, Quebec City, Quebec, May 29–May 30, 1989*, pp. 151–174 Available on line from http://www.cs.uleth.ca/~seldin under "Publications".

Seldin J (1991) From exhaustion to modern limit theory In Abeles FF, Katz VJ, Thomas RS, editors, *Canadian Society for History and Philosophy of Mathematics. Proceedings of the Sixteenth Annual Meeting, Victoria, British Columbia, May 31–June 1, 1990*, pp. 120–136 Available on line from http://www.cs.uleth.ca/~seldin under "Publications".