# The Mathematics of Plato's Academy

## A New Reconstruction

*Second Edition*

David Fowler

Raphael's fresco *Causarum Cognitio* in the Vatican's Stanza della Segnatura represents a high-water mark of the myths surrounding Plato. The scene is a lofty and imposing building, and the composition is dominated by Plato, holding his *Timaeus* and pointing upwards with his right hand (up to the understanding of causes?), in lively and public discussion with Aristotle, who holds his *Ethics*. Geometry is clearly represented by the group at the right front; astronomy is behind them, where the king and the globe illustrate a common confusion between the astronomer Ptolemy and the dynasty of Macedonian rulers of Egypt; and music occupies the left foreground, indicated by a tablet containing a harmonic system. Arithmetic is less easily identified. Perhaps the only popular feature of the tradition that is not to be found in the fresco is an inscription 'Let no one unskilled in geometry enter'. The common name of the fresco, 'The School of Athens', has no better authority than a seventeenth-century French guidebook.

# THE MATHEMATICS OF PLATO'S ACADEMY

A NEW RECONSTRUCTION

SECOND EDITION

DAVID FOWLER

*Mathematics Institute*
*University of Warwick*

CLARENDON PRESS · OXFORD
1999

# OXFORD

*For*
*Annette Gruner Schlumberger*
*Maryse Stroh*
*Denise Fowler Stroh*

*and dedicated to the memory of*
*Wilbur Richard Knorr (1945–1997)*

Plato greatly advanced mathematics in general and geometry in particular because of his zeal for these studies. It is well known that his writings are thickly sprinkled with mathematical terms and that he everywhere tries to arouse admiration for mathematics among students of philosophy.

Proclus quoting Eudemus, *Commentary on the First Book of Euclid's Elements*, tr. G. Morrow, p. 66.

In the 'Cloisters' of the Metropolitan Museum in New York there hangs a magnificent tapestry which tells the tale of the Unicorn. At the end we see the miraculous animal captured, gracefully resigned to his fate, standing in an enclosure surrounded by a neat little fence. This picture may serve as a simile for what we have attempted here. We have artfully erected from small bits of evidence the fence inside which we hope to have enclosed what may appear as a possible, living creature. Reality, however, may be vastly different from the product of our imagination; perhaps it is vain to hope for anything more than a picture which is pleasing to the constructive mind when we try to restore the past.

O. Neugebauer, *The Exact Sciences in Antiquity*, Chapter 6.

# PREFACE TO THE SECOND EDITION

The changes in this second edition are the inclusion of a substantial Appendix, Chapter 10, the rewriting of much of Chapter 5, the inclusion of three new Plates, a significantly expanded Bibliography, and many smaller additions, modifications, and corrections. Also the Addenda at the end of the chapters of the paperback edition have been incorporated into the text.

The original Preface, below, explained how the book can be read starting with any of its three Parts; I now take this further and *recommend* any reader new to the book to start with the first section of the new Appendix, which forms an introduction to a substantial and important part of the book and makes explicit some of the unspoken attitudes that underlie its treatment. I have adopted this arrangement in order to maintain the overall structure of the book and because, as I explain there, that section approaches the subject differently from the main text. Thereafter the new Appendix material is more in the nature of amplifications and extensions taking the argument further, and references to it have been inserted into the main text where appropriate. Its final section is a slightly revised version of the original Epilogue, and a new, brief, autobiographical Epilogue has been added to round off the book.

One very apt criticism of the book has been that, despite its title, it treated only a part of the mathematics of Plato and his associates. I was aware of this; indeed I said as much in the original Epilogue, now Section 10.5, and wanted the book to be called *Some of the Mathematical Topics of Plato's Academy?* However, the publisher vetoed this on the grounds that it was too long (though there is a good precedent for ignoring this kind of advice: W. C. Fields is said to have called a film *Never Give a Sucker an Even Break* since that was too long to fit on any cinema advertising board of that period), and because the Press tried to avoid punctuation in titles (though they seem to have relaxed this for the recently published *What Is Mathematics, Really?*). But they also said that it would drastically reduce the sales of the book, a clinching argument. Another suggestion for clarifying this issue has been to rename this second edition *A Mathematics of Plato's Academy*, thus making the point that it may be only one of many, and also distinguishing the first and second editions without greatly inconveniencing bibliographers and librarians, but this also has its disadvantages. I have therefore decided to do the conventional thing, and stick to 'Second Edition'.

There are two very different reasons why some topics were only fleetingly mentioned, if at all, in the text of the book: they may be spurious, untestified, or only found in manifestly unreliable sources; or, though very important in themselves, they may be irrelevant to the overall argument of the book—and it

may often be difficult for the reader to make the distinction between these two cases. Some examples of omitted but important topics are the interpretations of Plato's *Timaeus*, Plato and Aristotle's philosophies of mathematics, other aspects of the diversity of fourth-century mathematics, Greek logic, etc.; while examples of topics which seem to be untestified even in our best early sources are a possible Old Babylonian influence on the beginnings of Greek mathematics, the role of Eleatic philosophy, in particular Parmenides' use of *reductio ad absurdum*, the importance of ruler and compass constructions in early Greek mathematics—and the story of the discovery of incommensurability. In fact this last topic *is* treated at some length right at the end of the exposition of things Greek, in Section 8.3(c), and my conclusions are clearly stated at the end of this section, but only the most dedicated reader will get this far. One reason for not coming clean earlier with my point of view about this discovery, which is seen by most as a defining moment of Greek mathematics, was to minimise distraction from what I thought to be more important things, but this has clearly led to misunderstandings. Therefore I now reverse my earlier approach, which is why I recommend the reader to start with this introductory survey of the story of incommensurability in the Appendix.

While the main contents of the first edition of this book may be highly controversial, in general they avoided polemic as much as possible. In parts of the new Appendix, I go more on the offensive, and point out in more detail why I dissent from other more traditional views, and why I think that they may indeed sometimes be more speculative than my own.

Let me, at the outset, set out my position in this new polemical stance. Our evidence concerning pre-Euclidean mathematics is so indirect and fragmentary that the greatest part of the stories that are now told about the period must be invention. This book sets out an unconventional story in which the study of various kinds of ratio plays an important part: anthyphairetic ratios, associated with Theaetetus; astronomical ratios, associated with Eudoxus; musical ratios, associated with Archytas; and accountants' ratios, which grew out of the everyday Egyptian and Greek way of handling division. In this situation, where none of us has very much evidence to back up our accounts:

- Whose story contains the most invention?
- Whose story relies most on our own ways, today, of understanding arithmetic and the other basic intuitions of mathematics?
- Whose story has the greatest coherence?
- Whose story is the most interesting?

Again, readers must decide for themselves the answers to these questions, preferably after studying my proposals here; I am, of course, biased but my own answers will be perfectly obvious, and have if anything been reinforced during the period since the publication of the first edition.

Since writing the book more than fifteen years ago, I have come to know and appreciate the enormous contributions of many more scholars than I had then

read, let alone referred to, and am embarrassed by the brashness and ignorance of some parts of the text. (One small but pervasive example was an attempt at a bit of linguistic reform, fortunately unsuccessful, which has been corrected in this second edition. I had misunderstood the early meaning of 'algorism', and was not aware that 'algorithm' has an even longer and more noble ancestry than I had thought. For more details see p. 63.) However, had I read more before writing, this might have led me to replace some of the bounce and enthusiasm by a more sober and serious approach. Readers must also decide for themselves whether this would have been an improvement or not.

Those who would like to know more about the circumstances of the composition and execution of Raphael's fresco, reproduced here as the frontispiece, are recommended to E. H. Gombrich, *Symbolic Images*, and G. W. Most, 'Reading Raphael: The *School of Athens* and its pre-text'; publication details are given in the Bibliography.

Much of the material in the Appendix has been tried out at a series of meetings organised by Vassilis Karasmanis and Marinus Taisbak, and I would like to thank them and the other participants for their most helpful reactions and comments. While writing the Appendix, I have had a lively and stimulating correspondence with Reviel Netz, who was working on his own book, and I extract one of his many witty and apposite reflections: "We're like two pedlars caught in a derelict mining-town in the Old West; you carry a whole load of your 2nd edition, I a whole load of my *Shaping of Deduction in Greek Mathematics*, and we've got only each other to sell our wares to. Funny, isn't it?" Many of those who helped me in the writing of the first edition have continued to answer my questions over the succeeding years, and provide advice and material; and new people, not mentioned in that earlier Preface, include Benno Artmann, Gerhard Brey, Julio Gonzalez Cabillon, Menso Folkerts, Jens Høyrup, Henry Mendell and Eleanor Robson. Throughout, I have again benefited from the experienced and constructive help of John Fauvel. And all of my work on Greek mathematics leading up to the first edition of this book was inspired by W. R. Knorr's *Evolution of the Euclidean Elements* and his early articles. Wilbur died of melanoma, tragically early, while this second edition was being prepared. I dedicate it to his memory.

*Warwick*                                                                    D. H. F.
January 1998

# PREFACE TO THE FIRST EDITION

Theon of Smyrna, a neo-Platonic commentator of the third century AD, wrote a book *Expositio Rerum Mathematicarum ad Legendum Platonem Utilium*; with a little less hubris than Theon, I regard this book here as 'An Account of Mathematical Topics that May be Useful for Reading Some Bits of Plato'. But, at the outset, I must emphasise that, like Theon, I do not deal with Plato's philosophy. My objective is to provide a background, developed around a new interpretation of aspects of early Greek mathematics, against which to set some of our evidence about the activities of Plato and his colleagues in the Academy. The exposition is divided into three parts, each with its own introduction; the following brief preface, together with these introductions, will give a summary of the aims, contents, and organisation of the book.

The subject matter is early Greek mathematics, a phrase I shall use to denote the phase of developments that culminates in the works of Euclid and Archimedes. Part One (Chapters 1–5) sets out, in some detail, an exploration of some mathematical topics suggested by ideas found in the works of Plato, Aristotle, Euclid, and Archimedes and carried out using methods derived from their procedures; and it proposes that similar investigations might, indeed, have taken place and had some influence on the development of early Greek mathematics. The main theme, though not in its treatment here, is 'well known', as mathematician's jargon expresses it: it has been intensively studied since the seventeenth century, it can be found in readily available publications, and it has, from time to time, played an important role in the development of parts of mathematics. Yet it is possible to emerge from a modern training as a fully qualified member of the mathematical community without any knowledge whatsoever of this material. Every mathematician knows the algorithm which generates the greatest common factor of two numbers: at each step, we subtract the smaller as many times as possible from the larger to leave a still smaller remainder. When the procedure terminates, the last non-zero remainder is the greatest common factor. What we shall mainly be studying, usually in the context of two line segments rather than two numbers, is the relationship between the two original quantities and the sequence of numbers that describes how many subtractions have been performed at each stage. The original pair of numbers or lines clearly determines this pattern of subtractions; but conversely, we may ask, what is it about the original pair that is determined by this pattern? Can the same pattern arise from two different pairs and, if so, how are they related?

The algorithm and the theory associated with the pattern of subtractions are now generally called the Euclidean algorithm and continued fractions.

However, the word algorithm is a recent corruption of a fine old word algorism (derived from al-Khārizmi); and I shall eventually argue that Euclid's role in the development of our understanding of the process may have been far from positive. Moreover the treatment of the material within Greek mathematics will not permit the introduction of the fractions or real numbers that are needed for the standard descriptions of continued fractions; and so on. For these and other reasons, some to be described more fully in Chapter 2, I shall refer to the cluster of ideas and procedures around this Greek algorithm by the word 'anthyphairesis', the transcription of its Greek description by the perfectly ordinary word ἀνθυφαίρεςιϲ, which can be roughly translated as 'reciprocal subtraction', the subtraction of one thing from another.

The main theme of Part One is the definition and use of ratio, which is described by Euclid, at *Elements* V Definition 3, as "A sort of relation in respect of size between two magnitudes of the same kind". By meditating on the questions posed at the end of the last paragraph but one, we see that anthyphairesis provides one such definition of ratio; then astronomy suggests another, and music theory another. These ideas will be developed and related to each other, and this material will be connected to the description of a mathematical curriculum that Plato sets out in *Republic* VII.

Part Two, which could equally well have been placed before Part One, discusses some aspects of our evidence about the period. Chapter 6 takes a very general point of view; it examines the evidence for one well-known and representative story about Plato's Academy, and then looks at the transmission of Greek texts in general down to our time. Then, in Chapter 7, a single topic, the treatment of numbers and fractions, is examined in fine detail, for it is out of our underlying experience with arithmetic that much of mathematics grows. In order to highlight my argument that early Greek mathematics is non-arithmetised I shall examine its obverse, arithmetised mathematics and calculation with fractions, and will argue that there is a systematic and very widespread misunderstanding about Greek arithmetical calculations. What is often taken as proven fact is, at best, an interpretation not based on any clear evidence.

Parts One and Two combine, I believe, to set out a very plausible argument that the study of various different ways of describing ratio might have played an important role in the development of mathematics within Plato's Academy; yet there seems to be no explicit mention of this subject, nor any of its techniques, discoveries, problems, and pitfalls, in any surviving Greek source of any kind. In Part Three, Chapter 8, I discuss a few of the discrepancies between what later Greek commentators tell us about Greek mathematics and what we find in our surviving early evidence. And finally, in Chapter 9, I sketch the theory and story of the development, since the seventeenth century, of the mathematics underlying my reconstruction here of anthyphairetic ratio theory, and finish by reflecting on the fact that it, too, after having intrigued and influenced Fermat (perhaps), Brouncker, Wallis, Huygens, Euler, Lambert,

Lagrange, Legendre, Gauss, Galois, and many others, has subsequently been left to one side and omitted from the basic training of mathematicians, and so has been neglected by many mathematicians and historians of today.

The problems posed by trying to describe and develop mathematics that may be unfamiliar to many people, within a very novel, initially implied, historical context, have driven me, after a series of unsatisfactory rejected drafts, into adopting an unconventional expository technique. The book includes four dialogues that contain, in condensed form, the exposition of a lot of material. A substantial part of the text is then an elaboration of and commentary on material introduced in these dialogues.

The referencing system I adopt throughout does not conform to the different conventions of either classicists or mathematicians, but I hope that it can be immediately understood and used by everybody. Every book or article will normally be referred to by its author and an acronym of its title (in italic for a book, roman for an article) with, as necessary, volume, page, line, and note numbers; then these acronyms are expanded in the Bibliography. So, for example, Heath, *HGM* ii, 50–6, refers to pages 50 to 56 of volume 2 of T. L. Heath's *History of Greek Mathematics*, a discussion of Archimedes, *Measurement of a Circle*. Modern editions of ancient authors also fit into the same system, but with the refinement that the editor's name will be added, as in the example Archimedes–Heiberg, *Opera* i, 232–43, again a reference to *Measurement of a Circle*. Collected works such as this, which contain the texts in their original language, ancient or modern, with or without translations or commentaries, will be identified by a keyword: *Opera*, *Works*, *Werke*, etc. Where a particular translation of a text is referred to, the translator will be given in the same way; for example Archimedes–Heath, *WA*, 91–8, for this same work. If at any point in the text it seems that a fuller title will provide an extra grain of information, then it will be given; if an author's name requires initials for clarification, they will be given. Papyrological publications will be cited in their own standard way; for details of this, see Chapter 6, footnote 18. Euclid, Plato, and Aristotle, our major sources, need special mention. Of the nine volumes (in ten parts) of the critical edition of Euclid's surviving works, edited by J. L. Heiberg with H. Menge and M. Curtze, here called Euclid–Heiberg, *Opera*, five volumes (in six parts) have been reissued in a form described on its title page as "*Euclidis Elementa* post J. L. Heiberg edidit E. S. Stamatis"—this will be described as Euclid–Stamatis, *EE*; and there is a standard, exemplary, and readily available three-volume English translation and commentary to which I hope every half-way serious reader will have access, which will be called Heath, *TBEE* (= *The Thirteen Books of Euclid's Elements*), not Euclid–Heath, *TBEE*. 'The *Elements*' will always be Euclid's *Elements*, and most translations of passages from the *Elements* will be from Heath, *TBEE*, though some are taken from the recent comprehensive commentary, Mueller, *PMDSEE*. Similarly, Plato's and Aristotle's works will generally be referenced by title only, without editor or translator, though the translators of extended passages will be

identified, often in the notes; most of these translations will be taken from
Plato, *CD* and Aristotle–Ross, *WA*, now available in a convenient revised
two-volume version, Aristotle–Barnes, *CW*. Three basic English reference
works will be cited without author or editor: the *DSB*, *OCD*, and *OED*
(= *The Dictionary of Scientific Biography*, *The Oxford Classical Dictionary*,
and *The Oxford English Dictionary*). Further notes on this kind of referencing
system, which I first saw and appreciated in de Ste. Croix, *CSAGW*, are given
in the Bibliography. In this way, I hope that any text that is cited anywhere can
be quickly identified by anybody; but whether some of these texts can then be
located and consulted is quite a different matter, which may need the assistance
of a very specialised library. I must, at the outset, record my gratitude to my
own university library, the Ashmolean Library, and the Bodleian Library for
their generous and expert help in this respect.

The book finishes with an Index of Cited Passages, an Index of Names, and a
conventional General Index.

Direct quotations of texts are either displayed, or run into the text and
enclosed in *double* quotation marks ". . ."; single quotation marks are used to
set off words or phrases. I have used algebraic formulae, in contexts referring to
Greek mathematics, only as a shorthand for more long-winded *geometrical*
descriptions; this is explained in several places, for example Sections 1.2(e), 3.2,
4.5(b), and throughout Chapter 5. The examples just given illustrate the system
of cross-referencing used in the book, and the chapters and sections are
indicated throughout in the running heads.

I would like to thank the very many people, far more than I can possibly
name here, who have helped me write this book. If I do mention some by name,
I hope that they will not be inhibited thereby in their criticism, which has so
much helped in the past, and that those whom I do not name will not feel that
their help is less appreciated. My interest in Greek mathematics was kindled by
reading *The Evolution of the Euclidean Elements*, and its author Wilbur Knorr
has since provided further encouragement, information, and advice. Andrew
Barker, Alan Bowen, Malcolm Brown, Ivor Bulmer-Thomas, Harold Cherniss,
Roger Herz-Fischler, Geoffrey Lloyd, Bernard Teissier, and Tom Whiteside
have kept up a steady stream of help, encouragement, and material. It soon
became apparent that I needed to enter new and vast fields of scholarship, and I
wish to acknowledge warmly the help that has been given and the interest that
has been shown by many classicists, and how much I have come to appreciate
their information and insights. Very particular thanks are due to the papy-
rologists, especially Sir Eric Turner (1911–83) and his colleagues Walter
Cockle, Peter Parsons, and John Tait; and I have received special assistance
from Alain Blanchard, William Brashear, Myles Burnyeat, Patrice Cauderlier,
Jim Coulton, D. R. Dicks, Tiziano Dorandi, Brenda Farr, Hermann Harrauer,
Ludwig Koenen, Tom Pattie, Bob Sharples, David Thomas, and Nigel Wilson.
The community of mathematicians and historians of mathematics has looked
on, sometimes rather bemused, and offered advice and help of all kinds;

notable have been E. M. Bruins, J. W. S. Cassels, Donald Coxeter, Kay Dekker, John Fauvel, Bernard Goldstein, Bill Gosper, Ivor Grattan-Guinness, Jan Hogendijk, Jan van Maanen, John Mason, Joe and Barry Mazur, Ian Mueller, David Mumford, Mike Paterson, Paddy Patterson, Mark Rafter, Rolph Schwarzenberger, Jacques Sesiano, Jeff Smith, B. L. van der Waerden, and many of my immediate colleagues of whom I should like particularly to thank, for more than I could ever hope to enumerate, Elaine Greaves Coelho and Christopher Zeeman. The staff of Oxford University Press have given me comfort, encouragement, and expert help, and have indulged my idiosyncracies. The inclusion of several items, including the frontispiece, was made possible by a grant from the Fondation Les Treilles.

The book itself has been written in three very special places, each with its own presiding genius, to whom it is jointly and affectionately dedicated.

*Les Treilles, Tourtour,*                                                              D. H. F.
*Les Côtes du Plan, Ollioules,*                                                   *July* 1986
*and St. Nicholas', Warwick*

# ACKNOWLEDGEMENTS

# CONTENTS

PART THREE: LATER DEVELOPMENTS

## 8  LATER INTERPRETATIONS

## 9  CONTINUED FRACTIONS

# PLATES

## NOTE ON THE TRANSCRIPTIONS OF PAPYRI

The transcriptions of papyri (except for Plate 2) are made as follows: the texts are printed in modern form, with accents and punctuation, the lectional signs occurring in the papyri being noted in the *apparatus criticus*, where also faults of orthography, etc., are corrected. Iota adscript is printed where written, otherwise iota subscript is used. Square brackets [ ] indicate a lacuna, round brackets ( ) the resolution of a symbol or abbreviation, angular brackets ⟨ ⟩ a mistaken omission in the original, braces { } a superfluous letter or letters, double square brackets ⟦ ⟧ a deletion, the signs ` ´ an insertion above the line. Dots printed slightly below the line within brackets represent the estimated number of letters lost or deleted, dots slightly below the line outside brackets mutilated or otherwise illegible letters, and dots under letters indicate that the reading is doubtful.

(This description has been excerpted and adapted from recent volumes of *The Oxyrhynchus Papyri*. For further discussion, see E. G. Turner, *Greek Papyri* (2nd edn.) 187 n. 22 and 203 n. 22.)

# PART ONE

---

# INTERPRETATIONS

Finding out about ratios is a secret of deep logic.
Omar Khayyam, *Discussion of Difficulties in Euclid.*

This first part will develop the main mathematical and historical structure of the interpretation. I have tried to make it as direct and explicit as possible, and have used various expository devices to get across ideas and attitudes that may be unfamiliar or unconventional. It contains no footnotes and the minimum of distractions in the main text; so even the details of the ancient texts and their translations, other references, and additional incidental material are given, to begin with, in the notes at the end of each chapter. A new Appendix, Chapter 10, has been added to the second edition and readers are strongly recommended to read its first section, 10.1, for an overview of the argument of the book.

Chapter 1 sets the scene by describing the context, techniques, and themes that will dominate this book. I take as model, for both the kind of mathematics and the kind of mathematical exposition, the passage from Plato's *Meno* where Socrates persuades a slaveboy to double a square. Then, after analysing the concepts and techniques that will be at our disposal, I propose a sequel to this episode in which the slaveboy, now greatly increased in confidence and knowledge, faces the question of giving a definition of ratio.

Chapter 2 starts a commentary on my dialogue, in the form of an exposition of some of the techniques it introduced. Again I adopt a Platonic rather than a Euclidean style: problems are suggested rather than formally enunciated, the motivation is treated at length while sometimes the actual evaluation is left to the reader, the treatment is very personal and discursive, and I make no attempt to set out a definitive account of the subject. I have tried to illustrate the techniques on examples that arise in our ancient evidence and my aim, here and elsewhere, is one of comprehensiveness: I wish to describe a cluster of associated methods that will provide plausible reconstructions of not one but all of these ancient examples that sometimes seem so perplexing to a modern audience.

Chapter 3 is a reinterpretation of Book II of Euclid's *Elements* and some related material. Here I have attempted to present an exhaustive account. Rather than suggesting to the reader that something can be done, I have done it and set out the details, but I encourage my readers to skip these details where this is more palatable. Here again, the aim is of comprehensiveness: a successful

explanation of a coherent collection of results like Book II should account for everything that is found in the book. This chapter ends with a discussion of the texts on side and diagonals found in the Platonic commentaries by Theon of Smyrna and Proclus. This is the only substantial piece of evidence from later sources that is invoked in Part One, and my discussion indicates, incidentally, some of the difficulties with late and indirect material of this kind.

Chapter 4 presents an interpretation of Plato's curriculum in *Republic* VII. The central idea here is that Plato's *logistikē* refers to ratio theory, but that different definitions of ratio are appropriate for mathematics, astronomy, and music theory. Then the underlying unity of the curriculum comes, I suggest, from relating these different manifestations of ratio to each other. Again, the exposition of some of these ideas is done in dialogue form.

Chapter 5 gives a description of the classification of incommensurable lines found in *Elements* X and applied in *Elements* XIII. My main interest here is in the mathematical motivation of this classification, and I propose that it is a further extension of the problems suggested by the reconstructions of Chapters 3 and 4.

# 1

# THE PROPOSAL

## 1.1 SOCRATES MEETS MENO'S SLAVEBOY

Plato's *Meno* is an enquiry into virtue. As part of the discussion, Plato's Socrates digresses to illustrate his opinion that we cannot teach anything, but can only prompt recollection of the opinions of a disembodied soul. But Socrates' digression takes him into the subject of our enquiry here, so let us begin with this passage, *Meno* 81e–85d. The translation is by W. K. C. Guthrie; I have suppressed all editorial additions, made one very slight alteration throughout which will be explained in Section 1.2(c), below, and restored two speeches at the end of 83b that have unaccountably been omitted.

MENO: What do you mean you say that we don't learn anything, but that what we call learning is recollection? Can you teach me that it is so?

SOCRATES: I have just said that you're a rascal, and now you ask me if I can teach
82a   you, when I say there is no such thing as teaching, only recollection. Evidently you want to catch me contradicting myself straight away.

MENO: No, honestly, Socrates, I wasn't thinking of that. It was just habit. If you can in any way make clear to me that what you say is true, please do.

SOCRATES: It isn't an easy thing, but still I should like to do what I can since you
b   ask me. I see you have a large number of retainers here. Call one of them, anyone you like, and I will use him to demonstrate it to you.

MENO: Certainly. Come here.

SOCRATES: He is a Greek and speaks our language?

MENO: Indeed yes—born and bred in the house.

SOCRATES: Listen carefully then, and see whether it seems to you that he is learning from me or simply being reminded.

MENO: I will.

SOCRATES: Now boy, you know that a square is a figure like this?

BOY: Yes.

c   SOCRATES: It has all these four sides equal?

BOY: Yes.

SOCRATES: And these lines which go through the middle of it are also equal?

BOY: Yes.

SOCRATES: Such a figure could be either larger or smaller, could it not?

BOY: Yes.

SOCRATES: Now if this side is two feet long, and this side the same, how many feet will the whole be? Put it this way. If it were two feet in this direction and only one in that, must not the area be two feet taken once?

BOY: Yes.

d  SOCRATES: But since it is two feet this way also, does it not become twice two feet?

BOY: Yes.

SOCRATES: And how many feet is twice two? Work it out and tell me.

BOY: Four.

SOCRATES: Now could one draw another figure double the size of this, but similar, that is, with all its sides equal like this one?

BOY: Yes.

SOCRATES: How many feet will its area be?

BOY: Eight.

SOCRATES: Now then, try to tell me how long each of its sides will be. The
e    present figure has a side of two feet. What will be the side of the double-sized one?

BOY: It will be double, Socrates, obviously.

SOCRATES: You see, Meno, that I am not teaching him anything, only asking. Now he thinks he knows the length of the side of the eight-feet square.

MENO: Yes.

SOCRATES: But does he?

MENO: Certainly not.

SOCRATES: He thinks it is twice the length of the other.

MENO: Yes.

SOCRATES: Now watch how he recollects things in order—the proper way to recollect.
83a    You say that the side of double length produces the double-sized figure? Like this I mean, not long this way and short that. It must be equal on all sides like the first figure, only twice its size, that is, eight feet. Think a moment whether you still expect to get it from doubling the side.

BOY: Yes I do.

SOCRATES: Well now, shall we have a line double the length of this if we add another the same length at this end?

BOY: Yes.

SOCRATES: It is on this line then, according to you, that we shall make the eight-feet square, by taking four of the same length?
b  BOY: Yes.

SOCRATES: Let us draw in four equal lines, using the first as a base. Does this not give us what you call the eight-feet figure?

BOY: Certainly.

SOCRATES: But does it contain these four squares, each equal to the original four-feet one?

BOY: Yes.

SOCRATES: How big is it then? Won't it be four-times as big?

BOY: Of course.

SOCRATES: And is four-times the same as twice?

BOY: Of course not.

SOCRATES: But how much is it?

BOY: Fourfold.
c  SOCRATES: So doubling the side has given us not a double but a fourfold figure?

BOY: True.

SOCRATES: And four-times four are sixteen, are they not?

BOY: Yes.

SOCRATES: Then how big is the side of the eight-feet figure? This one has given us four-times the original area, hasn't it?

BOY: Yes.

SOCRATES: And a side half the length gave us a square of four feet?

BOY: Yes.

SOCRATES: Good. And isn't a square of eight feet double this one and half that?

BOY: Yes.

SOCRATES: Will it not have a side greater than this one but less than that?

d   BOY: I think it will.

SOCRATES: Right. Always answer what you think. Now tell me. Was not this side two feet long, and this one four?

BOY: Yes.

SOCRATES: Then the side of the eight-feet figure must be longer than two feet but shorter than four?

BOY: It must.

e   SOCRATES: Try to say how long you think it is.

BOY: Three feet.

SOCRATES: If so, shall we add half of this bit and make it three feet? Here are two, and this is one, and on this side similarly we have two plus one, and here is the figure you want.

BOY: Yes.

SOCRATES: If it is three feet this way and three that, will the whole area be three-times three feet?

BOY: It looks like it.

SOCRATES: And that is how many?

BOY: Nine.

SOCRATES: Whereas the square double our first square had to be how many?

BOY: Eight.

SOCRATES: But we haven't yet got the square of eight feet even from a three-feet side?

BOY: No.

SOCRATES: Then what length will give it? Try to tell us exactly. If you don't want to
84a   count it up, just show us on the diagram.

BOY: It's no use, Socrates, I just don't know.

SOCRATES: Observe, Meno, the stage he has reached on the path of recollection. At the beginning he did not know the side of the square of eight feet. Nor indeed does he know it now, but then he thought he knew it and answered boldly, as was appropriate—he felt no perplexity. Now however
b   he does feel perplexed. Not only does he not know the answer; he doesn't even think he knows.

MENO: Quite true.

SOCRATES: Isn't he in a better position now in relation to what he didn't know?

MENO: I admit that too.

SOCRATES: So in perplexing him and numbing him like the sting-ray, have we done him any harm?

MENO: I think not.

SOCRATES: In fact we have helped him to some extent toward finding out the right answer, for now not only is he ignorant of it but he will be quite glad to look for it. Up to now, he thought he could speak well and fluently, on many occasions and before large audiences, on the subject of a

c      square double the size of a given square, maintaining that it must have a side of double the length.

MENO: No doubt.

SOCRATES: Do you suppose then that he would have attempted to look for, or learn, what he thought he knew (though he did not) before he was thrown into perplexity, became aware of his ignorance, and felt a desire to know?

MENO: No.

SOCRATES: Then the numbing process was good for him?

MENO: I agree.

SOCRATES: Now notice what, starting from this state of perplexity, he will

d      discover by seeking the truth in company with me, though I simply ask him questions without teaching him. Be ready to catch me if I give him any instruction or explanation instead of simply interrogating him on his own opinions.

Tell me, boy, is not this our square of four feet? You understand?

BOY: Yes.

SOCRATES: Now we can add another equal to it like this?

BOY: Yes.

SOCRATES: And a third here, equal to each of the others?

BOY: Yes.

SOCRATES: And then we can fill in this one in the corner?

BOY: Yes.

SOCRATES: Then here we have four equal squares?

e      BOY: Yes.

SOCRATES: And how many times the size of the first square is the whole?

BOY: Four-times.

SOCRATES: And we want one double the size. You remember?

BOY: Yes.

SOCRATES: Now does this line going from corner to corner cut each of these squares in half?

BOY: Yes.

SOCRATES: And these are four equal lines enclosing this area?

BOY: They are.

SOCRATES: Now think. How big is this area?

BOY: I don't understand.

85a    SOCRATES: Here are four squares. Has not each line cut off the inner half of each of them?

BOY: Yes.

SOCRATES: And how many such halves are there in this figure?

BOY: Four.

SOCRATES: And how many in this one?

BOY: Two.

SOCRATES: And what is the relation of four to two?

BOY: Double.

1 P Herc 1061 photograph of columns 8, 9, and 10, and drawings and transcription of column 9 in J L Heiberg, Quelques papyrus traitant de mathématiques, *Oversigt over det Kgl Danske Videnskabernes Selskabs Forhandlinger* 2 (1900), 164–5 See Section 6 2
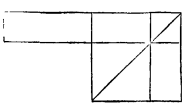
## XXIX. Euclid II. 5.

8·5 × 15·2 *cm.*

Fragment from the bottom of a column, containing the enunciation, with diagrams, of Euclid II. 5, and the last words of the preceding proposition.

From the character of the handwriting, which is a sloping rather irregular informal uncial, this papyrus may be assigned to the latter part of the third or the beginning of the fourth century. Diaereses are commonly placed over syllabic ι and υ. Iota adscript is not written. The corollary of Proposition 4 seems to have been omitted, while the two lines illustrating the division εἰς ἴσα καὶ ἄνισα in Proposition 5 are not found in ordinary texts. Otherwise the papyrus shows no variants from the text of the Oxford edition of 1703 or that of Peyrard, beyond the mistake of τετραγωνον for τετραγώνῳ in l. 9, and the spelling μετοξυ for μεταξύ in l. 6.

```
          .   .  .   .   .   .
    ΠΕΡΙΕΧΟΜΕ[ΝΩ ΟΡΘΟΓΩΝΙΩ . . . . . . . . .

   έ ΕΑΝ ΕΥΘΕΙΑ ΓΡΑΜΜΗ
     ΤΜΗΘΗ ΕΙC ΪCΑ ΚΑΙ ΑΝ       [_____ᴵ_____
     ΪCΑ ΤΟ ΫΠΟ ΤΩΝ ΑΝΙ            ——[———ᴵ——
   5 CΩΝ ΤΗC ΟΛΗC ΤΜΗΜ[ΑΤ]ΩΝ ΠΕΡΙΕΧΟΜΕΝΟΝ
     ΟΡΘΟΓΩΝΙΟΝ ΜΕΤΑ Τ[Ο]Υ ΑΠΟ ΤΗC ΜΕΤΟΞΥ
     ΤΩΝ ΤΟΜΩΝ ΤΕΤ[ΡΑ]ΓΩΝΟΥ ΪCΟΝ ΕCΤΙΝ
     ΤΩ ΑΠΟ ΤΗC ΗΜΙCΕΙ                        έ
     ΑC ΤΕΤΡΑΓΩΝΟΥ
```

5. ΤΗC Ο corrected from ΠΕΡΙ.    6. l. μεταξύ.    9. l. τετραγώνῳ.

1. If the reading is correct—and though the traces of letters after ΠΕΡ are scanty, there seems to be no alternative—the corollary of Prop. 4 was omitted. After ΟΡΘΟ-ΓΩΝΙΩ, too, there would not be room for more than about nine letters, so ὅπερ ἔδει δεῖξαι must have either been omitted or, more probably, abbreviated.

2-3. The shortness of these lines indicates that there were two horizontal strokes in the margin, the first showing the division into equal, the second that into unequal parts. The first is entirely broken away, and only the left-hand part of the second is preserved.

2. P. Oxy, i 29: photograph of text; and complete publication in B. P. Grenfell & A. S. Hunt, *The Oxyrhynchus Papyri* i (1898), 58. See Section 6.2.
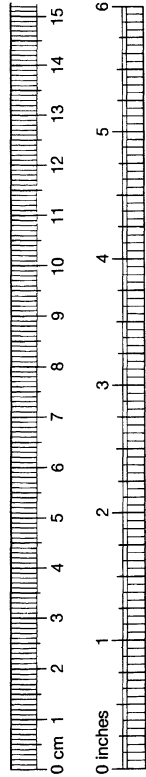
3. P. Fay. 9: photograph of text; and transcriptions in B. P. Grenfell, A. S. Hunt, & D. G. Hogarth, *Fayûm Towns and their Papyri* (1900), 96–9. See Section 6.2.

εϖ]ι της αυτη[ς βασεως
επεζευ]χθω η $\overline{αδ}$ [
5    ε]στιν τη [ $\overline{βγ}$
τη $\overline{β}$]γ παραλ[ληλος
] το $\overline{αβγ}$ [
της α]υτης β[ασεως
παρ]αλλη[λοις
10    ε]στιν ι[σ]ον το β[$\overline{δγ}$
μει]ζον τω ελ[ασ]σον[ι
η $\overline{α}$]ε τη $\overline{βγ}$ εστιν παρ[αλληλος
δειξο]μεν οτι [ο]υδε αλλη [
η ] $\overline{αδ}$ αρα τη $\overline{βγ}$ εστιν πα[ραλληλος
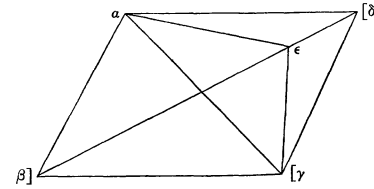


15 εαν παρα]λληλογραμμ[ον] τρι[γωνω
αυτη]ν και εν τ[αις] αυταις [
] εσται τ[ο πα]ραλληλ[ογραμμον
π]αρα[λληλογρ]αμμ[ον
βασιν] τε [
20    ] . [
διπλα]σιο[ν
επεζευ]χθω [
] τω [εβγ] τρ[ιγωνω
] της $\overline{βγ}$ κ[αι
25    ται]ς $\overline{βγ}$ [α]ε α[λλα
] παραλληλογρα[μμον
] και του $\overline{εβ}$[γ

[τὴν δοθ]εισαν γων[ιαν ευ]θυ
[γραμμ[ον διχα τεμειν
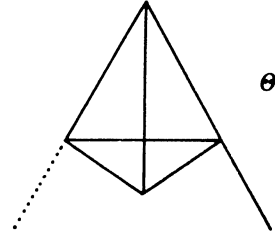
[τὴν δοθει]σαν ευθειαν πεπρασ
[μενην] διχα τεμειν
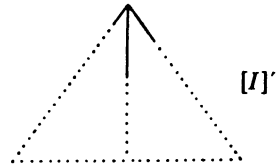
4. P. Berol [inv.] 17469: photograph of text; and transcription in W. Brashear, Vier neue Texte zum antiken Bildungswesen, *Archiv für Papyrusforschung* 40 (1994), 29–30. See Section 6.2.
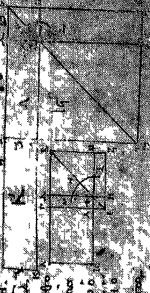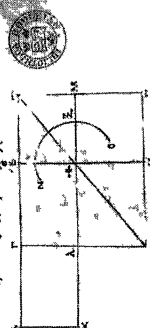
5. Bodleian manuscript D'Orville 301, folios 35ᵛ and 36¹, mid enunciation of *Elements* II 5 to mid proof of II 6, with scholia, extra inserted figures and annotations. See Section 6.3.
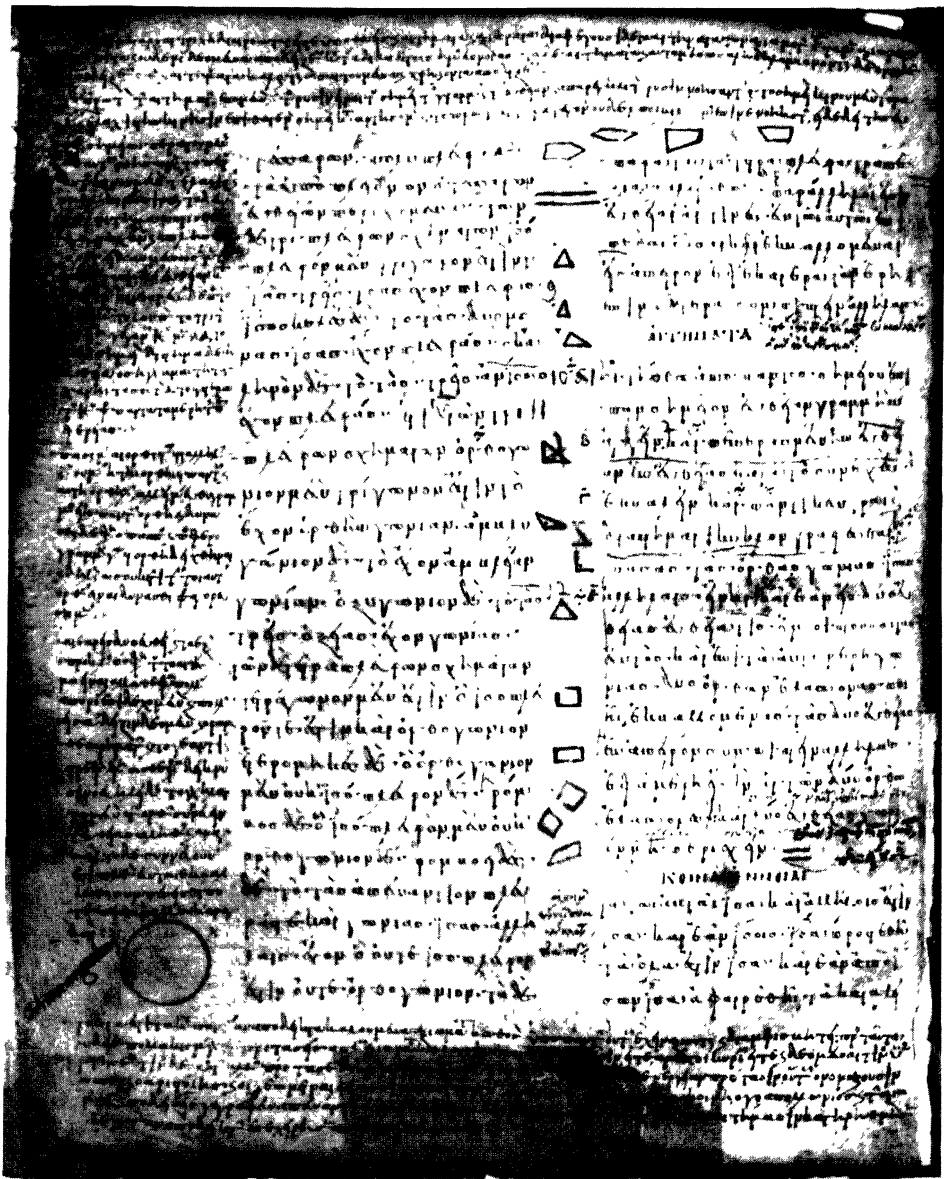
6. Vatican manuscript 190, folio 14ᵛ: *Elements* I, mid Definition 19 to mid Common Notion 3, with scholia, inserted figures, and annotations. See Sections 6.3 and 7.2, note 1.
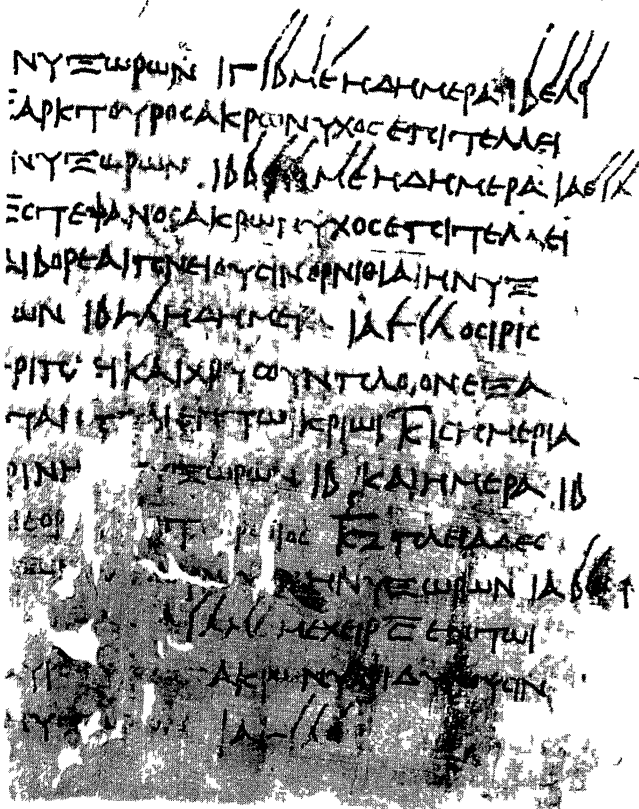
Fr. (*b*), 2nd hand.  Col. iv.

55 [ἡ] νὺξ ὡρῶν ιγιβ´μ´ε´, ἡ δ᾽ ἡμέρα ιβ´εχ´ϟ.
[ι]ϛ Ἀρκτοῦρος ἀκρώνυχος ἐπιτέλλει,
[ἡ] νὺξ ὡρῶν ιββ´ι´ε´μ´ε´, ἡ δ᾽ ἡμέρα ιαθ´ιχ´.
[κ]ϛ Στέφανος ἀκρώνυχος ἐπιτέλλει
[κ]αὶ βορέαι πνείουσιν ὀρνιθίαι, ἡ νὺξ
60 [ὡρ]ῶν ιβΖχ´, ἡ δ᾽ ἡμέρα ιαγίχ´. Ὄσιρις
[π]εριπλεῖ καὶ χρυσοῦν πλοῖον ἐξά-
[γε]ται. Τῦβι ⟨ε⟩ ἐν τῶι Κριῶι. κ ἰσημερία
[ἐα]ρινή, [ἡ] νὺξ ὡρῶν ιβ καὶ ἡμέρα ιβ,
[κ]αὶ ἑορ[τ]ὴ Φιτωρώιος. κζ Πλειάδες
65 [ἀκ]ρώνυχ[οι] δύνου[σ]ιν, ἡ νὺξ ὡρῶν ιαβ´ϛ´ϟ´,
[ἡ] δ᾽ ἡμέρα [ι]βίχ´μ´ε´. Μεχεὶρ ϛ ἐν τῶι
[Τ]αύρωι. Τάδες ἀκρώνυχοι δύνουσιν,
[ἡ] νὺξ ὡρῶν ιαΖίχ´ε´,

55. l. ι´ε´ for ι´β´.    57. ι´ε´ corr. from ε´.    65. ϛ´ corr.    68. l. χ´ϟ´ for χ´ε´.

7. P. Hib. i 27: photograph of column 4 and its transcription, both from B. P. Grenfell & A. S. Hunt, *The Hibeh Papryi* i (1906), 138–57 and Plate VIII. See Section 7.1 (c).
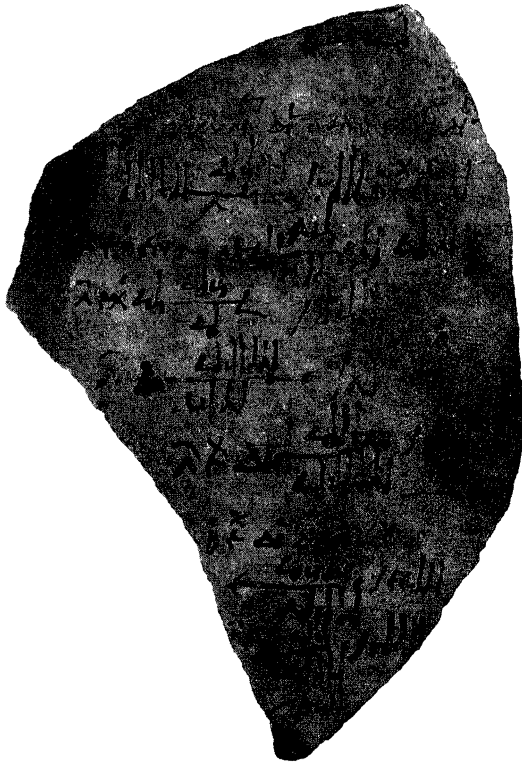
**1847** [1036].  30 B.C. - 14 A.D.

ε̣ ∠ d η ∠ d  ² *Γεω(μετρία) γῆ(ς) Καλλιε(δόντος) διὰ Πετεχ(ών-*
*σιος) σ⌐ αρ^χ Λιβ(ὸς)* ³ d' η' ι'ς' λ'β' $\frac{\angle d' \, \eta'}{a}$ η ι'ς' [[λβ]]/ *ἄ(ρο)υ(ραι)*
d' ι'ς'. *Βο(ρρᾶ) ἐχ(ομένη) ει^{ας} λ'β'.*  ⁴ *Ἀλ(λη) ἐχο(μένη) ἐν*
*τῶι ἐδά(φει)* ∠ d'η' $\frac{a \angle \eta'}{\iota'\varsigma'}$ a λ'β' / ∠ d λ'β' [[.'.']].  ⁵ *Ἀλ(λη)*
*ἐχο(μένη)* ∠ η' $\frac{\angle d' \, \eta'}{\angle d'}$ ∠/ d' η' ις'.  ⁶ *Ἀλ(λη)* δ $\frac{\angle d' \, \eta' \, \upsilon\varsigma' \, \lambda'\beta'}{a \, \eta' \, \lambda'\beta'}$
o / δ d'  ⁷ *Ἀλ(λη) ἐχομένη* δ ∠ d $\frac{\angle d \, \upsilon\varsigma' \, ..}{\angle d'\eta' \, \upsilon'\varsigma' \, \lambda'\beta'}$ / γ ∠ d η' λ'β'
⁸ / *ἄ(ρο)υ(ραι) θ* ∠ d' η' [ι'ς'].  ⁹ *Βο(ρρᾶ) ἐχ(ομένη) διὰ Λυσιμάχ(ου)*
¹⁰ ε $\frac{\angle \, d' \, \eta' \, \lambda'\beta'}{a \, d' \, \eta' \, \lambda'\beta'}$ o / ε ∠ d λ'β'.  ¹¹ *Ἀλ(λη) ἐχ(ομένη)* δ $\frac{\angle \, d' \, \eta'}{\angle d' \, \eta' \, \lambda'\beta'}$
o / γ d' η' λ'β'.  ¹² / θ ∠ η' ι' ς'.

3 Au quatrième côté du quadrilatère, le η n'est pas surmonté du trait qui indiquerait qu'il s'agit de 1/8.  Il faut cependant comprendre 1/8 et non 8, en raison du produit : 1/4 et 1/16 d'aroure.        7 Il manque un quatrième nombre aux dimensions des côtés données pour le calcul de la superficie. Le signe o (= ὁμοίως ?) indique que le quatrième côté considéré est égal au premier.        9 *Λυσίμαχος* : le même sans doute que dans W.O. 760 et 761.
2 Je ne comprends pas σ⌐ = σα(  ).  αρχ = probablement ἀρχ(ή).

8. O Bodl. ii 1847: photograph of text; and transcription in J. G. Tait & C. Préaux, *Greek Ostraca in the Bodleian Library at Oxford* ii (1955), 308. See Sections 7.1 (d), 7.4, and 8.1.

9. P. Lond. II 265 (p. 257), columns 2 and 3. See Section 7.3 (c).

10. Brit. Mus. Add. Ms. 37516, recto: photograph of text; transcription from F. G. Kenyon, Two Greek school-tablets, *Journal of the Hellenic Society* 29 (1909), 29–40 and Plate V; and translation. See Section 10.1(c). (Continued on the next page.)

ο Πυθαγορας φιλοσοφος αποβας και γραμματα διδασκων συνεβουλευεν τοις εαυτου μαθηταις εναιμονων απεχεσθαι

του Πυθαγορου φιλοσοφου αποβαντες και γραμματα διδασκοντος λογος απομνημονευεται συνβουλευοντος τοις εαυτου μαθηταις εναιμονων απ[ε]χεσθ[αι]

τω Πυθαγορα φιλοσοφω αποβαντι και γραμματα διδασκοντι ʽεδοξεν᾿ συνβουλευσαι τοις εαυτου μαθηταις εναιμονων απεχεσθαι

τον Πυθαγοραν φιλοσοφον αποβαντα και γραμματα διδασκοντα φασιν συνβουλευσαι τοις εαυτου μαθηταις εναιμονων απεχεσθαι

ω Πυθαγορε φιλοσοφε αποβας και γραμματα διδασκον συ ποτε συνεβουλευσατον τοις εαυτου μαθηταις εναιμονων απεχεσθαι

και δυκως

τω Πυθαγορα φιλοσοφω αποβαντην και γραμματα διδασκοντε συνβουλευετην τοις εαυτου μαθηταις εναιμονων απεχεσθαι

τοιν Πυθαγοροιν φιλοσοφοιν αποβαντοιν και γραμματα διδασκοντοιν λογος απομμνημονευεται συνβουλευοντοιν τοις εαυτου μαθηταις εναιμονω[ν] απεχεσθαι

τοιν Πυθαγοροιν φιλοσοφοιν αποβαντοιν και γραμματα διδασκοντοιν εδοξεν συνβουλευσαι τοις εαυτου μαθηταις εναιμονων απεχεσθαι

τω Πυθαγορα φιλοσοφω αποβαντην και γραμματα διδασκοντε φασιν συνβουλευσαι τοις εαυτου μαθηταις εναιμονων απεχεσθαι

ω Πυθαγορα φιλοσοφω αποβαντην και γραμματα διδασκοντε σφω ποτε συνβουλευσατην τοις εαυτου μαθηταις εναιμονων απεχεσθαι

και πληθυντικως

οι Πυθαγοραι φιλοσοφοι αποβαντες και γραμματα διδασκοντες συνεβουλευσθην τοις εαυτου μαθηταις εναιμονων απεχεσθαι

των Πυθαγορων φιλοσοφων αποβαντων και γραμματα διδασκοντων λογος απομνημονευεται συνβουλευσαντων τοις εαυτου μαθηταις εναιμονων απεχεσθαι

τοις Πυθαγοραις φιλοσοφοις αποβασι και γραμματα διδασκοντοις εδοξεν συνβουλευσαι τοις εαυτου μαθηταις εναιμονων απεχεσθαι

τους Πυθαγορας φιλοσοφους απο[βαν]τας και γραμματα διδασκοντας φασιν συνβουλευσαι τοις εαυτου μαθηταις εναιμονων απεχεσθαι

ω Πυθαγοραι φιλοσοφοι αποβαντες και γραμματα διδασκοντες υμεις ποτε συνβουλευετιν τοις εαυτου μαθηταις εναιμονων απεχεσ[θαι]

Pythagoras [nominative] the philosopher, when he had gone away and was teaching letters, advised his pupils to abstain from meat.
There is a story of Pythagoras [genitive] the philosopher, when he had gone away and was teaching letters, advising his pupils to abstain from meat.
It occurred to Pythagoras [dative] the philosopher, when he had gone away and was teaching letters, to advise his pupils to abstain from meat.
They say that Pythagoras [accusative] the philosopher, when he had gone away and was teaching letters, advised his pupils to abstain from meat.
O Pythagoras [vocative] the philosopher, when you had gone away and were teaching letters, you once advised your pupils to abstain from meat.
And in the dual:
The two Pythagorases [nominative dual], the philosophers, when they had gone away and were teaching letters, advised their pupils to abstain from meat.
There is a story of the two Pythagorases [genitive dual], the philosophers, when they had gone away and were teaching letters, advising their pupils to abstain from meat.
It occurred to the two Pythagorases [dative dual], the philosophers, when they had gone away and were teaching letters, to advise their pupils to abstain from meat.
They say that the two Pythagorases [accusative dual], the philosophers, when they had gone away and were teaching letters, advised their pupils to abstain from meat.
O you two Pythagorases [vocative dual], philosophers, when you had gone away and were teaching letters, you once advised your pupils to abstain from meat.
And in the plural:
The Pythagorases [nominative plural], the philosophers, when they had gone away and were teaching letters, advised their pupils to abstain from meat.
There is a story of the Pythagorases [genitive plural], the philosophers, when they had gone away and were teaching letters, advising their pupils to abstain from meat.
It occurred to the Pythagorases [dative plural], the philosophers, when they had gone away and were teaching letters, to advise their pupils to abstain from meat.
They say that the Pythagorases [accusative plural], the philosophers, when they had gone away and were teaching letters, advised their pupils to abstain from meat.
O Pythagorases [vocative plural], philosophers, when you had gone away and were teaching letters, you once advised your pupils to abstain from meat.
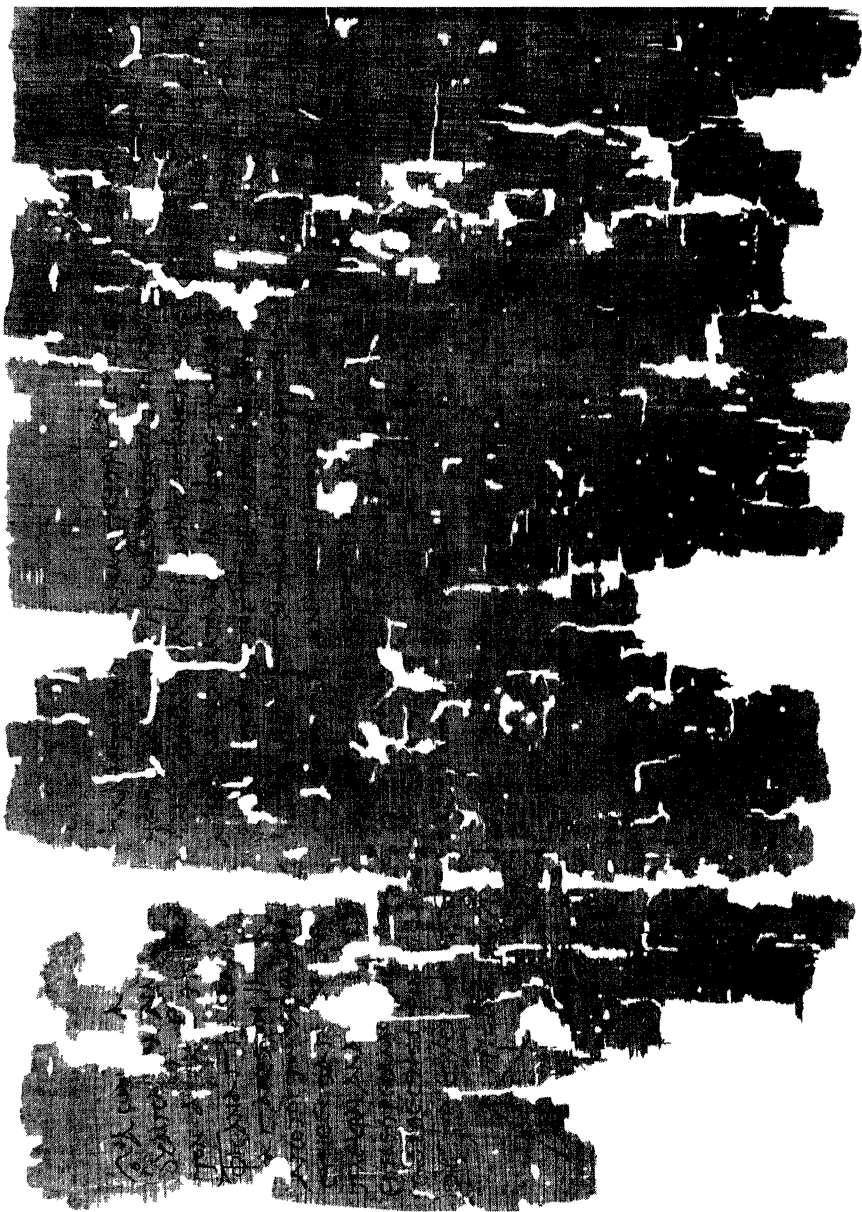
b   SOCRATES: How big is this figure then?

BOY: Eight feet.

SOCRATES: On what base?

BOY: This one.

SOCRATES: This line which goes from corner to corner of the square of four feet?

BOY: Yes.

SOCRATES: The technical name for it is 'diagonal'; so if we use that name, it is your personal opinion that the square on the diagonal of the original square is double its area.

BOY: That is so, Socrates.

SOCRATES: What do you think, Meno? Has he answered with any opinions that were not his own?

c   MENO: No, they were all his.

SOCRATES: Yet he did not know, as we agreed a few minutes ago.

MENO: True.

SOCRATES: But these opinions were somewhere in him, were they not?

MENO: Yes.

SOCRATES: So a man who does not know has in himself true opinions on a subject without having knowledge.

MENO: It would appear so.

SOCRATES: At present these opinions, being newly aroused, have a dreamlike quality. But if the same questions are put to him on many occasions and in different ways, you can see that in the end he will have a knowledge on the

d   subject as accurate as anybody's.


The passage is well known and frequently discussed, but I quote it here in full for special reasons:

(i) It is our *first* direct, explicit, extended piece of evidence about Greek mathematics; it probably dates from about 385 BC. By going forward from this time we may probe further back into the past, as in the following three typical examples: In commentaries, written in the sixth century AD, by Simplicius and Eutocius on Aristotle and Archimedes respectively, we find descriptions of work by Hippocrates (*c*.425 BC?) and Archytas (*c*.385 BC or earlier?) that come, directly or indirectly, from a now lost *History of Mathematics* written around 325 BC by Eudemus, a pupil of Aristotle. Second, Euclid's *Elements*, which is believed to have been compiled around 300 BC, gives us an edited and anonymous treatment of some of the mathematics of Euclid's predecessors, in a text that has passed through the hands of an unknown number of further scribes and editors before arriving at the versions that we now possess. Third, to find more than very short and isolated references to the mathematics of the Pythagoreans, and any information about Pythagoras himself, we have to go forward to the notoriously biased and uncritical hagiographical writings of Iamblichus in the third century AD. In this shifting and murky world of fragmentary, anonymous, secondary and tertiary accounts, this passage by Plato and other mathematical passages in his dialogues stand out as direct

testimonies from someone located at the centre of that remarkable group of creative mathematicians who initiated the style and some of the techniques and preoccupations that have since dominated the subject.

(ii) It serves as a lucid and straightforward example of some of the problems, styles, and techniques of early Greek mathematics. I shall discuss this in more detail later in this chapter.

(iii) It leads directly into my exposition, at the end of this chaper, of some of the themes and preoccupations of this book.

My aim throughout will be to try to emphasise mathematical texts that survive from the fourth and third centuries BC, including another long passage in Plato's *Republic* VII which describes a curriculum for the future guardians of the state, and to attempt to fit them together into some coherent account. To begin with, I shall ignore the accounts by the commentators in late antiquity; then, when the details of the proposed reconstruction have been presented, I shall consider, in Part Three of this book, a few of the different interpretations presented by the neo-Platonic and neo-Pythagorean commentators from the third century AD onwards. I now believe that these are later misunderstandings and that they can be explained and may be almost inevitable. I introduce this argument right at the end of the book, but a proper discussion of the issues will have to be deferred to future publications.

## 1.2 THE CHARACTERISTICS OF EARLY GREEK MATHEMATICS

### 1.2(a) *Arithmetised mathematics*

There is a style of mathematics that I shall call 'arithmetised'. It is characterised by the use of some idea of number that is sufficiently general to describe some model of what is now frequently called 'the positive number line' *and its arithmetic*. The range of numbers that can be handled may be more or less comprehensive. For example, that great part of Babylonian mathematics which is restricted to the so-called regular numbers, numbers with terminating sexagesimal expansions, deals, in effect, with numbers that can be written as common fractions with denominators containing powers of two, three, and five; while, at other times, the scope of arithmetic has been extended to include informal or formal descriptions of irrational, complex, infinitesimal, and infinite quantities under this general umbrella of 'numbers'. These numbers then permeate mathematics: in geometry, lines become endowed with a 'length'; the 'area' of a rectangle is the product of the lengths of its base and height, and this basic definition is extended to the area of more and more two-dimensional figures; three-dimensional figures have a numerical 'volume'; then, more recently, higher-dimensional spaces have been created, purely numerical constructs whose properties are then described by geometrical analogies. The ratio of two quantities becomes the quotient of the two numbers that represent them;

for example, the ratio of the circumference to diameter of a circle becomes a number that can be approximated by an integer 3, or by fractions such as $3\frac{1}{8}$, $\frac{22}{7}$, or $\frac{355}{113}$, or represented as a decimal number $3\cdot1416\ldots$, or an infinite product

$$2\left(\frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdots\right),$$

or a continued fraction

$$4\left(\frac{1}{1+}\frac{1^2}{2+}\frac{3^2}{2+}\frac{5^2}{2+}\cdots\right).$$

Familiarity with the manipulation of numbers can then lead to the experience and confidence out of which algebra can grow; a non-mathematician, looking at almost any mathematics text of today, might be inclined to describe mathematics as 'all algebra', a natural reaction to the result, over the last three hundred years, of our version of this programme of arithmetisation.

Many cultures have developed their own versions of arithmetised mathematics. Our evidence about Babylonian mathematics comes from contemporary clay tablets which contain practically nothing but numbers, expressed in an efficient floating-point, decimally coded, sexagesimal notation. Egyptian mathematics is arithmetised, though its description in terms of unit fractions seems to us to be unwieldy and inconvenient. Ptolemaic astronomy and the mathematics of Heron and Diophantus seem to be a powerful and profound blend of the earlier Greek geometrical and Babylonian arithmetised methods; and these were then further developed by Arab astronomers and mathematicians. Western mathematics since the sixteenth century has, as I have already suggested, been dominated by the arithmetised point of view, and this culminated in the developments of the nineteenth and twentieth centuries when the arithmetised formulation permeated analysis; the resulting minute investigations led to the first known coherent, formal descriptions of what might be meant by 'number' and 'arithmetic'. Dedekind tells us, in the introduction to his *Stetigkeit und irrationale Zahlen*, that he conceived his definition of 'real numbers' on Wednesday, 24 November 1858, and other versions were given by Méray, Weierstrass, and Cantor; other, different formal models for the number line have been proposed since. This recent activity shows a consciousness by mathematicians that the description of the number line and its arithmetic is at the root of our understanding and intuition about mathematics today. There may have been pockets of non-arithmetised mathematics such as, for example, nineteenth-century 'synthetic geometry', but the trend has been towards a more and more comprehensively arithmetised approach. And it seems that the power and scope of modern mathematics owes much to the completeness and effectiveness of this approach.

Greek mathematics up to the second century BC seems, to an extraordinary degree, to be different. I shall argue, throughout this book, that while the Greeks may originally have deployed techniques that could serve perfectly as

labelling systems for the positive number line, they did not and could not go on to consider arithmetical operations with these labels. Thus my first characteristic of early Greek mathematics is negative: it seems to be completely non-arithmetised.

### 1.2(b) *Non-arithmetised geometry*

Geometry, modelled on the properties of the three-dimensional space we seem to experience about us, provides the main ingredient of Greek mathematics. And Socrates gives a good illustration of his common-sense attitude to this geometry when he says, right at the beginning of his encounter with the slaveboy:

Such a figure could be either larger or smaller, could it not?

(I shall return to this assertion much later, at the end of Chapter 4.) A typical solution of a problem or proof of a proposition will consist of a figure and a collection of statements about the figure. Thus the first two incorrect solutions to Socrates' problem of doubling a square were two figures, which the reader was left to supply, together with statements about counting equal squares; while the third and correct solution consisted of a further figure with the statements like: "these squares are cut in half"; "these lines are equal"; "this square contains two (equal) halves"; and "four is twice two". Another good example of this demonstrative procedure, which also illustrates Euclid's use of the word 'equal' (*isos*), is *Elements* I 35:

Parallelograms which are on the same base and in the same parallels are equal to one another.

Here the figure is a straightforward instantiation of the enunciation of the proposition. Different configurations are possible, according to the amount of overlap, and Fig. 1.1 illustrates one case, slightly different from that treated by Euclid. The argument then runs: the two triangles are equal so that, when the trapezium is added to each, Common Notions 2 and 4:

If equals be added to equals, the wholes are equal.
Things which coincide with one another (*ta epharmozonta ep' allēa*) are equal to one another,

will confirm the equality of the parallelograms. We need some kind of assurance that this particular figure will ensure the universal truth of the
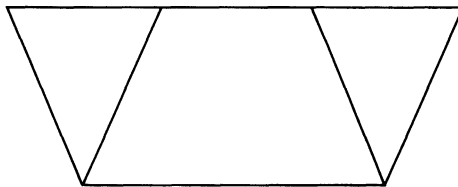


Fig. 1.1

proposition; while Aristotle worries about this kind of question, it does not seem to bother the mathematicians.

Equality, then, does not mean arithmetical equality of some numerical measure of area and it cannot mean congruence. In fact congruence does not seem to be a basic concept in the *Elements*, since Euclid has no single word for it; the idea of coincidence in Common Notion 4 might seem a candidate, but this piece of text may be a later interpolation (see Heath, *TBEE* i, 224 f.) and Euclid does not make systematic use of this terminology (it appears elsewhere only in I 4 & 8 and III 24). Another way of describing congruence is found, for the first time, in XI Definition 10.

Equal and similar (*isos kai homoios*) solid figures are those contained by similar planes equal in multitude and magnitude.

This 'definition' illustrates further how Euclid's geometry consists of assertions about space either that are taken as self-evident, as apparently this is, or that may be listed explicitly as Common Notions, such as the examples I have just quoted, or as postulates, for example Postulate 5:

That, if a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than two right angles,

or asserted without comment, as in I 1:

From the point C in which the circles cut one another . . . ,

or asserted because they are held to follow from some earlier assumptions or propositions, which may be implied, or paraphrased, or quoted word-for-word, but which are almost never located by exact references, so that the two occasions when distant locations are specified (possibly in later interpolations) are very surprising:

For it was proved in the first theorem of the tenth book that, if two unequal magnitudes be set out, and if from the greater there be subtracted a magnitude greater than the half, and from that which is left a greater than the half, and if this be done continually, there will be left some magnitude which will be less than the lesser magnitude set out [XII 2],

and

For this has been proved in the last theorem but one of the eleventh book [XIII 17].

The problems of elucidating the status of Euclid's geometry have occupied and perplexed commentators since late antiquity up to the present day. Of the examples I have just cited, XI Definition 10 would now seem to us to be not a definition but a proposition, indeed a proposition that is only true if sufficiently strong additional conditions, such as convexity, are imposed, and which otherwise can now be shown, by ingenious counter-examples, to be false. Postulate 5 is now perceived not simply as the postulation of the converse of I 17:

In any triangle two angles taken together in any manner are less than two right angles,

but as a profound statement about a particular kind of geometry; and the assertions in I 1 and XII 2 are interpreted in terms of conditions that need to be imposed on our mathematical model of space. We can have no clear idea of the point of view of Euclid and his sources; we simply do not have enough reliable information. But these are not the kinds of questions I shall be considering here. I shall rather use geometry as Socrates and Euclid do, as describing some non-arithmetised idea of space through which assertions and relations are made manifest, and I will leave to one side questions of the status of the geometrical objects and procedures that participate in the figures and constructions, and refer to the geometry in the way that Plato describes, at *Republic* VII, 527a:

Their [the geometers'] language is most ludicrous, though they cannot help it, for they speak as if they were doing something and as if all their words were directed towards action. For all their talk is of squaring and applying and adding and the like, whereas in fact the real object of the entire study is pure knowledge.

I shall go on to describe how geometry can be used to establish a very surprising kind of pure knowledge and will argue that this may be what Plato is referring to here.

   The objects of this geometry are points, lines, and simple figures which reside in two- and three-dimensional space, and which are manipulated in various ways; and one of the main preoccupations of Euclid's geometry is the transformation, combination, and comparison of these figures—see Plato's "squaring and applying and adding and the like". The idea behind Euclid's use of 'equality' within geometry is one of size, not shape, and his concern, at first, is to see if two lines, two plane figures, or two three-dimensional regions are equal in size or, when they are not equal, which is the bigger. Let us look at these operations in a little more detail.

   The comparison of two lines can be carried out using the first three propositions of Book I: these enable us to superimpose one line and its endpoint on the other, an operation that is used explicitly at the beginning of I 4. Then Book I develops some of the machinery for manipulating and comparing two rectilineal plane regions; it culminates in Proposition 44:

To a given straight line to apply, in a given rectilineal angle, a parallelogram equal to a given triangle,

and I 45 shows how to extend this to any given rectilineal figure, though Euclid here shows no concern about the one step that might today seem to require some detailed proof, the operation of triangulating the figure that is to be transformed. A crucial preliminary result is established in I 43, which asserts that the two shaded parallelograms in Fig. 1.2 are equal: the large parallelogram contains two smaller parallelograms, and each of these three is cut into equal halves by the diagonal. Hence when the two equal smaller halves are
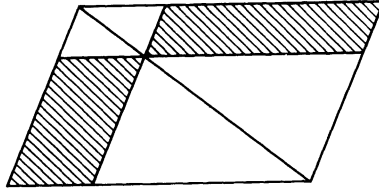
Fig. 1.2

subtracted from the equal larger halves, the results will be equal. The last proposition of Book II, II 14, describes how

To construct a square equal to a given rectilineal figure,

and this theory for rectilineal plane figures is further elaborated in Book VI. Problems that arise with more general plane magnitudes and three-dimensional figures are considered in Book XII but, by that time, new ingredients have introduced new mathematical preoccupations.

Let me return, for a final illustration of non-arithmetised geometry, to Socrates and the slaveboy. At 82c–d, they have to work out the area of a square two feet by two feet. In arithmetised geometry, we would immediately say that its square is two times two square feet. Socrates proceeds differently. He argues, with my interpolations:

If it were two feet in this direction and only one in that [it would comprise two juxtaposed one-foot squares, forming a rectangle, so that], must not the area be two feet taken once; but since it is two feet this way also [and so comprises two such rectangles juxtaposed], does it not become twice two feet?

This illustrates the basic geometrical operation of addition, represented by juxtaposition of the corresponding figures. Similar passages occur at 83b and e.

Greek mathematics does deploy some numbers, as in the passages just considered, and these are manipulated arithmetically, even in a geometrical context, as at *Meno* 83e, again with my interpolation:

If it is three feet this way and three that, will not the whole area be [three juxtaposed rectangles each comprising three juxtaposed squares, so that it is] three-times three feet.

It is to this ingredient that I now turn.

### 1.2(c) *Numbers and parts: the arithmoi and merē*

Greek mathematics makes free use of the cardinal numbers; but rather than our thinking of the sequence $1, 2, 3, \ldots$, *qua* cardinals, a much more faithful impression of the very concrete sense of the Greek *arithmoi* is given by the sequence:

duet, trio, quartet, quintet, . . . .

I have omitted the first term because in early Greek mathematics the unit, the *monas*, has a different status from the others, so that an argument may have to be reformulated when it applies to this case. These numbers are ordered by size ('a quartet is bigger than a trio'), and can be added by concatenation ('a trio plus a quartet makes a septet') and subtracted 'the less from the greater'.

The *arithmoi* also appear in other forms, such as the adverbial sequence:

once, twice, three-times, four-times, . . . .

These two particular kinds of numbers can act on each other to give multiplication, as in Socrates' *"dis duoin"* (twice two) or *"treis tris"* (three three-times). Division, or division with remainder when the less does not measure the greater, is described similarly. There are further numerical sequences, such as:

simple, double, triple, quadruple, . . . ,

which Socrates also draws on, and which are used in the same way; the particular choice seems to depend on grammatical or stylistic convenience. None of these other forms seem to have a common name in English, and I shall refer to them as 'repetition numbers'.

Another kind of numerical sequence is:

half, third, quarter, fifth, . . . ;

see, for example, Socrates' remark at the end of *Meno* 83c. These are occasionally found in formal mathematics and are ubiquitous in certain kinds of calculations. They are frequently confused with the ordinal numbers, the different sequence:

first, second, third, fourth, fifth, . . . ,

but I shall refer to them by their Greek name as the series of parts (*meros* or *morion*, plural *merē* or *moria*), and will denote them by $\acute{2}, \acute{3}, \acute{4}, \ldots$ or $2', 3', 4', \ldots$. Chapter 7 will be devoted to a discussion of these parts; meanwhile I urge that they should not be thought of as: one-half, one-third, one-quarter, one-fifth, . . . , nor be written: $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots$.

Some examples from the *Elements* will illustrate these features. The basic definitions are VII Definitions 1 to 5:

A unit is that with respect to which each existing thing is called one; and a number is a multitude composed of units.
A number is a part of a number, the less of the greater, when it measures the greater; but parts when it does not measure it.
The greater number is a multiple of the less when it is measured by the less.

Then the answer to a question such as 'What multiple?' or 'What part?' will be a repetition number, or a phrase which describes a repetition number. See, for example, VII Definition 8:

An even-times even number (*artiakis artios arithmos*) is that which is measured by an even number according to an even number (*kata artion arithmon*).

But Euclid never tells us how to answer the question 'What parts?' Heath's hyphen, "even-times even", conveys the nuance of the adverbial form. I have added similar such hyphens to Guthrie's translation of the *Meno* passage with which this chapter began; see, for example, 83c: *tettarōn tetrakis estin hekkaideka*, four four-times are sixteen.

The multiplication that arises when a repetition number acts on a cardinal number is explicitly described in terms of addition:

A number is said to multiply a number when that which is multiplied is added to itself as many times as there are units in the other, and thus some number is produced [VII Definition 15];

and here the phrase "as many times as there are units in the other" illustrates how a cardinal number can be converted into a repetition number. We shall see, in the next chapter, how division is similarly construed in terms of subtraction.

Greek *arithmētikē* (i.e. *arithmētikē technē*) is the art of the *arithmoi*, the study of the properties of these kinds of numbers. It is frequently translated as 'arithmetic', but 'number theory' will better convey the sense of the word provided it is always interpreted within the context of these kinds of manipulations of the Greek *arithmois*. However, I shall generally leave it in its transliterated form, and refer to *arithmētikē*.

### 1.2(d)  *Ratio (logos) and proportion (analogon)*

The two words, ratio and proportion, *logos* and *analogon*, refer to different kinds of mathematical entities, though they are frequently conflated and used as if they were interchangeable. A ratio is something; in modern terms, it is a function of two (or possibly more) variables, and if we want to make use of this concept, we should say what it is and how it is calculated and manipulated. I shall very often write $a:b$ as shorthand for the phrase 'the ratio of $a$ to $b$'; the symbol means no more and no less than the words, and it is up to us to say what they stand for if we want to go on to make use of the idea. Euclid's definition—though description might be a better word to use—occurs at V Definition 3:

A ratio is a sort of relation in respect of size between two magnitudes of the same kind. (*Logos esti duo megethōn homogenōn he kata pēlikotēta poia schesis.*)

This is elaborated in V Definition 4:

Magnitudes are said to have a ratio to one another which are capable, when multiplied, of exceeding one another,

but these do not tell us enough about ratios to enable us to prove any results about them. I shall explain this point in more detail later in this section.

On the other hand, a proportion is a condition that may or may not hold between four objects. It may appear as a question: 'Are $a$, $b$, $c$, and $d$ proportional?' to which, if it is properly posed, we have to answer 'Yes' or 'No'. There are other ways of expressing the same question: 'Is $a$ to $b$ as $c$ is to $d$?' (this being a question about a proportion, and not about two ratios), or the idea may be introduced as an assertion, not as a question. I shall often use '$a:b::c:d$' as shorthand for all of these forms of words. Euclid gives two different definitions of proportionality. One, at V Definition 5, is believed to be due to Eudoxus and to date from $c$. 350 BC (though the evidence for this attribution and date is indirect, or late and anonymous), at the end of the period I shall mainly be considering. Since my concern here will be with the precursors of this definition, I shall not quote or try to explain it here. The second definition occurs at VII Definition 20:

Numbers are proportional when the first is the same multiple, or the same part, or the same parts, of the second that the third is of the fourth.

I have earlier suggested that the answer to 'What multiple?' or 'What part?' is an adverbial number. This fits well within the restrictions on the *arithmoi* and leads to an idea of 'same multiple' and 'same part' which makes this definition workable in these cases; but Euclid nowhere explains how to treat the case when the first number is 'parts' of the second, neither explicitly in a definition, nor implicitly in a proposition, so we do not know how to interpret 'same parts' in this definition. In fact what information he does give seems inconsistent: VII Definitions 3 and 4, already quoted earlier but repeated here for emphasis, *tell* us that

A number is part of a number, the less of the greater, when it measures the greater; but parts when it does not measure it,

so that any number is either part or parts of a greater number; but VII 4 then *proves*, at some length, that

Any number is either a part or parts of any number, the less of the greater.

Something is curiously wrong. I shall discuss some of our evidence concerning these numerical parts, *merē*, in Chapter 7.

It is often asserted that the discovery of the phenomenon of incommensurability led to a situation in which the early Greek mathematicians were unable to set the theory of ratio or proportion on firm foundations, within the means at their disposal, until the development by Eudoxus, in the middle of the fourth century BC, of the proportion theory of Book V. I wish, throughout this book, to query everything in this sentence except that which relates to the identification, dating, and attribution of Eudoxan proportion theory; see Section 10.1. As a first step, I shall address here the mathematical component of this assertion, in the following interpretation: that the ingredients of early Greek mathematics, such as I have described them so far or, alternatively, as

can be inferred from what are believed to be the early books of the *Elements*, are inadequate to set up a theory of ratio or proportion sufficiently general to handle incommensurable magnitudes. At this point a long quotation from Aristotle's *Topics*, VIII 3, 158a31–159a2, is in order. The *Topics* is a manual of 'syllogistic dialectic', a kind of formal debate between a 'questioner' and an 'answerer':

There are certain hypotheses upon which it is at once difficult to bring, and easy to stand up to, an argument. Such, e.g., are those things which stand first and those which stand last in the order of nature. For the former require definition, while the latter have to be arrived at through many steps if one wishes to secure a continuous proof from first principles, or else all discussion about them wears the air of mere sophistry; for to prove anything is impossible unless one begins with the appropriate principles, and connects inference with inference till the last are reached. Now to define first principles is just what answerers do not care to do, nor do they pay any attention if the questioner makes a definition: and yet until it is clear what it is that is proposed, it is not easy to discuss it. This sort of thing happens particularly in the case of the first principles (*archē*): for while the other propositions are shown through these, these cannot be shown through anything else: we are obliged to understand every item of that sort by definition. The inferences, too, that lie too close to the first principle are hard to treat in argument . . . . The hardest, however, of all definitions to treat in argument are those that employ terms about which, in the first place, it is uncertain whether they are used in one sense or in several, and, further, whether they are used literally or metaphorically by the definer. For because of their obscurity, it is impossible to argue upon such terms; and because of the impossibility of saying whether this obscurity is due to their being used meta-phorically, it is impossible to refute them. . . . It often happens that a difficulty is found in discussing or arguing a given position because the definition has not been correctly rendered. . . . In mathematics, too, some things would seem to be not easily proved for want of a definition, e.g. that the straight line parallel to the side, which cuts a plane figure divides similarly (*homoiōs*) both the line and the area. But once the definition is stated, the said property is immediately manifest: for the areas and the lines have the same *antanairesis* and this is the definition of the same ratio. . . . But if the definitions of the principles are not laid down, it is difficult, and may be quite impossible, to apply them. There is a close resemblance between dialectical and geometrical processes.

Here Aristotle (i) makes some observations about method: to avoid sophistry you must define your terms unambiguously and connect inference to inference up to your conclusion; (ii) introduces an example from mathematics. His description is loose, but it clearly means that, in Fig. 1.3, where the rectangles might equally well be replaced by parallelograms (though other such examples always seem to refer to rectangles), the ratio of $A$ to $B$ is equal to the ratio of $a$
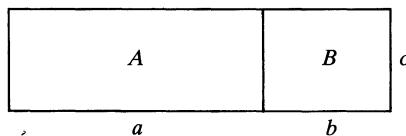


FIG. 1.3. The *Topics* proposition

to $b$. Note that this can be interpreted as a statement either about ratio or about proportion; and (iii) he sketches a proof of this statement. I shall now discuss the mathematical proposition (ii), while attempting to follow the precepts in (i); and will postpone the exploration of Aristotle's proof (iii) to the next chapter.

The proposition of Fig. 1.3, which I shall hereafter refer to as 'the *Topics* proposition', is a version of *Elements* VI 1:

Triangles and parallelograms which are under the same height are to one another as their bases.

Euclid proves it in the usual way of Greek mathematics: he draws an appropriate figure and makes an observation about this figure which refers, implicitly, to his definition of proportionality at V Definition 5 (which I have not quoted). This makes manifest that the four magnitudes satisfy this condition and so are proportional. I shall construct an alternative definition of proportionality out of early ingredients, and then give another proof. However, I am *not* offering this as a historical reconstruction of a definition and proof that may have been used in early Greek times; on this we have no evidence. Its roles are to refute the mathematical assertion that such a proof is impossible and to introduce an important kind of manipulation. Other definitions of ratio and proportion, also constructed out of early ingredients, will be described later, and each such definition will lead to a corresponding proof of the *Topics* proposition.

We perform a manipulation that is familiar within any deductive development of mathematics: we take a proposition that is equivalent to the definition we wish to replace, and use that instead as the definition. We have to proceed case by case. For lines, we have VI 16:

If four straight lines be proportional, the rectangle contained by the extremes is equal to the rectangle contained by the means; and if the rectangle contained by the extremes be equal to the rectangle contained by the means, the four straight lines will be proportional;

the second half will now serve as our definition. For numbers, we adapt the similar proposition VII 19. For two geometrical objects $a$ and $b$, and two numbers $n$ and $m$, we define $a:b::n:m$ if $ma = nb$ where $ma$ is the operation of multiplication as repeated addition, such as we have already seen in V Definition 4 and VII Definition 15. For two straight lines $a$ and $b$, and two plane regions $A$ and $B$, we define $a:b::A:B$ if the rectangular prism with base $A$ and edge $b$ is equal to the prism with base $B$ and edge $a$. Dimensional restrictions prevent us from giving a similar definition for four plane regions, but we can circumvent these by proving the *Topics* proposition and then applying this to reduce one or both of the pairs of plane regions to ratios of lines. A new approach will be needed to deal with the ratio of two three-dimensional regions, but, as it stands, this definition is adequate to handle all the rectilineal plane proportion theory of the *Elements* including the whole of Book X.

The proof of the *Topics* proposition now follows immediately: $a:b::A:B$ means that the rectangular prism with base $A$ and edge $B$ should be equal to that with base $B$ and edge $a$; but both of these are rectangular parallelepipeds with edges $a$, $b$, and $c$, where $c$ is the height of the original rectangle. (The reader is highly recommended to draw a figure.)                    QED

Note that this proof is expressed in terms of proportion theory; the ratios $a:b$ and $A:B$ are not here defined and they have no meaning outside a proportion. We have formal techniques, today, for passing from a definition of proportion to an idea of ratio—any reasonable definition of proportion will define an equivalence relation, and the resulting equivalence class *is* the ratio—but I believe that we cannot find anywhere, in Greek mathematics, anything that can be reasonably interpreted as implying any consideration of this kind of formal manipulation; see Section 10.2. Nor can we find any other definition of ratio, other than the statements I have quoted. Euclid's two definitions that concern ratio, V Definitions 3 and 4, do not give us enough information to construct a proof of the *Topics* proposition, which is the criterion I shall use for deciding whether a definition of ratio is mathematically adequate. Heath's comment sums up the situation:

The true explanation of its presence [sc. V Definition 3] would appear to be substantially that given by Barrow (*Lectiones Cantabrig*, London, 1684, Lect. III of 1666), namely that Euclid inserted it for completeness' sake, more for ornament than for use, intending to give the learner a general notion of ratio by means of a metaphysical, rather than a mathematical definition; "for metaphysical it is and not, properly speaking, mathematical, since nothing depends on it or is deduced from it by mathematicians, nor, as I think, can anything be deduced." This is confirmed by the fact that there is no definition of λόγος in Book VII, and it could equally have been dispensed with here [Heath, *EE* ii, 117].

As Heath observes here, we also do not find anywhere a definition of the ratio of two *arithmoi*, $n:m$. Moreover I do not believe that we have any good evidence that anybody—mathematicians, accountants, schoolmasters, or schoolchildren—deployed anything corresponding to our conception of a common fraction $n/m$; all that we do have is a minute quantity of very tenuous and ambiguous early evidence on papyrus, and then abundant examples, from the ninth century onwards, in manuscripts that have been subjected to drastic modernisation of script and, possibly, numerical notation. This is a very contentious opinion, so I shall later devote much of the long Chapter 7 to this topic. Yet, notwithstanding this absence of any apparent mathematically usable definition, the word *logos* frequently appears to refer to ratio, in isolation outside a proportion: Plato and Aristotle make many such references, and ratios are 'compounded' in the *Elements*, at VI 23:

Equiangular parallelograms have to one another the ratio compounded of (the ratios of) their sides.

and, in a similar proposition for the *arithmoi*, at VIII 5. The difficulties of interpretation of these last two propositions are augmented by the obscurities

of the definition of compounding, itself almost certainly an interpolation, at VI Definition 5:

A ratio is said to be compounded of ratios when the sizes of the ratios, having been multiplied together, make some ratio. (*Logos ek logōn sugkeisthai legetai, hotan ai tōn logōn pēlikotētes eph' hautas pollaplasiastheisai poiōsi tina.*)

I shall return to this topic in Section 4.5(c). Further, the massive and monolithic programme of construction and classification of incommensurable lines in Book X of the *Elements* seems to be connected with ratios of pairs of lines. This programme, though not necessarily the actual text of Book X, is attributed to Theaetetus, who died in 369 BC, before Eudoxus is supposed to have introduced the general theory of proportionality of Book V, and it is based on an idea of pairs of lines being either commensurable or incommensurable 'in square'; those that are incommensurable in square are called *alogos*, without ratio. I shall discuss the content and motivation of this book later, in Chapter 5.

Our evidence seems abundantly to indicate that some idea of ratio, in isolation, not occurring within a proportion, appears to be an important ingredient of early Greek mathematics.

### 1.2(e) *The language of Greek mathematics*

Greek mathematicians seemed to confront directly the objects with which they were concerned: their geometry dealt with the features of geometrical thought-experiments in which figures were drawn and manipulated, and their *arithmētikē* concerned itself ultimately with the evident properties of numbered collections of objects. Unlike the mathematics of today, there was no elaborate conceptual machinery, other than natural language, interposed between the mathematician and his problem. Today we tend to turn our geometry into arithmetic, and our arithmetic into algebra so that, for example, while *Elements* I 47:

In right-angled triangles the square on the side subtending the right angle is equal to the squares on the sides containing the right angle

means literally, to Euclid, that a square can be cut into two and manipulated into two other squares (see Fig. 1.4 for the proof of I 47, or Fig. 1.5 for a proof that may have been excised from between II 8 and 9), the result is now usually interpreted as

$$p^2 + q^2 = r^2,$$

where we now must explain just what the *p*s, *q*s and *r*s are and how they can be multiplied and added. To us, the literal squares have been replaced by some abstraction from an arithmetical analogy.

A passage from Plato's *Gorgias*, 451a–c, illustrates the point further, and introduces new features that will come to preoccupy us:

Suppose that somebody should ask me: . . . Socrates, what is *arithmētikē*? I should reply, as you did just now [about rhetoric], that it is one of the arts which secure their effect
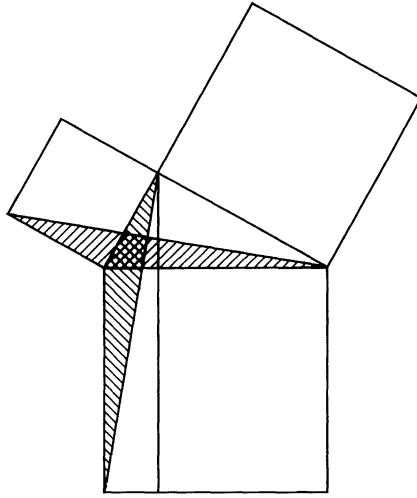
FIG. 1.4

through speech. and if he should further enquire in what field, I should reply that of the odd and the even, however great their respective numbers might be. And if he should next enquire, What art do you call *logistikē*? I should say that this art too is one of those that secure their entire effect through words. And if he should further demand in what field, I should reply, like the mover of an amendment in the Assembly, that in details 'hereinbefore mentioned' *logistikē* resembles *arithmētikē*—for its field is the same, the even and the odd—but that *logistikē* differs in this respect, that it investigates how numerous are the odd and the even both relative to themselves and relative to each other.
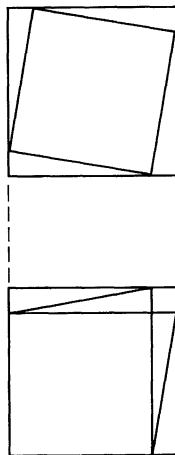


FIG. 1.5

Note how Plato emphasises speech, not writing (though this should also be seen in the context of Socrates' discussion here of the art of rhetoric); and add to this the barbed myth he tells at *Phaedrus* 274c–5b:

The story is that in the region of Naucratis in Egypt there dwelt one of the old gods of the country, the god to whom the bird called Ibis is sacred, his own name being Theuth. He it was that invented *arithmos te kai logismos*, geometry and astronomy, not to speak of draughts and dice, and above all writing. Now the king of the whole country at that time was Thamus, who dwelt in the great city of Upper Egypt which the Greeks call Egyptian Thebes, while Thamus they call Ammon. To him came Theuth, and revealed his arts, saying that they ought to be passed on to the Egyptians in general. Thamus asked what was the use of them all, and when Theuth explained, he condemned what he thought the bad points and praised what he thought the good. On each art, we are told, Thamus had plenty of views both for and against; it would take too long to give them in detail. But when it came to writing Theuth said, "Here, O king, is a branch of learning that will make the people of Egypt wiser and improve their memories; my discovery provides a recipe for memory and wisdom". But the king answered and said, "O man full of arts, to one it is given to create the things of art, and to another to judge what measure of harm and of profit they have for those that shall employ them. And so it is that you, by reason of your tender regard for the writing that is your offspring, have declared the very opposite of its true effect. If men learn this, it will implant forgetfulness in their souls; they will cease to exercise memory because they rely on that which is written, calling things to remembrance no longer from within themselves, but by means of external marks. You have found a charm not for remembering but for reminding, and you are providing your pupils with the semblance of wisdom, not the reality. For you will find that they have heard much without having been taught anything and that they will seem men of many judgements, though for the most part being without judgement and hard to live with into the bargain since they have become conceited instead of wise".

The transition in Athenian intellectual life from an oral to a written culture may not have been far in the past when Plato wrote these words, though we do not know enough to be absolutely sure about this; but Plato's sympathies lie clearly in the past, tremendous reader and writer though he was. I shall return to the question of the physical form of our guidance—just what texts we have, how reliable they might be, and how they have been edited in ancient and modern times—in Chapters 6 and 7.

To a mathematician, the point of an argument resides not in the words that make up the argument, but in the ideas that lie behind these words. In Greek mathematics, these ideas can always be expressed, and not merely described, in everyday language, although sometimes we may find this language tortured almost to extinction of its meaning. Also these words might have initially been carried within an oral tradition, and so been even more subject to the reorganisation, adaptation, adjustment, and revision that is commonplace even within the later written tradition of mathematics.

We shall therefore be concerned with a kind of mathematics that can be completely expressed in everyday language. Whenever I resort to any symbolic notation in the context of Greek mathematics, the reader should confirm that it is being used only as a convenient shorthand and that the symbols themselves

have neither more nor less meaning than the words that they replace. For example, I shall introduce the notation $\sqrt{n}$ to stand for the side of a square equal to $n$ times the square on a given line; and if $a$ and $b$ are two lines, then $a.b$ will denote the rectangle with sides $a$ and $b$, the construction introduced in the first proof of the *Topics* proposition given above. These ideas and notations are prompted by *Elements* II Definition 1 and Proposition 14:

Any rectangular parallelogram is said to be contained by the two straight lines containing the right angle

and

To construct a square equal to a given rectilineal area.

So, to write $\sqrt{2} . \sqrt{3}$ is to employ a convenient shorthand to describe a particular rectangle. But, within this interpretation, to write $\sqrt{2} . \sqrt{3} = \sqrt{6}$ is to make a meaningless assertion, while if we render it meaningful by writing, for instance, $\sqrt{2} . \sqrt{3} = \sqrt{6}.1$, then it becomes a non-trivial geometrical statement about the equality of two rectangles which must be established by some geometrical argument based, for example, on the configuration of I 43 (see Fig. 1.2, above, and the end of Sections 4.5(b) and (e)). If, however, we are tempted to interpret the statement arithmetically, then we should heed Dedekind's words, in his *Stetigkeit und Irrationale Zahlen*. After having defined addition, he goes on:

Just as addition is defined, so can the other operations of the so-called elementary arithmetic be defined, viz., the formation of differences, products, quotients, powers, roots, logarithms, and in this way we arrive at real proofs of theorems (as, e.g., $\sqrt{2} . \sqrt{3} = \sqrt{6}$) which, to the best of my knowledge have never been established before [p. 22].

Our familiarity, since the sixteenth century, with the convenient notation for decimal fractions and, since the nineteenth century, with well-founded arithmetised mathematics, has tended to dull our awareness of the machinery that must lie behind a formally correct account of the arithmetical manipulations of real numbers. These problems will not be our concern here, but we must be ever vigilant that any symbolism that is introduced as a geometrical shorthand does not involve any anachronistic and prohibited arithmetical sleight-of-hand.

## 1.3 SOCRATES MEETS THE SLAVEBOY AGAIN

Early Greek mathematics, I have suggested, is not arithmetised and did not even have at its disposal our notation or conception of common fractions. Yet, on the other hand, we find that mathematicians of the highest calibre, who elsewhere were exercised to attempt precise and explicit explorations of their basic concepts, seemed to find no difficulty in using freely the idea of the ratio of two numbers or of two geometrical magnitudes, and nobody at that time seemed to have expressed any concern about its meaning, not even in the face of knowledge of the phenomenon of incommensurability.

Can this picture possibly be consistent?

I shall now take seriously Socrates' final remark about Meno's slaveboy's potential as a mathematician and will make him the mouthpiece for some of my proposals. I will try to imagine how he could have answered if Socrates had asked him for the meaning of the ratio of two numbers. I shall suppose that Socrates' previous question in the *Meno* some time earlier has led the boy to do a substantial amount of private study and this has developed a latent ability in mathematics. And, to keep the exposition moving briskly, I shall make him a little more forthcoming than he was when Socrates first met him. So picture a talented youth, who advances towards learning and investigation smoothly and surely and successfully, with perfect gentleness, like a stream of oil that flows without a sound, so that one marvels how he accomplishes all this at his age (Plato's description of another mathematician, borrowed from *Theaetetus* 144b). I imagine the conversion set around 370 BC; at this time, Plato is involved with his highly philosophical and mathematical middle dialogues (*Republic, Theaetetus, Parmenides, Sophist, Statesman,* and *Timaeus*), Archytas and Theaetetus have made considerable contributions to mathematics, and Eudoxus' influence is beginning to be felt, though his proportion and exhaustion theories of *Elements* V and XII lie in the future; but the dialogue contains one blatantly anachronistic fantasy in Socrates' reference to 'my young friend Euclid'. Suppose, then, that Socrates meets the slaveboy again in Meno's house, and is tempted to resume his earlier conversation. (Each speech is numbered for later reference.)

SOCRATES$_1$: Tell me, boy, what is the relationship of size between this heap of sixty stones, and that heap of twenty-six stones?

BOY$_2$: Do you mean the number of times the smaller goes into the larger?

SOCRATES$_3$: Try it.

BOY$_4$: It goes more than twice, but less than three-times.

SOCRATES$_5$: Can you be more precise?

BOY$_6$: It goes twice with eight stones left over.

SOCRATES$_7$: Those 'eight stones' aren't related to anything now.

BOY$_8$: I just omitted to say that they're still in relation to the other heap of twenty-six stones.

SOCRATES$_9$: Go on.

BOY$_{10}$: Give me time. I can now describe that relationship in the same way. So let's say: twice at the first step, and now I'll tell you about the relationship between eight stones and twenty-six stones.

SOCRATES$_{11}$: Go on.

BOY$_{12}$: Again that will be the number of times the smaller goes into the larger: that's three-times and then the relation between two stones and eight stones; and that's four-times-exactly.

SOCRATES$_{13}$: So you've extracted a relationship expressed by: first step, twice; second step, three-times; third step, four-times-exactly. The technical name for that is a ratio.

$$(60 : 26 = [2, 3, 4])$$

BOY$_{14}$: So that's what a ratio is! I've always wondered what those mathematicians mean when they use the word *logos*.

SOCRATES$_{15}$: It's just one of the different possible definitions. Can you deduce anything from that last step?

BOY$_{16}$: Yes, that we could have manipulated throughout in twos. So heaps of thirty and thirteen stones will give the same ratio.

SOCRATES$_{17}$: Now let's compare this ratio with some other ratios. Is it bigger than the ratio: *three*-times, three-times, four-times-exactly?

BOY$_{18}$: You haven't said that ratios can always be compared or told me what it means; but I'd say that, however it's defined, it ought to imply that anything that starts 'first step, three-times' will be bigger than anything that starts 'first step twice'.

$$([2,\dots] < [3,\dots])$$

SOCRATES$_{19}$: Good. It's sometimes better in mathematics to leave those formal points of definition and verification until later, and first find out what kind of thing you want to do. Try another comparison, this time with the ratio twice, *four*-times, four-times-exactly. Think carefully.

BOY$_{20}$: That's strange. Since the second remainder comes from the second, originally smaller, pile, we can see that anything that starts 'twice, four-times' will be *less* than anything that starts 'twice, three-times'.

$$([2,4,\dots] < [2,3,\dots])$$

That kind of thing seems to happen generally, since it depends on which of the two original piles the remainder comes from.

SOCRATES$_{21}$: Yes. To see which of two ratios is the greater, you have to note whether their terms differ first at the first, third, fifth, . . . step, or the second, fourth, sixth, . . . step since, in the latter case, the reverse of the relationship of less and greater between the *terms* holds between the *ratios*; this will be the first way we shall find that the greater and less is connected with the odd and even. So do you see how we can sandwich our ratio: twice is less than twice, four-times, which is less than our ratio twice, three-times, four-times, which is less than twice, three-times, which is less than three-times.

$$([2] < [2,4] < [2,3,4] < [2,3] < [3])$$

BOY$_{22}$: Yes, and I can see how it will work in general. With each step you get closer and closer, alternating above and below.

SOCRATES$_{23}$: Will you always get there?

BOY$_{24}$: I suppose so. The heaps are always getting smaller and smaller, so the process can't go on for ever; I wonder how long it takes. Can I ask you a question?

SOCRATES$_{25}$: Is that your question? Ignore that remark; we're doing mathematics, not logic! Go ahead.

BOY$_{26}$: What two heaps will have the ratio twice, four-times, three-times-exactly?

SOCRATES$_{27}$: It's easy to work that out backwards: the final step will be three stones to one stone (or six to two, and so on). So the penultimate step will be four-times three plus one (that's thirteen) stones to three. So the answer will be twenty-nine stones to thirteen stones, or fifty-eight to twenty-six, and so on. But if you now want to work out the ratio twice, four-times, three-times, followed by anything else, you'll have to

start all over again. Now there's a very convenient and easy forward method, but my young friend Euclid is baffled by the problem of explaining it formally.

BOY$_{28}$: So we can now translate that chain of inequalities you just told me back into heaps of stones: the ratio of two stones to one stone is less than the ratio of nine stones to four stones, which is less than our ratio of sixty stones to twenty-six stones, which is less than the ratio of seven stones to three stones, which is less than the ratio of three stones to one stone.

$$([2] < [2,4] < [2,3,4] < [2,3] < [3] \text{ becomes}$$

$$2:1 < 9:4 < 60:26 < 7:3 < 3:1)$$

SOCRATES$_{29}$: Yes, and if you do a few more conversions that stretch out a bit further than three steps, and explore the idea a bit, you'll soon find out the forward method for yourself. Sometimes I think in terms of heaps of stones, and sometimes I convert them into ratios, since I find that gives me a lot more information.

BOY$_{30}$: What do you mean?

SOCRATES$_{31}$: Not too fast! I can't keep up with your questions.

BOY$_{32}$: Well here's another question: Can you add or multiply these ratios?

SOCRATES$_{33}$: What nonsense! Who ever thought of adding the ratio of twenty-nine to thirteen stones to the ratio of sixty to twenty-six stones. It doesn't make sense!

BOY$_{34}$: But suppose we have three heaps of stones; how can we calculate the ratio of the first to the third from the ratio of the first to the second and the second to the third? Or a similar question about adding: how can we get the ratio of the first and second combined to the third out of the ratios of the first to third and second to third? Surely these are sensible questions.

SOCRATES$_{35}$: You've got me there! This definition of ratio has the peculiar feature that these kinds of calculations seem curiously intractable; even Theaetetus can't do them directly and if he can't, I'm certainly not going to try. But if you think about ratios in quite a different way, the way that Archytas and the music theorists have been doing, then these kinds of manipulations make good sense in general and become quite easy; but their approach doesn't allow them to describe all ratios that can occur in geometry. Eudoxus and the theoretical astronomers have just started to think about yet another description of ratio, but they haven't been able to do much with their definition so far. Then linking these different but equivalent ideas poses further problems. There are even technical difficulties with the procedures used by scribes and taxmen when they work out their accounts, but there's no interest in cleaning up that mess! So you can see that there are lots of problems still to be solved with your 'adding and multiplying ratios'. Let's move on to a different topic: what about the ratio of two lines?

BOY$_{36}$: Surely I can just use the same process. I wonder if it can then go on forever.

SOCRATES$_{37}$: And of two plane regions or two solids?

BOY$_{38}$: I can use geometrical constructions to convert two rectilineal regions into two adjacent rectangles with equal vertical sides; then the ratio of the rectangles is the ratio of the bases, so we're back to the ratio of lines. But I foresee problems in doing these operations with curvilinear regions in general and most kinds of solids.

SOCRATES$_{39}$: Yes. Eudoxus is working on just these problems now. What about the ratio of the diagonal to a side?

BOY$_{40}$: Of what?

SOCRATES$_{41}$: A square, a pentagon, anything you like. Or the circumference to a diameter? Or the surface to a section? Or the circumdiameter to a side, a diagonal, or an edge? Or the dimensions of squares and cubes?

BOY$_{42}$: Not so fast! All of this is very intriguing. How can I find out more?

SOCRATES$_{43}$: Go and see Plato and his friends in the garden of Academia out in the north-western suburbs; they don't seem to talk about anything else now. But be warned: studies that demand more toil in the learning and practice than this you will not discover easily, nor find many of them; it looks like a long time before this idea of ratio will be understood.

This last sentence begins with another quotation from Plato (*Republic* VII, 526c); and my slaveboy should perhaps also be warned about the reception he is likely to get at the garden of Academia if the description at *Republic* VI, 495c–496a, a passage that I am not going to quote here, is a reliable indication of Plato's aristocratic attitude to other social classes and ungentlemanly activities. This is an important aspect of any evaluation of the beliefs and activities of Plato and his associates, but it will only make a very bizarre and fleeting appearance in my story here, in Chapter 6, note 11.

## 1.4 NOTES AND REFERENCES

For a description of the system of referencing books and articles, see the introduction. There is a detailed bibliography, arranged alphabetically, at the end of the book.

1.1 The translation by Guthrie of the *Meno* is taken from Plato, *CD*; this translation is reprinted, with an interesting collection of articles, in Brown, *PM*. For a critical edition of the Greek text with full commentary and discussion of the date of the dialogue, see Plato–Bluck, *PM*. There is almost complete unanimity over the figures, all of which are omitted from the manuscripts. For a recent discussion, see Boter, PM; Sharples MPM; and Fowler, YMPM.

1.2(a) Dedekind's *Stetigkeit* is translated in his *ETN*. This essay remains the most lucid description of the 'real numbers' that I know. Dedekind's opinions on Eudoxus and his assertion about arithmetic, quoted in Section 1.2(e), have often been challenged, but I believe they stand correct.

One by-product of the development of computers and the consequent study of algorithms has been to make evident some of the subtleties and problems that underlie arithmetised mathematics. An excellent general reference is Knuth, *ACP* ii; and, for an informal description of some of the difficulties underlying arithmetic with decimal fractions, see my *FHYDF*. I shall return to this topic in Sections 4.5(b) and (e), and elsewhere.

1.2(b) Heath's discussions of equality in the notes in *TBEE* i are unsatisfactory: up to I 35 he treats the word as if it meant congruence, and then he writes (*TBEE* i, 327): "Now, without any explicit reference to any change in the meaning of the term, figures are inferred to be *equal* which are equal in *area* or in *content* but need not be of the same *form*". Surely 'equal' throughout means 'equal in magnitude'. However, his remarks on the following page on 'content' are very valuable; they enable us to say, in modern terms, that Euclid is developing the results needed for a non-arithmetised theory of ordered content, for line magnitudes in Books I 1–3, rectilineal plane areas in I, II, and VI, and simple solids in XI. Eudoxus' theory of exhaustion enables this theory to be extended to circles, pyramids, cones, and spheres in Book XII; see my dialogue, Socrates$_{39}$, in reply to Boy$_{38}$. Using these theories, we can say when two geometrical

magnitudes are equal or, if they are unequal, which is the greater, which is exactly what is needed for the ratio theories I shall be describing.

There are four places in the *Elements* where a local reference such as "as was proved in the theorem preceding this" is given: VIII 19 and 21, IX 12, and X 30. There are other places where general references to earlier results are found: "It is then manifest from what was proved before that . . . (X 55)"; "For let the same construction be made as before . . . , then, in manner similar to the foregoing" (X 56; also see 57–9, 61–5, and 91); and the general and vague "for we have learnt how to do this" (X 10). These are the *only* specific references in the *Elements*. (I am grateful to Malcolm Brown for this information, which does not seem to have been pointed out or discussed in any of the literature on the *Elements*.) The various references added, in editorial square brackets, to Heath's translation *TBEE* are far from complete; a more detailed analysis of the logical structure of the *Elements* can be found in Mueller, *PMDSEE* which incorporates the earlier analyses of Neuenschwander, *EVBEE* and *SBEE*. Some such back-references are found in marginal annotations (scholia) to our earliest complete manuscript of the *Elements*; for an illustration, see the frontispiece of Heath, *TBEE* i with its brief description on p. xi.

A counter-example to *Elements* XI Definition 10 can be got by pushing inwards one of the pyramidal groups of faces round a vertex of an icosahedron; however Cauchy proved that the definition is unambiguous for convex polyhedra. For instructions on how to make a model of an exotic non-convex flexible counter-example, see the article by its discoverer, Connelly, *FS*.

1.2(c)   See Klein, *GMTOA*, for a thorough discussion of the *arithmoi, arithmetikē*, and related topics. I know of no place where the role of repetition numbers is considered, but have not made an exhaustive search; I was tempted to use the language of the schoolroom and call them 'gozinter numbers', as in 'two gozinter four twice', but have resisted. Greek numerical notation, and its extensions to fractional quantities, will be considered in Chapter 7.

1.2(d)   Book V Definitions 3 and 5 and Proposition 11 lead many mathematicians to believe that the early Greek idea of ratio is something like the equivalence class defined by proportionality. For a discussion, see the new Appendix, Section 10.2.

The translation of the *Topics* quotation is based on Aristotle–Ross, tr. Pickard–Cambridge, *WA* i, Heath, *MA*, 80, and Knorr, *EEE*, 257. The role of the *Topics* proposition in Book X is analysed in detail in Knorr, *EEE*, Chapter 8, and there is an excellent discussion of Book V proportion theory in Knorr, APEPT. The way in which the Greek word *parallēlogrammon* often refers to a rectangle is well illustrated in *Elements* X; this book will be described in detail in Chapter 5. *Elements* VI, Definition 5, quoted at the end of this section, is not listed by Heath with the other definitions on *TBEE* ii, 188, but it can be found on p. 189.

On the date of Theaetetus' death (see p. 211.11) there is now a lone dissenting voice of an expert on Platonic and Pythagorean matters. Thesleff, PC, p. 8 n. 67, writes: "I find it essential to note that historians of mathematics who take for granted that Theaitetos was still alive in the 370s must be wrong. He made some important discoveries as a young man, and Plato and his friends were deeply impressed by this. But he is likely to have died in 390 BC. . . .".

1.2(e)   Klein, *GMTOA*, discusses the conceptualisation and language of Greek mathematics; see especially p. 63: "Greek scientific arithmetic and logistic are founded on a 'natural' attitude to everything countable as we meet it in daily life. This closeness to its 'natural' basis is *never* betrayed in ancient science."

The translations of the *Gorgias* and *Phaedrus* quotations are adapted from those, by Woodhead and Hackforth respectively, in Plato, *CD*. I have corrected the last line of the first, and taken the last two pungent sentences of the second from Cherniss, *REA*, 3. The status of books, writing, and education in early Greek life is controversial; see Goody, *DSM* and Havelock, *LRG* for different points of view to that expressed here. Also, the use of 'book' may raise misunderstandings for a

modern user of books; see Eisenstein, *PREME*. I have followed Turner, *Athenian Books in the Fifth and Fourth Centuries B.C.*, and the final sentences of this brilliant analysis will serve as an envoi to my first chapter.

Plato . . . devotes some pages of the *Phaedrus* to an analysis of the deficiencies of books. Even if it be granted that the writer is disinterested and knows his theme inside out, a book suffers from many deficiencies. It inclines its reader to rely on someone else's bottled-up memory, and turns him into a pseudo-philosopher, a *doxosophos*. Its message, once frozen into writing, is rigid—a book cannot answer questions, if defective, or defend itself, if attacked. The author of a book is like a reckless gardener, writing in water or sowing in ink. It is impossible not to feel that Plato, tremendous reader though he was, is fighting a rearguard action against the written word's inhibiting effect on thought; that Plato realises the day is past when he could undo the harm done by a book by publicly convicting the author of ignorance. By the first thirty years of the fourth century books have established themselves and their tyranny lies ahead.

# ANTHYPHAIRETIC RATIO THEORY

## 2.1 INTRODUCTION

In response to the opening question of my dialogue, which was couched in the terms of *Elements* V Definition 3, my slaveboy develops a process of repeated reciprocal subtraction which is then used to generate a definition of ratio as a sequence of repetition numbers (see $S_{13}$ and $B_{14}$, i.e. the 13th speech, by Socrates, and the 14th, by the boy). As my Socrates says ($S_{15}$), and as I have already illustrated in Section 1.2(d), this is not the only possible way of defining ratio or proportion, even within the context of early Greek mathematics; so, to distinguish it, I shall often refer to this as 'the anthyphairetic ratio', and call the underlying subtraction process 'anthyphairesis'. These words are derived from the Greek verb *anthuphairein* used, at *Elements* VII 1 and 2, and X 2 and 3, to describe this operation of reciprocal subtraction. Euclid's only explicit reasons for introducing the process are to find the greatest common measure ($B_{16}$) and, in X 2, as a criterion for incommensurability which he then never uses explicitly; but I shall be arguing that substantial portions of the *Elements*, Plato's mathematical references, and other fourth- and third-century testimonies can be interpreted very plausibly in anthyphairetic terms. The sequence of repetition numbers that arises from anthyphairesis represents the result of repeated subtractions ($B_2$), not divisions, and this interpretation is corroborated by the Greek name, derived from *anti-hypo-hairesis*, 'reciprocal sub-traction'. I wish to avoid the modern names, the 'Euclidean algorithm' and 'continued fractions', by which the process and associated mathematical objects are now known, for the following reasons, among others that will underlie my reconstruction: the Euclidean algorithm is now generally construed as a division process, whereas anthyphairesis is based on repeated subtraction; continued fractions are now handled using the real numbers and a sophisticated generalisation of fractions, whereas I shall propose a different approach in which a heuristic stage, based on the properties of what are now called Farey series and making use only of manipulations of *arithmoi*, is followed by geometrically formulated proofs in the style of, but much more elaborate than, those we have seen in *Meno* 81–5; and Euclid's *Elements* may even have contributed to the decline of interest in and knowledge of the process.

It seems natural to expect that ratios should themselves be comparable in size though, as the boy points out ($B_{18}$), it is not obvious that the comparison of cardinal ($B_2$) and repetition ($B_4$) numbers should automatically imply a comparison of anthyphairetic ratios, sequences of repetition numbers. Euclid

makes a similar kind of silent assumption in his proportion theory at *Elements* V 10: he has defined proportion and disproportion (a condition that *a* to *b* should be less than or, implicitly, greater than *c* to *d*) in V Definitions 5 and 7, but in V 10 he assumes a trichotomy law, that any four magnitudes must be either proportional or disproportional. The comparison of anthyphairetic ratios reveals the curious feature that the lexicographic ordering of ratios is reversed at the second, fourth, sixth, etc. steps ($B_{20}$ and $S_{21}$). This is only the first of the instances we shall find in which an even–odd parity plays a fundamental role in anthyphairetic ratios, and, I shall argue, it is only one of the many examples of behaviour to which Plato might be referring when he talks of the 'mutual relationships' of 'the even and the odd' and, sometimes, the 'greater and the less'; one such passage, *Gorgias* 451a–c, was quoted in Section 1.2(e), and others will be considered in Section 4.2.

As my slaveboy observes ($B_{36}$), the same process can be applied to two lines and, more generally, to any two geometrical magnitudes which allow of the necessary operations of comparison and subtraction. The *Topics* proposition provides a very convenient configuration for evaluating the ratio of two plane regions ($B_{38}$), and, as Aristotle says at *Topics* 158b29 ff. (quoted in full in Section 1.2(d)),

The areas and the lines have the same *antanairesis* and this is the definition of the same ratio (*ton auton logon*).

(The importance of this passage was first pointed out in 1933 in Becker, ESI, since when every serious discussion of Greek proportion theory has considered it.) The word used here is derived from the compound *anti-ana-hairesis*, 'reciprocal re-traction'. Some have argued that the two words have different meanings, but I shall treat them as synonymous, as did Alexander of Aphrodisias, in his commentary on this passage (Alexander–Wallies, *IT*, 545):

The said [fact] expressed in such-like terms is not familiar, but it becomes familiar once the definition of 'in proportion' (*analogon*) is stated, that the line and the space are divided in proportion by the drawn parallel. Now this is the definition of proportionals which the ancients used: those magnitudes are in proportion to each other (and *homiōs* to each other) of which the *anthuphairesis* is the same. But he [sc. Aristotle] has called *anthuphairesis antanairesis*.

These words also seem to be used interchangeably in commercial documents on papyrus, to describe a deduction from a financial account. (There is here no suggestion of a process, the *repeated* subtraction, that is conveyed by the adverb *aei* in Euclid's usage; see $B_{10}$.) Of the many examples, here are two from the same collection of documents, the Zenon Archive: *anthuphairein* in P. Lond. vii 1994, lines 164, 176, 233, and 321, and vii 1995, line 333 (both dated Autumn 251 BC); and *antanairein* in P. Cair. Zen. iii 59355, lines 95 and 150 (dated Summer 243 BC). A third word, *antaphairein*, is found in Nicomachus, *Introduction to Arithmetic* I 13.11, where it is used with exactly the same sense.

Another comment of Aristotle, at *Metaphysics* 1021a3–8, has recently been explained as relating to anthyphairesis. This passage, taken out of its rather complicated context, reads

. . . But the exceeding relative to the exceeded generally is indefinite in respect of *arithmos*, for *arithmos* is commensurable, and *arithmos* is not said of something not commensurable; the exceeding relative to the exceeded is so much and still more, and this is indefinite; for it is by chance either equal or unequal. So these are the things spoken of relative to something in respect of *arithmos*.

This translation is taken from Pritchard, MDPEM, where the author argues, very plausibly, that Aristotle is here referring to the first subtraction in the anthyphairesis of two incommensurable magnitudes (for which "*arithmos* is not said"). Then the remainder and the originally smaller magnitude will also be incommensurable. The reader is recommended to consult the article for a full discussion of the passage.

It is a natural step within *arithmetised* geometry to assign a unit length to some important or basic line within a figure, and then calculate the length of the other lines. For example, we would say that the length of the diagonal of a square with unit side is $\sqrt{2}$ (a real number); the longer diagonal of a hexagon with unit side is 2, the shorter is $\sqrt{3}$; and so on. The equivalent step within *anthyphairetic* geometry is to evaluate the ratio of two lines, areas, or volumes that are related by some construction. My Socrates suggests several such problems ($S_{39}$ and $S_{41}$) and he also hints that this approach might reveal useful or interesting behaviour ($S_{29}$). The rest of this chapter will embark on the study of some of these examples. Since the mathematical approach may be novel, even to many mathematicians, I shall emphasise here this aspect of the subject, but will illustrate the procedures, wherever possible, with ancient examples. Until these mathematical techniques become more familiar, the historical questions of whether or not some of them might be plausible reconstructions of ancient practice cannot be evaluated.

My concern throughout this book will be to explore the properties of anthyphairetic and other kinds of ratios, and not to develop abstract theories of ratio or proportion, akin to those we find in *Elements* V and VII. Investigations of the formal aspects are very important, but a full treatment of the topic can be found in Knorr, *EEE*, Chapter 8 and Appendix B. I see no need to duplicate that work here.

## 2.2 SOME ANTHYPHAIRETIC CALCULATIONS

### 2.2(a) *The diagonal and side*

Consider first the problem of finding the ratio of a diagonal and side, proposed at $S_{39}$ and $S_{41}$. The general problem is as follows: given a line, construct a regular polygon with this line as side, take a diagonal (there will be a choice of different diagonals when there are six or more sides) and calculate the ratio of

this diagonal to the side. In each case, we needed some additional geometrical construction which will embody the information we seek, just as the first step of the solution of Socrates' problem in the *Meno* was to draw a figure which contained the obliquely placed square. (Note in passing that the word *diagramma* seems, in Plato and Aristotle, to refer ambiguously to either a geometrical figure or a proof.) The constructions for the square, pentagon, and hexagon, the three examples I shall consider here, are given in Fig. 2.1; these having been drawn, it remains to say the appropriate words about them.

First, the square. Figure 2.1(a) shows how, starting from the small square placed obliquely in the corner, we can construct a large square with side equal to the small side plus small diagonal; when the figure is completed, as shown, it becomes clear that the large diagonal is twice the small side plus the small diagonal. Observe that this figure is not unlike the successful third *Meno* construction, and it also resembles the proof of Pythagoras' theorem that may have been excised from between *Elements* II 8 and 9 (see Fig. 1.5): all consist of a square or rectangle placed obliquely inside a larger square. We now calculate the ratio of the large diagonal plus large side to the large side: this clearly is twice, followed by the ratio of the large diagonal minus side to the large side. Again, from the figure, this latter ratio, taken the greater to the less ($B_2$), will be the ratio of the small diagonal plus side (equal to the large side) to the small side (equal to the large diagonal minus side). Since the absolute size of the figure does not affect these ratios (cf. *Meno* 82c), we have proved that the ratio of a diagonal plus side to side of any square is equal to twice, followed by the same ratio; and so is equal to twice, twice, followed by the same ratio; and so is equal to twice, twice, twice, twice, followed by the same ratio; and so is equal to an unending sequence twice, twice, twice, etc. Now we can modify the first term to get the result that the ratio of the diagonal to side of a square is equal to once, twice, twice, twice, etc. (Any readers who find such a complete but informal description unsatisfactory or unconvincing should, here and elsewhere, rework the material to their own satisfaction, within the stylistic restrictions imposed by the historical context.) I shall discuss some evidence that may relate to Fig. 2.1(a) in Section 3.6, below.



(a)                    (b)                    (c)
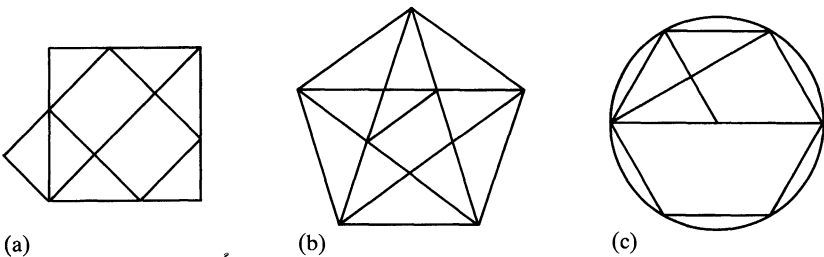
FIG. 2.1

This proof depends on the result that my slaveboy appreciated right from the start ($B_8$ and $B_{10}$), that some terms of a given ratio, of numbers or magnitudes, can be evaluated and then followed by the ratio of the remaining numbers or magnitudes. Here this gave rise to the recursive behaviour in which a ratio was equal to twice, followed by the same ratio, and so the original ratio, by eating its own tail, is completely evaluated. This remarkable kind of phenomenon will itself recur time and time again. The proof started from the ratio of diagonal plus side to side purely for reasons of brevity and convenience; the evaluation of the ratio of diagonal to side works equally well, as the reader is encouraged to verify; but is one step longer, since the recursive property does not then manifest itself until the second subtraction step has been performed.

Next the pentagon. It is easy to prove that the diagonals of a pentagon will cut each other so as to generate a smaller pentagon, in a configuration known as a pentagram; see Fig. 2.1(b). The many isosceles triangles in the figure enable us to see that the side of the large pentagon is equal to the side plus diagonal of the small, while the large diagonal is the small side plus twice the small diagonal. The evaluation of the ratio of the large diagonal to the large side then leads, after two steps, to the same kind of recursive behaviour; and this shows that their ratio is equal to the unending sequence once, once, once, etc. The reader is again strongly encouraged to work out the details.

There are many propositions on or involving the pentagon in the *Elements*. Of these, the closest in spirit to this figure seems to be XIII 8:

If in an equilateral and equiangular pentagon straight lines subtend two angles taken in order, they cut one another in extreme and mean ratio, and their greater segments are equal to the side of the pentagon,

which can be elucidated, pruned, and extended as follows:

If in an equilateral and equiangular pentagon [two adjacent diagonals] cut one another, . . . their greater segments are equal to the side of the pentagon [and their lesser segments are equal to the diagonals of the smaller pentagon formed by the larger diagonals].

Next the hexagon; see Fig. 2.1(c). Here there are two different diagonals: the larger is the diameter of the circumscribing circle and so is twice the side of a hexagon (this, of course, is a description of the ratio in any sense of the word), while the shorter is such that the square on it is three times the square on the side (this result is contained, explicitly, in the proof of XIII 12). The shorter diagonal is an example of a problem my Socrates called 'the dimensions of squares' ($S_{41}$): given a line $\rho$ and two numbers $n$ and $m$, construct squares equal to $n$ times and $m$ times the square on $\rho$; what then can we say about the ratio of their sides? I shall return to this example in Section 2.3(b) below, and discuss the solution of the problem in detail in the next chapter.

Other ratios associated with regular polygons and polyhedra will be considered in Chapter 5.

### 2.2(b) *The circumference and diameter*

One of the simplest of geometric figures is a circle. Clearly there are problems in treating the circumference of the circle as a geometrical magnitude, to be added to or compared with other line magnitudes, as Aristotle points out at *Physics* 248a10–13, but Archimedes, in his *Sphere and Cylinder* I, extends the scope of his geometry to include well-behaved curved one- and two-dimensional geometrical objects. The book is prefaced by some Axioms (*Axiōmata*), which start:

There are in a plane certain finite bent lines which either lie wholly on the same side of the straight lines joining their extremities or have no part on the other side. I call concave in the same direction a line such that, if any two points whatsoever are taken on it, either all of the straight lines joining the points fall on the same side of the line, or some fall on one and the same side while others fall on the line itself, but none falls on the other side;

and some Postulates (*Lambanomena*), which start:

Of all lines which have the same extremities, the straight line is the least. Of other lines lying in a plane and having the same extremities, [any two] such are unequal when both are concave in the same direction and one is either wholly included between the other and the straight line having the same extremities with it, or is partly included by and partly common with the other; and the included line is the lesser.

As a result of these principles, Archimedes is able to say, at the end of his postulates, that

With these premises, if a polygon be inscribed in a circle, it is clear that the perimeter (*perimetros*) of the inscribed polygon is less than the circumference (*peripherias*) of the circle; for each of the sides of the polygon is less than the arc of the circle cut off by it,

and, in Proposition 1, he then proves that the perimeter of a circumscribed polygon is greater than the circumference. The first six propositions then establish theoretical results about approximating the circumference by inscribed and circumscribed polygons; but the actual calculation of the magnitudes of their perimeters is carried out elsewhere, in his short work *Measurement of a Circle*. Proposition 3 of this treatise evaluates that:

The perimeter of every circle is three-times the diameter and further it exceeds by [a line] less than a seventh part of the diameter but greater than ten seventy-first [parts of the diameter].

I shall examine some textual details of Archimedes' argument in Section 7.3(a), where I shall describe some of the corruptions, additions, and emendations to our surviving manuscripts. I will illustrate there how Greek procedures may be based on manipulations that are very different in kind from, though equivalent to, our fractional manipulations; so that the last words here, "ten seventy-first [parts]" (*deka hebdomēkostomonois*), refer to lines, not numbers, so do not carry

the same connotations as our $\frac{10}{71}$. From the anthyphairetic point of view, the calculation shows explicitly that the ratio of circumference to diameter is less than three-times, seven-times, and greater than three-times followed by the ratio of seventy-one to ten (taken, as always, the greater to the less; see B$_2$), i.e. greater than three-times, seven-times, ten-times:

$$[3,7,10] < c:d < [3,7].$$

Other parts of the calculation also fit very well into an anthyphairetic context, as I shall illustrate in Section 2.4(d), below. We also have a short report of another closely related calculation by Archimedes, at Heron, *Metrica* I 25:

Archimedes proves in his work on plinthides and cylinders that of every circle the perimeter has to the diameter a greater ratio than 211 875 to 67 441, but a lesser ratio than 197 888 to 62 351. But since these numbers are not well-suited for practical measurements, they are brought down to very small numbers, such as 22 to 7.

Unfortunately Heron gives no further details, the work by Archimedes to which he refers is now lost, the alleged lower bound to the ratio is in fact a close upper bound, and the upper bound is a poorer estimate than 22 : 7.

Now to the illustration of some further properties of ratios, using this ratio $c:d$ as an example, as a first elucidation of my Socrates' remark (S$_{29}$) that these anthyphairetic ratios contain a lot of information. I encourage any reader who is daunted by the material in this or any subsequent section of this chapter and the next to skip-read or cut to the beginning of the following section, and return later, when the material becomes necessary, or more accessible, or more appealing. Each section starts from a modest level; some rise to encounter quite subtle phenomena.

The expansion of this ratio of circumference to diameter actually starts

$$c:d = [3,7,15,1,292,1,1,1, \ldots];$$

I shall extract some of the information that is coded in these numbers. Let us label the lines involved as follows:

$$c - 3d = r$$
$$d - 7r = s$$
$$r - 15s = t$$
$$s - t = u$$
$$t - 292u = v$$
$$u - v = w$$
$$\text{etc.}$$

(This is not algebra, i.e. symbolic manipulation, but mere shorthand; and a similar comment applies to the later manipulations in which we count
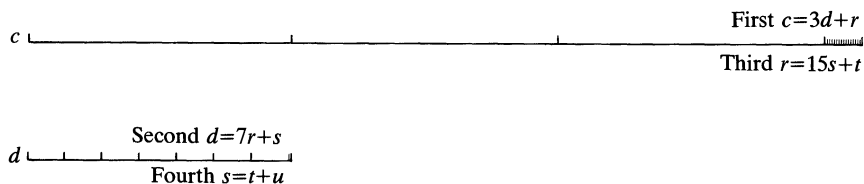
FIG. 2.2. Four steps in the anthphairesis of $c:d$

intervals). The lines $r$, $t$, and $v$ first arise as subintervals of $c$; while $s$, $u$, and $w$ are subintervals of $d$; see Fig. 2.2. Now consider the question of estimating the sizes of $r$, $s$, $t$, etc. We can make a series of increasingly precise statements:

(i) If we ignore the remainder at any step, and so treat the subtraction process as if it terminated at that step, we will get an approximation to the ratio. The remainders seem to be getting very small very quickly, so this should yield increasingly good approximations. The remainders at the first, third, fifth, etc. steps all arise as subintervals of the larger original magnitude $c$, so ignoring them should give underestimates since it will reduce the size of $c$; while ignoring the second, fourth, etc. remainders should similarly give overestimates ($B_{20}$ and $S_{21}$). Hence ($S_{21}$ and $B_{22}$):

$$[3] < [3, 7, 15] < [3, 7, 15, 1, 292] < \cdots < c:d < \cdots < [3, 7, 15, 1] < [3, 7];$$

then ($B_{28}$ and $S_{29}$) we can convert this into a string of ratios of numbers:

$$3:1 < 333:106 < 103\,993:33\,102 < \cdots < c:d < \cdots < 355:113 < 22:7.$$

(ii) The instruction to ignore a remainder at a certain step is rather vague, since this would seem to alter all of the preceding steps of the process. What we can do, instead, is to alter the *original* larger or smaller interval so that the subtraction process performed with this modified interval will terminate at the required step. We can see by how much to decrease or increase by counting intervals, as in the following analysis of Fig. 2.2. Consider the first two steps:

$$c = 3d + r = 3(7r + s) + r$$
$$= 22r + 3s$$
$$d = 7r + s.$$

Look at $7c$ and $22d$:

$$7c = 7 \times 22r + 21s$$
$$22d = 22 \times 7r + 22s;$$

they are very close to each other. If we increase $7c$ by $s$, or decrease $22d$ by $s$, they will be equal:

$$(c + s/7) : d = c : (d - s/22) = 22 : 7.$$

A similar calculation will hold at the third step, where the decomposition $r = 15s + t$ is introduced:

$$c = 22r + 3s = 22(15s + t) + 3s = 333s + 22t$$

$$d = 7r + s = 7(15s + t) + s = 106s + 7t.$$

Now look at $106c$ and $333d$. Since $106 \times 2 = 333 \times 7 + 1$, this time we have

$$(c - t/106) : d = c : (d + t/333) = 333 : 106;$$

and so on. The process seems to be perfectly general and seems to hold for any ratio; if the $n$th step of $c : d$ introduces a new interval $x$ as remainder, then the associated ratio $p : q$ which terminates on the $n$th step is got by adding ($n$ even) or subtracting ($n$ odd) the $q$th part of $x$ to the antecedent $c$, or subtracting ($n$ even) or adding ($n$ odd) the $p$th part of $x$ to the consequent $d$:

$$(c \pm x/q) : d = c : (d \mp x/p) = p : q,$$

where the signs alternate as $n$ increases. In Section 2.3(c) I shall address the further problem of calculating these ratios $p : q$.

Again, let me emphasise that the symbolism introduced here is merely a convenient abbreviation for long-winded descriptions of how the intervals can be decomposed into more and more, smaller and smaller intervals.

(iii) We can now estimate the error in these approximations. A first approach yields the following: we want to evaluate $c : d$, and have shown that $(c + s/7) : d = 22 : 7$; also, since $d = 7r + s$, we know that $s < d/7$. Hence $22 : 7$ is greater than $c : d$ by an amount that can be described by increasing the antecedent $c$ by less than $d/7^2$. Alternatively, we write $c : (d - s/22) = 22 : 7$, so that we can decrease the consequent $d$ by an interval less than $d/7 \times 22$.

At the next step, we have $(c - t/106) : d = 333 : 106$, and $t$ is less than $s$ (since it is a remainder at the third step when $s$ is subtracted), while $s < d/106$ (from the decomposition $d = 106s + 7t$). Hence $333 : 106$ is greater than $c : d$ by an amount that can be described by decreasing the antecedent $c$ by less than $d/106^2$. Alternatively we can increase the consequent; this gives a slightly less simple and convenient result.

This result also appears to be perfectly general: at the $k$th stage, if we estimate the remainder $x$ in this way we will get

$$(c \pm x/q) : d = c : (d \mp x/p) = p : q,$$

where $x$ is less than $d/q$; so the error corresponds to decreasing or increasing the consequent $c$ by an amount less than $d/q^2$.

(iv) We get a more refined error analysis by exploiting the following heuristic principle: If, at any stage of the evaluation of an anthyphairetic ratio, the lesser term goes a large number of times into the greater, then it must be very much smaller than the greater and the remainder it will leave will be even smaller. So if we neglect this remainder by modifying the antecedent or consequent as in (i) or (ii), and so truncate the anthyphairesis before this large term, we should expect to get a particularly good approximation. Consider, for example, the third step of this ratio $c : d$,

$$r - 15s = t.$$

This shows that $s < r/15$, which combines with the description

$$d = 7r + s$$

to yield that $s < d/15 \times 7$. Hence the first estimate of (iii) can be improved to the conclusion that $22 : 7 = [3, 7]$ is greater than $c : d$ by an amount that can be described by increasing the antecedent $c$ by less than $d/15 \times 7^2$.

Again, this result is perfectly general: if the $(k + 1)$st stage of the process gives

$$w - nx = y$$

(where the case of large $n$ is of particular interest), then

$$(c \pm x/q) : d = p : q \quad \text{where} \quad x < d/nq,$$

so the error corresponds to decreasing or increasing the consequent $c$ by an amount less than $d/nq^2$. For example, the ratio $355 : 113 = [3, 7, 15, 1]$ is in excess by an amount that corresponds to increasing the circumference less than $1/292 \times 113^2$ of the diameter.

It is possible to extend this method further to give a *lower* bound for the error, since if

$$w - nx = y \quad \text{with} \quad y < x$$

we know that $x$ will not go $(n + 1)$ times into $w$. Hence

$$w/(n + 1) < x < w/n,$$

and this leads to the estimate

$$d/(n + 2)q < x < d/nq.$$

(v) Two historical remarks are in order concerning the calculations I have just evaluated. First, the only fractional quantities involved are the unit fractions or 'parts', $2', 3', 4', \ldots$, described briefly in Section 1.2(c) and to be analysed in Chapter 7. Second, we today would tend to describe the difference between a smaller and a larger ratio, $p : q$ and $r : s$, by giving some arithmetical expression for their difference $r : s - p : q$, either exactly or by specifying some bound such as 'less than one part in a thousand' or the like. Greek practice is different. To my knowledge, ratios, even ratios of numbers, are never

subtracted like this; instead a typical manipulation would be to use a fourth proportional, write $r:s = x:q$, and then compare $x$ and $\overset{\circ}{p}$, just as we have been doing.

(vi) The argument of (iv) developed the heuristic principle that a large term in the anthyphairetic expansion indicates the existence of an unusually accurate approximation to the given ratio by the ratio of the relatively small numbers got by truncating the anthyphairesis just before the appearance of this large term. Conversely, if we are calculating, somehow, term by term, the anthyphairesis of a given ratio, then the amount of work needed to calculate a given term cannot be predicted in advance, since it will depend on the size of that term. Suppose, then, that someone, possibly Archimedes, set out to calculate the anthyphairesis of $c:d$; he would already find difficulty in calculating the third term, 15, of the expansion, and would have enormous problems in evaluating the fifth term, 292.

(vii) Consider, further, what will happen if we try to calculate the anthyphairesis of some commensurable ratio by finding closer and closer approximations to it; an illustration, to fix the idea, would be to calculate the ratio of the surface of a sphere to an equatorial section by approximating these by figures that are more amenable to manipulation. As the approximations get better and better, so the remainder corresponding to the last step of the exact calculation will get smaller and smaller, and so the corresponding term of the anthyphairesis will grow bigger and bigger. (In fact, a terminating anthyphairesis should often be considered as finishing with an additional infinite term; see Section 2.3(b) below.) So, for example, let us try to imagine the reaction of anyone (again possibly Archimedes) who set out to calculate the anthyphairesis of circumference to diameter of a circle with a view to formulating a hypothesis, for example on whether it is commensurable or not: as the fifth term appears to grow in size, they might be tempted to conclude that the ratio was indeed commensurable.

If we, today, are tempted to feel patronising about such ignorance, then we should reflect that $\pi$ was only proved to be irrational (that being our arithmetised formulation of this problem) by Lambert, in 1768, and to be transcendental, by Lindemann, in 1882; that numbers of the form $a^{\beta}$, such as $2^{\sqrt{3}}$, which look as if they ought to be transcendental, were only proved to be so by Gelfond in 1934, and this important result was generalised to $a_1^{\beta_1} + a_2^{\beta_2} + \cdots + a_k^{\beta_k}$ by Baker only in 1966; that, while Euler investigated the zeta function

$$\zeta(s) = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \cdots$$

and found that $\zeta(2) = \pi^2/6$, $\zeta(4) = \pi^4/90$ and, generally, that $\zeta(2n) = \pi^{2n}/k$ (a result described in Weil, *NT*, 184 and 256–76, as "one of Euler's most sensational early discoveries, perhaps the one which established his growing reputation most firmly . . . . Characteristically, before solving it, Euler had

engaged in extensive numerical calculations"), it was not until 1978 that it was proved, by Apéry, that $\zeta(3)$ is irrational, and it is still not known if it is transcendental; and that it is not yet known whether Euler's constant

$$\gamma = \lim\left[1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n\right]$$

is rational or irrational. Some questions like this are extraordinarily difficult to answer.

### 2.2(c)  *The surface and section*

The obvious two-dimensional generalisation of the problem of finding the ratio of circumference to diameter is to find the ratio of the surface of a sphere to its equatorial section. At first sight, this ratio of two curved regions would seem to be of even greater difficulty (see Section 2.2(b)(vii), above); in fact the answer is as simple as it could be. Archimedes proves, at *Sphere and Cylinder* I, Proposition 33:

The surface of any sphere is four-times the greatest of the circles in it.

Many other results by Archimedes yield similar examples of ratios which are commensurable, and have anthyphairetic expansions containing only one or two terms.

## 2.3  ANTHYPHAIRETIC ALGORITHMS

### 2.3(a)  *The Parmenides proposition*

Plato gives, at *Parmenides* 154b–d, a special case of a proposition that we shall use:

If one thing actually is older than another, it cannot be becoming older still, nor the younger younger still, by any more than their original difference in age, for if equals be added to unequals, the difference that results, in time or any other magnitude, will always be the same as the original difference. . . . [But] if an equal time is added to a greater time and to a less, the greater will exceed the less by a smaller part (*morion*).

To paraphrase and simplify slightly Plato's homely paradox: as two people grow older together, the difference between their ages will remain constant, but the ratio of their ages will get closer to $1:1$. Let $r$ and $s$ denote the different ages, and $p$ the equal time added to both, then

$$\text{if } p:p < r:s \text{ (i.e. } r > s) \quad \text{then} \quad p:p < (p+r):(p+s) < r:s.$$

A slightly more general form of this result will be basic to my reconstruction:

$$\text{if } p:q \leqslant r:s \quad \text{then} \quad p:q < (p+r):(q+s) < r:s.$$

I shall call this 'the *Parmenides* proposition'.

An illustration of this *Parmenides* proposition, in modern usage, would be that the average speed for a day's journey must lie between the average speeds for the morning and afternoon. Note, however, that this illustration does not apply to early Greek mathematics, and indeed such an inhomogeneous ratio, of distance to time, is prohibited by *Elements* V Definition 3. We can see a possible context for this prohibition by observing how this ratio, as now conceived, depends on arithmetised foundation of mathematics, in which it is interpreted as something like (distance in miles) $\div$ (time in hours). Also observe that we cannot evaluate the anthyphairesis of such a ratio of incomparable magnitudes without first converting it into some homogeneous ratio, so that we can apply the operation of $B_2$; typically we now arithmetise, though we could equally well geometrise and represent both distance and time by line segments. Similar comments will apply to other definitions of ratio that I shall give in Chapter 4 (cf. $S_{35}$).

We shall use the proposition in the case where $p$, $q$, $r$, and $s$ are numbers, *arithmoi*. Euclid's theory of proposition for numbers in *Elements* VII contains no definitions or propositions about disproportion, '$p$ to $q$ is greater or less than $r$ to $s$', but to explore the *Parmenides* proposition in this context, we can appeal to the device I described in Section 1.2(d), and extend a proposition into a new definition. We take *Elements* VII 19, which asserts that

$$\text{if } p:q::r:s \quad \text{then} \quad ps = qr, \text{ and conversely,}$$

and make some extension of the theory to enable us to assert, either as a definition or a proposition, that

$$\text{if } p:q < r:s \quad \text{then} \quad ps < qr, \text{ and conversely.}$$

The *Parmenides* proposition now follows easily: we wish to prove that

$$p:q < (p+r):(q+s),$$

which is equivalent to

$$p(q+s) < q(p+r), \quad \text{i.e.} \quad ps < qr,$$

and this now is equivalent to our assumption that $p:q < r:s$. The second inequality follows by a similar argument. The result can be conceived as belonging either within a theory of ratio, or within a theory of proportion that has been extended to include disproportion.

It is worth looking at this result in the case when $p$, $q$, $r$, and $s$ are magnitudes. The proof I have just given cannot then apply, since the product of two magnitudes cannot be defined in many concrete cases, such as for two plane regions, two solid regions, two intervals of time, or even two non-rectilineal lines. But all the ingredients for a different proof can be found in Book V, and while Euclid does not actually formulate the

proposition for magnitudes, he does prove an important related result on equality, at V 12:

$$\text{if } p:q::r:s, \quad \text{and} \quad p:q::t:u, \text{etc.,}$$

$$\text{then } p:q::(p+r+t+\text{etc.}):(q+s+u+\text{etc.}).$$

We shall use this in our proof. We proceed as follows: if $p:q < r:s$ then, by a step that Euclid assumes in V 18, the so-called existence of a fourth proportional, there is a magnitude $x$ such that

$$p:q::r:x.$$

(Note that this step is not true for numbers; for example we cannot find a number $x$ such that $3:4::5:x$.) By V 8, we know that $x > s$. An application of V 12 gives

$$p:q::(p+r):(q+x),$$

and another application of V 8 yields the result that

$$p:q < (p+r):(q+s).$$

The other inequality is proved similarly.

The *Parmenides* proposition, together with the proof I have just given, is found in Pappus, *Collection* VII 8. It was rediscovered by Chuquet, in his *Triparty en la Sciences des Nombres* of 1484, where it is called "la règle des nombres moyens". It was used by Wallis for his approximations to $\pi$ in his *Treatise of Algebra* of 1685. It is a basic property of Farey series. It is the first theorem enunciated by Cauchy in his *Cours d'Analyse* of 1821. And—the end to which I shall put it here—it can be used to explore many of the properties of anthyphairesis.

### 2.3(b)  *An algorithm for calculating anthyphaireses*

Two basic problems within anthyphairetic ratio theory are as follows. First, given some ratio $\theta$, find a procedure for calculating the anthyphairesis $[n_0, n_1, n_2, \ldots]$ of $\theta$; and, conversely, given such an anthyphairetic ratio, find some convenient procedure for calculating the sequence of approximations $[n_0, n_1, \ldots, n_k] = p_k : q_k$, called the *convergents* of the ratio ($B_{26}$ to $S_{29}$). I shall now describe an algorithm to solve the first of these problems, and illustrate it with the example that arose as the ratio of the shorter diagonal of the hexagon to the side, a typical example of the following general class of the ratio of sides of squares: given a line $\rho$ and numbers $n$ and $m$, construct squares equal to $n$ and $m$ times the square on $\rho$ and call their sides $\sqrt{n}$ and $\sqrt{m}$ respectively (so this notation stands for the lines, the geometrical objects, and not their arithmetical interpretation as lengths, real numbers); now evaluate $\sqrt{n}:\sqrt{m}$. The example considered here will be the ratio $\theta = \sqrt{3}:1$.

The procedure must be formulated in terms of, and only depend on, the kind of interpretations and manipulations of the *arithmoi* that were described in

1.2(c). For example, it is not appropriate to arithmetise in terms of decimal numbers and proceed as we usually do today:

$$\sqrt{3} - 1 \times 1 = 1 \cdot 732\ldots - 1 \times 1 = 0 \cdot 732\ldots$$

$$1 - 1 \times 0 \cdot 732\ldots = 0 \cdot 268\ldots$$

$$0 \cdot 732\ldots - 2 \times 0 \cdot 268\ldots = 0 \cdot 196\ldots$$

$$0 \cdot 268\ldots - 1 \times 0 \cdot 196\ldots = 0 \cdot 072\ldots$$

etc.

so that $\sqrt{3} : 1 = [1, 1, 2, 1, \ldots]$; or to use common fractions and write $[1, 1, 2, 1]$ as

$$1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1}}} = 1 + \cfrac{1}{1 + \cfrac{1}{3}} = 1 + \frac{3}{4} = \frac{7}{4}.$$

These kinds of manipulations occur in the modern theory of continued fractions, and they can be adapted into convenient ways of proceeding with numerical calculations on a pocket calculating machine, but they do not correspond to anything that we know of the style and techniques of early Greek mathematics.

Instead we use the *Parmenides* proposition. To operate it, we shall need two ingredients: a test for whether $p : q$ is less than, equal to, or greater than the given ratio $\sqrt{3} : 1$, and the initial under- and overestimates. The test will be

$$p : q \text{ is } <, =, \text{ or } > \sqrt{3} : 1 \quad \text{according as} \quad p^2 \text{ is } <, =, \text{ or } > 3q^2.$$

I shall not stop here to consider the formal status of this proposition within ratio theory, but will rather adopt the attitude of my Socrates, at $S_{19}$. The initial estimates will be the most natural ones, like those that my slaveboy produced at $B_4$:

$$1 : 1 < \sqrt{3} : 1 < 2 : 1$$

or, in general,

$$n : 1 < \theta < (n + 1) : 1,$$

which I shall refer to as the 'standard' under- and overestimates. Also I have taken the given ratio $\theta$ as 'the greater to the less'. There is never any difficulty in either setting up any problem so as to consider this case, or handling the reciprocal ratio, 'the less to the greater'; I shall give an illustration of this in Section 2.4(a), below.

Now proceed as follows. By the *Parmenides* proposition, $1 : 1 < 3 : 2 < 2 : 1$ and, since $3^2 < 3 \times 2^2$, $3 : 2$ is an underestimate. Hence $3 : 2 < \sqrt{3} : 1 < 2 : 1$. A second application of the proposition leads to the underestimate $5 : 3$, hence

TABLE 2.1. Approximations to $\sqrt{3}:1$

| Under-estimate | Over-estimate | New estimate $p:q$ | $p^2$ | $3q^2$ | Under/equal/over | Run length |
|---|---|---|---|---|---|---|
| 1:1 | 2:1 | 3:2 | 9 | 12 | 3 under | |
| 3:2 | 2:1 | 5:3 | 25 | 27 | 2 under | 2 |
| 5:3 | 2:1 | 7:4 | 49 | 48 | 1 over | 1 |
| 5:3 | 7:4 | 12:7 | 144 | 147 | 3 under | |
| 12:7 | 7:4 | 19:11 | 361 | 363 | 2 under | 2 |
| 19:11 | 7:4 | 26:15 | 676 | 675 | 1 over | 1 |
| 19:11 | 26:15 | 45:26 | 2025 | 2028 | 3 under | |
| etc. | | | | | | |

$5:3 < \sqrt{3}:1 < 2:1$. This then generates an overestimate of $7:4$; and so on. The operation is conveniently set out as in Table 2.1.

The process will generate runs of improving under- and overestimates. If, for convenience and simplicity, we retain only the estimate at the end of each run (they are underlined in the table), we get

$$1:1 < 5:3 < 19:11 < \sqrt{3}:1 < 26:15 < 7:4 < 2:1$$

which we can now convert into anthyphairetic ratios,

$$[1] < [1,1,2] < [1,1,2,1,2] < \sqrt{3}:1 < [1,1,2,1,2,1] < [1,1,2,1] < [2].$$

Anybody who performs these calculations—and everyone who would understand should—will notice an ambiguity of a terminating expansion that arises by holding over the very last subtraction to a new step:

$$[n_0, n_1, \ldots, n_k + 1] = [n_0, n_1, \ldots, n_k, 1].$$

This is the only possible ambiguity of anthyphairetic expansions. Also there is a subtlety of inequalities between terminating ratios, beyond that described in $S_{21}$. Consider the two relationships

$$[2,1] = [3] < [3,2],$$

neither of which seems to conform to $S_{21}$. The equality has just been explained; the inequality arises because any terminating expansion should be conceived as finishing with a large term, bigger than any number that will occur elsewhere; let us write this as '$m$' for 'many' or 'millions' or 'myriads'. Then

$$[2,m] < [2,1,m]$$

now does satisfy Socrates' description. So $S_{21}$ should be modified to:

$S'_{21}$: To see which of the two ratios is the greater, proceed as follows. Express any terminating ratio in the standard form, in which the last term is not 'once', and add an

extra notional very large step of 'many-times'. Now note whether their terms differ first at the first, third, fifth, ... step, or the second, fourth, sixth, ... step, since, in the latter case, the reverse of the relationship of less and greater between the *terms* holds between the *ratios*. ...

Now consider the string of inequalities that we have just calculated. It is reminiscent of those that appeared in my dialogue, at $S_{21}$ and $B_{28}$, and this suggests that it may be related to the anthyphairesis. In fact, the connection is very close indeed, since we can deduce immediately from our innermost inequalities,

$$[1, 1, 2, 1, 2] < \sqrt{3} : 1 < [1, 1, 2, 1, 2, 1],$$

that

$$\sqrt{3} : 1 = [1, 1, 2, 1, 2, 1 \text{ or more}, \ldots].$$

Alternatively, we can extract this information directly from Table 2.1, by reading off the lengths of the runs of under- and overestimates, except for some slight difficulty over the first two terms of the expansion. This difficulty, and its resolution, can best be explained to a modern audience by starting off the algorithm with the universal, but un-Greek, approximations

$$0 : 1 < \theta < 1 : 0,$$

as in Table 2.2 where the expansion of a ratio $\theta = [n_0, n_1, \ldots]$ is illustrated. The first term $n_0$ accounts for the first $n_0$ lines of Table 2.2; the next line generates the overestimate $(n_0 + 1) : 1$; and the $(n_0 + 2)$nd line, marked by an asterisk, corresponds to the first line of Table 2.1. Hence $n_1$ is one plus the number of overestimates at the beginning of Table 2.2. In the case of $\sqrt{3} : 1$, there are no initial *over*estimates, so this yields $\sqrt{3} : 1 = [1, 1, \ldots]$. Thereafter, the lengths of

TABLE 2.2. Expansion of $\theta = [n_0, n_1, \ldots]$ starting from $0 : 1 < \theta < 1 : 0$. The schema of Table 2.1 starts at the line marked *

| Under-estimate | Over-estimate | New estimate | Type | Run length |
|---|---|---|---|---|
| $0 : 1$ | $1 : 0$ | $1 : 1$ | under | |
| $1 : 1$ | $1 : 0$ | $2 : 1$ | under | |
| | | $\ldots$ | | |
| $(n_0 - 1) : 1$ | $1 : 0$ | $n_0 : 1$ | under | $n_0$ |
| $n_0 : 1$ | $1 : 0$ | $(n_0 + 1) : 1$ | over | |
| $n_0 : 1$ | $(n_0 + 1) : 1$ | $(2n_0 + 1) : 2$ | over | |
| | | $\ldots$ | | |
| $n_0 : 1$ | $((n_1 - 1)n_0 + 1) : (n_0 - 1)$ | $n_1 n_0 + 1 : n_1$ | over | $n_1$ |
| $n_0 : 1$ | $(n_1 n_0 + 1) : n_0$ | $((n_1 + 1)n_0 + 1) : (n_1 + 1)$ | under | |

etc.

* (left margin asterisk)    * (right margin asterisk)

the runs of under- and overestimates of Table 2.1 give the successive terms of the anthyphairesis.

I recommend anybody who is unfamiliar with any of these manipulations to calculate a few terms of the anthyphairesis of $\sqrt{n}:1$ for $2 \leqslant n \leqslant 10$.

### 2.3(c) *An algorithm for calculating convergents*

As my Socrates explains at S27, there is no problem in evaluating individual convergents: $[n_0] = n_0:1$, $[n_0, n_1] = (n_0 n_1 + 1):n_1$, $[n_0, n_1, n_2] = (n_0 n_1 n_2 + n_0 + n_2):(n_1 n_2 + 1)$, and so on. What we now develop is a much more convenient forward method. I shall describe and illustrate it on the examples of convergents of the two ratios $[2, 3, 4, 5, \ldots]$ and $[5, 5, 10, 5, 10, \ldots]$.

One problem with $[2, 3, 4, 5, \ldots]$ is that we do not know if any ratio $\theta$ exists that has this as anthyphairesis. This, I feel, is not a problem that would have had any meaning for the Greeks, even within the most extreme developments of my reconstruction, so I shall ignore it here; its solution depends either on the sort of eighteenth-century manipulations at which Euler excelled, or on nineteenth-century ideas of the foundations of mathematics. I choose this example since the numbers involved illustrate clearly and unambiguously the manipulations to be performed. This problem does not arise with $[5, 5, 10, 5, 10, \ldots]$, since this is the anthyphairesis of some explicit ratio that I shall identify later in this section, and establish by rigorous geometrical proof in the next chapter.

So suppose that $\theta = [2, 3, 4, 5, \ldots]$. This means that, if we expand $\theta$ using the algorithm of the previous section, then the pattern of runs of under- and overestimates will be determined by this expansion. For convenience, I shall adopt, in Table 2.3, the format of Table 2.2, and start with the initial estimates

$$0:1 < \theta < 1:0;$$

the more strictly historical algorithm will start instead (at the line in Table 2.3 marked with an asterisk) with the initial estimates

$$2:1 < \theta < 3:1.$$

With this information, we can set out the skeleton format of Table 2.3, which contains enough information for us now to insert the remaining information by an automatic procedure. In particular we are most interested in the derivation of the run-end estimates, the ratios that go into the boxes in column 3; these, it can be seen, are the successive convergents. Clearly the first box contains $(2 \times 1 + 0):(2 \times 0 + 1) = 2:1$. Then the next line starts with the estimates $2:1 < \theta < 1:0$ and so, after three steps of overestimates, the next box will contain $(3 \times 2 + 1):(3 \times 1 + 0) = 7:3$. Then the next line starts with the estimates $2:1 < \theta < 7:3$ and so, after four steps of underestimates, the next box will contain $(4 \times 7 + 2):(4 \times 3 + 1) = 30:13$; and so on. We can abstract this information in Table 2.4 where, for convenience, it is written out in a compact horizontal format.

TABLE 2.3. The approximation procedure for the ratio $\theta = [2, 3, 4, \ldots]$

| Under-estimate | Over-estimate | New estimate $p : q$ | Type | Run length |
|---|---|---|---|---|
| 0 : 1 | 1 : 0 | | under | |
| | | ▭ | under | 2 |
| | | | over | |
| *  2 : 1 | 3 : 1 | | over | * |
| | | ▭ | over | 3 |
| | | | under | |
| | | | under | |
| | | | under | |
| | | ▭ | under | 4 |
| etc. | | | | |

Once two columns of Table 2.4 are known, the rest of the table can be calculated easily and quickly. The table can be started from the first two columns, which correspond to the estimates

$$0 : 1 < \theta < 1 : 0,$$

or the next two columns can be written down, since $[2] = 2 : 1$ and $[2, 3] = 7 : 3$, and the algorithm applied from this point.

As a second example, consider $\phi = [5, 5, 10, 5, 10, \ldots]$; the reader is encouraged to compute the first five convergents of this ratio and confirm immediately that

$$5 : 1 < 265 : 51 < 13\,775 : 2651 < \phi < 1351 : 260 < 26 : 5.$$

The point of this example is that early results in what I have called the problem of dimensions of squares (see $S_{41}$, and Sections 2.2(a) and 2.3(b) above) will be that

$$\sqrt{(n^2 + 1)} : 1 = [n, 2n, 2n, 2n, \ldots]$$

$$\text{and } \sqrt{(n^2 + 2)} : 1 = [n, n, 2n, n, 2n, n, 2n, \ldots],$$

where the expansions are periodic, repeating indefinitely—this will be proved in the next chapter. Hence $\sqrt{27} : 1 = [5, 5, 10, 5, 10, \ldots]$, and the calculation of the

TABLE 2.4. The convergents of $[2, 3, 4, 5, 6, \ldots]$

| $n_k$ | | | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $p_k$ | 0 | 1 | 2 | 7 | 30 | 157 | $(6 \times 157 + 30) = 972$ |
| $q_k$ | 1 | 0 | 1 | 3 | 13 | 68 | $(6 \times 68 + 13) = 421$ |
| Type | under | over | under | over | under | over | under |

convergents then tells us that

$$265:51 < \sqrt{27}:1 < 1351:260,$$

in which the error estimates of Section 2.2(b) show that

if $\sqrt{(27+x)}:1 = 265:51$,    then    $x < 1/(5 \times 51^2)$,    and
if $\sqrt{(27-y)}:1 = 1351:260$,   then   $y < 1/(10 \times 260^2)$.

But since $\sqrt{27} = 3\sqrt{3}$ (this is a geometrical statement, about fitting together nine squares each equal to three times a given square, something in the style of Plato's slaveboy's first two attempts at doubling a square), this inequality implies that

$$265:153 < \sqrt{3}:1 < 1351:780,$$

with the corresponding error estimates also being divided by 3. I shall return to this statement in Section 2.4(d), below.

There is one place in the *Elements* where one looks, in vain, for details of this kind of calculation of convergents. At the beginning of Book X we find anthyphairesis introduced as a criterion of incommensurability:

If, when the less of two unequal magnitudes is continually subtracted in turn (*anthuphairein aei*) from the greater, that which is left never measures the one before it, the magnitudes will be incommensurable [X 2],

and as a construction to solve the problems

Given two (or three) commensurable magnitudes, to find their greatest common measure [X 3 (or 4)].

These are then followed by a block of four propositions that are variations on X 5:

Commensurable magnitudes have to one another the ratio which a number has to a number.

At this point, it would surely be an interesting and obvious step to relate the pattern of the anthyphairesis to the numbers involved in this ratio, but on this the text is silent. There would be no difficulty in describing the backwards procedure (Socrates' explanation at the beginning of $S_{27}$ can easily be translated into Euclidean idiom), but it is not easy to see how to express the much more convenient forward procedure within the stylistic restrictions of Euclid's *Elements*, where the enunciations of propositions cannot contain any reference to particular figures or numbers, and where no symbolic manipulation is permitted (see the end of $S_{27}$). Nevertheless, at no point does either algorithm transcend any of the restrictions on *arithmētikē*. The absence of any such exploration from the *Elements* has meant that very few of those who have learned their mathematics from Euclid, either directly or indirectly, have any intimate experience of the operation of anthyphairesis, in any guise.

## 2.4 FURTHER ANTHYPHAIRETIC CALCULATIONS

### 2.4(a) *Eratosthenes' ratio for the obliquity of the ecliptic*

Ptolemy explains, at *Syntaxis* I 12, how to measure the obliquity of the ecliptic, and then concludes:

We found that the arc between the northernmost and southernmost points, which is the arc between the solsticial points, is always greater than $47\frac{2}{3}°$ and less than $47\frac{3}{4}°$. From this we derive very much the same ratio as Eratosthenes, which Hipparchus also used. For [according to this] the arc between the solstices is approximately 11 parts where the meridian is 83 [Ptolemy–Toomer, *PA*, 63].

This asserts that $11 : 83$ lies between $47\frac{2}{3} : 360$ and $47\frac{3}{4} : 360$; and this curious ratio $11 : 83$ has provoked occasional perplexed comment. Let us apply the approximating algorithm. We start with the standard under- and overestimates to the interval,

$$1 : 8 < 47\tfrac{2}{3} : 360 < 47\tfrac{3}{4} : 360 < 1 : 7,$$

and apply the *Parmenides* proposition until the first intermediate ratio is generated. The calculation is set out in Table 2.5; the result, $11 : 83$, arises after four lines of calculation.

By-products of this calculation are the evaluations:

$$47\tfrac{2}{3} : 360 = [0, 7, 1, 1, 1, 4, \ldots] \text{ and}$$
$$47\tfrac{3}{4} : 360 = [0, 7, 1, 1, 5 \text{ or more}, \ldots].$$

In fact,

$$47\tfrac{2}{3} : 360 = [0, 7, 1, 1, 4, 3, 1, 3],$$
$$11 : 83 = [0, 7, 1, 1, 5], \text{ and}$$
$$47\tfrac{3}{4} : 360 = [0, 7, 1, 1, 5, 1, 6, 2].$$

This example is of a ratio taken 'the less to the greater' which has the effect of introducing an initial step of zero which reverses antecedent and consequent;

TABLE 2.5. Eratosthenes' ratio for the obliquity of the ecliptic

| Under-estimate | Over-estimate | New estimate $p : q$ | $360p$ | $47\frac{2}{3}q$ | $47\frac{3}{4}q$ | Type |
|---|---|---|---|---|---|---|
| $1 : 8$ | $1 : 7$ | $2 : 15$ | 720 | 715 | $716\frac{1}{4}$ | over both |
| $1 : 8$ | $2 : 15$ | $3 : 23$ | 1080 | $1096\frac{1}{3}$ | $1098\frac{1}{4}$ | under both |
| $3 : 23$ | $2 : 15$ | $5 : 38$ | 1800 | $1811\frac{1}{3}$ | $1814\frac{1}{4}$ | under both |
| $5 : 38$ | $2 : 15$ | $7 : 53$ | 2520 | $2526\frac{1}{3}$ | $2530\frac{1}{4}$ | under both |
| $7 : 53$ | $2 : 15$ | $9 : 68$ | 3240 | $3241\frac{1}{3}$ | 3247 | under both |
| $9 : 68$ | $2 : 15$ | $11 : 83$ | 3960 | $3956\frac{1}{3}$ | $3963\frac{1}{4}$ | between |

then the ratio 'the greater to the less' follows. We can handle the situation in words, by saying something like: 'the antecedent of the given ratio is less than the consequent, and the anthyphairesis of the ratio is seven-times, once, . . .'. So the use of a zero here carries no implications about the introduction of a zero into the Greek number system.

Reciprocation seems to be a process that is handled fluently and with confidence in Greek mathematics.

### 2.4(b) *The Metonic cycle*

Astronomers in Babylonia from the time of Darius (*c.*490 BC), and in Greece after the mid-fifth century BC, used a cyclical calendar in which 235 months, some (called 'hollow') of 29 days, the rest of 30 days ('full'), are made equal to 19 years. In Greece, this astronomical (not civil) lunar–solar calendar is attributed to Euctemon and Meton, and is said to have started on the summer solstice, 432 BC. It was then modified by Callippus, in the fourth century, into a cycle of $76 = 4 \times 19$ years.

Although this calendar may seem to embody some very precise astronomical parameters, it is shown in Goldstein, NMC, how it can in fact be deduced from the following relatively crude data:

(i) The year is a little more than 365 days long,

(ii) The length of 12 synodic months (i.e. new moon to new moon) is a little more than 354 days (note that $\frac{354}{12} = 29\frac{1}{2}$), and

(iii) The average length of a synodic month (the 'mean' synodic month) is a little more than $29\frac{1}{2}$ days.

We now look for a cycle of $p$ years in which the $365 - 354 = 11$ days' discrepancy, called the 'epact', approximates an integral number $q$ of months; so we have to find $p$ and $q$ such that

$$29\tfrac{1}{2}q < 11p < 30q,$$

i.e.

$$29\tfrac{1}{2} : 11 < p : q < 30 : 11.$$

This can easily be done by trial and error; or it can be used as another illustration of the algorithm. We take the standard under- and overestimates to the interval:

$$2 : 1 < 29\tfrac{1}{2} : 11 < 30 : 11 < 3 : 1$$

and apply the algorithm to generate a ratio which lies within the interval, exactly as in the previous example. This gives

$$29\tfrac{1}{2} : 11 < 19 : 7 < 30 : 11,$$

the cycle in which 19 years are approximately equal to $19 \times 12 + 7 = 235$ months and 6940 days.

The Metonic year is $365\frac{1}{4} + \frac{1}{19}$ days long. Callippus' calendar incorporates the more accurate parameter that a year is approximately $365\frac{1}{4}$ days long, so that 76 years are approximately 27 759 days; and so it relates the three principal astronomical periodic rotations: the earth around its axis, the moon around the earth, and the earth and moon around the sun. Another similar, slightly less accurate but more convenient, cyclical calendar in which 25 years, 9125 days, are made equal to 309 lunar months, will be described in Section 4.4.

### 2.4(c) *Aristarchus' reduction of ratios*

Aristarchus observes, within the proof of Proposition 13 of his *On the Sizes and Distances of the Sun and Moon*,

but 7921 has to 4050 a ratio greater than that which 88 has to 45,

and in Proposition 15,

but 71 755 875 has to 61 735 500 a ratio greater than that which 43 has to 37.

We could generate these results by applying the *Parmenides* proposition directly, as we did with Eratosthenes' ratio. But since we do not know, in advance, how long this will take, apart from the very rough bound that, starting with $1:1 < 7921:4050 < 2:1$, we must arrive at a ratio of $88:45$ at the latest within about forty steps (also see my slaveboy's unanswered reflection in B$_{24}$), it will generally be quicker, whenever possible, to calculate the anthyphairesis and use this to compute the convergents. An easy evaluation gives

$$7921:4050 = [1, 1, 21, 1, 1, 1, 2, 22],$$

from which we can calculate:

| $n$ | 1 | 1 | 21 | 1 | 1 | 1 | 2 | 22 |
|------|-------|------|-------|------|-------|------|---|----|
| $p$ | 1 | 2 | 43 | 45 | 88 | etc. | | |
| $q$ | 1 | 1 | 22 | 23 | 45 | | | |
| Type | under | over | under | over | under | | | |

and so the first of Aristarchus' results. (At this point we can see that the *Parmenides* proposition would have required 23 lines and 46 multiplications to arrive at this answer.)

The second inequality follows similarly. In this particular case, it might be sensible to remove the obvious factors $3^3 \times 5^3$; but, in general, it is a notorious fact that factorisation can be an enormously long process, while finding the common measure of two numbers using anthyphairesis, as in *Elements* VII 2, is extremely rapid. This means that the very process for removing common factors efficiently is precisely the same process which generates the anthyphairetic ratio and so, with very little extra effort, the approximations we are seeking.

I remark, in passing, that Aristarchus' treatise is an excellent illustration of what I have called 'non-arithmetised' mathematics. Relationships are expressed as ratios of *arithmoi*, and the only fractional quantities are simple parts. I shall

go into this question in more detail in Section 7.3(b), where I shall describe the two mild exceptions to the statements I have just made.

### 2.4(d) *Archimedes' calculation of circumference to diameter*

Proposition 3 of Archimedes' *Measurement of a Circle* is a mine of calculations, many of which can be interpreted anthyphairetically. I have already discussed its enunciation in Section 2.2(b). The first half of the calculation, which evaluates an upper bound for $c:d$, starts from an underestimate of the ratio of the diameter of the circle to the side of a circumscribing hexagon, $\sqrt{3}:1 > 265:153$. I showed, in Section 2.3(c), how this can be easily and quickly evaluated from the result that

$$\sqrt{27}:1 = [5, 5, 10\ldots].$$

The second half, the calculation of a lower bound, starts from an overestimate of the ratio of the shorter diagonal to side of an inscribed hexagon; this (see Section 2.2(a)) is again $\sqrt{3}:1$, and Archimedes' overestimate of $1351:780$ is again a third of the next convergent of $\sqrt{27}:1$, also calculated in Section 2.3(c). Both of these approximations are also convergents of $\sqrt{3}:1$, the ninth and twelfth respectively, but the fact that Archimedes did not take the eighth and ninth or the ninth and tenth convergents to $\sqrt{3}:1$, viz.:

$$265:153 < \sqrt{3}:1 < 97:56 < 362:209 < 1351:780,$$
$$\quad\text{9th} \qquad\qquad\qquad \text{8th} \qquad \text{10th} \qquad\quad \text{12th}$$

any of which would provide sufficient accuracy for his subsequent calculation, has caused some perplexity. A derivation that starts from the expansion of $\sqrt{27}:1$ accounts for this anomaly. Also, the terms of $\sqrt{27}:1$ are relatively large, and this will guarantee the very rapidly improving accuracy of the convergents; see Section 2.2(b)(iv). (For an endorsement of this procedure, see Knorr, AD, p. 435: "Thus, even accepting that the extant evidence does not permit certainty, one may well wonder what a more satisfactory solution of the quotation could possibly look like.")

Archimedes performs each half of the calculation by successively doubling the number of sides of each hexagon four times, and each time estimating the ratio of the side of the new polygon to the diameter of the circle, always rounding any subsequent approximation upwards for the circumscribed polygon and downwards for the inscribed polygon. Each doubling of the polygons (except the final step of the circumscribing polygon) involves estimating a square root. For example, the first step which replaces the circumscribing hexagon by a dodecagon, requires an underestimate for $\sqrt{349\,450}:153$. Since

$$\sqrt{349\,450}:1 = [591, 6, 1, 196, \ldots],$$

this has the extremel̃y accurate underestimate $591\frac{1}{7}:153$. But Archimedes uses instead the approximation $591\frac{1}{8}:153$, either for reasons of computational

convenience, or as a result of his computational techniques, or because he can estimate that his initial approximations to $\sqrt{3}:1$ allow him some latitude, or for other reasons; or, as I shall illustrate later, in Section 7.3(a), it is not impossible that this particular fractional part might be a later editorial addition. We shall see there how the text has been re-worked and altered so much that we cannot, any longer, be sure of some of the details of the original calculations.

The calculation of the overestimate ends with the following passage in which I have adapted the translation so as to imitate, in the English text, the different ways in which the numbers are expressed.

The ratio of $A\Gamma$ [the diameter] to the perimeter of the 96-sided polygon is greater than the ratio of 4673 2′ to 14 688, which [now taken 'the greater to the less'] is greater than threefold and there are remaining 667 2′, which [number] is less than the seventh part of 4673 2′; so that the circumscribed polygon is threefold the diameter and bigger by less than the seventh part; so therefore the perimeter of the circle is less by more than threefold and more by a seventh part.

The description here is reminiscent of the opening lines of my dialogue, but so also would be a conventional arithmetical description of a ratio as an integer part plus a unit-fractional part; any difference would become evident if the ratio needed a more accurate or less simple description. The corresponding passage at the end of the second part of the calculation, the evaluation of the underestimate, reads:

Conversely then the perimeter of the polygon bears to the diameter a ratio greater than 6336 has to 2017 4′, which is bigger than 2017 4′ or threefold and ten 71st [parts]. And so the perimeter of the inscribed 96-sided polygon is threefold the diameter and greater than the 10 71st [parts], so that also the circle is still more than threefold and greater than the 10 71st [parts].

Unfortunately this passage which, if more explicit, might have shed some light on the way Archimedes conceived the ratios involved, is tantalisingly brief; nor is it made any more informative by being repeated three times. What is more, each occurrence of the expression '10 71st [parts]' shows a wide range of variation between apparently incorrect or meaningless alternatives in the surviving manuscripts. Indeed, in many cases throughout the text the 'correct' numbers cannot be traced back any earlier than early sixteenth-century annotations to the manuscript; for example, the best manuscripts do not contain here the ratio $6336:2017\frac{1}{4}$, but $6301\frac{1}{6}:7017\frac{1}{4}$. More details will be given in Section 7.3(a).

Both of these calculations illustrate again the process that Heron called, in the passage from *Metrica* I 25 quoted in Section 2.2(b), "bringing down to small numbers". For example we can evaluate that

$$14\,688:4673\tfrac{1}{2} = [3,7,667,2]$$

and

$$6336:2017\tfrac{1}{4} = [3,7,10,2,1,36],$$

and hence the overestimate $[3, 7] = 22 : 7$ and underestimate $[3, 7, 10] = 223 : 71$. However, it is here more illuminating to describe the same calculation in terms of the *Parmenides* proposition. The starting estimates

$$3 : 1 < c : d < 4 : 1$$

follow from the figure of a hexagon inscribed in, and a square circumscribed about a circle. The initial run of overestimates is (read from the right):

$$3 : 1 < c : d < 22 : 7 < 19 : 6 < 16 : 5 < 13 : 4 < 10 : 3 < 7 : 2 < 4 : 1;$$

all can be verified from the inequalities

$$c : d < 14\,688 : 4673\tfrac{1}{2} < 22 : 7$$

since

$$14\,688 \times 7 = 102\,816 < 4673\tfrac{1}{2} \times 22 = 102\,817.$$

The next run of underestimates is:

$$3 : 1 < 25 : 8 < 47 : 15 < 69 : 22 < 91 : 29$$
$$< 113 : 36 < 135 : 43 < 157 : 50 < 179 : 57$$
$$< 201 : 64 < 223 : 71 < 6336 : 2017\tfrac{1}{4} < c : d < 22 : 7.$$

However, since the next estimate satisfies

$$6336 : 2917\tfrac{1}{4} < 245 : 768 < 14\,688 : 4673\tfrac{1}{2},$$

there is not enough information to decide whether $245 : 78$ is less than, equal to, or greater than $c : d$. In this sense, the ratios given in the enunciation extract all the anthyphairetic information from Archimedes' surviving calculation.

### 2.4(e) *Pell's equation*

The problem is best explained in the following two texts. First, a challenge to European mathematicians proposed by Fermat in 1657. The translation is from Diophantus–Heath, *DA*, 285–6:

There is hardly anyone who propounds arithmetical questions, hardly anyone who understands them. Is this due to the fact that up to now arithmetic has been treated geometrically rather than arithmetically? This has indeed generally been the case both in ancient and modern works; even Diophantus is an instance. For, although he has freed himself from geometry a little more than others in that he confines his analysis to the consideration of rational numbers yet even there geometry is not entirely absent, as is sufficiently proved by the *Zetetica* of Viete, where the method of Diophantus is extended to continuous magnitude and therefore to geometry.

Now arithmetic has, so to speak, a special domain of its own, the theory of integral numbers. This was only lightly touched upon by Euclid in his *Elements*, and was not sufficiently studied by those who followed him (unless, perchance, it is contained in those Books of Diophantus of which the ravages of time have robbed us); arithmeticians have therefore now to develop it or restore it.

To arithmeticians therefore, by way of lighting up the road to be followed, I propose the following theorem to be proved or problem to be solved. If they succeed in finding the proof or solution, they will admit that questions of this kind are not inferior to the more celebrated questions in geometry in respect of beauty, difficulty, or method of proof.

Given any number whatever which is not a square, there are also given an infinite number of squares such that, if the square is multiplied into the given number and unity is added to the product, the result is a square.

Example. Let 3, which is not a square, be the given number; when it is multiplied into the square 1, and 1 is added to the product, the result is 4, being a square.

The same 3 multiplied by the square 16 gives a product which, if increased by 1, becomes 49, a square.

And an infinite number of squares besides 1 and 16 can be found which have the same property.

But I ask for a general rule of solution which any number not a square is given.

E.g. let it be required to find a square such that, if the product of the square and the number 149, or 109, or 433 etc. be increased by 1 the result is a square.

Second, a passage from Theon of Smyrna–Hiller, *ERMLPU*, 42–5, adapted from the translation of Thomas, *SIHGM* i, 132–7, to bring out the evocation on the word *logos* at the beginning:

Just as numbers potentially contain triangular, square, and pentagonal ratios (*logoi*), and ones corresponding to the remaining figures, so also we can find side and diagonal ratios (*logoi*) appearing in numbers in accordance with the generative principles (*logoi*); for it is from these that the figures acquire balance. Therefore since the unit, according to the supreme generative principle (*logos*), is the starting-point of all the figures, so also in the unit will be found the ratio (*logos*) of the diagonal to the side. For instance, two units are set out, of which we set one to be a diagonal and the other a side, since the unit, as the beginning of all things, must have it in its capacity to be both side and diagonal. Now there are added to the side a diagonal and to the diagonal two sides, for as great as is the square on the side, taken twice, [so great is] the square on the diagonal taken once. The diagonal therefore became the greater and the side became the less. Now in the case of the first side and diagonal the square on the unit diagonal will be less by a unit than twice the square on the unit side; for units are equal, and 1 is less by a unit than twice 1. Let us add to the side a diagonal, that is, to the unit let us add a unit; therefore the [second] side will be two units. To the diagonal let us now add two sides, that is, to the unit let us add two units; the [second] diagonal will therefore be three units. Now the square on the side of two units will be 4, while the square on the diagonal of three units will be 9; and 9 is greater by a unit than twice the square on the side 2.

Again, let us add to the side 2 the diagonal of three units; the [third] side will be 5. To the diagonal of three units let us add two sides, that is, twice 2; there will be 7. Now the square from the side 5 will be 25, while that from the diagonal 7 will be 49; and 49 is less by a unit than twice 25. Again, if you add to the side 5 the diagonal 7, there will be 12. And if to the diagonal 7 you add twice the side 5, there will be 17. And the square of 17 is greater by a unit than twice the square of 12. When the addition goes on in the same way in sequence, the proportion will alternate; the square on the diagonal will be now greater by a unit, now less by a unit, than twice the square on the side; and such sides and diagonals are both expressible (*rhētos*).

The squares on the diagonals, alternating one by one, are now greater by a unit than double the squares on the sides, now less than double by a unit, and the alternation is

regular. All the squares on the diagonals will therefore become double the squares on the sides, equality being produced by the alternation of excess and deficiency by the same unit, regularly distributed among them; with the result that in their totality they do not fall short of nor exceed the double. For what falls short in the square on the preceding diagonal exceeds in the next one.

There is also a similar passage in Proclus' commentary on Plato's *Republic*, which contains an additional description of 'side and diagonal lines'; this will be quoted in Section 3.6(b), where the passages will be discussed in more detail.

   Both of these texts show an interest in finding, for a given number $n$ (Theon and Proclus deal only with the case $n = 2$), those numbers $q$ for which $nq^2 + 1$ is a square number $p^2$. Theon also refers to the further cases where $nq^2 - 1$ is to be a square and, while Fermat does not explicitly mention this case, the choice of the three examples he proposes shows without doubt that he was also interested in this related problem. Euler named the equation

$$x^2 - ny^2 = 1,$$

to be solved in integers, after John Pell (1611–85), a minor British mathematician who had nothing to do with the problem, and the name has stuck. More details of the recent history of the topic will be given in Chapter 9; my concern here is to point out that our basic approximating algorithm of Section 2.3(b) gives a practical method for solving the general problem that would be compatible with our knowledge of early Greek mathematics, and to introduce some further Greek evidence. Again I shall concentrate mainly on the mathematical issues; textual and historical problems will be considered in Section 3.6(b), and developments since the seventeenth century will be surveyed in Section 9.1(c).

   Suppose we apply the basic algorithm for calculating convergents to $\sqrt{n}:1$. At each step of the process, each application of the *Parmenides* proposition, we need to compare the new approximation $p:q$ with $\sqrt{n}:1$, and this is done in columns 4 and 5 of Table 2.1, by comparing $p^2$ and $nq^2$. In the example of $\sqrt{3}$, calculated in Table 2.1, we see that this gives a pattern of 3 under, 2 under, 1 over, 3 under, 2 under, 1 over, and this pattern will appear to continue indefinitely; it is a further example of the periodic behaviour exhibited by the process which I hope has been obvious to everybody, and which will soon come to dominate our discussion. This same kind of periodic behaviour will seem to appear whatever value of $n$ is chosen, and moreover the value '1 over' will always seem to occur, sometimes alternately with the value '1 under'. In other words, the algorithm set out in Table 2.1 will always seem to solve the problem. What is more, the solutions will always seem to be the penultimate convergents before the end of the period of $\sqrt{n}:1$; if the period contains an even number of terms, they all will be solutions of $x^2 - ny^2 = 1$; if an odd number, the solutions of $x^2 - ny^2 = \mp 1$ will alternate. Hence, if we already know in advance the anthyphairesis of $\sqrt{n}:1$, we can proceed directly to the calculation of the convergents and hence the solution. For example, the convergents of $\sqrt{27}:1$, calculated in Section 2.3(c), suggest that the solutions of $x^2 - 27y^2 = 1$ may be

$x = 26$; $y = 5$; $x = 1351$, $y = 260$; and so on. We can then check that this is so.

The heuristic arguments of Section 2.2(b)(iv) again hint at what may be happening: a solution of $x^2 - ny^2 = \pm 1$ will correspond to a very good approximation, $x : y$, to $\sqrt{n} : 1$. Now the run-end approximations that occur just before the appearance of a long run of under- or overestimates, being the convergents that precede a large term in the anthyphairesis, will be particularly good approximations. Moreover the largest term in the expansion of $\sqrt{n} : 1$ (but not of $\sqrt{n} : \sqrt{m}$) always seems to be the term '$2n_0$' that marks the end of the period. Hence we might expect that the penultimate convergent of each period would give rise to a small value of $x^2 - ny^2$; in fact it turns out to generate the smallest possible value—it cannot be zero, but it is $\pm 1$. If the period contains an even number of terms, the first penultimate convergent and all subsequent penultimate convergents will be overestimates, and hence the positive signs; if odd, they will alternate between under- and overestimates, and so the signs will also alternate.

The other piece of Greek evidence that gives rise to further speculation concerning Pell's equation in antiquity is Archimedes' remarkable *Cattle Problem*, an epigrammatic puzzle involving the sizes of four herds of cattle. The problem has two possible interpretations, which then give rise to the following two mathematical problems: either solve a collection of linear indeterminate equations

$$nx - my = 1,$$

where $n$ and $m$ satisfy

$$nm = 23\,014\,894 = 2\,.\,7\,.\,353\,.\,4657;$$

or solve Pell's equation

$$x^2 - 410\,286\,423\,278\,424y^2 = 1$$

(in which the coefficient of $y^2$ factorises to $2^3\,.\,3\,.\,7\,.\,11\,.\,29\,.\,353\,.\,4657^2$), an equation whose smallest solution is the stupendous

$$x = 7{\cdot}760\ldots \times 10^{206\,544}$$

$$y = 3{\cdot}831\ldots \times 10^{206\,537}.$$

The existence of this problem may imply that either the interpretation which leads to Pell's equation, or the interpretation as a linear problem, or both interpretations, could be solved, at least 'in principle', by the setter of the problem. So let us therefore look at the mathematics that lies behind the alternative interpretation.

2.4(f) *The alternative interpretation of Archimedes' Cattle Problem*

Consider the linear indeterminate problem expressed by the equation

$$nx - my = 1,$$

where, as always, $n$ and $m$ are integers, and the equation is to be solved in integers. Since $my$ will then be less than but close to $nx$, this means that $y : x$ will have to be less than but close to $n : m$; and this again suggests that we look at approximations to $n : m$. Again we can proceed in two similar ways, either by working with the approximating algorithm of Section 2.3(b), or by proceeding directly to the evaluation of the anthyphairesis and convergents of 2.3(c). In addition, each approach splits into two distinct cases.

First, look again at a typical sequence of approximation $p : q$, generated by the basic algorithm of Section 2.3(b); take, for example, those that arise in column 3 of Table 2.1, and compare successive pairs. We see that

$$3.3 - 2.5 = -1$$
$$5.4 - 3.7 = -1$$
$$7.7 - 4.12 = +1$$
$$12.11 - 7.19 = -1$$

$$\text{etc.}$$

and, in general, if $p : q$ and $p' : q'$ are consecutive ratios, they seem to satisfy

$$pq' - p'q = \pm 1,$$

the negative sign occurring when $p : q$ increases to $p' : q'$, the positive sign when it decreases; column 6 tells us which of these cases applies.

So suppose we wish to solve

$$26x - 15y = +1.$$

We execute the approximation algorithm on $26 : 15$, which will then generate all but the last line of Table 2.1 (at which point it will terminate with an equality). The penultimate estimate $19 : 11$ will be less than $26 : 15$. Hence, we guess (and can verify) that

$$26.11 - 15.19 = 1,$$

and the equation is solved. Other solutions (in fact, every other solution) then arise from the identity

$$26(11 + 15k) - 15(19 + 26k) = 1.$$

Suppose, however, that we wish to solve the dual equation

$$26x - 15y = -1.$$

We now want an approximation that will be *greater* than $26 : 15$; so we look back in Table 2.1 to the previous overestimate $7 : 4$ and check that it is a solution:

$$26.4 - 15.7 = -1.$$

Then, as before, the general solution is given by

$$26(4 + 15k) - 15(7 + 26k) = -1.$$

This second case has led us to the previous convergent, and this suggests that we look at the behaviour of the convergents and ignore the rest of the intermediate approximations. So we can repeat this kind of exploration on Table 2.4, where we see that the cross-products of adjacent convergents again satisfy the same kind of relationship. This leads to the second technique for solving these equations, in which we proceed directly to the anthyphairesis and convergents, without passing via the calculations of Table 2.1. Hence, for example, to solve

$$972x - 421y = -1,$$

we evaluate that $972 : 421 = [2, 3, 4, 5, 6]$; work out the convergents (exactly as in Table 2.4, except that the table will terminate with an equality); and so read off and check that the penultimate convergent, which is an overestimate, satisfies

$$972 . 68 - 421 . 157 = -1,$$

from which we can write down the general solution, as before. If, however, we want to solve the dual equation

$$972x - 421y = +1,$$

we adapt this method by exploiting the ambiguity of the terminating expansion (see Section 2.2(b)): We now write $972 : 421 = [2, 3, 4, 5, 5, 1]$, and take the penultimate convergent of this modified expansion. This gives the solution

$$972 . 353 - 421 . 815 = +1.$$

These results appear to be perfectly general, and appear to solve every example of such problems. They have been expressed, purely for convenience of exposition, in the form of equations, but they are, in fact, straightforward statements about differences of products of integers that can easily be expressed and manipulated in everyday language.

## 2.5 NOTES AND REFERENCES

2.1     Aspects of the modern theory of continued fractions, together with a sketch of the history of its development, will be given in Chapter 9. But for someone who is merely interested in understanding my proposals here, I do not think it is necessary and it may even be distracting to know this modern theory in advance. (In common with perhaps nine-tenths of mathematicians of today, I had never met the topic in my undergraduate or graduate mathematical training, and was only vaguely aware of a revival of interest in it among certain mathematicians when I accidentally stumbled on the path which eventually led to this anthyphairetic interpretation of early Greek mathematics.)

The translation of the passage of Alexander of Aphrodisias, *IT*, is taken from Knorr, *EEE*, 258. This book contains an excellent summary of previous proposals concerning anthyphairesis, in particular, the work of Allman, Becker, von Fritz, Heller, van der Waerden, and Zeuthen, together with Knorr's own interpretations. (I have also taken the brief description of the etymology of *anthuphairesis* and *antanairesis* from it; see p. 290, n. 26.) My reconstruction will be different. In brief, all preceding proposals have reconstructed an anthyphairetic theory of *proportion*, but I shall argue that there may have been a highly developed exploration of some remarkable properties of anthyphairetic *ratios*; see Section 1.2(d), above. Also, previous anthyphairetic interpretations have invoked the modern arithmetised theory of continued fractions, while the version developed here is based on the non-arithmetised *Parmenides* algorithm; see Section 9.1(b) and note 13.

Knorr, *EEE*, Chapter 8 and Appendix B, gives a very useful discussion of the formal development of the proportion theories of the *Elements*; but his interpretation makes crucial use of the surprising difficulties in an anthyphairetic proof of *Elements* V 9: if $A:C::B:C$ then $A = B$. Surely a practising mathematician, faced with the difficulties that Knorr has uncovered, would proceed indirectly, via the *alternando* property that he has just proved. For then $A:C::B:C$ is equivalent to $A:B::C:C$. Now the anthyphairesis of $C:C$ is [1]; so also must be that of $B:C$; so $B = C$.

For an example of an interpretation in which *anthuphairesis* and *antanairesis* do not mean the same thing, see Heron–Bruins, *CC* iii, 60 ff.

For details of abbreviations used to refer to papyri, see Liddell, Scott, & Jones, *GEL*, pp. xl ff.; or Turner, *GP*. Several aspects of the Zenon archive are discussed by Turner; see the index, s.v. Zenon. Some topics in papyrology will be discussed more systematically in Chapters 6 and 7 below.

2.2(a)   The evaluation of the ratio of diagonal to side of a square given here is based on an interpretation of the description of side and diagonal lines found in Proclus' commentary on Plato's *Republic*, to be described in detail in Sections 2.4(e) and 3.6(b). It is the one important place where my reconstruction draws in a crucial way on mathematical evidence that is only found in uncorroborated reports in later commentators—here, in Theon of Smyrna and Proclus.

The anthyphairesis of the side and diagonal of a pentagon is made the basis of a reconstruction of early Greek mathematics in von Fritz, DIHM; for a discussion, see Knorr, *EEE*, 29–32.

The usual Greek word for both diagonal and diameter is *diametros*, 'the through measure'. It can refer to the diagonal of a parallelogram in contexts where there is no corresponding diameter of a circumcircle; see for example, *Elements* I 34: "The opposite sides and angles of parallelogrammic areas (*parallēlogramma chōria*) are equal to one another and the diagonal (*diametros*) bisects the areas"; here *diametros* cannot mean the diameter of a circle, in any interpretation. There is a very rare word *diagōnios* which is only used twice in the *Elements*, at XI 28 and 38, and in very few places elsewhere (I am grateful to Malcolm Brown for this information). I shall adapt translations to read 'diagonal' or 'diameter' as appropriate. For example, the usual translation as 'side and *diameter* lines (or numbers)' is, I believe, inappropriate.

2.2(b)   The translations of Archimedes' *Sphere and Cylinder* are taken from Thomas, *SIHGM* ii, 40 ff. There is a photographic reproduction of our sole surviving manuscript of Heron's *Metrica* in Heron–Bruins, *CC* i (with transcription in vol. ii

and annotated translation in vol. iii); the translation of the passage quoted here comes from Knorr, *AMC*. Also see the notes to Sections 2.3(c) and 2.4(d), below.

An anecdote from Reid, *H*, 164, illustrates the difficulty and unpredictability of problems associated with irrationality and transcendence: "[Siegel] was always to remember a lecture on number theory which he heard [as a student in 1919] from Hilbert . . . . Hilbert wanted to give his listeners examples of the characteristic problems of the theory of numbers which seem at first glance so very simple but turn out to be incredibly difficult to solve. He mentioned Riemann's hypothesis [concerning the zeros of $\zeta(s)$], Fermat's theorem [the hypothesis concerning the insolubility of $x^n + y^n = z^n$ in integers if $n \geqslant 3$], and the transcendence of $2^{\sqrt{2}}$ (which he had listed as his seventh problem at Paris) as examples of this type of problem. Then he went on to say that there had recently been much progress on Riemann's hypothesis and he was very hopeful that he himself would live to see it proved. Fermat's problem had been around for a long time and apparently needed entirely new methods for its solution—perhaps the youngest members of his audience would live to see it solved. But as for establishing the transcendence of $2^{\sqrt{2}}$, no one present in the lecture hall would live to see that! The first two problems which Hilbert mentioned are still unsolved. But less than ten years later, a young Russian named Gelfond established the transcendence of $2^{\sqrt{-2}}$. Utilising this work, Siegel himself was shortly able to establish the desired transcendence of $2^{\sqrt{2}}$." Fermat's Last Theorem was proved by Andrew Wiles in 1994 but there still does not seem to have been any significant progress on the Riemann Hypothesis, so the problems have succumbed in precisely the opposite order to that predicted by Hilbert.

2.3    (a), (b), & (c) The '*Parmenides* proposition' is described in Steele, MRCP, on p. 179: "The *Parmenides* (154b–d) notes that the ratio $(a + x) : (b + x)$ lies nearer to 1 : 1 than does $a : b$ (all terms being positive) and enunciates a strangely explicit principle of continuity." There is also a useful discussion in Plato–Allen, *PP*, 258–9 which concludes: "It has been said that Plato was not an original mathematician. But it is worth remarking that the theorems he here uses are not attested before Hellenistic times."

My original treatment, in REGM, of the approximation algorithm and the evaluation of convergents was more diffident than the account given here; see, e.g., p. 823: "the algorithm described here is not proposed as a reconstruction of an original procedure" (in italics there!). At that stage I had not appreciated the implications of the *Parmenides* passage, I was unaware of the existence of the proposition in Pappus, and my treatment of the question of the starting estimates was offhand and unsatisfactory. I now believe that it is not unreasonable to propose some appropriate formulation of this algorithm as a reconstruction of an early Greek procedure. Also see, for example, Knorr, AMC, 137 ff., where a much more sophisticated variant of the algorithm, there expressed in terms of the manipulation of common fractions, is proposed as an explanation of Archimedes' approximation to $\sqrt{3} : 1$.

The *Parmenides* proposition is an example of a mean: given a smaller and larger ratio, it will generate an intermediate ratio. We have many reports of early interest in different kinds of means among Greek mathematicians and music theorists (see e.g. Heath, *HGM* i, 85–9); but in all these cited examples, the mean generates an intermediate between two magnitudes or, in special cases, between two numbers, while the *Parmenides* proposition generates an intermediate between two ratios. There is a connection between the arithmetic mean and

the *Parmenides* proposition: if $c = \frac{1}{2}(a+b)$ is the arithmetic mean between two magnitudes $a$ and $b$ with $a < b$, the *Parmenides* proposition, applied to $a:1$ and $b:1$ (with $a:1 < b:1$) will generate the ratio $a + b : 2 = c : 1$, and this kind of description can be extended to the general case of weighted arithmetic means. I shall return to this question in Section 4.5(b); see especially B$_{74}$ there.

The first edition of this book attempted to reintroduce the word 'algorism' to replace 'algorithm', a word castigated by the *Oxford English Dictionary* (1933), s.v. algorism, as one of the "many pseudo-etymological perversions, including a recent *algorithm* in which [*algorism*] is learnedly confused with the Greek $\alpha\rho\iota\tau\theta\mu\acute{o}\varsigma$, number". The *Supplement* to the *OED* (in vol. 1, 1972, incorporated into the *OED* second edition of 1989), admitted the word and cited, as its first example, the first edition of Hardy & Wright, *ITN* (1938), a passage that was slightly changed in the second edition of 1945. In fact, the early meaning of 'algorism' is actually 'calculation with Hindu–Arabic numerals', as opposed to calculation with the abacus. During the twelfth to fifteenth centuries, there were hundreds of treatises whose titles or incipits were variations on *De arte numerandi algoristica*, where the abacus treatises might be called *Practica aritmetica cum denariis proiectilibus* or *Algorithmus linealis*. This last description might be seen as a precursor of our modern use of 'algorithm'; in any case, the word has an even more established and distinguished history than I had thought. Its first indubitably modern use as far as I know—and what must surely be its first significant mathematical use—was in Latin in 1684, in Leibniz' first publication on the calculus in *Acta Eruditorum*, indeed in the first publication of the calculus, 'Nova Methodvs pro maximis et minimis . . .': "Ex cognito hoc velut *Algorithmo*, ut ita dicam, calculi hujus, quem voco *differentialem*, . . .". This short article of six pages was translated by Joseph Raphson in 1715, in *The Theory of Fluxions* . . . , where this passage is rendered "Now from this being known as the *Algorithm*, as I may say of this *Calculus*, which I call differential, . . ." (p. 23). The word was then taken up by Euler, for instance in his article 'De usu novi algorithmi in problemate Pelliano solvendo' (and the 'Pell' here gave rise to the misnomer described in Section 2.4(c), above), and its use was then firmly established. (Details of these publications are listed in the bibliography, and I would like to thank Gerhard Brey, Julio Gonzalez Cabillon, Menso Folkerts, Ivor Grattan-Guinness, and Steve Russ for information and help in compiling this note.)

2.4(a)   For discussions of Eratosthenes' ratio, see Ptolemy–Toomer, *PA*, 63 n. 75; Dicks, *GFH*, 167 ff. and *DSB* s.v. Eratosthenes; Neugebauer, *HAMA* ii, 734 (especially n. 15); Rawlins, EGU with my subsequent note EROE; Goldstein, OEAGE; and Taisbak, EET. This translation uses common fractions, but the Greek text follows the practice described briefly in Section 1.2(c) and to be examined in detail in Chapter 7: Ptolemy describes Eratosthenes' ratio as *μζ kai meizonos men ē dimoirou tmēmatos, elassonos de hēmisous tetartou*, "47 and more than the two parts of a division, but less than a half fourth".

2.4(b)   There is a full discussion of calendric cycles in Neugebauer, *HAMA* ii, 615–29. Of the precise data that can be extracted from some of them, he observes drily (p. 616): "We may, of course, say that [Callippus' calendar] implies that 1 month = 29; 31, 53, 3, . . . days and be pleased by the accuracy of this value but one may well ask whether Callippus was ever interested (or able) to carry out the division . . ." (this sexagesimal value of the mean synodic month is equal to the decimal 29·530 85 . . . days; the actual value is 29·530 59 days), and (p. 623):

"Neither the Metonic nor the Callippic cycle have anything to do with an accurate determination of the length of the mean synodic month".

2.4(c)  There is a critical edition of *On the Sizes and Distances of the Sun and Moon* in Aristarchus–Heath, *AS*. I shall describe the style of Aristarchus' calculations in Section 7.3(b).

2.4(d)  Also see 2.2(b), above. There is a complete text of Archimedes' *Measurement of a Circle* with translation, filled out with extracts from Eutocius' commentary, in Thomas, *SIHGM* i, 316–33, but this omits the apparatus to be found in the critical edition, Archimedes–Heiberg, *Opera* i, 232–43; and Heiberg's apparatus needs to be corrected by and collated with the Latin translation by William of Moerbeke in Clagett, *AMA* ii. See Section 7.3(a) for more details, and a discussion of the treatment of the fractions in the original.

On the excessive accuracy of Archimedes' approximations of $\sqrt{3}:1$, see Knorr, AMC, 137 f. This article contains the best overall analysis of Archimedes' general strategy that I know; my only reservation concerns the arithmetised context that is assumed for the calculation. I shall discuss the text in more detail in Section 7.3(a).

2.4(e)  Fermat's challenge, together with an exchange of letters between Frenicle de Bessy, Brouncker, and Wallis, is published in Wallis, *CE*; there is a discussion of the different methods of solving the problem in Edwards, *FLT*, 25 ff. The translation of Theon of Smyrna, on side and diagonal numbers, the last paragraph excepted, is taken from Thomas, *SIHGM* i, 132–7, except that I have again replaced 'diameter' by 'diagonal' throughout.

2.4(f)  I know of no discussion of the implications of this alternative interpretation of Archimedes' *Cattle Problem*. My own exploration of these mathematical aspects of the problem, ACPPCM, has never been published.

# *ELEMENTS* II:
# THE DIMENSION OF SQUARES

## 3.1 INTRODUCTION

This chapter will be entirely devoted to the problem of the dimension of squares (see $S_{41}$ and Section 2.2(a)): given a line $\rho$, and two numbers $n$ and $m$, what can we say about the anthyphairesis of $\sqrt{n}$ and $\sqrt{m}$, the ratio of the sides of the two squares equal to $n$ and $m$ times the square on $\rho$? We shall only consider the case where $\sqrt{n}$ and $\sqrt{m}$ are incommensurable, and so the anthyphairesis does not terminate.

The algorithms of the last chapter have begun to reveal some remarkable properties of the anthyphairetic ratio $\sqrt{3} : 1$, and a similar arithmetical exploration, leading to the formation of a hypothesis, will be the first step of the investigation of any ratio $\sqrt{n} : \sqrt{m}$. This hypothesis will then determine the second step, the construction of a geometric figure. Although some of these figures are very elaborate, they are built up out of simple ingredients such as squares, rectangles, and parallelograms; their geometrical properties are, on the whole, straightforward and will be taken for granted, just as Socrates, in the *Meno*, omitted to verify the details of the various geometrical figures he described. Finally, in each case, we shall see that the hypothesis with which we started will correspond to some property exhibited by the figure.

These detailed mathematical arguments underlie a historical argument: I propose that this whole programme may indeed correspond to an early Greek investigation of which only the basic geometrical figures survive. Thus the situation is complementary to that which we find in the *Meno*: our surviving manuscripts of Plato's dialogues do not—and perhaps none of them ever did—contain any of the figures on which the mathematical arguments depend. (In Socrates' conversation with the slaveboy, the information in the text is so detailed that we have no difficulty in reconstructing the figures with confidence but, at *Meno* 86e–87b, there is another mathematical passage that has provoked an enormous variety of different geometrical reconstructions.) For the problem of the dimension of squares, I shall propose that the basic figures needed in this investigation of the dimension of squares survive in the form of a series of simple though convoluted geometrical exercises: the substantial, coherent, and idiosyncratic material comprising Book II of the *Elements* and the block of cognate propositions at XIII 1–5. It is now a much harder and more hazardous process to proceed in the opposite direction, to reconstruct the mathematical investigation which gave rise to these geometrical figures. One

important methodological principle can be stated at the outset, and will be observed hereafter throughout this book: that the evidence must be taken as a whole. Since the geometrical techniques set out in the propositions of Book II can be adapted to other purposes in other contexts, very many applications of some of these propositions can be found elsewhere. Because of this versatility, and because of the coherence of this material within the *Elements*, if we wish to try to explain its role within the development of pre-Euclidean mathematics, it is not sufficient to provide some retrospective justification in which *some* of these propositions are related to topics studied *somewhere* throughout the long history of the development of mathematics, from ancient Babylonian arithmetical techniques or before up to modern preoccupations with algebraic structures; we must rather try to provide some acceptable reconstruction which leads forward, up to *all* of them. And since this material also seems to relate to the construction and classification of incommensurable lines in Book X, the interpretation should connect with the developments of this book also.

The arguments of this chapter will be arduous. They will involve novel mathematical procedures, elaborate numerical explorations, and detailed textual examinations. If it is any consolation to the non-mathematician, I think that a conventional mathematical training may generate difficulties in following the line of reasoning, for much of the work fits uneasily into our arithmetised viewpoint of today; to the non-specialist in ancient mathematics, a similar consolation applies, since my explanation requires that a great deal of the received interpretation be put aside. It is not important to follow the mathematical argument in detail to understand the proposals that follow in later chapters, and I again encourage my reader to skip many of the technical details that arise. However, historical considerations now start moving into the foreground, and I believe it is necessary to appreciate the kinds of mathematical techniques I shall be describing here in order to evaluate the plausibility of some of my later historical proposals.

## 3.2 BOOK II OF THE *ELEMENTS*

Book II comprises two definitions:

Any rectangular parallelogram is said to be contained by the two straight lines containing the right angles. And in any parallelogramic area let any one whatever of the parallelograms about its diameter with the two complements be called a gnomon,

and fourteen propositions about squares, rectangles, and gnomons. The subject of each proposition is best conveyed by its figure (and it is these figures, not what is made of them in their enunciations or proofs, that will enter my proposed reconstruction); these are given here as Figs 3.1–3.14, each corresponding to Book II Propositions 1–14. (The additional Figs 3.9(a) and

3.10(a) will be explained presently.) The caption of each figure gives an abbreviated statement of the enunciation. For example the enunciation of II 4:

If a straight line [AB] be cut at random [at C], the square on the whole [AB$^2$] is equal to the squares on the segments [AC$^2$ + CB$^2$] and twice the rectangle contained by the segments [2AC . CB]
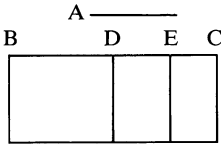


Fig. 3.1.
$A . BC = A . BD + A . DE + A . EC$
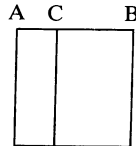


Fig. 3.2.
$AB . BC + BA . AC = AB^2$
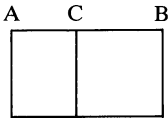


Fig. 3.3. $AB . BC = AC . CB + BC^2$
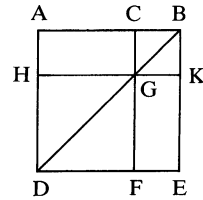


Fig. 3.4.
$AB^2 = AC^2 + CB^2 + 2AC . CB$
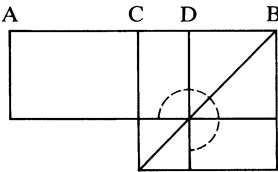


Fig. 3.5. $(AC = CB)$
$AD . DB + CD^2 = CB^2$
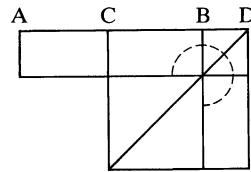


Fig. 3.6. $(AC = CB)$
$AD . DB + CB^2 = CD^2$
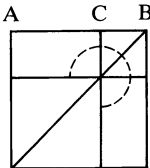


Fig. 3.7.
$AB^2 + BC^2 = 2AB . BC + CA^2$
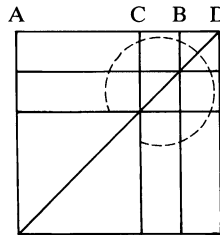


Fig. 3.8. $(BC = BD)$
$4AB . BC + AC^2 = (AB + BC)^2 = AD^2$
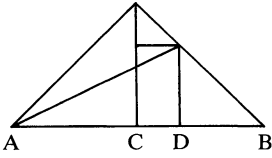
FIG. 3.9.  (AC = CB)
$$AD^2 + DB^2 = 2AC^2 + 2DC^2$$
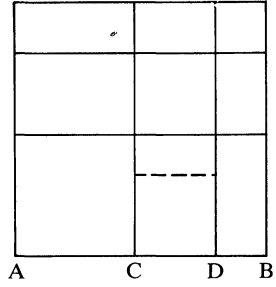
FIG. 3.9(a).  (AC = CB)
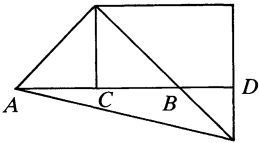
FIG. 3.10.  (AC = CB)
$$AD^2 + DB^2 = 2AC^2 + 2CD^2$$
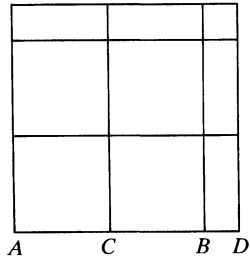
FIG. 3.10(a).  (AC = CB)
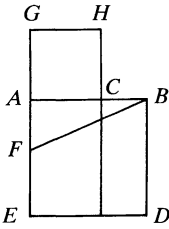
FIG. 3.11.  $AB . BC = AC^2$
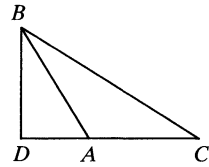
FIG. 3.12.
$$BC^2 = BA^2 + AC^2 + 2CA . AD$$
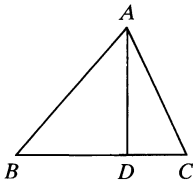
FIG. 3.13.
$$AC^2 + 2CB . DB = CB^2 + BA^2$$
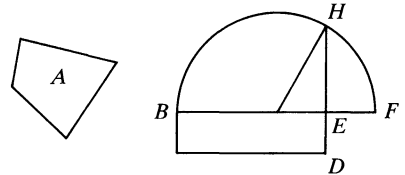
FIG. 3.14.  (EF = ED)
$$A = BE . EF = EH^2$$

has become the caption of Fig. 3.4:

$$AB^2 = AC^2 + CB^2 + 2AC \cdot CB.$$

It is important to realise that all these operations are geometrical; neither Euclid's propositions, nor my notation, makes or needs any reference to the manipulation of any kind of number other than the *arithmoi*, as with the 'twice' that appears in this enunciation.

The demonstration of Proposition 4 is typical of Propositions 1–8. The setting-out is brief (lines 4 to 7 in Heath's translation); it identifies the various components of the figure: "For let the straight line AB be cut at random at C", then "The square on the whole" is "the square on AB", and so on. The construction follows (lines 8 to 14): the square ADEB, the diagonal BD, CF parallel to AD or EB, and HK through G parallel to AB or DE. The proof itself is a long and meticulous verification of the consequences of the construction: that BC equals CG (lines 15 to 22), that CGKB is a square, the square on CB (lines 23 to 35), and that HDFG is the square on AC (lines 36 to 38, repeated on line 39). Then it is asserted that AG (i.e. the rectangle AHGC) equals GE (line 40), and that AG and GE are equal to the rectangle contained by AC and CB (lines 41 to 42). The results so far are summarised (lines 43 to 45), and this summary is repeated (lines 46 to 48):

Therefore the four areas HF, CK, AG, GE are equal to the squares on AC, CB, and twice the rectangle contained by AC, CB,

and the fact that this gives a decomposition of the square, the ostensible point of the proposition, is merely stated (lines 49 to 50):

But HF, CK, AG, GE are the whole ADEB, which is the square on AB.

This is then followed by a repetition of the setting-out (lines 51 to 52) and the enunciation (abbreviated, as usual, to "Therefore etc." in Heath's translation). The proposition has an alternative proof and is followed by a porism:

From the demonstration it is manifest that in square areas the parallelograms about the diameter are squares,

but both are believed to be interpolations by Theon of Alexandria; see Heath, *TBEE* i, 381.

The elaborate verifications in this proof, for example that CBGK is a square, are not repeated in the later propositions, so that the proofs of Propositions 5, 6, and 7 are shorter. Proposition 8 again has a long proof, since the verification that the double gnomon gives rise to equal squares in the corner and equal rectangles along the arms is carried out in detail, again with repetitions. In all of these propositions, almost all of the text of the demonstrations concerns the construction of the figures, while the substantive content of each enunciation is merely read off from the constructed figures, at the end of the proofs, as in lines 49 to 50 of II 4, the proposition just considered.

Proposition 9 starts in the same style; its enunciation follows the well-established pattern:

If a straight line be cut into equal and unequal segments, the squares on the unequal segments of the whole are double of the square on the half and of the square on the straight line between the points of section;

it has the same kind of setting-out, but the associated figure and proof differs from the previous propositions. Euclid's proof proceeds by drawing two isosceles right-angled triangles which fit together into a larger isosceles right-angled triangle (Fig. 3.9), and applying I 47, the so-called Pythagoras' theorem; alternative proofs are easily constructed which avoid the use of I 47 and continue to apply the techniques of dissecting squares. One such proof is given in Heath's note to II 9, and an even simpler version of it follows directly from Fig. 3.9(a). A similar situation occurs with Proposition 10: Euclid's proof is based on Fig. 3.10, and an alternative proof can be read off from Fig. 3.10(a). Proposition 11 solves a specific problem:

To cut a straight line so that the rectangle contained by the whole and one of the segments is equal to the square on the remaining segment;

Propositions 12 and 13 give versions of a generalisation of I 47 which correspond to what we now call the cosine law for obtuse- and acute-angled triangles respectively; and Proposition 14, which concludes the book, shows how

To construct a square equal to a given rectilineal figure.

This construction starts with a typical Euclidean back-reference, to I 45:

For let there be constructed the rectangular parallelogram BD equal to the rectilineal figure A,

and the new and important contribution of the rest of the proposition is to show how to convert this rectangle into an equal square.

With the exception of implied uses of I 47 and 45, Book II is virtually self-contained in the sense that it only uses straightforward manipulations of lines and squares of the kind assumed without comment by Socrates in the *Meno*. Moreover the only reference to I 45, just quoted above, occurs in the last proposition, II 14, where it is tacked on to contribute extraneous generality, out of keeping with the style of the rest of the book. Also, it can be argued that a proof of 'Pythagoras' theorem', referred to in Section 1.2(e), has been excised from between Propositions 8 and 9. This proof, which I shall call Proposition 8a, exploits the manipulation of gnomons, the basic technique of Book II, and, with its obliquely placed square, is reminiscent of the successful third *Meno* figure; so the proof also conforms in style with the testified ingredients of early Greek mathematics. Moreover the stylistic variations in the proof of II 9 and 10 could then arise from the wish of the book's author, compiler, or editor to integrate this new result further into the logical development before it is

required in II 11 and generalised in II 12 and 13; then a later editor might have excised the proposition to remove the duplication of I 47 and II 8a. The argument is plausible, and further considerations in favour of it can be found in Knorr, *EEE*, 174–9. I shall return, in Section 3.6, to further discussion of Propositions 12 and 13.

## 3.3 THE HYPOTHESES

The algorithms of Chapter 2 enable us to evaluate as many terms as we want of the anthyphairesis of any ratio $\sqrt{n} : \sqrt{m}$, the only restriction being our ability to express or evaluate the squares of the integers involved. We also evaluate, at the same time, the convergents of these ratios. For example:

$$\sqrt{2} : 1 = [1, 2, 2, 2, 2, \ldots],$$

with convergents $1 : 1, \ 3 : 2, \ 7 : 5, \ 17 : 12, \ 41 : 29, \ldots$;

$$\sqrt{3} : 1 = [1, 1, 2, 1, 2, 1, 2, \ldots],$$

with convergents $1 : 1, \ 2 : 1, \ 5 : 3, \ 7 : 4, \ 19 : 11, \ 26 : 15, \ 71 : 41, \ldots$;

$$\sqrt{3} : \sqrt{2} = [1, 4, 2, 4, 2, 4, 2, \ldots],$$

with convergents $1 : 1, \ 5 : 4, \ 11 : 9, \ 49 : 40, \ 109 : 89, \ 485 : 396, \ 1079 : 881, \ldots$;

$$\sqrt{4} : 1 = 2 : 1 = [2];$$

$$\sqrt{4} : \sqrt{2} = \sqrt{2} : 1, \quad \text{given above;}$$

$$\sqrt{4} : \sqrt{3} = [1, 6, 2, 6, 2, 6, 2, \ldots],$$

with convergents $1 : 1, \ 7 : 6, \ 97 : 84, \ 209 : 181, \ 1351 : 1170, \ 2911 : 2521, \ldots$; etc.

The obvious hypothesis to draw from these and further examples is that the expansion of $\sqrt{n} : \sqrt{m}$ is periodic, with a period that starts with the second term; and we have already verified this for the example $\sqrt{2} : 1$, the ratio of the diagonal to side of a square. Table 3.1 lists the expansions, but not the convergents, of $\sqrt{n} : 1$ and $\sqrt{(2n+1)} : \sqrt{2}$ for $1 \leqslant n \leqslant 50$. (When $n$ is a square, the ratio $\sqrt{n} : 1$ is commensurable and the entry for the anthyphairetic ratio contains only one term; otherwise the entry is terminated by a comma after one apparent period. Hereafter I shall always silently assume that $n$ and $m$ are such that $\sqrt{n}$ and $\sqrt{m}$ are incommensurable, so that all examples such as $\sqrt{8} : \sqrt{2}$ have been suppressed, as in the second column of this table.)

The examples in the table lead us to formulate a slightly more refined hypothesis, that

$$\sqrt{n} : \sqrt{m} = [n_0, \overline{n_1, n_2, \ldots, n_k, 2n_0}].$$

where the period, identified by the superior bar, terminates with a term equal to twice the first term of the expansion. It also seems that we can guess the general behaviour of some examples that lie around the commensurable ratio $\sqrt{n^2} : 1$,

TABLE 3.1

| $n$ | $\sqrt{n}:1$ | $\sqrt{2n+1}:\sqrt{2}$ |
|---|---|---|
| 1 | 1 | 1, 4, 2, |
| 2 | 1, 2, | 1, 1, 1, 2, |
| 3 | 1, 1, 2, | 1, 1, 6, 1, 2, |
| 4 | 2 | 2, 8, 4, |
| 5 | 2, 4, | 2, 2, 1, 8, 1, 2, 4, |
| 6 | 2, 2, 4, | 2, 1, 1, 4, |
| 7 | 2, 1, 1, 1, 4, | 2, 1, 2, 1, 4, |
| 8 | 2, 1, 4, | 2, 1, 10, 1, 4, |
| 9 | 3 | 3, 12, 6, |
| 10 | 3, 6, | 3, 4, 6, |
| 11 | 3, 3, 6, | 3, 2, 1, 1, 3, 1, 12, 1, 3, 1, 1, 2, 6, |
| 12 | 3, 2, 6, | 3, 1, 1, 6, |
| 13 | 3, 1, 1, 1, 1, 6, | 3, 1, 2, 14, 2, 1, 6, |
| 14 | 3, 1, 2, 1, 6, | 3, 1, 4, 4, 1, 6, |
| 15 | 3, 1, 6, | 3, 1, 14, 1, 6, |
| 16 | 4 | 4, 16, 8, |
| 17 | 4, 8, | 4, 5, 2, 5, 8, |
| 18 | 4, 4, 8, | 4, 3, 3, 8, |
| 19 | 4, 2, 1, 3, 1, 2, 8, | 4, 2, 2, 2, 8, |
| 20 | 4, 2, 8, | 4, 1, 1, 8, |
| 21 | 4, 1, 1, 2, 1, 1, 8, | 4, 1, 1, 1, 3, 18, 3, 1, 1, 1, 8, |
| 22 | 4, 1, 2, 4, 2, 1, 8, | 4, 1, 2, 1, 8, |
| 23 | 4, 1, 3, 1, 8, | 4, 1, 5, 1, 1, 3, 2, 1, 18, 1, 2, 3, 1, 1, 5, 1, 8, |
| 24 | 4, 1, 8, | 4, 1, 18, 1, 8, |
| 25 | 5 | 5, 20, 10, |
| 26 | 5, 10, | 5, 6, 1, 3, 3, 1, 6, 10, |
| 27 | 5, 5, 10, | 5, 4, 10, |
| 28 | 5, 3, 2, 3, 10, | 5, 2, 1, 20, 1, 2, 10, |
| 29 | 5, 2, 1, 1, 2, 10, | 5, 2, 3, 6, 1, 20, 1, 6, 3, 2, 10, |
| 30 | 5, 2, 10, | 5, 1, 1, 10, |
| 31 | 5, 1, 1, 3, 5, 3, 1, 1, 10, | 5, 1, 1, 1, 1, 2, 1, 1, 1, 1, 10, |
| 32 | 5, 1, 1, 1, 10, | 5, 1, 2, 2, 1, 10, 1, 2, 2, 1, 10, |
| 33 | 5, 1, 2, 1, 10, | 5, 1, 3, 1, 2, 1, 1, 22, 1, 1, 2, 1, 3, 1, 10, |
| 34 | 5, 1, 4, 1, 10, | 5, 1, 6, 1, 10, |
| 35 | 5, 1, 10, | 5, 1, 22, 1, 10, |
| 36 | 6 | 6, 24, 12, |
| 37 | 6, 12, | 6, 8, 12, |
| 38 | 6, 6, 12, | 6, 4, 1, 7, 2, 7, 1, 4, 12, |
| 39 | 6, 4, 12, | 6, 3, 1, 1, 24, 1, 1, 3, 12, |
| 40 | 6, 3, 12, | 6, 2, 1, 2, 1, 24, 1, 2, 1, 2, 12, |

TABLE 3.1 (*cont.*)

| $n$ | $\sqrt{n}:1$ | $\sqrt{2n+1}:\sqrt{2}$ |
|---|---|---|
| 41 | 6, 2, 2, 12, | 6, 2, 3, 1, 4, 2, 1, 1, 1, 7, 1, 24, 1, 7, 1, 1, 1, 2, 4, 1, 3, 2, 12, |
| 42 | 6, 2, 12, | 6, 1, 1, 12, |
| 43 | 6, 1, 1, 3, 1, 5, 1, 3, 1, 1, 12, | 6, 1, 1, 2, 8, 2, 1, 1, 12, |
| 44 | 6, 1, 1, 1, 2, 1, 1, 1, 12, | 6, 1, 2, 26, 2, 1, 12, |
| 45 | 6, 1, 2, 2, 2, 1, 12, | 6, 1, 2, 1, 12, |
| 46 | 6, 1, 3, 1, 1, 2, 6, 2, 1, 1, 3, 1, 12, | 6, 1, 4, 1, 1, 8, 1, 1, 4, 1, 12, |
| 47 | 6, 1, 5, 1, 12, | 6, 1, 8, 3, 1, 4, 1, 3, 8, 1, 12, |
| 48 | 6, 1, 12, | 6, 1, 26, 1, 12, |
| 49 | 7 | 7, 28, 14, |
| 50 | 7, 14, | 7, 9, 2, 2, 9, 14, |

for example:

$$\sqrt{(n^2-1)}:1 = [(n-1), \overline{1, 2(n-1)}],$$

$$\sqrt{(n^2+1)}:1 = [n, \overline{2n}], \quad \text{and}$$

$$\sqrt{(n^2+2)}:1 = [n, \overline{n, 2n}].$$

This chapter will be dedicated to the geometrical verification of statements of this sort.

While the computation of the whole of Table 3.1 using the algorithms of Chapter 2 might be straightforward, it would be very time-consuming and tedious, as can be seen by calculating the example of $\sqrt{19}:1$, the first example of a ratio $\sqrt{n}:1$ which involves a long period containing several terms greater than one in its period. (However, a skilled and resourceful calculator would notice further patterns of behaviour that could be exploited to reduce greatly the amount of routine calculation.) In fact I calculated the results of this table using a pocket calculating machine and a simple arithmetical algorithm:

$$x = x_0;$$

$$x_k = n_k + \varphi_k \quad \text{with} \quad 0 \leqslant \varphi_k < 1;$$

$$\text{if } \varphi_k \neq 0 \quad \text{then} \quad \varphi_k^{-1} = x_{k+1};$$

$$\text{then} \quad x:1 = [n_0, n_1, n_2, \dots],$$

and I verified the one doubtful example, $\sqrt{83}:\sqrt{2}$, which lay beyond the limits of the accuracy of the machine, using the algorithm to be described in Chapter 9. So I do not wish to propose that such a systematic and large-scale set of calculations as is embodied in Table 3.1 might ever have been performed in

antiquity. It would seem more reasonable to assume that only some very restricted corpus of examples was ever explored, at most, for example, the ratios $\sqrt{n}:1$ for $2 \leqslant n \leqslant 20$, and a few other isolated examples of $\sqrt{n}:\sqrt{m}$. Therefore the quite startling general hypothesis, that the period comprises a *palindromic* block terminated by twice the first term, would not present itself so forcefully from this limited exploration. (In the case of $\sqrt{(n^2 + 1)}:1$, this palindromic block is void, while in some other cases it is reduced to only one term.) I know of no evidence, of any kind whatsoever, that leads me to suggest that this hypothesis might have any part in a reconstruction of Greek mathematics; all of the abundant Greek evidence about palindromes seems to be later in time and connected with charms, magic, and gratuitous word-play.

An example of a false hypothesis which results from an inference made from an insufficiently universal collection of examples is that the period has been evaluated when the first occurrence of a term equal to $2n_0$ is reached. A more extensive search would reveal examples such as

$$\sqrt{15}:\sqrt{8} = [1, 2, 1, 2, 2, \ldots].$$

However, it is true that in any ratio of the form $\sqrt{n}:1$, the first occurrence of a term $2n_0$ will mark the end of the period, though I shall not prove this here in this book and I know of only one obscure place in the mathematical literature (Muir, *EQSCF*) where this type of result is investigated.

It seems curiously difficult to prove any of these arithmetical conjectures directly from the properties of the *Parmenides* proposition, from the operation of the algorithm that generated them, the more so if the proof is restricted to the ingredients of Greek *arithmētikē* such as I have described them in Section 1.2(c). But we can reformulate the problem in various ways and seek indirect proofs; for example, in modern style, we can arithmetise these conjectures by converting them into statements about the manipulation of real numbers, and then explore them algebraically. I shall proceed differently, and geometrise. For us here, the hypotheses that $\sqrt{2}:1 = [1, \overline{2}]$ will be a statement about the anthyphairesis of the diagonal and side of a square, and one version of such an interpretation was proved in 2.2(a). So we now turn to the problem of constructing further figures which may help in establishing and explaining some of the hypotheses that arise from our heuristic investigations of these ratios $\sqrt{n}:\sqrt{m}$.

### 3.4 THE FIRST ATTEMPT: THE METHOD OF GNOMONS

The first method is a direct confrontation of the problem: we build up squares of the appropriate sizes by adding gnomons to a given square on an assigned line $\rho$, introduce some additional construction lines, and then convert the statements about the anthyphairesis into statements about these figures which we then verify. We shall always be dealing with geometrical lines, never their

arithmetical lengths, but for brevity and clarity we will usually omit mention of $\rho$. For example, P will denote the square on $\rho$ and $\sqrt{2}$ the line $\sqrt{2}\rho$, the side of the square equal to 2P.

PROPOSITION   $\sqrt{2}:1 = [1,\bar{2}]$.

PROOF   We prove the equivalent result that $(1 + \sqrt{2}):1 = [\bar{2}]$. Start with the square P; then the square 2P with the same bottom left-hand corner differs from P by a gnomon $Q + R + S$, equal to P (Fig. 3.15). This figure now contains the sides $AB = 1$ and $AC = \sqrt{2}$. We can check that BC, the width of the gnomon, is less than AB, since a gnomon of width AB is equal to 3P, which is too big. Add a line AD, equal to AB; then $DC = 1 + \sqrt{2}$, and all the lines needed in the enunciation of the proposition have been constructed. The first step of the anthyphairesis of $DC:AB = (1 + \sqrt{2}):1$ is

$$DC - 2AB = BC \quad \text{with} \quad BC < AB.$$

Hence

$$DC:AB = [2, AB:BC].$$

We now prove that $AB:BC = DC:AB$. This is equivalent to the statement that the square on AB is equal to the rectangle contained by DC and BC; hence we must verify that

$$P = T + Q + R$$

where T is the appended rectangle contained by AD and BC. But these two plane regions are obviously equal, since $T = Q = S$, and $Q + R + S = P$. Hence, by the same recursive property that was described and exploited in Section 2.2(a), we see that

$$DC:AB = [2, DC:AB] = [\bar{2}]. \qquad\qquad \text{QED}$$

The same technique, starting from a square of side $n$ and size $n^2 P$, will prove the generalisation:

PROPOSITION   For any number $n$, $\sqrt{(n^2 + 1)} = [n, \overline{2n}]$.

I shall consider such general statements later. Meanwhile here is its first case.
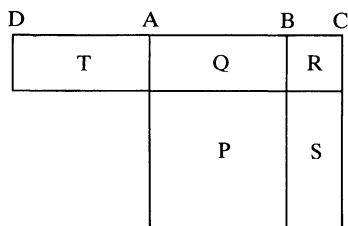


FIG. 3.15

PROPOSITION   $\sqrt{3}:1 = [1,\overline{1,2}]$.

PROOF   We prove the equivalent result that $(1 + \sqrt{3}):1 = [\overline{2,1}]$. Start with the square P, a gnomon $Q + R + S$ of size 2P, and the square with side $AC = \sqrt{3}$, size 3P (Fig. 3.16). We can check that $\frac{1}{2}AB < BC < AB$, since a gnomon with width $\frac{1}{2}AB$ is too small, while a gnomon of width AB is too big. Add a new gnomon $T + U + V + W + X$ of size P to make the total width of added gnomons, BD, equal to AB; and append a rectangle Y with sides $EA = AB$ and CD. We can see that $Q = S$, $T = X = Y$, and $U = W$. We now proceed to the anthyphairesis of $EC:AB = (1 + \sqrt{3}):1$.

$$EC - 2AB = BC \quad \text{with} \quad BC < AB,$$

and, since $AB = BD$,

$$BD - 1BC = CD \quad \text{with} \quad CD < BC.$$

Hence

$$EC:AB = [2, 1, BC:CD].$$

We now prove that $BC:CD = EC:AB$. This is equivalent to the statement that the rectangle with sides AB and BC is equal to the rectangle with sides EC and CD; hence we must verify that

$$R + U = Y + T + U,$$

that is,

$$R = 2T.$$

But

$$2Q + R = 2P,$$

hence

$$2Q + 2T + R = 2P + 2T,$$

and

$$Q + T = P.$$

Thus

$$R = 2T,$$

and the equality is established.
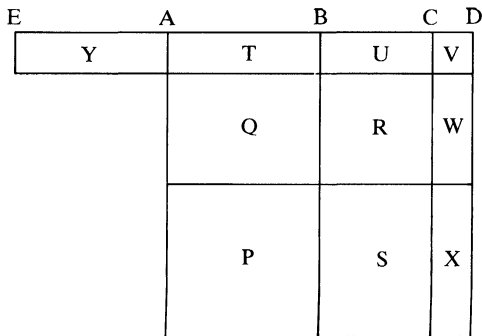


FIG. 3.16

Therefore

$$(1 + \sqrt{3}) : 1 = EC : AB = [2, 1, BC : CD]$$
$$= [2, 1, EC : AB]$$
$$= [\overline{2, 1}],$$

and so

$$\sqrt{3} : 1 = [1, \overline{1, 2}]. \qquad\qquad \text{QED}$$

This result can be generalised many ways, as can be guessed from an inspection of Table 3.1. The following results suggest themselves almost immediately:

$$\sqrt{(n^2 - 1)} : 1 = [(n - 1), \overline{1, 2(n - 1)}],$$

$$\sqrt{(n^2 + 2)} : 1 = [n, \overline{n, 2n}],$$

$$\sqrt{(n^2 + n)} : 1 = [n, \overline{2, 2n}],$$

$$\sqrt{(2n^2 + n)} : \sqrt{2} = [n, \overline{4, 2n}].$$

The evaluation of further expansions suggests yet more results; for example, a table of $\sqrt{n} : \sqrt{3}$ will yield

$$\sqrt{(3n^2 + 2n)} : \sqrt{3} = [n, \overline{3, 2n}], \quad \text{and}$$

$$\sqrt{(3n^2 + n)} : \sqrt{3} = [n, \overline{6, 2n}].$$

These expansions are particular cases of the result that

$$\sqrt{(p^2 q + 2p)} : \sqrt{q} = [p, \overline{q, 2p}],$$

the general form of the expansion which has two terms in its period; I shall give a geometrical derivation of this formula in the next section. All of these results can be verified with elaborations of the procedure used to evaluate $\sqrt{3} : 1$; I shall go on to consider a few examples, continuing as long as the procedures exhibit new features, and will discuss only the new features that enter each calculation. But first, let us relate what we have done so far to the propositions of Book II.

The verification of $\sqrt{2} : 1 = [1, \overline{2}]$ started with the construction of a particular case of a figure of the type of II 4 or 7, to which a rectangle was appended to make a figure of type II 5 and 6. The lines $\sqrt{n}$ needed in this and other examples can all be constructed using II 14. The statement, in Fig. 3.15, that the square on AC is equal to 2P is a particular example of II 4, and that the rectangle contained by DC and BC can be decomposed into T + Q + R is an example of II 1. Figure 3.16 introduces the configuration of double gnomons of II 8 (except that, in this proposition, the gnomons are of equal width; we shall soon encounter examples of multiple equal gnomons), and the decomposition there of P into Q + T is an example of II 2, while the decomposition of the rectangle with sides AB and BC into R + U is an example of II 3.

Propositions II 1 to 10 refer to properties of the general configurations of Figs 3.1–10, each of which is based on "a line cut at random" (*hos etuchen*), while the configurations of Figs 3.15 and 3.16, which arise from particular anthyphairetic ratios, are precisely specified. (In fact, while Figs 3.15 and 3.16 are drawn approximately to scale, in the following Figs 3.17–19 the correctly proportioned gnomons will be too narrow to be labelled conveniently.) My proposed interpretation is that the author of the archetype of Book II is introducing the type of figure and kind of argument that are needed for anthyphairetic calculations, but adapted into self-contained propositions or exercises that do not refer to anthyphairesis. Therefore general statements have been chosen to illustrate the particular kinds of manipulations that arise in anthyphairesis. So, for example, the very particular assertion, in Fig. 3.15, that

$$AC^2 = P + Q + R + S = P + 2P = 3P$$

has become, in II 4, the general statement that

$$AC^2 = P + Q + R + S,$$

and a variation, which introduces overlapping regions such as arise later in these anthyphairetic examples, occurs in II 7

$$AC^2 + R = 2(Q + R) + P.$$

(This proposition reads:

If a straight line be cut at random, the square on the whole and that on one of the segments both together are equal to twice the rectangle contained by the whole and the said segment and the square on the remaining segment.)

Similarly, the configuration of Fig. 3.16 gives rise to the two variations of II 5 and 6, according as whether the gnomon is added to or subtracted from the initial square. The correspondence between these anthyphairetic arguments and Book II will be of this nature, and my arguments can only be sustained if every proposition of Book II can eventually be accounted for in terms of some reasonable anthyphairetic procedure. The first method of gnomons will, in this sense, account for II 1–10 and 14.

Let us return to further examples of this method of gnomons.

PROPOSITION     $\sqrt{(n^2 - 1)} : 1 = [(n - 1), \overline{1, 2(n - 1)}]$.

PROOF     The Euclidean technique for handling a general proposition such as this, which I shall follow in this proof, is to consider the result for a small fixed typical value of $n$, e.g. $n = 4$. We construct Fig. 3.17 as follows: $P_1$ is the square on $\rho$, $P_9$ a square of size $(n - 1)^2 P_1 = 9P_1$, $Q_1 + Q_2 + Q_3 + R + S$ a gnomon of size $2(n - 1)P_1 = 6P_1$, so that $AC = \sqrt{15}$, the side of a square of size $(n - 1)^2 P_1 + 2(n - 1)P_1 = (n^2 - 1)P_1 = 15P_1$; $BD = AE = 1$; and $FA = AB = (n - 1) = 3$. As before, we verify that $\frac{1}{2}AE < BC < AE$.

| F | | | A | E | | B | C | D |
|---|---|---|---|---|---|---|---|---|
| | | | $T_1$ | $T_2$ | $T_3$ | U | V | |
| | | | $Q_1$ | $Q_2$ | $Q_3$ | R | W | |
| | | | | $P_9$ | | | S | X |
| | | | $P_1$ | | | | | |
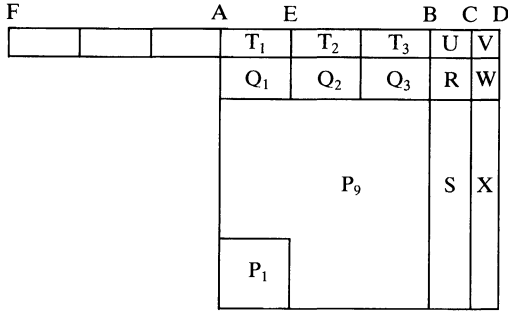
FIG. 3.17

The first two steps of the anthyphairesis of $FC : AE = (3 + \sqrt{15}) : 1$ are

$$FC - 6AE = BC$$

and, since $AE = BD$,

$$BD - 1BC = CD.$$

Hence

$$FC : AE = [6, 1, BC : CD].$$

We now prove that $BC : CD = FC : AE$, i.e. that $BC \cdot AE = FC \cdot CD$, i.e. that

$$R + U = 6T + U$$

where T denotes the equal rectangles $T_1$, $T_2$, and $T_3$. But

$$6Q + R = 6P_1.$$

So, adding 6T and observing that $Q + T = P_1$, we get the result.      QED

PROPOSITION    $\sqrt{(n^2 + 2)} : 1 = [n, \overline{n, 2n}]$.

REMARK    Instead of following the Euclidean technique of choosing a fixed $n$, e.g. $n = 3$, I shall here illustrate the point that a general proof involves little more than the device of writing $Q_1$, $Q_2$, $Q_3$ as $Q_1, \ldots, Q_n$ and recognising where to put the dots and the $n$s; this kind of notational trick is a modern way of automatically indicating generality of an argument. Euclid's use of everyday language, which does not permit the symbolic manipulation involved, does not allow him the possibility of that kind of expression of generality, but he seems aware of the underlying technique. For a clear and celebrated example, see his proof of IX 20.

PROOF    Construct Fig. 3.18 as follows: Start with a square $P_{n^2}$ of size $n^2 P_1$; add a gnomon $Q_1 + Q_2 + \cdots + Q_n + R + S$ of size $2P_1$; add $(n - 1)$ further gnomons of the same width; the total width of the gnomons will then be less than the unit (see below); so add a final gnomon $T_1 + T_2 + \cdots + T_n + U_1 +$
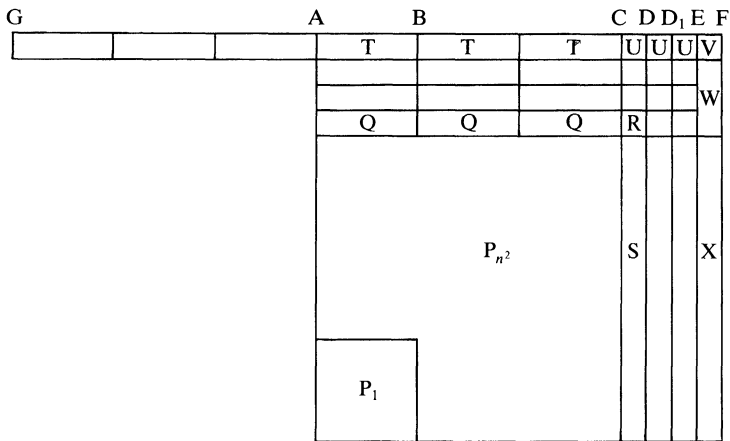
FIG. 3.18

$U_2 + \cdots + U_n + V + W + X$ to bring up the total width of added gnomons to the unit AB. Finally adjoin GA = AC.

Thus the following lines have been constructed: AB, the assigned line; AC = $n$; AD = $\sqrt{(n^2 + 2)}$; GD = $n + \sqrt{(n^2 + 2)}$; CD = $DD_1 = D_1D_2 = \cdots = D_{n-2}E$; and CF = AB.

We need to check that $(1/(n + 1))AB < CD < (1/n)AB$; this follows by the usual kind of argument, since $(2n/(n + 1))AB^2 + (1/(n + 1)^2)AB^2 < 2P_1$ but $(2n/n)AB^2 + (1/n^2)AB^2 > 2P_1$.

The first two steps of the anthyphairesis are:

$$GD - 2nAB = CD, \quad \text{where} \quad CD < (1/n)AB,$$

and, since AB = CF,

$$CF - nCD = EF, \quad \text{where} \quad EF = AB - nCD < CD.$$

Hence

$$GD : AB = [2n, n, CD : EF].$$

Finally, we prove that CD : EF = GD : AB, i.e. since AB = CF, CD . CF = GD . EF. As before, this is equivalent to

$$nR + U = 2nT + U, \quad \text{i.e.} \quad 2T = R.$$

But, by the original construction,

$$2nQ + R = 2P_1,$$

and

$$nQ + T = P_1,$$

from which the conclusion follows.      QED
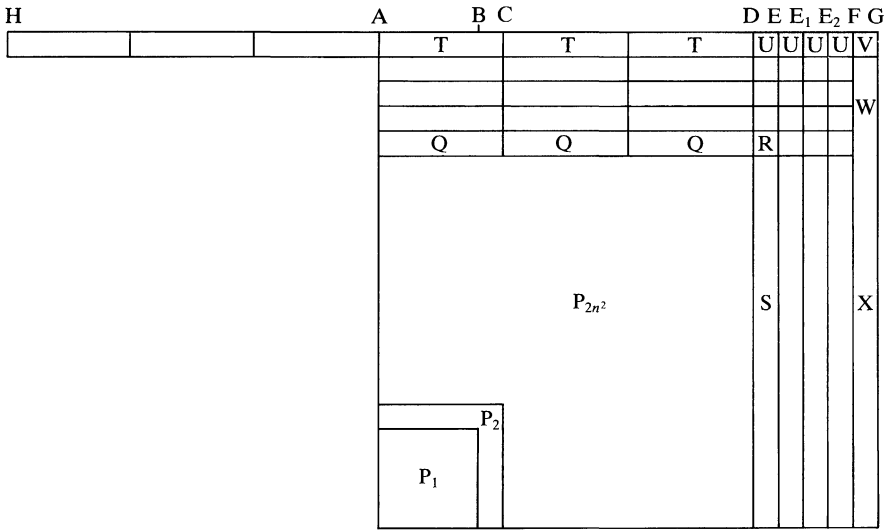
FIG. 3.19

PROPOSITION  $\sqrt{(2n^2 + n)} : \sqrt{2} = [n, \overline{4, 2n}]$.

When $n$ is even, this case reduces to the result

$$\sqrt{(4n^2 + n)} : 1 = [2n, \overline{4, 4n}]$$

which can be evaluated using the techniques already described. Therefore we need only consider the case of odd $n$, and Fig. 3.19 illustrates the case of $n = 3$, around which the reader is invited to construct a proof.

The only new idea of the construction of the figure for this proposition is to start with a square $P_2$ other than the given square $P_1$; in fact, ratios of the form $\sqrt{n} : \sqrt{m}$ are no more difficult to handle than ratios $\sqrt{n} : 1$. So the reader is invited to investigate, in either the Euclidean or a more general geometric interpretation, the result:

PROPOSITION  $\sqrt{(mn^2 + 2n)} : \sqrt{m} = [n, \overline{m, 2n}]$.

This shows that this method of gnomons can verify every expansion of the form $\sqrt{n} : \sqrt{m}$ which has one or two terms in its period. We now move on to the case of periods of length three, which must therefore have the form $[n, m, m, 2n]$ since the first two terms of the period are palindromic. The first such example of a ratio $\sqrt{n} : 1$ occurs with $n = 41$:

$$\sqrt{41} : 1 = [6, \overline{2, 2, 12}].$$

Since it is no more difficult to handle ratios of the form $\sqrt{n} : \sqrt{m}$, I shall illustrate the technique on a ratio involving smaller numbers.
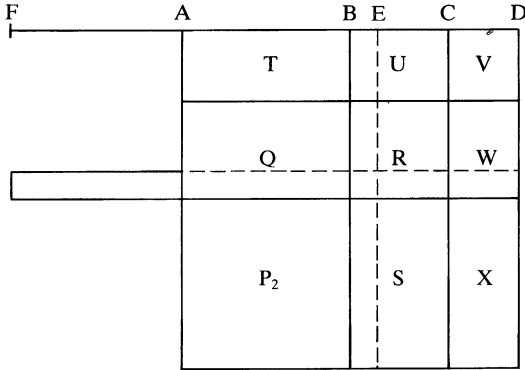
FIG. 3.20

PROPOSITION    $\sqrt{5}:\sqrt{2} = [1,\overline{1,1,2}]$.

PROOF    See Fig. 3.20: Start with a square $P_2 = 2P_1$; add a gnomon $Q + R + S$ of size $3P_1$. By the usual argument, this gnomon will have a width less than $\sqrt{2}$, so we can add a further gnomon $T + U + V + W + X$ to bring the total width of added gnomons up to $\sqrt{2}$. Adjoin the line $FA = AB$. Hence $AB = \sqrt{2}$; $AC = \sqrt{5}$; $FC = \sqrt{2} + \sqrt{5}$; and $AB = BD$. We can check, as usual, that $\frac{1}{2}AB < BC < AB$.

The first two steps of the anthyphairesis of $FC:AB = (\sqrt{2}+\sqrt{5}):\sqrt{2}$ are

$$FC - 2AB = BC, \quad \text{where} \quad BC < AB,$$

and, since $AB = BD$,

$$BD - 1BC = CD \quad \text{where} \quad \tfrac{1}{2}BC < CD < BC.$$

Let $CE = CD$; the third step is then

$$BC - 1CE = BE.$$

Hence

$$FC:AB = (\sqrt{2}+\sqrt{5}):\sqrt{2} = [2,1,1,CE:BE],$$

and, as usual, we must now prove that

$$FC:AB = CE:BE, \text{ i.e. } FC.BE = AB.CE$$

which is equivalent to

$$2(Q - T) + (R - U) = U + V,$$

i.e.

$$2Q + R = 2T + 2U + V.$$

But $2Q + R$ is the first gnomon, of size $3P_1$, and $2T + 2U + V$ is the second gnomon, also of size $3P_1$.                    QED

The new feature of this proof is the third step of the anthyphairesis, which introduces a complication in which the second gnomon folds back in a comparison with the first. The most important details of this new figure occur in the square on BD, in the top right-hand corner, in which "a straight line is bisected, and a straight line is added to it in a straight line" exactly as in II 10. In fact my Fig. 3.10(a), from which can be read off a proof of Euclid's Proposition 10 that conforms to the style of the proofs of II 4–8, is precisely the same as the top right-hand corner of Fig. 3.20, rotated through 180°.

Most of the examples in Table 3.1 of ratios which have three terms in their period are special cases of the general result:

PROPOSITION    $\sqrt{(2n^2 + 2n + 1)} : \sqrt{2} = [n, \overline{1, 1, 2n}],$

that can be established by the same technique. The remaining cases are

$$\sqrt{41} : 1 = [6, \overline{2, 2, 12}], \quad \text{and}$$

$$\sqrt{37} : \sqrt{2} = [4, \overline{3, 3, 8}],$$

which can also be verified in the same way. However it is not easy, from an arithmetical search, as from an extension of Table 3.1, to find general formulae to describe these cases. In fact, with a bit of algebra (which I do not propose was available in any form to the Greeks) we can now easily derive the following general result:

PROPOSITION    $\sqrt{(n^2 m^2 + n^2 + 2nm + 1)} : \sqrt{(m^2 + 1)} = [n, \overline{m, m, 2n}].$

It is possible to verify this for small values of $n$ and $m$ by the same method, and even to see that this verification is perfectly general.

The procedure can still be used for examples with four terms within the period, provided they involve only small numbers. For example, the table yields

$$\sqrt{7} : 1 = [2, \overline{1, 1, 1, 4}]$$

and a short further search reveals

$$\sqrt{8} : \sqrt{3} = [1, \overline{1, 1, 1, 2}],$$

both examples of the general result:

PROPOSITION    $\sqrt{(3n^2 + 4n + 1)} : \sqrt{3} = [n, \overline{1, 1, 1, 2n}].$

It is even just possible to verify the

PROPOSITION    $\sqrt{13} : 1 = [3, \overline{1, 1, 1, 1, 6}],$

which has five terms in its period, but anything that involves a longer period, or larger numbers within the period, becomes unfeasible. For example, the conjecture:

HYPOTHESIS    $\sqrt{19} : 1 = [4, \overline{2, 1, 3, 1, 2, 8}]$

generates a figure that seems to me to defy analysis by this method.

The new Appendix 10.3 takes up the story at this point.

## 3.5 THE SECOND ATTEMPT:
## SYNTHESISING RATIOS

### 3.5(a) *Introduction*

Since the method of gnomons has run into difficulties, we now try a different kind of attack on the problem: we change direction, and approach a solution from the other side. Hitherto we had started in each case from a hypothesis that

$$\sqrt{n} : \sqrt{m} = [n_0, \overline{n_1, n_2, \ldots, n_k, 2n_0}]$$

which was then actually handled in the equivalent form

$$(\sqrt{n} + n_0\sqrt{m}) : \sqrt{m} = [\overline{2n_0, n_1, \ldots, n_k}],$$

and we exploited in an essential way the feature that this anthyphairesis was now purely periodic. Let us now instead try to explore the questions: can we construct the ratio whose anthyphairesis is some purely periodic expansion $[\overline{n_0, n_1, \ldots, n_k}]$, and can we then show how, or explain why, if $n_0$ is even and $n_1, n_2, \ldots, n_k$ is palindromic, then this ratio will be of the form $(\sqrt{n} + \frac{1}{2}n_0\sqrt{m}) : \sqrt{m}$? The arguments that can be plausibly reconstructed within the Greek context will only give a very partial answer to these general problems; indeed the ultimate resolution of this question, to be described in Section 9.1(c), will require contributions from some of the most celebrated mathematicians of the seventeenth, eighteenth, and nineteenth centuries AD.

### 3.5(b) *The extreme and mean ratio*

These questions about purely periodic expansions have been phrased very generally in order to describe, at the outset, the approach and objective. Much more reasonable, as a reconstruction of ancient explorations, would be to study the particular ratios $[\overline{1}], [\overline{2}], \ldots, [\overline{1,2}], [\overline{1,3}], \ldots, [\overline{2,1}], [\overline{2,3}], \ldots, [\overline{1,2,3}], \ldots$.

The explicit surviving Greek evidence then relates only to the simplest case of $[\overline{1}]$: if a line AB is to be cut at the point C such that

$$AB : AC = [\overline{1}],$$

then

$$AB - 1AC = CB \quad \text{with} \quad CB < AC$$

and

$$AB : AC = AC : CB.$$

Hence, by the usual transformation of this proportion, the problem of finding C is precisely that solved by the construction of II 11:

To cut a straight line [AB] so that the rectangle contained by the whole and one of the segments [CB] is equal to the square on the remaining segment [AC].

Euclid's construction is illustrated in Fig. 3.11: ABDE is the square on AB, F bisects AE, FG = FB, and AGHC is the square on AG. The proof is a

straightforward verification, using II 6 applied to GAFE, and Pythagoras' theorem applied to the triangle ABF.

This construction is then invoked in IV 10:

To construct an isosceles triangle having each of the angles at the base double of the remaining one. Let any straight line AB be set out, and let it be cut at the point C so that the rectangle contained by AB, BC is equal to the square on CA; . . . . Therefore the isosceles triangle ABD [with AB = AD, and BD = AC] has been constructed having each of the angles at the base DB double of the remaining one,

and this particular triangle is then used in IV 11:

In a given circle to inscribe an equiangular and equilateral pentagon. . . . Let the isosceles triangle FGH be set out having each of the angles at G, H double the angle at F; . . .

Then the same construction, now described using proportion theory, is named at VI Definition 3:

A straight line is said to have been cut in extreme and mean ratio (*akron kai meson logon tetmēsthai*) when, as the whole line is to the greater segment, so is the greater to the less,

and its metrical properties are explored in a series of propositions at the beginning of Book XIII. Finally, the properties of the extreme and mean ratio and the pentagon are used in XIII 8, 9, 11, 16, and 17. The only other surviving references to extreme and mean ratio are found in Ptolemy's *Syntaxis*, Pappus' *Collection*, Hypsicles' 'Book XIV' of the *Elements*, and an anonymous scholium on Book II of the *Elements*; no other Greek text of any kind makes any explicit allusion to the construction.

### 3.5(c) *The nth order extreme and mean ratio*

Consider now the case of $[\bar{2}]$. The ratio $\sqrt{2}:1$, the diagonal to side of a square, is $[1,\bar{2}]$; hence, if ABDE is the square on AB, then $AB:(EB - AB) = [\bar{2}]$. We can construct the point $C_1$ on AB with $AC_1 = EB - AB$, as in Fig. 3.21, where $EG = EB$. If $C_1C_2 = AC_1$, then

$$AB:AC_1 = [2, AC_1:C_2B];$$

so the rectangle contained by AB and $C_2B$ is equal to the square on $AC_1$.

This construction is reminiscent of II 11, and it prompts us to explore the following generalisations:

DEFINITION    The point $C_1$ is said to divide AB in the noem ratio (read: $n$th order extreme and mean ratio) if, taking $n$ points $C_1, C_2, \ldots, C_{n-1}, C_n$ on AB such that $AC_1 = C_1C_2 = \cdots = C_{n-1}C_n$, then $C_n$ lies between A and B, and $AB . C_nB = AC_1^2$. (See Fig. 3.22.)

PROPOSITION    If $C_1$ cuts AB in the noem ratio, then $AB:AC_1 = [\bar{n}]$.

PROOF    $AB:AC_1 = [n, AC_1:C_nB]$, and $AB:AC_1 = AC_1:C_nB$.          QED
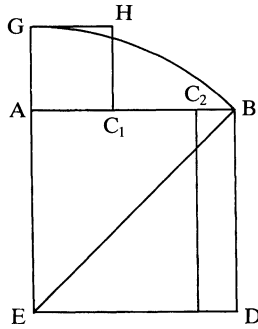
FIG. 3.21

The definition of the noem ratio implies that $C_nB$ is less than $AC_1$; it will be called the 'lesser segment' of the noem ratio. The 'greater segment' of the extreme and mean ratio generalises in two ways: to $AC_1$, which will be called the 'initial segment' of the noem ratio; and to $AC_n$, which will again be called the 'greater segment'. Care must sometimes be exercised in making the appropriate choice. I shall describe two typical results about the 3rd order ratio, but will indicate the generality of these kinds of procedures by labelling the division points $C_1$, $C_{n-1}$, and $C_n$, thus following a compromise path between the different Euclidean and modern styles of handling general statements.

CONSTRUCTION    To divide a line AB at $C_1$ in the noem ratio, construct the square ABDE (see Fig. 3.23) and on AE, produced as necessary, take points $F_1, F_2(=E), F_3, F_4, \ldots$, with $AF_1 = F_1F_2 = F_2F_3, \ldots$; then these points will be used in the construction of the 1st, 2nd, 3rd, 4th, . . . order extreme and mean ratios. To construct the $n$th order ratio, take G on EA produced with $F_nG = F_nB$, and $C_1$ on AB with $AC_1 = AG$. Let $AC_1 = C_1C_2 = \cdots = C_{n-1}C_n$; we must prove that $C_n$ lies between A and B and that $AB . C_nB = AC_1^2$. Now

$$F_nG^2 = (AF_n + AG)^2 = (AF_n + AC_1)^2$$

$$= AF_n^2 + AC_1^2 + 2AF_n . AC_1, \quad \text{(by II 4)}$$

and

$$2AF_n . AC_1 = nAE . AC_1 = AB . AC_n.$$

But

$$F_nG^2 = F_nB^2 = AF_n^2 + AB^2. \quad \text{(by I 48 or II 8a)}$$
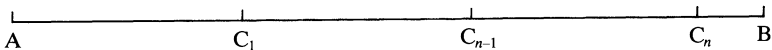


FIG. 3.22
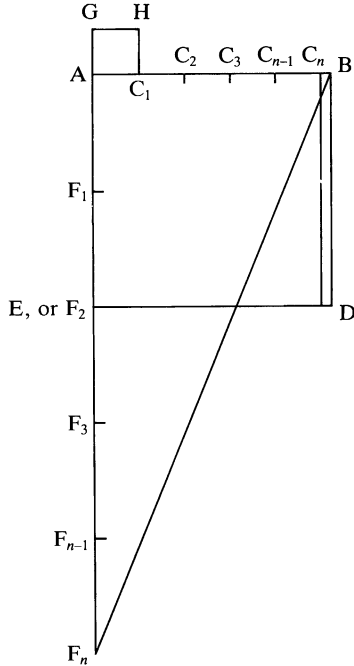
FIG. 3.23

Hence

$$AC_1^2 + AB . AC_n = AB^2.$$

Therefore $AC_n < AB$ and, subtracting $AB . AC_n$ from both sides, we see that $AC_1^2 = AB . C_nB$. 　　　　　　　　　　　　　　　　　　　　　　　　QED

The case where $n$ is even, $n = 2m$, is connected with the dimension of squares: here $AF_n = mAB$, an integral multiple of $AB$, so $F_nB^2 = (m^2 + 1)AB$, and $AC_1 = \sqrt{(m^2 + 1)}AB - mAB$. If $AB$ is the assigned line, then

$$AB : AC_1 = 1 : (\sqrt{(m^2 + 1)} - m) = [\overline{2m}].$$

Thus we have again arrived, by a circuitous route, at the result that

$$\sqrt{(m^2 + 1)} : 1 = [m, 1 : (\sqrt{(m^2 + 1)} - m)] = [m, \overline{2m}].$$

### 3.5(d) *Elements* XIII, 1–5

The metrical properties of the extreme and mean ratio are considered in XIII 1–5, a series of propositions reminiscent of the style and techniques of Book II. All of these propositions can be easily generalised to the noem ratio; here, as an illustration, is XIII 1 followed by its generalisation:

If a straight line be cut in extreme and mean ratio, the square on the greater segment added to half of the whole is five-times the square on the half. °

Euclid's proof is based on Fig. 3.24; the lettering of the vertices is his but the labelling of the regions has been added, for our convenience. He proceeds as follows. If AB is cut in extreme and mean ratio at C, and $DA = \frac{1}{2}AB$, we must prove $DC^2 = 5AD^2$. Draw the squares on DC and AB and complete the figure as shown. We then know that

$$AB \cdot CB = AC^2, \quad \text{i.e.} \quad P = Q,$$

and that

$$AB \cdot AC = 2AD \cdot AC, \quad \text{i.e.} \quad R = S_1 + S_2.$$

Hence

$$P + R = Q + S_1 + S_2.$$

Add $AD^2 = T$, and assemble the result into squares:

$$DC^2 = AB^2 + AD^2.$$

But $AB^2 = 4AD^2$, so $DC^2 = 5AD^2$.           QED

We tend, today, to refer to the extreme and mean ratio by its modern name, the 'golden section', and to arithmetise and identify it with the 'golden number', $\frac{1}{2}(\sqrt{5} + 1)$. The proposition expresses this result in the equivalent Greek idiom: if $AB = 2$, twice some assigned line, then $CD = \sqrt{5}$, the side of a five-fold square, so we have proved that

$$2 : (\sqrt{5} - 1) = [\overline{1}].$$

Note that this is not, nor can it be converted into, a ratio of two sides of integral squares.

The generalisation of this proposition explains the role of the number 5:
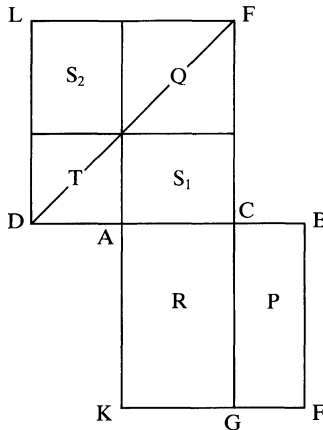


FIG. 3.24

PROPOSITION    If a straight line be cut in the noem ratio, the square on the initial segment added to $n$ times the half of the whole is $(n^2 + 4)$ times the square on the half.

PROOF    If AB is cut in noem ratio at $C_1$, and $D_n A = \frac{n}{2} AB$, we must prove that $D_n C_1^2 = (n^2 + 4) AD_1^2$ (see Fig. 3. 25). Draw the squares on $D_n C_1$ and AB, and complete the figure as shown. We know that

$$AB \cdot C_n B = AC_1^2, \quad \text{i.e.} \quad P = Q$$

and that

$$AB \cdot AC_1 = 2AD_1 \cdot AC_1, \quad \text{i.e.} \quad R = 2S.$$

Hence

$$P + nR = Q + 2nS.$$

Add $AD_n^2 = n^2 T$, and assemble the result into squares:

$$D_n C_1^2 = AB^2 + n^2 T$$
$$= (n^2 + 4)T \quad (\text{since } AB^2 = 4T)$$
$$= (n^2 + 4)AD_1^2. \qquad\qquad \text{QED}$$

Propositions 2 to 5 of Book XIII can easily be generalised in this way; complete details are given in my article GGS.

### 3.5(e) *Further generalisations*

Now consider the problem of synthesising a ratio containing two terms in its period. In advance, I must say that I feel that this will be implausible as a reconstruction of a Greek procedure.
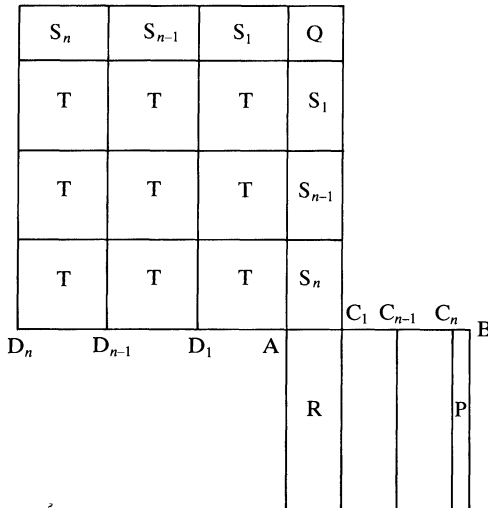


FIG. 3.25

PROPOSITION   If
$$a_0 - na_1 = a_2$$
and
$$a_1 - ma_2 = a_3$$
are the first two steps of the anthyphairesis of the ratio $a_0 : a_1$ of two lines, and $a_0 : a_1 = [\overline{n,m}]$, then
$$ma_0 . a_2 = na_1^2.$$

PROOF   Since $[\overline{n,m}]$ is an anthyphairetic ratio, $n \geqslant 1$, $m \geqslant 1$, and $a_0 > a_1$. We have
$$a_0 : a_1 = [n, m, a_2 : a_3]$$
and hence, since the anthyphairesis is periodic,
$$a_2 : a_3 = a_0 : a_1 \quad \text{so} \quad a_0 . a_3 = a_1 . a_2.$$
Hence
$$a_0 . a_1 = a_0 . (ma_2 + a_3)$$
$$= ma_0 . a_2 + a_0 . a_3$$
$$= ma_0 . a_2 + a_1 . a_2,$$
and so
$$ma_0 . a_2 = a_0 . a_1 - a_1 . a_2$$
$$= a_1 . (a_0 - a_2)$$
$$= na_1^2. \qquad\qquad \text{QED}$$

Note that, although this proposition is expressed symbolically, everything can be interpreted as a shorthand for the geometrical manipulations of rectangles and squares. However, the algebraic content and the amount of autonomous identity that is attributed to the symbols and their manipulations are much greater here than anything we have so far encountered.

We are really more concerned with the converse result:

PROPOSITION   If $a_0 - na_1 = a_2$, where $a_0 > a_1 > a_2$ and $ma_0 . a_2 = na_1^2$, then $a_0 : a_1 = [\overline{n,m}]$.

PROOF   Write $a_1 - ma_2 = a_3$; we want to prove that $a_3 < a_2$ and that $a_2 : a_3 = a_0 : a_1$, i.e. that $a_0 . a_3 = a_1 . a_2$. Now
$$a_0 . a_3 = a_0 . (a_1 - ma_2)$$
$$= a_0 . a_1 - ma_0 . a_2$$
$$= a_0 . a_1 - na_1^2$$
$$= a_1 . (a_0 - na_1) = a_1 . a_2.$$

Also $a_0 . a_3 = a_1 . a_2 < a_0 . a_2$, and so $a_3 < a_2$. $\qquad\qquad$ QED

CONSTRUCTION    To divide AB at $C_1$ so that $AB:AC_1 = \overline{[n,m]}$ (see Fig. 3.26, which illustrates the case of $\overline{[4,3]}$), first carry out a preliminary construction to locate $B'$ on AB with $nAB'^2 = mAB^2$; this can be done as follows, using II 14. Draw the square ABDE and, on EA produced, take J with $nAJ = mAB$. On EJ as diameter, draw a semicircle meeting AB at $B'$; then $AJ.AE = AB'^2$, so $nAB'^2 = nAJ.AE = mAB^2$. Next, on AE, produced if necessary, take the point $F_m$ with $AF_m = \frac{m}{2}AB$; on EA produced, take G with $F_mG = F_mB'$; and $C_1$ on AB with $AC_1 = AG$. Take $C_2,\ldots,C_n$ with $AC_1 = C_1C_2 = \cdots = C_{n-1}C_n$; we shall prove that $mAB.C_nB = nAC_1^2$. This verification follows the same pattern as the earlier constructions:

$$F_mG^2 = (AF_m + AG)^2 = (AF_m + AC_1)^2$$

$$= AF_m^2 + AC_1^2 + 2AF_m.AC_1$$

$$= AF_m^2 + AC_1^2 + mAB.AC_1.$$

But

$$F_mG^2 = F_mB'^2 = AF_m^2 + AB'^2;$$

hence

$$AC_1^2 + mAB.AC_1 = AB'^2.$$
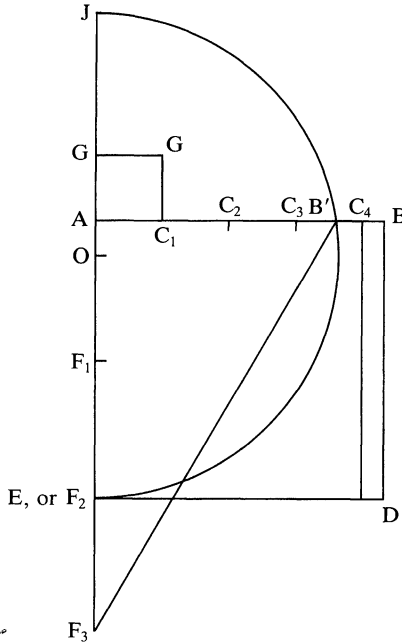


FIG. 3.26

Now multiply by $n$, and write $nAB'^2 = mAB^2$ and $nAC_1 = AC_n$; then

$$nAC_1^2 + mAB \cdot AC_n = mAB^2.$$

Hence

$$nAC_1^2 = mAB^2 - mnAB \cdot AC_1$$
$$= mAB \cdot (AB - nAC_1)$$
$$= mAB \cdot C_nB,$$

and so $AB:AC_1 = [\overline{n,m}]$.      QED

If we take $AB$ as the unit, we can evaluate $AC_1$ as follows:

$$AC_1 = GF_m - AF_m$$
$$= \sqrt{(AF_m^2 + AB'^2)} - AF_m$$
$$= \sqrt{\left(\frac{m^2}{4} + \frac{m}{n}\right)} - \frac{m}{2}.$$

With this evaluation, we can now synthesise the general ratio having the form $[p,\overline{q,2p}]$. We start by writing $q$ for $n$ and $2p$ for $m$, to get

$$[\overline{q,2p}] = 1 : \left(\sqrt{\left(p^2 + \frac{2p}{q}\right)} - p\right),$$

and hence

$$[p,\overline{q,2p}] = \sqrt{\left(p^2 + \frac{2p}{q}\right)} : 1$$
$$= \sqrt{(p^2q + 2p)} : \sqrt{p}.$$

Thus we have proved that every ratio of the form $[p,\overline{q,2p}]$ is indeed a ratio of sides of squares.

It is possible to generalise this construction to the problem of cutting a line $AB$ at a point $C$ such that $AB:AC$ has an arbitrary periodic expansion. We start again, with a similar result:

PROPOSITION    If $AB:AC = [\overline{n_1, n_2, \dots, n_k}]$, then there exist numbers $p$, $q$, and $r$, determined by $n_1, n_2, \dots, n_k$, such that

$$pAB^2 = qAB \cdot AC + rAC^2,$$

and conversely.

Then a construction similar to that of Fig. 3.26, in which $rAB'^2 = pAB^2$ and $AF_3 = \frac{p}{r}AB$, will yield the required division point. But the evaluation of $p$ and $q$ requires so much detailed symbolic manipulation that it is quite unthinkable as a reconstruction of an ancient procedure. More details of this construction, which is due to Christopher Zeeman, are given in my BTEE.

## 3.6 THE THIRD ATTEMPT: GENERALISED SIDES AND DIAGONALS

### 3.6(a) *The method*

We now take the procedure of sides and diagonals that was used in Section 2.2(a) to compute the ratio $\sqrt{2}:1$ and generalise it so that it will apply to a ratio $\sqrt{n}:\sqrt{m}$ of any sides of squares. I shall discuss in more detail the historical evidence relating to the procedure in the next section; first, I describe its operation.

Behind the construction of Fig. 2.1(a), repeated here as Fig. 3.27 (with a slight addition that makes explicit how the side of the larger square is constructed out of the diagonal plus side of the smaller), there lay a relation

$$S = s + d$$
$$D = 2s + d.$$

There is no problem in drawing a square corresponding to a prescription $S = as + bd$, for any values of $a$ and $b$, and deriving the corresponding relation $D = 2bs + ad$; but only very special values of $a$ and $b$ will lead to the recursive behaviour which will confirm the periodicity of $D:S = \sqrt{2}:1 = [1,\overline{2}]$ (or, equivalently and more conveniently, of $(S + D):S = (1 + \sqrt{2}):1 = [\overline{2}]$); the reader is encouraged to draw the figure for $a = 2$, $b = 3$, and then try to use this to evaluate $D:S$, to see how the periodic behaviour fails to manifest itself, in contrast with the case of $a = 3$, $b = 2$. The historical evidence, plus some experience with anthyphairetic ratios, suggests that we look at the behaviour of the convergents to identify what is so special about the successful side and diagonal relations.

The anthyphairesis and convergents can always be calculated using the algorithms of Section 2.3, though I shall now, for convenience, transpose the second and third lines of the standard format of Table 2.4 and will refer to the convergents as $d_k:s_k$ (for diagonal and side, of course) rather than $p_k:q_k$. For the example of $\sqrt{2}:1$, we have:

$$
\begin{array}{llllll}
\sqrt{2} = [1 & 2 & 2 & 2 & 2 & \ldots] \\
s_k \quad 1 & 2 & 5 & 12 & 29 \\
d_k \quad 1 & 3 & 7 & 17 & 41
\end{array}
$$
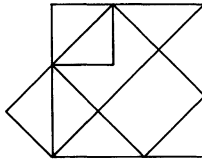


FIG. 3.27

The algorithm of 2.3(c) described how to generate the successive convergents by the relations

$$s_{k+1} = n_{k+1}s_k + s_{k-1}.$$

$$d_{k+1} = n_{k+1}d_k + d_{k-1}.$$

The side and diagonal *numbers* described by Theon (see Section 2.4(e)) and Proclus (see below) appear to be the different relations that, for the case of $\sqrt{2} : \sqrt{1}$,

$$s_{k+1} = s_k + d_k$$

$$d_{k+1} = 2s_k + d_k$$

with

$$s_0 = d_0 = 1;$$

these also seem to generate the convergents. It is a plausible assertion that these two relations are equivalent; however, we do not attempt to prove this here, but immediately convert the arithmetical side and diagonal relation for $\sqrt{2} : 1$ into a geometrical prescription for manipulating squares, which is then used to construct the particular figure which embodies the proof of our hypothesis about the periodicity of the anthyphairesis of $D : S$. The first stage of heuristic investigation of the convergents is then again left behind, incomplete but superseded.

I encourage the reader to explore further how this new kind of relation is apparently satisfied by the convergents of any ratio, by looking at the sequence of convergents to ratios whose anthyphaireses have successively longer and longer periods which (for convenience only) contain only small terms: $\sqrt{5} : 1$, $\sqrt{3} : 1$, $\sqrt{5} : \sqrt{2}$, .... Here I shall present some of the conclusions of such an exploration, illustrated in the case of $\sqrt{7} : 1$.

The algorithm of 2.3(b) applied to $\sqrt{7} : 1$ gives, after fourteen lines of calculation:

| $\sqrt{7} =$ | [2 | 1 | 1 | 1 | 4 | 1 | 1 | 1 | 4 | ...] |
|---|---|---|---|---|---|---|---|---|---|---|
| $s_k$ | 1 | 1 | 2 | 3 | 14 | 17 | 31 | 48 | 372 | |
| $d_k$ | 2 | 3 | 5 | 8 | 37 | 45 | 82 | 127 | 717 | |

and so the main hypothesis we wish to prove is that $\sqrt{7} : 1 = [2, \overline{1, 1, 1, 4}]$. Since the period is four terms long, we are led to seek a side and diagonal relation of the form

$$s_{k+4} = as_k + bd_k$$

$$d_{k+4} = 7bs_k + ad_k,$$

and further exploration suggests that the appropriate values of $a$ and $b$ always seem to be given by the convergent just before the occurrence of the end-of-

period term '$2n_0$'; here the end-of-period term appears to be $n_4 = 4 = 2n_0$, which suggests that $a = 8$ and $b = 3$. Hence we guess that

$$s_{k+4} = 8s_k + 3d_k$$

$$d_{k+4} = 21s_k + 8d_k.$$

(This 'penultimate' convergent $8:3$ has a very special relation to the ratio $\sqrt{7}:1$, some aspects of which have already been described in Section 2.4(e). The side and diagonal relation will thus always be built around solutions of Pell's equation; here $8^2 - 7.3^2 = 1$.)

We now convert this into a prescription for a geometrical construction:

$$S = 8s + 3d$$

$$D = 21s + 8d,$$

which we apply to an appropriate figure built out of the lines 1 and $\sqrt{7}$. A convenient choice, which can be generalised to any ratio of sides of squares $\sqrt{n}:\sqrt{m}$, is a small parallelogram with one side $s$ equal to the assigned line (or, in general, to $\sqrt{m}$), one diagonal $d$ equal to $\sqrt{7}$ (or, in general, to $\sqrt{n}$), and the other side an integral multiple $ps$ of $s$. The basic inequalities for the sides of the triangle (*Elements* I 20) tell us that $p$ must satisfy $d - s \leqslant ps \leqslant d + s$, and hence, when $d = \sqrt{7}$ and $s = 1$, $p = 2$ or 3. Let us choose $p = 2$. So we take the parallelogram with sides $s = 1$ and $2s = 2$, and (longer) diagonal $d = \sqrt{7}$, and apply the prescription to construct a larger similar parallelogram with shorter side $S = 8s + 3d$, and longer side $2S = 16s + 6d$. Figure 3.28 illustrates this case of $n = 7$, $m = 1$, $p = 2$, $a = 8$, and $b = 3$, and once its complexities have been explained it can be seen to be a straightforward generalisation of Fig. 3.27 with $S = as + bd$.
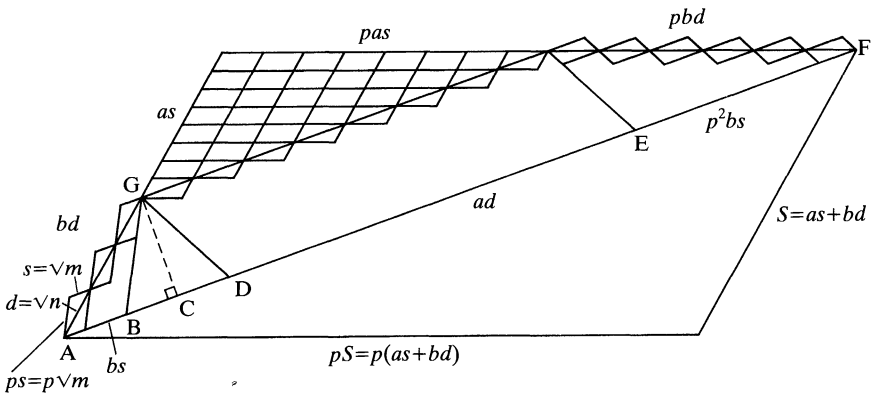


FIG. 3.28

We now evaluate the longer diagonal $D$ of this larger parallelogram and verify that $D = nbs + ad$. But, from the figure, we see that

$$AB = bs, \qquad DE = ad, \qquad EF = p^2 bs,$$

so it remains only to calculate BD. We invoke *Elements* II 12:

In obtuse-angled triangles the square on the side subtending the obtuse angle is greater than the squares on the sides containing the obtuse angle by twice the rectangle contained by one of the sides about the obtuse angle, namely that on which the perpendicular falls, and the straight line cut off outside by the perpendicular towards the obtuse angle,

and apply it to ABG. Note that this triangle is obtuse-angled at B (since $d$ is the longer diagonal of the parallelogram) and BD is equal to "twice [BC], ..., the straight line cut off outside [the triangle ABG] by the perpendicular [GC] towards the obtuse angle [at B]". Hence

$$b^2 d^2 = b^2 s^2 + p^2 b^2 s^2 + bs \cdot \mathrm{BD},$$

and so, since $d^2 = ns^2$, we get

$$\mathrm{BD} = (n - p^2 - 1)bs.$$

(The derivation of this last statement should be interpreted in terms of the converse of II 1.) We now add AB, BD, DE, and EF, which gives

$$D = nbs + ad,$$

and the geometric prescription is confirmed. A different kind of figure, which generates an acute-angled triangle and which the reader is encouraged to explore, will occur when $d$ is the shorter diagonal of the parallelogram; in this case the verification requires *Elements* II 13.

Note that this construction can be carried out for any values of $a$ and $b$. However, the next and final step depends in a subtle way on the underlying relation $a^2 - nb^2 = \pm 1$, but the general elucidation of this kind of result will not be achieved before the eighteenth and nineteenth centuries. So here we must leave the general case $\sqrt{n} : \sqrt{m}$ and revert to the verification of the particular case of $\sqrt{7} : 1$ for which $S = 8s + 3d$ and $D = 21s + 8d$. We evaluate $(2S + D) : S$, as follows:

$$\begin{aligned}
(2S + D) : S &= [4, S : (D - 2S)] \\
&= [4, (8s + 3d) : (5s + 2d)] \\
&= [4, 1, (5s + 2d) : (3s + d)] \\
&= [4, 1, 1, (3s + d) : (2s + d)] \\
&= [4, 1, 1, 1, (2s + d) : s] \\
&= [4, 1, 1, 1, (2S + D) : S] \\
&= [\overline{4, 1, 1, 1}].
\end{aligned}$$

Hence $\sqrt{7} : 1 = [2, \overline{1, 1, 1, 4}]$.　　　　　　　　　　　　　　　　QED

This method will appear to verify any example of a ratio of sides of squares, no matter how long its period may be and, in this sense, the problem is solved. However, I cannot conceive of a general proof which could be formulated within the techniques of Greek mathematics that this procedure will always work, no matter what values of $n$ and $m$ are chosen. Nor, as I observed earlier, do I know of any evidence that Greek mathematicians observed or explored the palindromic behaviour of these anthyphairetic ratios $\sqrt{n} : \sqrt{m}$, nor can I conceive of how Greek techniques could elucidate this result, even in any particular case. The eventual explanation of this general behaviour will take us into the nineteenth century, and will combine contributions by, among others, Fermat, Brouncker, Wallis, Euler, Lagrange, Legendre, and Galois. Details and references will be given in Section 9.1(c).

### 3.6(b) *Historical observations*

It is, at the very least, a noteworthy coincidence that the programme of the verification of anthyphairesis of sides of squares can be completed by a procedure whose only significant mathematical prerequisites, other than some very basic results of Book I of the *Elements*, are those two slightly incongruous Propositions 12 and 13 of Book II that were left over from our earlier partially successful assaults on the same problem. This correlation between mathematical problems and ancient evidence could immediately be carried further if I were to proceed, as I now could, to consideration of Book X, because just as the material of Book II seems to be connected with and preparatory for Book X (though this in a way that is difficult to accommodate in detail within many interpretations of Book II) so does this anthyphairetic interpretation of Book II lead into an anthyphairetic motivation of Book X. But to embark straightaway on this arduous topic would be to tax cruelly the stamina of even my most sympathetic reader, so I shall leave this next step until later (see Chapter 5) and will, in the next chapter, turn to a discussion of Plato's mathematics within this interpretation.

Consider, now, the historical evidence concerning this procedure of 'sides and diagonals'. In the previous chapter (at 2.4(e)), I quoted, in its entirety, what Theon of Smyrna writes about side and diagonal numbers. Our other sources of information on the topic are Proclus' commentary on Plato's *Republic* (Proclus–Kroll, *IR* ii, 24–9), written in the fifth century AD, in which Proclus makes a wide variety of observations concerning Plato's mysterious nuptial number, *Republic* VIII, 546b–d, and Iamblichus-Pistelli, *NAI*, 91.3–93.6, which does not add anything further, so which I shall leave to one side here. Proclus' discussion of side and diagonal numbers starts at ii 24.16:

The Pythagoreans demonstrate by numbers (*dia tōn arithmōn*) that the expressible (*rhētos*) [squares] constructed on the inexpressible (*arrhētos*) diagonals are greater or less than double by a unit. For since the unit is, generatively, everything, it is clear, they say, that it is both side and diagonal. So let there be two units, the one as side, the other as diagonal, . . .

and the text continues in a very similar vein to Theon's description, with the addition only of a customary reference to Plato:

Whence Plato said that the number forty-eight is a square of the expressible diagonal of five minus one, and of the inexpressible minus two, since the square of the diagonal is double that of the side [i.e. $48 = (\sqrt{(50-1)})^2 - 1 = (\sqrt{50})^2 - 2$ where $(\sqrt{50})^2 = 2 \times 5^2$].

It seems apparent that both passages derive ultimately from a common archetype. I leave a detailed analysis of this textual question to others more competent than me (I know of no published discussion of this issue), and start with the following three observations:

(i) The information that this construction derives from the Pythagoreans is found only in Proclus, not in Theon, though Theon does elsewhere frequently mention the Pythagoreans. Therefore, since both authors seem to be drawing on a common source, it is not improbable that this is an addition, either by Proclus himself or by some intermediate source of his text.

(ii) Neither Theon nor Proclus actually demonstrates the arithmetical result that $d_k - 2s_k^2 = \pm 1$ in any sense beyond verifying that it seems to hold for the first few cases.

(iii) The mathematical implications of the words (*ar*)*rhētos* are not entirely clear. (I shall return to this question in Section 5.2(a).)

All of these points are further illustrated by the sentences with which this first passage from Proclus ends (ii 25.9–13):

And if we take all the [squares] on the diagonals of this kind, of which each is greater or less than double by a unit, they will be really double: for instance nine and 49 [is double] of 25 and 4. This is why the Pythagoreans too had confidence in this method.

After this inconsequential description of side and diagonal numbers, and after some obscure and sometimes corrupt passages in which Proclus discusses various apparently unrelated views of Dercyllides, Nicolas, Magnus, and Pythagoras, Proclus returns to considerations of sides and diagonals (*IR* ii, 27.1–28.12):

Since it is impossible that the diagonal be expressible when the side is expressible (since there does not exist a square number double of a square number, from which it is also clear that there are incommensurable (*asummetros*) magnitudes, and that Epicurus was wrong to make the atom a measure of every body, and Xenocrates was wrong to make the indivisible line the measure of lines) the Pythagoreans and Plato thought to say that, the side being expressible, the diagonal is not absolutely (*haplōs*) expressible, but, in the squares whose sides they are, [the square of the diagonal] is either less by a unit or more by a unit than the double ratio which the diagonal ought to make: more, as for instance is 9 than 4, less as is for instance 49 than 25. The Pythagoreans put forward the following kind of elegant theorem of this, about the diagonals and sides, that when the diagonal receives the side of which it is diagonal it becomes a side, while the side, added to itself and receiving in addition its own diagonal, becomes a diagonal. And this is demonstrated by lines (*grammikōs*) through the things in the second [book] of *Elements* by him. If a straight line be bisected and a straight line be added to it, the square on the whole line with the added straight line and the square on the latter by itself are together double

the square on the half and of the square on the straight line made up of the half and the added straight line. Let $AB$ be a side and let $B\Gamma$ be equal to it, and let $\Gamma\Delta$ be the diagonal of $AB$, having a square double that of it [i.e. $AB$]; by the theorem, the square on $A\Delta$ with that on $\Delta G$, will be double that on $AB$ and on $B\Delta$. Of these, the square on $\Delta\Gamma$ is double that on $AB$; and it remains that the square on $A\Delta$ is double that on $B\Delta$, for if as whole is to whole, so is what is taken away to what is taken away, the remainder, also, will be to the remainder as the whole is to the whole. Then the diagonal $\Gamma\Delta$, receiving in addition the side $B\Gamma$, is a side; and $AB$, taking in addition itself, [i.e.] the [side] $B\Gamma$, and its own diagonal $\Gamma\Delta$, is a diagonal; for it has a square double that of the side [$\Delta B$]. These things, then, this way. Let us now prove by numbers (*arithmētikōs*) the result on the expressible diagonals which are, as we said earlier, by a unit greater or less. Let there be a unit and let there be another unit around it; ... ;

and with this Proclus reverts to another inconsequential description of side and diagonal numbers not dissimilar to Theon's and his own earlier description at ii 24.11–25.6, but carried now up to a side of 12 and diagonal of 17. Let me continue the list of observations on these passages:

(iv) Proclus, only, gives a description of an "elegant theorem" about side and diagonal *lines*.

(v) This elegant theorem is said to be "demonstrated by lines through the things in the second book of *Elements*"—surely Euclid's *Elements*. (Alternatively, the end of this phrase can be construed to mean "through the elements in the second book".)

(vi) There is no proof of, or reference to, this elegant theorem in *Elements* II as we have it.

(vii) Proclus cites the enunciation of *Elements* II 10 with slight, but surely unimportant, variations in the wording, and he uses this to verify the result.

(viii) The manuscript of Proclus' commentary contains no figure but it seems certain that the configuration to which his text refers is that provided by the editor Kroll in his discussion of the passage (at ii 393–400), and given here as Fig. 3.29, through Kroll and most modern commentators continue the diagram further to the right.

(ix) Proclus makes no further use of this elegant geometrical theorem but reverts abruptly, at the end, to a further inconsequential discussion of side and diagonal numbers.
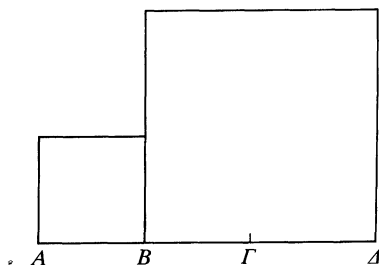


Fig. 3.29

(x) Anybody seeking a connection between the elegant theorem and *Elements* II, as we now have it, would be led to Propositions 6, 8 (perhaps), and 10, since all of these propositions deal, more or less explicitly, with one line added to another (like $s + d$), then with the line added to itself and added to the other (like $2s + d$); and, of these, Propositions 6 and 8 involve rectangles, while 10 involves only squares. Thus the elegant theorem seems, in the allusive style in which such references are made in Greek mathematics, to be referring to II 10, and this proposition does indeed provide a quick proof; but it leads no further.

(xi) Proclus' completion of the proof by invoking a general result on proportions, in fact *Elements* V 19 or VII 11, is unnecessarily and irrelevantly heavy-handed; all that is needed is Common Notion 3.

To these, let me add the subjective opinion:

(xii) Proclus' proof of the elegant theorem is pedestrian, unilluminating, and inconsequential, in contrast to the perspicuous proof based on Fig. 3.27.

Now let me fit these observations together. Proclus and Theon are drawing on some earlier source, now lost to us. Either Theon's source is already less complete than Proclus', or Theon only abstracts a portion of it; in either case, having made observation (i) above, we can set Theon's testimony aside, as being contained within Proclus' text. Next, from (ii), we can infer that this earlier source did not contain a complete proof of the arithmetical results, though such is easy to provide since it follows from any acceptable description of the manipulation

$$d_{k+1}^2 - 2s_{k+1}^2 = (2s_k + d_k)^2 - 2(s_k + d_k)^2 = -(d_k^2 - 2s_k^2)$$
$$= \pm(d_1^2 - 2s_1^2) = \pm 1.$$

Hence, we infer, this source may not have been primarily concerned with proving this relation, but with illustrating or investigating the connection between its essential ingredients: both side and diagonal equal to one as starting values; the recurrence relation for $s_k$ and $d_k$; and the relation $d_k^2 - 2s_k^2 = \pm 1$. This opinion is reinforced by the way the same discussion is repeated, each time at length. (Note also how, in my proposed reconstruction, it is not the proofs of these kinds of relations that is pertinent, but their association with a geometrical figure, and thence to the anthyphairesis of some ratio. So it would be more important to explore and understand these relations than to prove one particular example.) Observation (iv) was included to emphasise a feature that often passes without comment: that there are two distinct relations, one involving side and diagonal *lines*, the other side and diagonal *numbers*, and these are often conflated in discussions of the topic (see, for example, Heath, *TBEE* i, 398 ff.). Observations (v) and (vi) are compatible with a very wide range of explanations concerning the form and role of *Elements* II or its archetype, too wide to be described and analysed here; but observations (vii) to (xi) seem to indicate that the citation of II 10, along with the figure (see (viii)), and the details of the proof, was a later addition to the original text, possibly by Proclus or his source.

This leaves (xii). Underlying this observation is an inconsistency within Proclus' testimony, that the theorem as he describes and demonstrates it is *not* elegant. This is my subjective judgement but, as an illustration of how it does seem to be more widely held, see how the alternative proof by Fig. 3.27 is described by Heath, *TBEE* i, 400–1, as "an ingenious suggestion", and by Burkert, *LSAP*, 430, n. 16 (in a footnote which contains the only geometrical figure in this long book) as "still more perspicuous". Just as Meno's slaveboy must have marvelled at Socrates' ability to draw the figure which led to the solution of the problem of doubling a square (*Meno* 84d–e), so we can still marvel at the construction of Fig. 3.27, and the way this diagram makes evident the anthyphairesis of $\sqrt{2} : 1$. Moreover, Fig. 3.27 fits well in the style of attested fourth-century Greek mathematics, since it bears a striking resemblance to the successful *Meno* figure, as well as to the alternative reconstructed proof of Pythagoras' theorem II 8a (see Fig. 1.5 and its associated text, and Section 3.2, above); all consist of obliquely placed squares or rectangles within larger squares. And further, once this figure has been drawn, it is not too difficult to see some of the remarkable ways in which this simple construction can be exploited and generalised within the context of anthyphairetic mathematics. These qualities of simplicity, surprise, power, and generality are, today, some of the features that lead mathematicians to use words like 'elegant', and I can conceive of it being so also in antiquity. But, removed from its fertile context, for example as we find it in Proclus, the theorem scarcely seems to merit the description.

I suggest therefore that Proclus' description of a figure in which "when the diagonal receives the side of which it is diagonal it becomes a side" may indeed originally have been a reference to the construction of Fig. 3.27, and that this construction was investigated in the context of its generalisation to diagrams like Fig. 3.28 with a view to proving the observed periodicity in the problem of the dimension of squares. Then, as the original anthyphairetic content was excised from the material to be included in Euclid's *Elements*, all that remained of this procedure were *Elements* II, Propositions 12 and 13, and a haunting echo of an elegant theorem about diagonals and sides.

Finally, it is worth returning to the opening paragraph of this Section 3.6(b) and remarking that the detailed evaluations in Saito, IPPC and its extensions, have not yet thrown up any uses of II Propositions 12 and 13.

## 3.7 SUMMARY

My proposal is that Book II of the *Elements* may be a collection of figures and results whose connecting feature is that they arose in three different attempts to verify and explore the periodicity of the anthyphairetic ratios of incommensurable sides of squares, $\sqrt{n} : \sqrt{m}$. Ultimately, the third and most involved method succeeded in establishing a proof of periodicity in any given case, and each method shed some light on the phenomena involved.

## 3.8 NOTES AND REFERENCES

3.2     The most common interpretation of Book II is in terms of what is now called 'geometrical algebra', a name that was apparently coined by Tannery and Zeuthen for a much older point of view. This kind of interpretation has recently been the focus of vigorous argument; for a general discussion see Mueller, *PMDSEE*, 41–53 and 301–3, and Berggren, HGM, 397–8. My proposals here provide an entirely different reconstruction of the antecedents and original context of the book. For an illuminating and informed discussion of later uses of the propositions of Book II, see Saito, BTEE, and his catalogue in IPPC.

3.3     The hypotheses about the periodic palindromic behaviour of $\sqrt{n} : \sqrt{m}$ will be proved in Section 9.1(c). Lack of space has precluded any discussion of the Greek evidence on palindromes; I recommend the curious reader to start with Horsfall, SB, and the references contained therein. I cannot see any evident connection between this material and mathematics. I do not know of any palindromes in Plato's writings or other fourth-century texts beyond isolated words like αρα, coφoc, etc.

3.4     There is a discussion that relates incommensurability, anthyphairesis, and Book II-style arguments in Zeuthen, CLAEE, but this only deals with a few simple cases; it does not involve the figures of II 5 and 6; and the arguments seem to be set in an arithmetised context.

3.5(b)   The extreme and mean ratio has been known under a variety of names. There is a brief history of these in the introduction to my GGS; in particular the most common name, the golden section, first seems to appear in a book by Martin Ohm, the younger brother of the physicist Georg Simon Ohm, in 1834 or 1835. The dramatically simple construction of the extreme and mean ratio of Fig. 3.30, based on the mid-points of two sides of an equilateral triangle inscribed in a circle, was published, perhaps for the first time, in 1984; see Odom, EP. I am indebted to Roger Herz–Fischler for much detailed information about the extreme and mean ratio; see Curchin & Fischler, HANTDEMR, DQDPR, and his monograph *MHDEMR*.
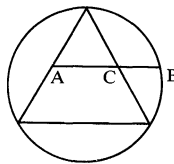


FIG. 3.30

3.6     I know of nowhere in the mathematical literature where the underlying property of the generalised method of sides and diagonals (that the convergents of a ratio $\sqrt{m} : \sqrt{n}$ whose period contains $p$ terms satisfies a recurrence relation of the form $s_{k+p} = nas_k + bd_k$, $d_{k+p} = mbs_k + ad_k$, where $a : b$ is the convergent that occurs just before the end-of-period term $2n_0$) is proved. I am grateful to J. W. S. Cassels for sketching a proof of this result.

On the contrast between arguments that proceed *dia tōn arithmōn* or *dia tōn grammōn*, see Neugebauer, *HAMA* ii, 771 n. 1.

# PLATO'S MATHEMATICS
# CURRICULUM IN *REPUBLIC* VII

## 4.1 PLATO AS MATHEMATICIAN

My attitude to Plato and the Academy has been greatly influenced by the sceptical and rigorous analysis described in Cherniss, *The Riddle of the Early Academy*, especially Lectures 1 and 3 therein. And, to a historian of ancient mathematics, an additional riddle is raised by Cherniss's scrupulous exploration of the fragmentary evidence, in the tension between his conclusion, on the one hand, that:

All the evidence points unmistakably to the same conclusion: the Academy was not a school in which an orthodox metaphysical doctrine was taught, or an association the members of which were expected to subscribe to the theory of ideas. . . . The metaphysical theories of the director were not in any way 'official' and the formal instruction in the Academy was restricted to mathematics [pp. 81 f.],

and

Philodemus says that mathematics made great progress under the direction of Plato, who formulated problems which the mathematicians zealously investigated. Proclus, too, in his famous summary, which appears to derive ultimately from the *History of Mathematics* written by Eudemus, credits Plato's concern for mathematics with the great progress of these studies and particularly of geometry. Besides Theaetetus, Leodamas, and Phillip of Opus, he names six specialists in mathematics who, he says, passed their time together in the Academy and pursued their investigations in common. [The next sentence will be quoted below.] [Plato] is said to have induced Phillip to turn his attention to the subject, to have originated the theorems about the section, the number of which Eudoxus increased, and to have communicated to Leodamas the method of analysis [p. 65],

while, on the other hand, Cherniss is forced to add, in the sentence omitted from the previous quotation, that:

It cannot be imagined that Plato *taught* any of these men mathematics, though he is said . . . [p. 65, original italics].

This opinion is then amplified:

Plato's influence on these men, then, was that of an intelligent critic of method, not that of a technical mathematician with the skill to make great discoveries of his own; and it was by his criticism of method, by his formulation of the broader problems to which the mathematician should address himself, and, as the summary of Proclus says, by arousing in those who took up philosophy an interest in mathematics that he gave such a great impulse to the development of the science [p. 66].

(I have omitted the detailed references to ancient literature and modern scholarship that fill out these quotations; the reader is strongly encouraged to consult the original book and the cognate review article, Cherniss, PM.) I believe that, within my new interpretation, we may argue that while Plato's principal interest was in dialectic, for which he regarded mathematics only as a preliminary, he does none the less show detailed knowledge of important characteristics and problems of technical mathematics, and there is no indication that he could not communicate on equal terms with the mathematicians who seem to have dominated, if not comprised, the group of friends and associates that assembled round him. This thesis I shall now attempt to illustrate by sketching an interpretation of his mathematics curriculum in *Republic* VII.

This curriculum (*Republic* VII 521c–531c) is proposed as part of the education of future guardians of the state; it will occupy men from age twenty (537b–c) to thirty (537d). It will be preceded by early training in childhood, imparted through play (536d–e), by sightseeing trips to the battlefield (537a), and by a two- or three-year break for gymnastics (537b), and only those who show promise will go on to study mathematics. At age thirty, and after a second selection, the students will pass on to a training in dialectic (537d–e) for five years (539e). They will then be compelled "to hold commands in war and other offices suitable to youth" for a further fifteen years, up to the age of fifty (539e–40a); and those who come through this programme of preparation will, from time to time for the rest of their lives, be called on to rule the state (540a–b). In this way, the objective will be achieved:

What we require is that those who take office should not be lovers of rule. Otherwise there will be a contest with rival lovers. What others, then, will you compel to undertake the guardianship of the city than those who have most intelligence of the principles that are the means of good government and who possess distinctions of another kind and a life that is preferable to the political life? [521b]

This life is that of the true philosopher as conceived by Plato. A related programme of training for citizens of the state, at a lower technical level and excluding the training in dialectic, is described in *Laws* VII, especially at 817e–822c.

If this state is not to remain a day-dream, then the first generation of true philosophers must be trained. Plato finished the book with a chilling description of the speediest and easiest way of doing this (540e–541a).

The curriculum is in five parts: *arithmētikē te kai logistikē* (524d–526c); plane geometry (526c–527c); three-dimensional geometry (528a–d); astronomy, which was prematurely introduced as the third subject (527d–528a), then replaced in this position by solid geometry and reintroduced as the fourth topic (528d–530c); and, finally, music theory (530d–531c). Plato emphasises, at the end of his description, that this is not a disparate collection of topics (531c–d):

I take it that if the investigation of all these [mathematical] studies goes far enough to bring out their community and kinship with one another, and to infer their affinities, then to busy ourselves with them contributes to our desired end and the labour taken is not lost; but otherwise is in vain, . . . , but it is a huge task. . . . Do we not know that all this is but the preamble of the law itself, the prelude of the strain we have to apprehend?

It is such a unity that I shall emphasise, by illustrating how each topic can contribute to our understanding of ratio. (All quotations from the *Republic* will be given in the translation of P. Shorey in the Loeb Classical Library and reprinted in Plato, *CD*.)

## 4.2 *ARITHMĒTIKĒ TE KAI LOGISTIKĒ*

From the aphorisms of Heraclitus, for example Fragment 1:

*Tou de logou toud' eontos aei axunetoi ginontai anthrōpoi kai prosthen ē akousai kai akousantes to prōton.* . . . (Of the *logos* which is as I describe it men always prove to be uncomprehending, both before they have heard it and when once they have heard it. . . . ),

and Fragment 50:

*Ouk emou alla tou logou akousantas homologein sophon estin hen panta einai* (Listening not to me but to the *logos* it is wise to agree that all things are one),

up to the opening sentence of St John's gospel:

*En archē ēn ho logos, kai ho logos ēn pros ton theon, kai theos ho logos* (In the beginning was the word, and the word was with God, and God was the word (*sic*)),

and beyond, the word *logos* has carried an enormous range of connotations. It occupies five dense columns of text in the Greek–*English Lexicon* by Liddell, Scott, & Jones, with a substantial entry in the *Supplement*; this is one of the longest entries in the whole lexicon. I shall here be concerned only with its very restricted range of mathematical meanings and uses.

An unambiguous ingredient of our surviving contemporary references to Greek mathematical and scientific thought in the first half of the fourth century BC, especially in Plato's and Aristotle's writings, is the frequent appeal to the idea of *logos* as ratio, and the use of the derived words that may even have been coined by Plato and his associates: *logistikē*, the art (i.e. *technē*, understood) of *logos*, usually translated as calculation, reckoning, computation, etc., and *logismos*, also used for calculation, reckoning, computation, etc. or, more abstractly, for judgement and rationality; an index to the occurrences in Plato, Archytas, Aristotle, and the pre-Socratic philosophers is given in an appendix to this chapter. While these words are used in a range of contexts and with a range of meanings that may be irrelevant to my mathematical enquiry here, a very substantial number of explicitly technical occurrences remain. And while, to many readers with other interests in mind, it often does not matter whether Plato is making a precise allusion to mathematics or not, the identification and interpretation of these references is crucial to a specialised

investigation of early Greek mathematics. So, to take one minute example, that the sentence:

*Logistikē pou tis hēmin hēn technē* [*Statesman* 259e]

can be translated (by Skemp in Plato, *CD*) as

There exists, we agree, an art of counting,

or (by H. N. Fowler, in the Loeb Classical Library edition) as

We recognised a sort of art of calculation

may be a matter of little general concern; but we must here make a more serious attempt to identify just what Plato meant by *logistikē* in this mathematical reference. This issue is complicated by three factors:

(i) Plato frequently refers to both *logistikē* and *arithmētikē* together, sometimes treating them as if they are overlapping parts of the same mathematical study; see, for example, the first part of the curriculum, which we shall be discussing in more detail below. Then, at other times, he is prepared to draw a fine distinction between them; see, in particular, *Gorgias* 451a–c, quoted in 1.2(e), and the related passage, *Charmides* 165e–166a, which does not even mention *arithmētikē*. In other contexts, for example *Protagoras* 356a–357b, he talks of *metrētikē*, the art of measurement, in similar terms:

If the saving of our life depended on the choice of odd and even, and on knowing when to make the right choice of the greater and when of the less—taking each by itself or comparing it with the other, and whether near or distant—what would save our life? Would it not be knowledge; and knowledge of *metrētikē*, since the art here is concerned with excess and defect, and of *arithmētikē* as it has to do with the odd and even?

(ii) *Logistikē* has an everyday meaning of practical calculation and financial accounts, and Plato himself uses it with this sense, for example at *Philebus* 56d–57a, where he again talks of *arithmētikē*, *logistikē*, and *metrētikē*:

Socrates: Are there not two kinds of *arithmētikē*, that of the people and that of philosophers? . . . And how about *logistikē* and *metrētikē* as used in building and trade when compared with the geometry and *logismoi* studied (or practised) in philosophy (*hē kata philosophian geōmetria te kai logismoi katameletōmenoi*)—shall we speak of each of them as one or two?
Protarchus: I should say that each of them was two.

Similar references to everyday *arithmētikē* and *logistikē* can be found in the first part of the curriculum. I shall distinguish the two different types by talking of 'practical' or 'theoretical' *arithmētikē* and *logistikē*.

(iii) By the first century BC, at the latest, theoretical *logistikē* had come to refer to that kind of artificial calculation that mathematicians tend to devise about numbered collections such as 'pebbles' or 'cattle on the isle of Sicily'; and neo-Platonic commentators, when discussing *logistikē*, seem to dwell exclusively on this interpretation. See, for example, Proclus–Morrow, *CFBEE*, 38–40:

Others, like Geminus, think that mathematics should be divided differently; they think of one part as concerned with intelligibles only and of another as working with perceptibles and in contact with them. . . . Of the mathematics that deals with intelligibles they posit *arithmētikē* and geometry as the two primary and most authentic parts, while the mathematics that attends to sensibles contains six sciences: mechanics, astronomy, optics, geodesy, canonics, and *logistikē*. . . . Geodesy and *logistikē* are analogous to these sciences [sc. geometry and *arithmētikē*], since they discourse not about intelligible but about sensible numbers and figures. . . . Nor does the student of *logistikē* consider the properties of number as such, but of numbers as present in sensible objects; and hence he gives them names from the things being numbered, calling them sheep numbers or cup numbers.

Another such description, credited to Anatolius, is cited in the Heronian *Definitions* (Heron–Heiberg, *Opera* iv, 164); and an anonymous scholium to the passage at *Charmides* 165e–166a gives a similar description of *logistikē* as the science of numbered objects. (Both of these passages are translated in Thomas, *SIHGM* i, 16–19.) But in his discussion of the curriculum in *Republic* VII, Plato is emphatic that this theoretical *arithmētikē te kai logistikē*, his astronomy, and his music theory are theoretical sciences which do *not* concern themselves with sensibles. There is a great gulf between what Plato is describing and what the neo-Platonic commentators understand by these terms. So, for the moment, I shall continue to ignore the later evidence.

   A comprehensive discussion of *logistikē* can be found in J. Klein, *Greek Mathematical Thought and the Origin of Algebra*. Since I agree in large measure with Klein's conclusion about the meaning of Plato's *logistikē*, but wish to propose a different reconstruction of the mathematics that lies behind it, I would like to proceed by making explicit some of these agreements and divergences, and recommend this book for an immensely detailed enumeration and evaluation of the evidence; and also the summary, with comprehensive references, in Burkert, *LSAP*, 446 n. 119. Klein concludes, and I concur, that

The ['theory of ratios and proportions'] does seem identical with the 'theoretical logistic' postulated by Plato [p. 6],

though I disagree with an implication of the word 'postulated', that while the possibility of this theory was assumed, little of it had actually been achieved. The next passages indicate Klein's conception of this theoretical logistic. (I have occasionally silently arranged the syntax of these quotations slightly to fit the context here, without, I hope, disturbing the sense of the originals. The italics and quotation marks are all Klein's. The first passage is followed by my annotations.)

In the face of definite multitudes of things we habitually *determine their exact number—* we number, i.e., count, the things . . . . In order to be able to count we must know and distinguish the single numbers, we must "distinguish the one and the two and the three" (*to hen te kai ta duo kai ta tria diagignōskein—Republic* VII, 522c). Plato calls the totality

of this science of all possible numbers the 'art of number'—'arithmetic'.[a] But we are also in the habit of *multiplying or dividing these multitudes*. This means that we are no longer satisfied with the number by which we have enumerated the things in question,[b] but that we bring to bear on this number new 'numbers', whether we wish to separate off a 'third' part of the respective quantity or wish to produce a multitude which amounts to 'four' times the given one.[c] In such multiplications and divisions, or, more generally, in all *calculations* which we impose on multitudes, we must *know beforehand* how the different numbers *are related to one another* and how they are constituted *in themselves*.[d] This whole science . . . is called the 'art of calculation'— 'logistic' [pp. 18–19].

*Notes:*
   [a] In fact, at *Republic* VII, 522c, Plato actually writes: "I mean, in sum, *arithmos te kai logismos*". See (i) above.
   [b] I take this and the following sentence to be a reference to the unit in which we are counting, which is then cut up or multiplied to make a new unit. But Plato insists, explicitly and emphatically, that this is not permitted in theoretical *logistikē*; see *Republic* VII, 525e–526a. Klein is well aware of this, as the quotations below will indicate.
   [c] Klein rightly insists throughout on the concreteness of the Greek *arithmoi* as numbered collections of objects, as cardinal numbers; but he does not analyse how repetition numbers describe relations between these cardinal numbers. So here he must put 'third' and 'four times' in quotation marks, presumably because they are not cardinal numbers, and so do no fit within the scope of his *arithmoi* without some further explanation. See Section 1.2(c).
   [d] This is, without doubt, a reference to the explanations of *logistikē* at *Gorgias* 451a–c (quoted in 1.2(e)) and *Charmides* 165e–166a (quoted above).

The reason why Klein needs such an elaborate description of multiplication and division is, I believe, partly because he has not identified the role of repetition numbers in Greek mathematics, but mainly because he cannot get away from the common conception that a ratio of two numbers *must* ultimately be manipulated and conceived mathematically as a fraction, or some more acceptable formulation of our idea of a fraction, and this contravenes the Platonic prohibition of "cutting up the unit" at *Republic* VII, 525e. This is made quite explicit in his Chapter 5, 'Theoretical Logistic and the Problem of Fractions', which opens:

The question before us is: What prevents later writers from interpreting the arithmetical theory of relations, i.e., proportions, as theoretical logistic? [p. 37],

and his answer follows:

The crucial obstacle to theoretical logistic—keeping in mind its connection with calculation—arises from *fractions*, or more exactly, from the fractionalisation of the unit of calculation [p. 39],

and

A calculation which is intended to be as exact as possible simply cannot be effected within the realm of these pure numbers [whose noetic character is expressed precisely in

the indivisibility of the units]. The immediate consequence of this insight, at least within the Platonic tradition, is the exclusion of all computational problems from the realm of the 'pure' sciences [p. 43].

A further difficulty, suggested in the introduction, is not developed in detail:

An additional reason was the elaboration of the theory of ratios into a *general* theory of proportion, which depended on the discovery of incommensurable magnitudes and which led altogether beyond the realm of counted collections [p. 7].

I shall postpone discussion of the topic of this last quotation until much later (see Sections 8.3 and 10.1) and here only comment that I find the choice of 'depended' (*beruhen*, in the original German) puzzling.

It is, I hope, now clear that it *is* possible to handle exact calculations, ratios of incommensurable magnitudes, approximations, and other such manipulations without transcending any of the restrictions on the *arithmoi* given in 1.2(c); many anthyphairetic illustrations of such manipulations were given in Chapters 2 and 3, and further procedures based on other definitions will be described below, in Sections 4.4 and 4.5. The essential liberating step comes from rejecting the techniques and preoccupations of our own present-day brand of arithmetised mathematics, which *is* founded on the arithmetical manipulations of fractions, with its further extensions to real numbers and beyond, and developing in its place techniques more suited to the Greek context.

I propose, then, that we should conceive of *logistikē* (*technē*) and *logismos* as 'ratio theory'. There are different varieties of ratio theory: I have, so far, developed only some of the theory of anthyphairetic *logistikē*, but I shall also illustrate astronomical and music-theoretical *logistikē* below and will introduce an accountant's ratio theory, with several variations, in Section 7.4(b). The unity of Plato's curriculum will then depend on being able to pass freely between these different ideas of ratio; we shall see how this will pose tricky problems for the mathematicians and will direct their attention away from the particular contexts of astronomy and music, towards abstract problems. Let us start by looking at Plato's description of *arithmētikē te kai logistikē* (*Republic* VII, 522c–526c), to see how well it corresponds to this point of view.

The study of ratio is intimately bound up with the use and manipulations of the *arithmoi*; in fact ratios are characterised by various kinds of patterns that can then be described by sequences of numbers. In this manner *arithmētikē* and *logistikē* enter in a fundamental way. As Plato says:

Is it not true of them that every art and science must necessarily partake of them? [sc. *arithmos te kai logismos*] [522c].

Plato describes how this study leads naturally to the awakening of thought by illustrating how the comparison of three lines (three fingers is the example he chooses in 523c–524c) leads to a true comparison, not of how they seem to appear—which may depend on how near or far each is—or of their irrelevant

characteristics of colour or touch or weight, but purely in relation to the "great and small". Now the theories of geometrical magnitudes, abstract magnitudes, and number that we find in the *Elements* also develop and exploit this kind of comparison—"the less of the greater" (*ho elassōn tou meizonos*) is an ubiquitous phrase in *Elements* V and VII—and this comparison of *arithmoi* and magnitudes is precisely the ingredient needed to define any kind of ratio (see, for example, my dialogues, $B_2$–$S_{13}$ and $B_{36-8}$, and $E_{67}$ and $B_{85}$–$A_{88}$, below). We shall see, indeed, that:

It is one of those studies which we are seeking that naturally conduce to the awakening of thought [523a],

and we will go on to see just how valid is the assertion ($S_{43}$) that

Studies that demand more toil in the learning and practice than this we shall not discover easily nor find many of them [526b],

though the difficulty of the subject is offset by its extraordinary attractiveness and charm (cf. 528c–d).

One feature of these ratios that has provoked a lot of subsequent toil is very strikingly evocative of the assertion that

It occurs to me, now that the study of *logismos* has been mentioned, that there is something fine in it . . . that it strongly directs the soul upwards and compels it to discourse about pure numbers, never acquiescing if anybody proffers to it in the discussion numbers attached to visible and tangible bodies [525c–d].

Practical *logistikē*—the activities of merchants and hucksters (525c), *arithmētikē*, *logistikē*, and *metrētikē* as used in building and trade (*Philebus*, 56e), and the calculation of financial accounts and taxes (see Chapter 7)—is mainly concerned with arithmetical calculations. But let us look more closely at what is implied by this arithmetic. The ontological status of arithmetic with *ratios* is far from evident, and my slaveboy's question at $B_{32}$ is a blatant anachronism: however fluent we might now be in performing these kinds of operations, and however basic they might now be to the development of our kind of mathematics, they have presented enormous obstacles in the understanding of mathematics, both phylogenically and ontogenically. (It is not too difficult to explain to a beginner what we might want a fraction like 'five-eighths' to mean; but to explain, from basic principles, what a calculation like 'five-eighths divided by three-quarters' might reasonably signify is of a different order of difficulty, and I suspect that many who can perform such manipulations would have serious problems in elucidating what they are doing.) Moreover, we find little or no explicit trace of such fractional arithmetic in our evidence on Greek mathematics. As an illustration of this, observe how reactions to the text of the 'arithmetical' Book VII of the *Elements* can vary from a recognition that this book does not seem to deal with the arithmetic of fractions in any form (see, for example, Mueller, *PMDSEE*, Chapters 2 and 3) to what I believe is a wholesale rewriting of this book in terms of fractional calculations (see, for example, van

der Waerden, *SA*, 48–9 and 110–16, quoted in Section 4.5(c), below), which then sometimes involves a comprehensive apparatus of equivalence relations, purely formal tricks of the last hundred years (see, for example, Taisbak, *DL*). I believe that my Socrates' reply $S_{33}$ is also anachronistic: these topics are not found in the *Elements*, or elsewhere in early Greek mathematics, because they are not a part of the non-arithmetised mathematics which is being developed there, and neither my boy's question $B_{32}$, nor the reply $S_{33}$, fits naturally within this non-arithmetised approach. But if we excise these two speeches, and modify the dialogue at this point to:

BOY$'_{32-4}$: Well here's another question. Suppose we have three heaps of stones; how can we calculate the ratio of the first to the third from the ratio of the first to the second and the second to the third? Or a similar question about adding . . . ,

we now have something that is historically acceptable and mathematically interesting; my slaveboy is now asking about the operation known to Euclid as 'compounding' ratios. Not only is Euclid's text corrupt and confused (see the brief comments at the end of 1.2(d), and the more detailed analysis to follow in Section 4.5(c), below) but also the underlying mathematical problem is curiously difficult. It is a remarkable feature of anthyphairetic and astronomical ratios that their arithmetic is intractable; it is difficult to pass directly from the information that $\sqrt{2}:1 = [1, 2, 2, 2, \ldots]$ and $\sqrt{3}:1 = [1, 1, 2, 1, 2, 1, 2, \ldots]$ to the results that $\sqrt{6}:1 = [2, 2, 4, 2, 4, 2, 4, \ldots]$ and that $(\sqrt{2} + \sqrt{3}):1 = [3, 6, 1, 5, 7, 1, 1, \ldots]$, where this latter ratio seems to exhibit no evident pattern and its evaluation, using the algorithms of Chapter 2, is a great deal more difficult than all of the ratios we have met so far except for the circumference and diameter of a circle. (This 'binomial' line $\sqrt{2} + \sqrt{3}$ is one of the *alogoi* of the classification of *Elements* X; see Chapter 5, below.) Even the relation between the ratios $b:c$ and $2b:c$ is far from obvious. As a modern textbook puts it:

There is, however, another and yet more significant practical demand that the apparatus of continued fractions [i.e. anthyphairetic ratios, modern style] does not satisfy at all. Knowing the representations of several numbers we would like to be able, with relative ease, to find the representations of the simpler functions of these numbers (especially, their sum and product). In brief, for an apparatus to be suitable from a practical standpoint, it must admit sufficiently simple rules for arithmetical operations; otherwise it cannot serve as a tool for calculation. We know how convenient systematic fractions are in this respect. On the other hand, for continued fractions there are no practically applicable rules for arithmetical operations; even the problem of finding the continued fraction for a sum from the continued fractions representing the addends is exceedingly complicated, and unworkable in computational practice [Khinchin, *CF*, p. 20].

The situation is quite extraordinary: on the one hand, in the world of the sensibles, these everyday operations of arithmetic with general kinds of numbers are commonplace. We have many, many examples of financial accounts on papyrus from Graeco-Roman Egypt which contain explicit

examples of such arithmetic, expressed throughout in unit fractions. This is practical *logistikē*; there seems to be no insuperable difficulty in doing it, but it is very difficult to describe in any correct and complete way how it is done, or, at a more subtle level, to say just what is being done. (See the last part of $S_{35}$; more details of our evidence about these calculations will be given in Chapter 7.) In some kinds of theoretical *logistikē*, on the other hand, we can define our terms precisely, but we cannot now perform some of these everyday manipulations. (Now see the beginning of $S_{35}$). The reaction of a practical calculator to a description of the insights and power of this theoretical *logistikē* is likely, even today, to be not dissimilar to the reception that Plato describes, at the beginning of *Republic* VII, 514a–518b, that will be accorded to someone who returns to the cave of shadows after seeing the reality of the world of sunlight outside! The final ironic twist in this story of the search for algorithms for arithmetic—that anthyphairetic arithmetic is actually not very difficult while decimal arithmetic is much more subtle than is generally thought—will be described in Section 9.3(a).

So, on the one hand, theoretical *logistikē* seems, in the ancient evidence, to show little preoccupation with the arithmetic of ratios and, even today, the corresponding mathematical problem of arithmetic with continued fractions is still little discussed because of its supposed difficulty. On the other hand, the cluster of mathematical ideas round the subject as described by Plato seems to involve the ingredients of 'the odd and the even', 'the greater and the less' (or 'the excess and the defect') and 'their mutual relationships'; does this resonate with anything of significance in connection with ratio theories? In fact, the associations, in particular with anthyphairetic ratios, are now so rich and diverse that any attempt to enumerate them completely would spread through many pages, but here, in summary, are a few instances:

(i) At the ground level of the definition and evaluation of any kind of ratio of two numbers or two magnitudes we shall need to manipulate 'the greater and the less' in some kind of comparison process (see $S_1$ to $S_{13}$ and the further definitions to follow); and this comparison tends to implicate *arithmētikē* to describe the pattern of the process. In anthyphairesis, the even-numbered terms, $n_0, n_2, \ldots$, describe subtractions from the initially larger pile, the odd, $n_1, n_3, \ldots$, from the initially smaller.

(ii) Now ascend to the level of understanding some details of anthyphairetic ratios. Here the order relationship ('the less and greater') between ratios involves taking account, in a surprising and subtle way, of 'the less and the greater' numbers that arise in 'the odd and the even' steps; see $S_{17}$ to $B_{22}$, and $S'_{21}$ in Section 2.3(b).

(iii) Look now at the algorithms for calculating anthyphairetic ratios, described in Section 2.3. These are built around the alternating manipulations of under- and overestimates, 'the mutual relationship of the greater and the less and the odd and the even'.

(iv) Pass to the properties of anthyphairetic ratios as explored in Section 2.4. Again we find an increasingly subtle exploration of the greater and the less; and the odd–even parity enters in the behaviour of Pell's equation (see the generalised side and diameter numbers and Archimedes' *Cattle Problem*) in a way that still has not received a satisfactory characterisation.

(v) There is an odd–even parity underlying the astronomical ratios to be described in Section 4.4, below, but it plays a less prominent role there.

When I try to identify the mathematical features which underlie and permeate the definition, development, and application of anthyphairetic ratio theory, I find it hard to better these allusive descriptions that can be borrowed from Plato.

## 4.3  PLANE AND SOLID GEOMETRY

Plato does not tell us what his plane geometry is, but what it does (*Republic* VII 526d–527c):

[SOCRATES:]  What we have to consider is whether the greater and more advanced part of it tends to facilitate the apprehension of the idea of good. . . . If it compels the soul to contemplate essence, it is suitable; if genesis, it is not. . . . [It] will not be disputed by those who have even a slight acquaintance with geometry, that this science is in direct contradiction with the language employed in it by its adepts. . . . Their language is most ludicrous, though they cannot help it, for they speak as if they were doing something and as if all their words were directed towards action. For all their talk is of squaring and applying and adding and the like, whereas in fact the real object of the entire study is pure knowledge . . . the knowledge of that which always is, and not of a something which at some time comes into being and passes away.
[GLAUCON:]  . . . Geometry is the knowledge of the eternally existent. . . .
[SOCRATES:]  It would tend to draw the soul to truth, and would be productive of a philosophical attitude of mind, directing upward the faculties that now wrongly are turned earthward. . . . We must require that the men of your Fair City shall never neglect geometry, for even the by-products of such study are not slight.

We can only infer from this that geometry is thought by Plato to be very important, but it is difficult to tell, from this report alone, just what kind of geometry is being referred to.

About solid geometry, Plato is much more informative (528a–d):

[SOCRATES:]  After plane surfaces, we went on to solids in revolution [sc. astronomy] before studying them in themselves. The right way is next in order after the second dimension to take the third. This, I suppose, is the dimension of cubes (*auxē kubōn*) and of everything that has depth.
[GLAUCON:]  . . . But this subject does not appear to have been investigated yet.
[SOCRATES:]  There are two causes of that: first, inasmuch as no city holds them in honour, these enquiries are languidly pursued owing to their difficulty. And secondly, the investigators need a director, who is indispensable for success and who, to begin with, is not easy to find, and then, if he could be found, as things are now, seekers in

this field would be too arrogant to submit to his guidance. But if the state as a whole should join in superintending these studies and honour them, these specialists would accept advice, and continuous and strenuous investigation would bring out the truth. Since even now, lightly esteemed as they are by the multitude and hampered by the ignorance of their students as to the true reasons for pursuing them, they nevertheless in the face of all of these obstacles force their way by their inherent charm and it would not surprise us if the truth about them were made apparent.

[GLAUCON:] It is true that they do possess an extraordinary attractiveness and charm.

Note particularly how Plato insists that his solid geometry has "not been investigated", is "languidly pursued", is "lightly esteemed", is "hampered by ignorance", and is "absurdly neglected".

The standard interpretations of this passage are that Plato is referring either to the construction of the regular solids, or to solid geometry in general, or to the duplication of the cube. However, this does not seem compatible with our evidence about Hippocrates, Archytas, and Theaetetus, which indicates that these subjects *do* "appear to have been investigated" by the specialists and are *not* "lightly esteemed by the multitude". Rather than develop these objections in detail here, I shall proceed directly to my alternative interpretation.

The first stage of the problem of duplicating the cube is to find some geometrical configuration from which we can construct the double cube. This generalises the successful duplication of a square in the third section of Socrates' encounter with the slaveboy in the *Meno*, and we have evidence that this aspect of the problem for cubes had been resolved by the contributions of Hippocrates and Archytas, and perhaps even of Eudoxus and Plato, at the time of the composition of the *Republic*. We now try to pass to the next stage of the problem of duplicating the cube, and try to 'count up' the solution. This will correspond to the further stages in the problem of doubling (or tripling, etc.) the square, in which we evaluate the ratio $\sqrt{2}:1$ (or, more generally, the ratio $\sqrt{n}:\sqrt{m}$)—what I have called 'the dimension of squares'—and it will be my proposed interpretation of Plato's problem of the "dimension of cubes and everything that has depth" (*Esti de pou touto peri tēn tōn kubōn auxēn kai to bathous metechon*, 528b). Since I have so heavily loaded the terminology by referring, in my $S_{41}$ and throughout Chapter 3, to the earlier problem as the 'dimension of squares', an explanation of the Greek text is perhaps in order at the outset. The word used here is *auxē*, as in *deutera auxē*, 'second dimension' or *auxē kubon*, 'dimension of cubes'. This carries connotations of growth and increase, as in the example at *Republic* VI 509b: *tēn genesin kai auxēn kai trophēn*, "generation and growth and nurture"; also see the other examples in Liddell, Scott, & Jones, *GEL*. The first step in the plane problem is the generation (*genesis*) of the square (see *Meno* 82b–c or *Elements* I 46). Then we pass on to the 'growth' of larger squares; for example, two juxtaposed squares can be converted into a double square (see *Meno* 82b–5c or *Elements* II 14) whose side is then 'counted up' (*arithmeō*, *Meno* 84a) in relation to the original square; three juxtaposed squares are converted into a triple square,

whose side is 'counted up' against the original and the double square, etc. In such a way, all squares can be systematically 'grown' and 'counted up' against each other in a programme that uncovers some remarkable mathematical phenomena, some of which were described in the previous chapter. (The interesting proposal that Plato is not completely satisfied with the conclusion of Socrates' encounter with the slaveboy in the *Meno*, since they do not attempt to 'count up' the side of the double square, is developed in Brown, PDSA.)

Let us try a similar procedure for cubes, and so pass to this next step of trying to 'count up' the double cube. As with the dimension of squares (see Section 3.3), we start with a heuristic exploration. The generalised *Parmenides* proposition and a table of cubes up to 99 enable us to apply the algorithm of 2.3(c) to yield the expansion

$$\sqrt[3]{2} : 1 = [1, 3, 1, 5, 1, 1, \ldots].$$

No pattern is yet evident. A table of cubes up to 999 enables us to extend this calculation to

$$\sqrt[3]{2} : 1 = [1, 3, 1, 5, 1, 1, 4, 1, 1, \ldots].$$

TABLE 4.1. Evaluation of $\sqrt[3]{2} : 1 = [1, 3, 1, 5, 1, 1, 4, 1, 1, \ldots]$

| Under-estimate | Over-estimate | New estimate $p : q$ | $p^3$ | $2q^3$ | Under/over | Anthy-phairesis |
|---|---|---|---|---|---|---|
| 0 : 1 | 1 : 0 | 1 : 1 | 1 | 2 | under | 1 |
| 1 : 1 | 1 : 0 | 2 : 1 | 8 | 2 | over | |
| 1 : 1 | 2 : 1 | 3 : 2 | 27 | 16 | over | |
| 1 : 1 | 3 : 2 | 4 : 3 | 64 | 54 | over | 3 |
| 1 : 1 | 4 : 3 | 5 : 4 | 125 | 128 | under | 1 |
| 5 : 4 | 4 : 3 | 9 : 7 | 729 | 686 | over | |
| 5 : 4 | 9 : 7 | 14 : 11 | 2744 | 2662 | over | |
| 5 : 4 | 14 : 1 | 19 : 15 | 6859 | 6750 | over | |
| 5 : 4 | 19 : 15 | 24 : 19 | 13824 | 13718 | over | |
| 5 : 4 | 24 : 19 | 29 : 23 | 24389 | 24334 | over | 5 |
| 5 : 4 | 29 : 23 | 34 : 27 | 39304 | 39366 | under | 1 |
| 34 : 27 | 29 : 23 | 63 : 50 | 250047 | 250000 | over | 1 |
| 34 : 27 | 63 : 50 | 97 : 77 | 912673 | 913066 | under | |
| 97 : 77 | 63 : 50 | 160 : 127 | 4096000 | 4096766 | under | |
| 160 : 127 | 63 : 50 | 223 : 177 | 11089567 | 11090466 | under | |
| 223 : 177 | 63 : 50 | 286 : 227 | 23393656 | 23394167 | under | 4 |
| 286 : 227 | 63 : 50 | 349 : 277 | 42508549 | 42507866 | over | 1 |
| 286 : 227 | 349 : 277 | 635 : 504 | 256047875 | 256048128 | under | 1 |
| 635 : 504 | 349 : 277 | 984 : 781 | 952763904 | 952759082 | over | |
| etc. | | | | | | |

Still no pattern manifests itself. (The calculation so far is set out in Table 4.1. Note that it has already exceeded the standard Greek numeral system, to be described in Chapter 8.) The evaluation can be continued as long as the arithmetical operation of cubing the numbers is practically feasible, or we can go on to develop new and more elaborate procedures for extending the calculation further. The first 75 terms of the ratio are:

$$\sqrt[3]{2} : 1 = [1, 3, 1, 5, 1, 1, 4, 1, 1, 8, 1, 14, 1, 10, 2, 1, 4, 12, 2, 3,$$
$$2, 1, 3, 4, 1, 1, 2, 14, 3, 12, 1, 15, 3, 1, 4, 534, 1, 1, 5, 1,$$
$$1, 121, 1, 2, 2, 4, 10, 3, 2, 2, 41, 1, 1, 1, 3, 7, 2, 2, 9, 4,$$
$$1, 3, 7, 6, 1, 1, 2, 2, 9, 3, 1, 1, 69, 4, 4, \ldots].$$

Still no hypothesis manifests itself; there is neither any apparent periodicity, nor any associated apparent systematic behaviour of $p^3 - 2q^3$ that might lead to a generalisation of Pell's equation. The only regularity that persists, since it is a fundamental property of the *Parmenides* proposition and not of the particular example to which it is applied, is the behaviour of the cross-product of consecutive approximations in column 3 of Table 4.1:

$$3^2 - 2 \cdot 4 = 1$$
$$4^2 - 3 \cdot 5 = 1$$
$$5 \cdot 7 - 4 \cdot 9 = -1$$

etc.

The difficulty with the problem of the dimension of cubes is that we cannot even begin, at this stage, to formulate any hypothesis that we could attempt to prove. The hunt for results cannot start until we have identified the quarry.

This problem is still only partially understood. It was not until Gauss's work on the continued fraction expansion of a randomly chosen real number, a problem that he abandoned incomplete because of its difficulty, that any reasonable general hypothesis could be formulated, and only very partial results have so far been established, notwithstanding more and more extended explorations made possible by the development of electronic computers; see, for example, von Neumann and Tuckerman, CFE, and Churchhouse and Muir, CFANMI. Alternatively we might ask the simpler question: are the terms bounded or, if not, can we obtain some idea of their rate of growth? Again, only partial results have been found, very weak bounds like $n_k \leqslant a^{b^k}$ where $a > 1$ and $b > 1$. More explanations and details will be given in Section 9.2(c).

Within this interpretation, Plato's words about the problem of the dimension of cubes can apply, in their entirety, to this problem today; and we still have no informed idea of its possible ramifications, though the charm of the problem still attracts a few mathematicians. But I have tried, in my description, to show how this problem also fits naturally within an anthyphairetic formulation of

early Greek mathematics, and Plato's words can fit no less appropriately there also, without any anachronism. Moreover we can now go back to Plato's plane geometry. We read later that

It is by means of problems, then, as in the study of geometry, that we will pursue astronomy . . . [*Republic* VII, 530b].

Plato does not appeal to the deductive method, the proving of theorems, or the development of formal theories, but to "problems". I propose that one of the things that has so impressed Plato about geometry is the initial success there of the anthyphairetic programme: the fruitful pursuit of problems such as my Socrates outlined at $S_{41}$. I shall describe, in Chapter 5, how a clever kind of argument enables this programme to be extended further to yield some information about the circumdiameter and side of the pentagon and the circumdiameter and edge of all the regular solids, and so provide an interpretation of, and motivation for, Book X of the *Elements* and bring out its relationship to Books II, IV, and XIII. In this way Books X and XIII, which surely deserve the description as "the greater and more advanced parts" of geometry (*Republic* VII, 526e), will be seen to fit together within the anthyphairetic programme. But I know of no technique, even today, that will yield any anthyphairetic result about any other problem in solid geometry.

## 4.4 ACADEMIC ASTRONOMY

### 4.4(a) *Introduction*

Plato's views on theoretical astronomy have earned him the contempt of observational and computational astronomers. In answer to Glaucon's query about

How . . . ought astronomy to be taught contrary to the present fashion if it is to be learned in a way to conduce to our purpose? [529c],

Socrates explains (529c–530c):

These sparks that paint the sky, since they are decorations on a visible surface, we must regard, to be sure, as the fairest and most exact of material things; but we must recognise that they fall far short of the truth, the movements, namely, of real speed and real slowness in true number and in all true figures both in relation to one another and as vehicles of the things they carry and contain. These can be apprehended only by reason (*logos*) and thought, but not by sight. . . . We must use the blazonry of the heavens as patterns to aid in the study of those realities, just as one would do who chanced upon diagrams drawn with special care and elaboration by Daedalus or some other craftsman or painter. For anyone acquainted with geometry who saw such designs would admit the beauty of the workmanship, but would think it absurd to examine them seriously in the expectation of finding in them the absolute truth with regard to equals or doubles or any other ratio (*isōn ē diplasiōn ē allēs tinos summetrias*). . . . Do you not think that one who was an astronomer in very truth would feel in the same way when he turned his eyes upon the movements of the stars? He will be willing to concede that the artisan of heaven

fashioned it and all that it contains in the best possible manner for such a fabric; but when it comes to the proportions of day and night, and of their relation to the month, and that of the month to the year, and of the other stars to these and one another, do you not suppose that he will regard as a very strange fellow the man who believes that these things go on forever without change or the least deviation. . . . It is by means of problems, then, as in the study of geometry, that we will pursue astronomy too, and we will let be the things in the heavens, if we are to have a part in the true science of astronomy.
[Glaucon:] You enjoin a task that will multiply the labour of our present study of astronomy many times.

There is a related passage at *Timaeus*, 37c–39e, where Plato describes how the universe was made out of uniformly rotating motions which generate time and the pattern of eternal nature. See, for example, 39b–c:

In order that there might be a clear measure of the relative speeds, slow and quick, . . . God kindled a light which we now call the Sun, to the end that it might shine, so far as possible, throughout the whole Heaven, and that all the living creatures entitled thereto might participate in number, learning it from the revolution of the Same and Similar. In this wise and for these reasons were generated Night and Day, which are the revolution of the one and most intelligent circuit; and Month, every time that the Moon having completed her own orbit overtakes the Sun; and Year, as often as the Sun has completed his own orbit. . . . [Translation by R. G. Bury, in the Loeb Classical Library.]

I shall again describe my interpretation by a dialogue in which my slaveboy shows further evidence of his considerable mathematical abilities. The subject will again be "the absolute truth in regard to equals or doubles or any other ratio", but now in the context of "the relation [of the day] to the month, and that of the month to the year".

### 4.4(b)  *The slaveboy meets Eudoxus*

BOY$_{44}$: Excuse me, Eudoxus, Socrates told me [see S$_{35}$] that you astronomers have your own way of handling ratios. Would you explain it to me?

EUDOXUS$_{45}$: How do you think the relationship of day to month can best be described?

BOY$_{46}$: Well, I could do the reciprocal subtraction process, but I can't now point to heaps of stones or manipulations with geometric figures while I'm doing it.

EUDOXUS$_{47}$: Yes, that technique is well adapted to *arithmētikē* and geometry, but it doesn't seem to be the right thing for astronomy. Tell me: what would happen if the lunar month was always thirty days long?

BOY$_{48}$: I suppose that the next new moon would always appear sometime on the thirty-first day after the previous one.

EUDOXUS$_{49}$: Anytime?

BOY$_{50}$: That's possible. I've heard talk of astronomical anomalies, variations in the regularity of the heavens, so that the month could sometimes be a bit shorter, sometimes a bit longer, provided the new moon always appeared sometime on the thirty-first day after the previous new moon.

EUDOXUS$_{51}$: Let's do some theoretical astronomy. Suppose that the motions are uniform, and go on forever without change or the least deviation, and that the month is exactly thirty days long.

BOY$_{52}$: Then the moment of new moon must always take place at the same time on the thirty-first day after the previous new moon.

EUDOXUS$_{53}$: Why?

BOY$_{54}$: If it didn't, it would slowly creep backwards or forwards through the day. Now the first time the new moon can actually be observed is at dusk just after sunset. So, if the new month starts at the sunset of this observation, and if the moment of new moon is actually creeping forwards in time, then eventually one month will start with a thin observable crescent and the next month will start with a thick crescent, thirty-one days later. Or, if the month were a bit shorter than thirty days, it would eventually give rise to a twenty-nine day month, in the same kind of way.

EUDOXUS$_{55}$: Let's draw this out schematically on a line which represents time flowing by uniformly. We'll mark each sunset, the beginning of each day, to give a scale of equal time intervals:

$$D_1\ D_2\ D_3\ D_4\ D_5\ D_6\ D_7\ D_8\ D_9\ D_{10}\ D_{11}\ D_{12}\ D_{13}\ D_{14}\ D_{15}\ D_{16}\ D_{17}\ D_{18}\ D_{19}\ D_{20} \cdots$$
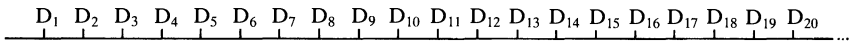
Then, on the same line, we mark off the conjunction of sun and moon, the theoretical moment of new moon. For convenience, I'll suppose the 'month' is very short, just one or two days long, or the line will have to stretch out a very long way for us to see what's happening:

$$D_1\ D_2\ D_3\ D_4\ D_5\ D_6\ D_7\ D_8\ D_9\ D_{10}\ D_{11}\ D_{12}\ D_{13}\ D_{14}\ D_{15}\ D_{16}\ D_{17}\ D_{18}\ D_{19}\ D_{20}$$
$$M_1\ M_2\ M_3\ M_4\ M_5\ M_6\ M_7\ M_8\ M_9\ M_{10}\ M_{11}\ M_{12}\ M_{13}\ M_{14}$$
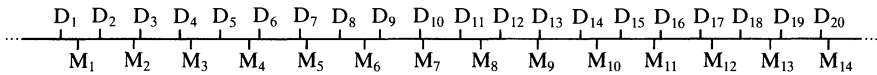
BOY$_{56}$: It's beginning to look like geometry now.

EUDOXUS$_{57}$: Yes, you can do the same thing with two lines or two numbers. But don't be misled by first appearances; I haven't been able to find any problems of interest in geometry with this yet—just a few formal definitions and propositions about ratio and proportion.

BOY$_{58}$: I see now how you're using the day like the dripping of a water-clock, to mark the rhythm of the passage of time.

EUDOXUS$_{59}$: Then describe the rhythm of the months.

BOY$_{60}$: You mean the pattern of Ds and Ms? I can simply say:

$$D_0\&M_0\ D_1\ M_1\ D_2\ M_2\ D_3\ D_4\ M_3\ D_5\ M_4\ D_6\ D_7\ M_5\ D_8,\ \text{etc.},$$

where D&M denotes a conjunction at sunset that I can suppose starts the sequence; you did say that we were doing theoretical astronomy! Or I can say: a one-day month, a one-day month, a two-day month, a one-day month, a two-day month, etc., and then I can code it as $\{\ldots, *1*, 1, 2, 1, 2, \ldots\}$ where those asterisks again indicate the coincidence. I'm also supposing here that the new month always starts at the sunset coinciding with or following conjunction, and that's another highly theoretical supposition.

EUDOXUS$_{61}$: These sequences are the astronomical ratios, of course. Can any sequence of numbers arise in this way?

BOY$_{62}$: Clearly not. You couldn't have $\{\ldots, 1, 2, 3, \ldots\}$ or, for subtler reasons, $\{\ldots, 1, 1, 2, 2, \ldots\}$; I wonder how you can distinguish a permitted sequence? I suppose, just like anthyphairetic ratios, that there must be lots of information coded in those numbers.

EUDOXUS$_{63}$: Quite so. Here's an example of a longer sequence; see if you can decode it:

$$\{\ldots, *1*, 1, 2, 1, 2, 1, 2, 1, 1, 2, 1, 2, 1, 2, 1, 1, 2, 1, 2, 1, 2, 1, 1, 2, 1, 2, 1, 2, 1, *1*, \ldots\}.$$

BOY$_{64}$: Since it contains two pairs of asterisks it must be a commensurable ratio and then—because I can count the numbers and the commas and do a bit of fiddling at the ends—I can see that twenty-nine commas represent thirty months, while the ones and twos, added together, make forty-two which represents forty-three days. So it's the ratio $43 : 30$.

Incidentally I can represent that $43 : 30$ anthyphairetically as [1, 2, 3, 4]; and that seems to me to be shorter, neater, and clearer.

EUDOXUS$_{65}$: Socrates told me that you were quick! But I did help you a lot by that gift of two coincidences; what would you have done without them? And you are right: these astronomical ratios are very dilute, small beer compared with the pure spirit of anthyphairetic ratios.

BOY$_{66}$: I can't yet see what to do if there aren't any coincidences. But I can see how that first anthyphairetic term $n_0 = 1$ comes from the way that the astronomical sequence contains only ones and twos, since that means that the month is more than one but less than two days long. So the interval between $D_1$ and $M_1$ will give the difference between a day and a month, here less than a day; then the interval between $D_2$ and $M_2$ is twice that difference, still less than a day; but the interval between $D_3$ and $M_3$, three-times that difference, is bigger than a day, since the third month contains an extra day. So $n_1 = 2$. I suppose that now that new remainder, the interval between $M_2$ and $D_3$, between two months and three days—that's a convergent!—can be measured against the previous remainder, the difference between $D_1$ and $M_1$, in one of the larger patterns that seem to occur in the sequence.

EUDOXUS$_{67}$: Not bad! Since you're obviously so keen on anthyphairetic ratios, I'll show you how to generate the astronomical ratio which corresponds to the anthyphairetic ratio $[1, 2, 3, 4, \ldots]$. I'll describe longer and longer sequences which start from an initial coincidence $M_0\&D_0$. I, too, can only see how to do this starting from a coincidence, though I can also see that lots of these astronomical ratios won't ever have a coincidence in them, however far you go backwards or forwards through them.

Since $n_0 = 1$, we must start M&D D M, .... Let's call that initial pattern $S_0$—that doesn't include the coincidence M&D; so $S_0 = $ D M. In fact, it's better, for the moment, to forget that initial coincidence.

Since $n_1 = 2$, we now repeat this pattern $S_0$ twice, and then we'll have to slip in an extra day: <u>DM DM</u> D. Let's call that $S_1$ and write $S_1 = S_0^{n_1} D$.

Now we look at the interval from $M_2$ to $D_3$, and compare that with $D_1$ to $M_1$. Since $n_2 = 3$, when $M_2$ to $D_3$ has been repeated three-times it will still be less than $D_1$ to $M_1$ but, on the next time, we'll have to slip in an extra $D_1$ to $M_1$. So we get $S_2 = $ <u>DMDMD</u> <u>DMDMD</u> <u>DMDMD</u> <u>DM</u> $= S_1^{n_2} S_0$. At the end of that, we have the interval from $D_{10}$ to $M_7$—another convergent!

Next, since $n_3 = 4$, we repeat $S_2$ four-times and then, on the next time round, we have to slip in an extra $S_1$ sequence; so $S_3 = S_2^{n_3} S_1$. Work it out, and you'll see we get the sequence I've been talking about [in E$_{55}$ and E$_{63}$, above].

If the anthyphairetic ratio terminated there, [1, 2, 3, 4], this would end with a coincidence. Now you know that terminating anthyphairetic sequences behave as if they finish with very, very big numbers [see $S_{21}'$ in Section 2.3(b)]. So if it wasn't

exactly terminating, but actually was $[1, 2, 3, 4, 1\,0000\,0000, \ldots]$—that myriad-myriad is an everyday approximation to infinity—then, as this cycle $S_3$ was repeated, so the gap between $M_{30}$ and $D_{43}$ would be doubled to $M_{60}$ to $D_{86}$, tripled to $M_{90}$ to $D_{129}$, and so on, and would eventually manifest itself. (If, on the other hand, $D_{43}$ came before $M_{30}$, then we'd be dealing with a ratio like $[1, 2, 3, 3, 1, 1\,0000\,0000, \ldots]$, which is close to but bigger than $[1, 2, 3, 4,]$—this time the month is slightly longer.) This is another illustration of how these large terms in anthyphairetic expansions code large amounts of information; an astronomical ratio would have to go on for scores of myriads of myriads of terms to encode this same information. But theoretical astronomy doesn't face any problems with detecting exact coincidences; it exploits the repetition of precisely uniform events and the 'before, coincident, or after' in exactly the same way that geometry exploits addition and the 'less, equal, or greater'. Complaining that the heavens aren't like that, or saying that these kinds of measurements aren't possible, is like objecting that a geometer's figure of a circle isn't exactly round.

    Do you follow?

BOY$_{68}$: I'm sorry, Eudoxus, I wasn't listening! Take that long sequence you gave me and count the number of one-day months lying between consecutive two-day months and add one to each. (That's probably just the usual trouble with counting points and intervals, that there's always one more point than intervals.) This gives a sequence:

$$\{\ldots, *3*, 2, 2, 3, 2, 2, 3, 2, 2, 3, 2, 2, *3*, \ldots\}.$$

Note that this sequence contains just twos and threes. Now count the number of twos between the threes and add one to each: $\{\ldots, *3*, 3, 3, *3*, \ldots\}$. I think this tells you that the anthyphairetic ratio is $[1\ 2, 3, \ldots]$ and, since this contains four terms—those coincidences introduce complications, and I think incommensurable ratios without coincidences may be easier to handle—I think I can deduce from this that $n_4 = 4$. And I think that the sequences that arise as astronomical ratios are precisely those that, when you go on transforming them in this way, will always generate new sequences only containing $n$s and $(n+1)$s, where that $n$ will be the next term of the corresponding anthyphairetic ratio.

EUDOXUS$_{69}$: That's brilliant! Let me think.

[*And they both walk off, thinking.*]

### 4.4(c) *Egyptian and early Greek astronomy*

One feature of the early phase of astronomy in Egypt and Greece, parallel to and distinct from the cosmological speculations of pre-Socratic philosophers, is the construction of cyclical calendars. The earliest Greek examples, known as *parapegmata*, are associated with the names of Meton, Euctemon, Democritus, Eudoxus, and Callippus, and several examples survive, either directly on papyrus or stone, or indirectly, as described by later authors. There is an enormous literature on these parapegmata which I shall not attempt to describe here; it is summarised and reviewed in Neugebauer, *History of Ancient Mathematical Astronomy* ii, 615–29. Lunar calendars and the phenomena of first and last visibility of the moon (see B$_{54}$) are very clearly described and analysed in Parker, *The Calendars of Ancient Egypt*. Also see

Bulmer-Thomas, PA; Goldstein & Bowen, NVEGA; Riddell, EMES; and van der Waerden, GAC.

Any discussion of calendars is complicated by the distinction that must be made between solar and lunar calendars, a distinction that frequently corresponded, in ancient astronomy, to the difference between civil and astronomical calendars. The solar calendar is based on the movement of the sun through the zodiac: the year therefore has 12 zodiacal months, each 30 or 31 days long. In a lunar calendar, the first day of the month is tied to some lunar phenomenon, for example the appearance of the new crescent after sunset, the last appearance or first disappearance of the old crescent before sunrise, or the observation or estimation of the moment of full moon; the lunar month is 29 or 30 days long; and the year will contain either 12 or 13 lunar months.

Another complication is caused by astronomical anomaly. Since the motions of the heavenly bodies vary in speed, the intervals they determine will vary in duration. Here are two examples. First, while the *mean* synodic month (conjunction to conjunction, averaged) is $29 \cdot 530\,59$ days, the synodic months themselves may vary between about $29\frac{1}{4}$ and $29\frac{4}{5}$ days long, and the observations of crescent visibility in Egypt can give up to three 29-day or five 30-day lunar months in a row. Second, because the sun's motion is not uniform, the intervals between the solstices and equinoxes will vary during the year. These periods, starting with the summer solstice, are 92, 89, 90, and 94 days long, and they are so given in Calippus' parapegma as reported in the 'Eudoxus papyrus', P. Par. 1. (Geminus' later description of Calippus' parapegma is slightly different; see Neugebauer, *HAMA* ii, 627, n. 9.) On the other hand, Eudoxus' parapegma, as described in this same text, apparently makes these intervals into [91], 92, 91, and 91 days (the slight doubt is caused by restoration of damaged text), in clear defiance of the other recorded calendric evidence and astronomical observations. In my interpretation of Plato's astronomy, I consider a kind of theoretical astronomy in which anomaly is ignored (see $E_{47-51}$), and I suggest that Eudoxus may indeed have been involved in this kind of study.

Here, briefly, are some examples of ancient cyclical calendars. First a Greek text which contains an early lunar calendar: P. Ryl. [inv.] 666 = P. Ryl. iv 589, dated 180 BC, first published in Turner & Neugebauer, GDNM. This starts with a list of financial accounts, perhaps owed by the members of a gymnasium, on which interest is charged at the implied rate of two per cent per month; then the second half of the text is a lunar calendar. This calendar omits the meteorological, astronomical, and religious data that is a feature of most of the surviving Greek parapegmata and, in the preserved section at least, and despite what it says in its introduction, it simply lists lunar months. The text reads:

Year 1 of Queen Cleopatra and King Ptolemy the son, gods Epiphaneis. Parapegma of lunar new moons, showing how they are related to the days of the Egyptian twelvemonth. The period of the table is twenty-five years, three hundred and nine

months [including intercalary months], nine thousand one hundred and twenty-five days. It indicates the lunar months and which of them are full [i.e. 30 days long], which hollow [i.e. 29 days long], which intercalated [i.e. the 13th month]; and in what sign of the Zodiac the sun will be during each month. When the sun has traversed the twenty-five years it will return to the same starting point and revolve in the same manner. . . . The lunar new moons in the first year are

| | | |
|---|---|---|
| Thoth 20 | [29 days] | |
| Phaophi 19 | [30 days] | |
| Hathyr 19 | [30 days] | |
| Choiak 19 | [29 days] | |

. . . .

The text containing the month lengths, given at the end in square brackets, is lost at this point through damage to the manuscript, but a later similar fragment allows it to be reconstructed with confidence. If the calendar had continued for twenty-five years, it would have stretched over some two metres of roll; only about half a metre survives. Another, more extensive, example of this same twenty-five year lunar calendar is found in the demotic papyrus P. Carlsberg 9 (see Neugebauer & Parker, *EAT* iii, 220–5, Neugebauer, *HAMA* i, 563f., and Parker, *CAE*). There are good reasons for believing that this cycle was developed in Egypt in the fourth century BC (see Parker, *CAE*). Ptolemy refers to it much later, at *Almagest* VI 2:

Now 25 Egyptian years less 0; 2, 47, 5 days contain approximately [in fact, exactly!] an integer number [309] of [mean synodic] months, and [in 25 years] the mean notions are . . . , [translation and annotation from Ptolemy–Toomer, *PA*, 276]

and Ptolemy's table of conjunctions and oppositions which then follows, together with some calculations, is based on this cycle. Other more accurate but less convenient cyclical calendars, the Metonic cycle and its refinement by Callippus, were described in Section 2.4(b), and an extract from another parapegma is illustrated in Plate 7 and described in Section 7.1(c).

These calendars are precisely the raw material of what my Eudoxus called 'astronomical ratios' (see $B_{60}$ and $E_{61}$): any sequence that describes the patterns generated by interacting periodic phenomena. My dialogue and these examples both deploy two equivalent ways of describing the lengths of successive lunar months: either as a sequence which describes the succession of new days and new moons, as in the left-hand column of P. Ryl. iv 589, which is a condensed form of my kind of description as

. . . MDDD . . . (29 times) . . . DDMDD . . . (30 times) . . . DDMDD . . . ,

or in terms of the length of the months, as in the right-hand column, for which I introduced the notation $\{ \ldots, 29, 30, 30, 29, \ldots \}$. There is not enough text in P. Ryl. iv 589 for us to continue this sequence uninterrupted for very long, but the example of P. Carlsberg 9 is complete. From it we can reconstruct the

relationship between this lunar calendar (in which a period of 9125 days is split into 309 mean synodic months) and a civil zodiacal calendar. (It is skilfully argued in Parker, *CAE*, that the lunar calendar was originally closely tied to actual lunar observations, though anomaly and the underlying approximation, of $\frac{9125}{309} = 29.5307\ldots$ for the mean synodic month of 29.530 59 days, mean that the fit can never have been exact and that it will drift with the passage of time.) From this, a complete enumeration of the astronomical ratio of day to month that it embodies could be extracted, in the form of a list of the lengths of the 309 successive months.

Our evidence about early Egyptian and Greek astronomy contains many other such examples of enumerations of almost-periodic astronomical phenomena; I shall not press this point further, but again refer the reader to Neugebauer, *HAMA*. Let us now look to the next stage.

There are two main directions in which we can turn. The first is to explore, observationally and computationally, the irregularities of the heavens: the anomalies, the erratic behaviour of the planets, the fine details of the observable universe. The tradition that stretches from the astronomers of Babylonia, through Hipparchus and Ptolemy, then the Arabs, up to the sixteenth century and on to today, one of the gigantic intellectual achievements of mankind, is a massive programme of such an exploration. Note how this kind of astronomy has, throughout, been carried out within an arithmetised tradition, and indeed it can scarcely be conceived without the use of such arithmetic for its records and calculations. In this approach the cyclical phenomena which so dominate the early stages of observation atrophy and fall away, just as the original lunar calendars give way to the much more abstract but observationally and computationally convenient solar calendars. But there is a second route which beckons. Since the astronomical data so often presents itself in cyclical form, an understanding of the behaviour of interacting cycles might lead directly to the heart of the phenomena of astronomy themselves. Within the programme I have been developing, this suggests that we look for some *anthyphairetic* understanding of *astronomical* ratios. The first step in such an understanding would be merely to relate these two ways of describing the same thing. Until this first step is achieved, we do not know what further illumination is possible.

What stifles this second enterprise is its subtlety and difficulty. The problem is not hard to formulate ($E_{55}$–$B_{60}$), the initial explorations are promising ($E_{61-6}$), but the next steps, for example the two algorithms described in the dialogue at $E_{67}$ and $B_{68}$, are far too ingenious for everyday use. (The algorithm in $E_{67}$ is found, independently and almost simultaneously, in E. B. Christoffel, OA (1875) and H. J. S. Smith, NCF (1876) and is also described in my AREP; the algorithm in $B_{68}$ is due to Christopher Zeeman; see Zeeman, GG, and Series, GMN.) I suggest that Plato's and Eudoxus' approach to astronomy may have been to attempt to explore further in this direction. We have little direct evidence about whether they might have achieved any results, such as $E_{67}$ and

$B_{68}$, but these ideas and procedures are very closely related to the techniques of *Elements* V and their applications in Book VI.

We retire from these tentative explorations of this kind of investigation of interacting, almost-periodic phenomena with profound insights but few results beyond the "formal definitions and propositions about ratio and proportion" ($E_{57}$, my unconventional but defensible description of *Elements* V). The original motivating problems (to understand and manipulate cyclical phenomena via their astronomical and anthyphairetic ratios) recede, unanswered; the formal machinery becomes the theory that survives and takes on an independent life of its own, and a new phase of development begins. For a good illustration of how this formal theory can then be used in geometry, see the proof of *Elements* VI 1.

This is only a sketch of the first steps of a proposal for reconstructing some of the astronomy of Plato's circle, a conceptual framework within which to fit our fragmentary and tentative evidence. (Some of this evidence is negative: for example, we perhaps need to account for the lack of any further astronomy of consequence in Egypt and in the post-Eudoxan tradition.) I have not here begun to introduce any geometrical or kinematic considerations that might enrich and complicate further the point of view, as Glaucon recognises at *Republic* VII, 530b–c (a passage that was quoted at the beginning of this section); for an illustration of what might be involved here, see Riddell, EMES. But I believe that it does offer a plausible interpretation that is compatible with Plato's remarks on astronomy within the Academy, our evidence about Eudoxus' achievements, and other early developments in both Greek and Egyptian astronomy.

## 4.5 ACADEMIC MUSIC THEORY

### 4.5(a) *Introduction*

Plato's astronomy and music theory (*harmonikē*) are "kindred sciences", both of which should be studied not by experiment and observation but by "ascending to generalised problems". Music theory is the last topic in the curriculum, *Republic* VII, 530d–531c:

[SOCRATES:] We may venture to suppose that as the eyes are framed for astronomy so the ears are framed for the movements of harmony; and these are in some sort kindred sciences, as the Pythagoreans affirm and we admit, do we not. . . . Then, since the task is so great, shall we not enquire of them what their opinion is and whether they have anything to add? And we in all this will be on the watch for what concerns us, [which is] to prevent our fosterlings from attempting to learn anything that does not conduce to the end we have in view, and does not always come out at what we said ought to be the goal of everything, as we were just now saying about astronomy. Or do you not know that they repeat the same procedure in the case of harmonies? They transfer it to hearing and measure audible concords and sounds against one another, expending much useless labour just as the astronomers do.

[GLAUCON:] Yes, by heaven, and most absurdly too. They talk of something they call minims and, laying their ears alongside, as if trying to catch a voice from next door, some affirm that they can hear a note between and that this is the least interval and the unit of measurement, while others insist that the strings now render identical sounds, both preferring their ears to their minds.

[SOCRATES:] You are speaking of the worthies who vex and torture the strings and rack them on the pegs; but—not to draw out the comparison with strokes of the plectrum and the musician's complaints of too responsive and too reluctant strings—I drop the figure, and tell you that I do not mean these people, but those others whom we just now said we would interrogate about harmony. Their method exactly corresponds to that of the astronomer; for the numbers they seek are those found in these heard concords, but they do not ascend to generalised problems and the consideration which numbers are inherently concordant and which not and why in each case.

[GLAUCON:] A superhuman task.

[SOCRATES:] Say, rather, useful for the investigation of the beautiful and the good, but if otherwise pursued, useless.

[GLAUCON:] That is likely.

Plato distances himself even from the Pythagorean tradition of music theory in a way that leaves us with very little evidence of any kind on which to base our reconstruction; at most we have a single text, the *Sectio Canonis*, which I shall describe, that may give us a later and flawed example of "generalised problems and the consideration of which numbers are inherently concordant and which not and why in each case". Any interpretation must therefore be very speculative; I shall present here a collection of ideas that relate the basic objects and operations of music theory to the development of *logistikē*, and will also include some related topics of wider interest. The exposition will be in another dialogue, between Archytas and the slaveboy, and again I beg the reader to ignore any distortions of chronology or social order that this introduces.

### 4.5(b)  *Archytas meets the slaveboy*

ARCHYTAS70: Do you have a few minutes to spare? Ever since Plato started sending me details of your discussions on anthyphairesis, I've been wanting to talk to you. I'd like to see what you think about a suggestion of mine on how music theory may help us to understand more about the behaviour of ratios.

BOY71: I'd be delighted. Meno lets me spend as much time as I want on mathematics now.

ARCHYTAS72: Have you been able to make any progress with the problem of compounding anthyphairetic ratios? [See $B_{32}$–$S_{35}$, with its modification $B'_{32-4}$ in Section 4.2 above.]

BOY73: Not much. I can only work out examples by ugly and crude methods that rest on some rather dubious operations, but I can't understand much or prove anything.

ARCHYTAS74: Would you tell me more?

BOY75: Let's take an example like compounding $\sqrt{2}:1 = [1,\overline{2}]$ and $\sqrt{3}:1 = [1,\overline{1,2}]$. First work out some convergents of the ratios:

$$1:1 < 7:5 < 41:29 < 239:169 < \cdots \sqrt{2}:1 \cdots < 577:408 < 99:70 < 17:12 < 3:2$$
$$1:1 < 5:3 < 19:11 < 71:41 < \cdots \sqrt{3}:1 \cdots < 97:56 < 26:15 < 7:4 < 2:1.$$

Now multiply together the antecedents and the consequents of any two under-estimates and work out the anthyphairesis of the result; and then do the same for two overestimates. We might as well ignore all but the best approximations, so this gives

$$239 \times 71 : 169 \times 41 < ? < 577 \times 97 : 408 \times 56$$

i.e. $$16969 : 6929 < ? < 55969 : 22848$$

i.e. $$[2,2,4,2,2,\ldots] < ? < [2,2,4,2,6,\ldots].$$

Hence the new unknown ratio in the middle looks as if it really will be $[2,\overline{2,4}] = \sqrt{6}:1$. But I can't explain what's going on, let alone prove anything. What's more, I can't really see what it means to compound two commensurable ratios $p:q$ and $r:s$ to give $(p \times r):(q \times s)$. If we add, giving $(p+r):(q+s)$, then we get the basic operation of the *Parmenides* proposition which I can understand much better. For example, it is a mean, while compounding isn't; and it is fundamental to anthyphairesis, which I now understand better and better, but compounding doesn't seem to have much connection with anthyphairesis. If both of the ratios are multiples, $n:1$ and $m:1$, or both are parts, $1:n$ and $1:m$, then compounding involves the product of the *arithmoi* $n$ and $m$, and it is possible that, in the more general case, it may be an extension of this kind of result. Next, I can use compounding to say something about the ratio of two rectangular numbers [see *Elements* VIII 5], with a similar geometrical result about two rectangles, or about two equiangular parallelograms [see *Elements* VI 23]; but these seem isolated results of no great importance. Finally, while I can prove the particular geometrical relations like $\sqrt{2} \cdot \sqrt{3} = \sqrt{6} \cdot 1$ that are suggested by the manipulation above, there's something curious about the underlying geometry that I don't understand and can't prove. That's a summary of just about everything I can say about compounding.

ARCHYTAS$_{76}$: Have you tried exploring what happens with commensurable ratios?

BOY$_{77}$: No. So I'm glad to have met you because they've told me that you've done a lot with these ratios of *arithmoi*.

ARCHYTAS$_{78}$: Well, my idea is that if you look at this operation in the context of the simplest kind of commensurable ratios, then you might get some help from music theory. Let me explain. When Plato first started to tell me about the need for a mathematical 'definition' of ratio, I couldn't understand what he was talking about. Surely, I then thought, a ratio $b:c$ is just a convenient word for 'the relationship between two lines (or numbers) $b$ and $c$' or 'the separation between two sounds'; does it really need any more definition than that? And if $b$ can be expressed as $m$ parts of something, and $c$ as $n$ parts of the same thing, then isn't the ratio $b:c$ just $m:n$?

BOY$_{79}$: If you'd said that in Athens, you'd have had people like Socrates asking you awkward but interesting and fruitful questions! For instance, you haven't really said what $m:n$ is.

ARCHYTAS$_{80}$: So they keep telling me, and so I learned from the account of your second meeting with Socrates! Now, although the music theorists don't usually use the word

*logos*, but talk instead of *diastēma*, 'intervals', I'll treat musical intervals as if they are ratios. The most important intervals in music are the octave $2:1$, twelfth $3:1$, and double octave $4:1$, and the fifth $3:2$, fourth $4:3$, and tone $9:8$; and all of them except the last are consonant intervals. These all correspond either to multiple ratios $n:1 = [n]$ or to the so-called epimoric ratios $(n+1):n = [1,n]$, so they're particularly simple kinds of anthyphairetic ratios, of a form that is specified by just one number $n$. I might also add that the next such kind of simple anthyphairetic ratio is $[1,1,n] = (2n+1):(n+1)$ (where $n \geqslant 2$, since $[1,1,1] = [1,2]$ is an epimoric ratio), and this doesn't seem to be of any importance in music.

BOY[81]: Much as I am a champion of anthyphairesis, I feel here that there is a more commonplace explanation for this word *epimorion*, 'a part in addition', that it describes the simplest kinds of fractional quantities greater than one, namely $1 + n'$ [see Chapter 7]. However, we can push your anthyphairetic explanation a bit further. Take a general ratio $b:c = [n_0, n_1, n_2, \dots]$ and try to work out just what it is about $b$ and $c$ that is described by each of the $n_k$s. You'll see that the connection is very direct in the case of $n_0$ and $n_1$, but it starts to get rather remote for $n_2$ and beyond. For example, it's immediately clear why $[2] = 2:1$, $[1,3] = 4:3$, and $[2,3] = 7:3$, etc., but it's rather harder to see why $[2,3,4] = 30:13$ without thinking in detail of the subtraction process. Do you see what I mean? So these ratios from music theory are also simple in the sense that the relations between the antecedent, the consequent, and the terms of the anthyphairesis are all absolutely transparent. In this sense, ratios like $[1,1,n]$ are also getting rather complicated.

ARCHYTAS[82]: Yes, when Plato sent me the account of your second conversation with Socrates, I had to think for a long time to work out how you could compare the ratios just by looking at the terms of the anthyphairesis [see $S_{17-21}$ and $S'_{21}$ in Section 2.3(b)], since I couldn't fathom out what these later $n_k$ meant. Then I saw how each was associated with a remainder that came from either the antecedent or the consequent, according as $k$ was even or odd, just as you said there, and that made all the difference.

Back to music theory: we see that these basic intervals in music theory all correspond to these simplest kinds of anthyphairetic ratios, each of them described by a single number $n$, in either the fractional or the anthyphairetic sense. Next let's look at what the basic operations of music theory do to these intervals. In music we adjoin or subtract two intervals to make a new interval, rather as lines or other magnitudes are added or subtracted in geometry, and in this way we can generate all the basic intervals of music out of fourths and fifths. For instance:

| | |
|---|---|
| a fourth and fifth give an octave, | i.e. $4:3$ and $3:2$ give $2:1$, |
| a fifth and an octave give a twelfth, | i.e. $3:2$ and $2:1$ give $3:1$, |
| two octaves together give a double octave, | i.e. $2:1$ and $2:1$ give $4:1$, |
| and a fifth less a fourth gives a tone, | i.e. $3:2$ less $4:3$ gives $9:8$. |

Now these operations correspond to compounding the ratios (or the opposite of compounding when the intervals are subtracted) and all of these examples of compounding (but not its opposite) give rise to concordant ratios.

BOY$_{83}$: Are all concordant ratios multiple or epimoric?

ARCHYTAS$_{84}$: An interesting question that I'll leave to one side, since I've now got to my main point. I've seen how Eudoxus' ideas on astronomical ratios have led to mathematical insights into the relation between astronomical and anthyphairetic ratios, and to some elegant mathematical definitions and proofs [see E$_{57}$ and *Elements* V]; is there any possibility that these musical analogies might give some other such insights into properties of ratios? For example, consider the problem of compounding anthyphairetic ratios. Surely it must be simpler to guess or work out what the rule should be by looking at simple examples like

$$[1, 3] \text{ and } [1, 2] \text{ give } [2],$$
$$[1, 2] \text{ and } [2] \text{ give } [3], \text{ and}$$
$$[1, 3] \text{ and } [1, 8] \text{ give } [1, 2],$$

which are precisely the examples arising from music theory, and perhaps also using some further insights from music theory, than by studying your example

$$[1, 2, 2, 2, \dots] \text{ and } [1, 1, 2, 1, 2, \dots] \text{ give } [2, 2, 4, 2, 4, \dots].$$

So if we investigate music to see why certain basic intervals are concordant, and how and why new concordant intervals are generated out of this basic set, we may get closer to understanding the operation of compounding ratios.

BOY$_{85}$: I'm not so sure it will help, but it's certainly worth trying. Let me explain further why this particular problem may be of special interest. Plato is looking for some kind of absolute comparison against a standard [see *Statesman* 283c–284d], and I think that ratios might give some mathematical illustration of this, if only we could find how to add and multiply them. For example, our basic results about inequalities between ratios tell us that

$$[1, \overline{2}] < [2, \overline{1, 4}] < [4, \overline{4, 8}] < [5, \overline{1, 1, 1, 10}],$$

but, until we find out how to relate the anthyphairesis of $2b : c$ with $b : c$, we don't have any general way of seeing that the second ratio is *exactly* twice the first, the third is *exactly* three times the first, and so on. Or again, if we could find out how to add and subtract ratios in general then, since we can already compare two ratios for size, we could define ratios of ratios. I don't know quite what this would mean, but I think our hope of understanding would increase if we could first perform the operations and explore their behaviour. So I'm particularly interested in arithmetic with ratios.

ARCHYTAS$_{86}$: What do you mean here by 'arithmetic'? It doesn't seem to have much to do with *arithmētikē*, the study of the *arithmoi*.

BOY$_{87}$: It's part of a more general point of view. Just as *arithmētikē* grows out of the properties of the *arithmoi*—their addition and multiplication, the subtraction of the less from the greater, the unit, oddness, evenness, and primality, and so on—so by analogy I'm coming to think of the study of adding, subtracting, multiplying, etc. of lots of other kinds of things as looking for their own characteristic underlying 'arithmetic'. This arranges the sciences into some kind of hierarchy. For example, all aspects of the arithmetic of anthyphairetic ratios seem problematic while, for geometrical magnitudes, especially for straight lines, addition and subtraction are straightforward, but multiplication seems to be a different kind of operation with its own subtleties. That's connected with what I was saying at the beginning about

$\sqrt{2} \cdot \sqrt{3} = \sqrt{6} \cdot 1$, or the general result about $\sqrt{n} \cdot \sqrt{m} = \sqrt{nm} \cdot 1$ for any two *arithmoi* $n$ and $m$; and there are similar problems about arithmetic with edges of cubes, and perhaps yet other kinds of things.

ARCHYTAS[88]: And in music theory we can also talk about the greater and less, and we can put together and take away musical intervals in an operation that we describe as if it were addition and subtraction, though it seems to be more analogous to multiplication and a more general kind of division than that of addition and subtraction of *arithmoi* or lines; indeed there seems to be nothing in music theory that would correspond to that addition and subtraction. (So, incidentally, we can define the anthyphairesis of musical ratios, just as you were asking above [in B[85]]. Here's the example of $2:1$ and $3:2$.

> $2:1$ less $3:2$ once leaves $4:3$, smaller than $3:2$,
>
> $3:2$ less $4:3$ once leaves $9:8$, smaller than $4:3$,
>
> $4:3$ less $9:8$ twice leaves $256:243$, smaller than $9:8$,
>
> $9:8$ less $256:243$ twice leaves $531\,441:524\,288$, smaller than $256:243$,
>
> etc.,

so the musical anthyphairetic ratio of the ratio $2:1$ and $3:2$ is $[\![1,1,2,2,\ldots]\!]$. I can easily prove that this won't terminate, since the remainder will always be either $3 \times 3 \times \cdots \times 3:2 \times 2 \times \cdots \times 2$, or the other way round, so it can't be $1:1$; but I can't say anything more about it. Or we could define a musical version of Eudoxus' astronomical ratios. For example $3:2$ could be described by the pattern of

$$1 \quad 2 \quad 4 \quad 8 \quad 16 \quad 32 \quad 64 \quad 128 \qquad 256 \quad 512 \qquad 1024 \quad 2048 \quad \ldots$$
$$1 \quad 3 \qquad 9 \quad 27 \qquad 81 \qquad 243 \qquad\qquad 729 \qquad\qquad\qquad 2187 \ldots$$

So now $3:2 = \{\{*1*, 2, 1, 2, 1, 2, 2, 1, \ldots\}\}$. Can you make anything of these manipulations?) On the other hand, music only involves ratios of *arithmoi*, and that sometimes restricts the possibilities; for example, it means that the tone $9:8$ cannot be divided into two equal musical intervals. I see how you are saying that these different subjects—geometry, astronomy, music theory, and I'm sure we'll find others—seem to be growing out of a considerable body of overlapping ideas about ratio, and information about one branch seems to feed back into the other topics.

But I'm completely baffled about your remarks about $\sqrt{2} \cdot \sqrt{3} = \sqrt{6} \cdot 1$. What are you talking about there?

BOY[89]: Let me spell that out, since you might be unfamiliar with the definitions and shorthand that I'm using. We choose some assigned line $a$, and let $b, c, d, \ldots$ be any other lines. Geometrical addition $b + c$ corresponds to adjoining lines, and we can similarly adjoin areas and volumes. Now consider geometrical multiplication of lines. Let $b \cdot c$ denote the rectangle with sides $b$ and $c$; it's one kind of geometrical multiplication that is inhomogeneous (that is, $b \cdot c$ is a different kind of geometrical magnitude from $b$ and $c$) and restricted (since we can't define the product of four lines, or two rectangles). There is another kind of geometrical multiplication, where we apply $b \cdot c$ as a rectangle to the assigned line $a$; the width of this rectangle will be a homogeneous unrestricted product of $b$ and $c$. This satisfies all the same kinds of properties of multiplication as the multiplication of *arithmoi*, and much more, but it now depends on an arbitrary choice of $a$.
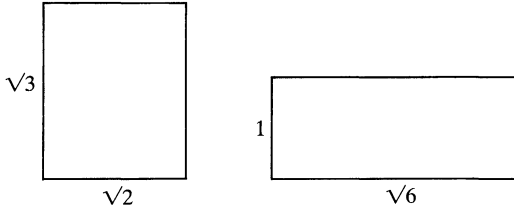
FIG. 4.1

ARCHYTAS[90]: Surely that's just the fourth proportional to $b$, $c$, and $a$ [see *Elements* VI 12]?

BOY[91]: Yes! And you see here how the assigned line $a$ is now analogous to the unit in *arithmētikē*, and how this kind of multiplication of lines is more general than the multiplication of *arithmoi*, since the geometrical fourth proportion always exists, while it doesn't necessarily exist in *arithmētikē*. There are lots of interesting developments relating these two kinds of geometrical multiplication that Theaetetus has started exploring in order to extend the scope of our anthyphairetic knowledge of the dimensions of squares, but it soon gets very complicated. [See *Elements* X, to be described in Chapter 5.] Here, I'll just consider the rectangular multiplication $b.c$. Given $b$ and $c$, we can define $\sqrt{(b.c)}$ to be the side of the square equal to $b.c$—call it the 'square side' of $b$ and $c$; it's the mean proportional (or geometric mean) of $b$ and $c$. Now, in a straightforward example or illustration, I tend to refer to the assigned line and the square on it as the unit line and the unit square, or even omit mention of them altogether, and so just say 'the square side of 2' or '$\sqrt{2}$' for $\sqrt{(2a.a)}$, etc. I've studied closely your amazing geometrical configuration for doubling, tripling, etc. the cube, in connection with the seemingly intractable problem of the dimensions of cubes; you could use this notation there and write $\sqrt[3]{2}$ for the 'cube edge' of $2a.a.a$.

ARCHYTAS[92]: So this result $\sqrt{2}.\sqrt{3} = \sqrt{6}.1$ is just a statement about rectangles, that the rectangle with sides $\sqrt{2}$ and $\sqrt{3}$ is equal to the rectangle with sides $\sqrt{6}$ and 1 [see Fig. 4.1]. Or we could say that the homogeneous product of $\sqrt{2}$ and $\sqrt{3}$ is $\sqrt{6}$.
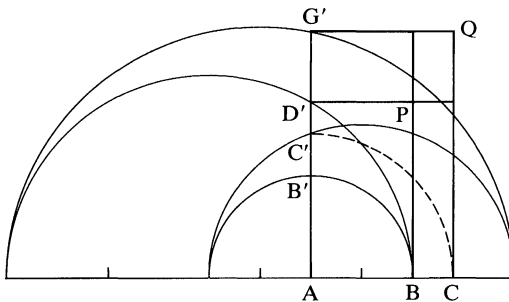


FIG. 4.2
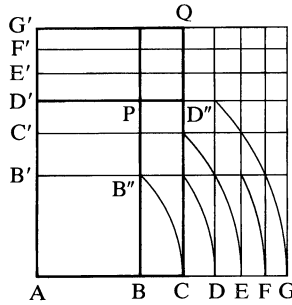
FIG. 4.3

BOY93: Exactly; and I'll explain how this result gives yet another illustration of a profound link between *arithmētikē* and *geōmetrikē*. One direct assault on the problem leads to this interesting configuration [see Fig. 4.2], in which we start with the assigned line AB = 1 and construct [using *Elements* II 14] $AC' = AC = \sqrt{2}$ and $\sqrt{AD'} = \sqrt{3}$; then [using *Elements* III 35], we construct $AG' = \sqrt{6}$. Hence we want to show that $AC . AD' = AG' . AB$. If we subtract the rectangle ABPD' from both and apply the standard result on complements about a diagonal [*Elements* I 43; see Fig. 1.2, above], we see that $\sqrt{2} . \sqrt{3} = \sqrt{6} . 1$ means that the point P lies on the diagonal AQ. But I can't prove anything from this figure, though I'm still looking. A more successful proof uses gnomons of a square [see Fig. 4.3]. We start with the unit line AB and unit square ABB''B', and put unit gnomons round it, in the usual way, thus generating lines $AC = \sqrt{2}$, $AD = \sqrt{3}$, etc. (There's an easy way of doing that by setting $AC = AB''$, etc., or we can use many other such relations like $AD'' = AG$.) Then $\sqrt{2} . \sqrt{3}$ is a rectangle with sides AC and AD', while $\sqrt{6} : 1$ is a rectangle with sides AB and AG', for instance. If we subtract the overlapping rectangle ABPD' from both, then again we see that we have to prove that the point P lies on the diagonal AQ. This will hold if the angles PAD' and QAG' are equal. Now these triangles have their sides proportional: we start with a line (AB or AC), construct the square side of three-times the square on these lines (AD' or AG') and construct triangles with these two sides around the right angle (AD'P or AG'Q). Hence these triangles have their sides proportional, and so [see *Elements* VI 5] their angles will be equal. But there's something strange about this kind of argument. To take the easiest possible example, why is it that the angles of any equilateral triangle are always equal to a third of two right angles? Perhaps we have to postulate this.

ARCHYTAS94: Surely the angles of an equilateral triangle are all equal and therefore, since the sum of . . .

BOY95: Yes, but just why is the sum of the angles of a triangle two right angles? The astronomers have started describing geometrical models of the earth and heavens in terms of nested spheres. Suppose we look at equilateral triangles whose edges are great circles on a sphere, precisely the kind of thing these astronomers sometimes consider. We can still talk about angles and triangles, and the angles of a very small equilateral triangle are indeed each still approximately two-thirds of a right angle but, as the triangle grows, so also the angles grow. If one vertex is at a pole and the other two are on the equator, the angles will each be a right angle; then as the triangle expands over

the sphere, so the angles can grow to almost three and a third right angles each. This example doesn't *prove* anything, but it certainly makes you think. I recall that the first thing Socrates said to me was that a square could be "either larger or smaller", and I was naïve enough to let it pass [see *Meno* 82c]. I wonder what he would have answered if I'd questioned him on this point.

ARCHYTAS[96]: I doubt if even all of your evident mathematical skills could persuade Socrates to answer questions, rather than ask them. Alas, I must be getting on. Farewell.

### 4.5(c)   *Compounding ratios*

Addition and multiplication of common fractions,

$$\frac{m}{n} + \frac{p}{q} = \frac{mq + np}{nq} \quad \text{and} \quad \frac{m}{n} \times \frac{p}{q} = \frac{mp}{nq},$$

are operations so basic to arithmetised mathematics that it may be difficult for us today to conceive of a mathematics of any sophistication in which they are unknown or unimportant. But, as my dialogue explains—see especially B[75]—I believe that this may indeed have been the case for early Greek mathematics. There is a manipulation in the *Elements*, generally described in translation by the word 'compounding', that corresponds to the multiplication of two common fractions, but Euclid does not develop or exploit its properties; and there is nothing that deals with the addition of common fractions. In the next section I shall suggest that the *Sectio Canonis* might have been an attempt, albeit obscure and irretrievably flawed, to explore the behaviour of compounding through the study of music, and music may have been the one place in ancient science and mathematics where the operation arose naturally; the idea behind this proposal is set out in A[76–84]. Then, in Chapter 7, I shall describe our evidence about arithmetical calculations, and argue that a commonly accepted conclusion, that Greek mathematicians conceived and sometimes even expressed fractional manipulations in the form of common fractions, written upside-down, is far from being an established fact. In this section, I shall now discuss some questions of the interpretation of some relevant passages of the *Elements*.

Our assumptions about the role of these manipulations with common fractions within mathematics are pervasive and often not articulated, so I shall start with a brief statement of the point of view against which I am arguing. This is taken from van der Waerden, *SA*, where the interpretations are always set out with exemplary clarity. We read there, with the original emphases:

It is probable that it was *calculation with fractions* which led to the setting up of this theory [sc. *Elements* VII]. Fractions do not occur in the official Greek mathematics before Archimedes, but, in practice, commercial calculations had of course to use them. The reason why fractions were eliminated from the theory is the theoretical *indivisibility of unity*. In the *Republic* (525e), Plato says: "For you are doubtless aware

that experts in this study, if anyone attempts to cut up the 'one' in argument, laugh at him and refuse to allow it; but if *you* mince it up, *they* multiply, always on guard lest the one should appear to be not one but a multiplicity of parts." . . . When fractions are thus thrown out of the pure theory of numbers, the question arises whether it is not possible to create a mathematical equivalent of the concept fraction and thus to establish a theoretical foundation for computing with fractions. Indeed, this equivalent is found in the ratio of numbers [again, sc. *Elements* VII]. . . . The art of computing, including the calculation with fractions, is called 'logistics' by the Greeks [pp. 115–16; also see pp. 49–50].

Contrast this description with an illustration of a Euclidean manipulation with proportions, taken from *Elements* XIII 11. I have replaced the references to the particular lines involved by the letters $b$, $c$, $d$, and $e$; otherwise the text is taken verbatim from Heath, *TBEE* iii, 462–3:

. . . therefore, proportionally, as $b$ is to $c$, so is $d$ to $e$. And doubles of the antecedents may be taken; therefore, as the double of $b$ is to $c$, so is the double of $d$ to $e$. But, as the double of $d$ is to $e$, so is $d$ to the half of $e$; therefore also, as the double of $b$ is to $c$, so is $d$ to the half of $e$. And the halves of the consequents may be taken; therefore, as the double of $b$ is to the half of $c$, so is $d$ to the fourth of $e$. And $b_1$ is double of $b$, $c_1$ is half of $c$, and $e_1$ a fourth part of $e$; therefore as $b_1$ is to $c_1$, so is $d$ to $e_1$. *Componendo* also, as the sum of $b_1$, $c_1$ is to $c_1$, so is [the sum of $d$ and $e_1$] to $e_1$; therefore also, as the square on the sum of $b_1$, $c_1$ is to the square on $c_1$, so is the square on [the sum of $d$ and $e_1$] to the square on $e_1$.

This is taken from a proposition in which the theory of proportion is applied fluently, as we see from the last two steps and the further passage quoted below. Yet Euclid takes four steps to deduce from $b:c::d:e$ that $2b:\frac{1}{2}c::d:\frac{1}{4}e$, in a context where the fractional quantities $\frac{1}{2}c$ and $\frac{1}{4}e$ pose no problems since they correspond to labelled lines in the figure. Compare this with the manipulation *componendo*, that

$$\text{if } x:y::z:w \quad \text{then} \quad (x+y):y::(z+w):w,$$

(see V Definition 14 and Propositions 17 and 18; this operation is also known as 'composition') and the even more drastic step that

$$\text{if } x:y::z:w \quad \text{then} \quad x^2:y^2::z^2:w^2$$

(see VI 22, with the long note in Heath, *TBEE* ii, 242 ff.) which go through without comment. We see a similar confident manipulation later in the same proof (Heath, *TBEE* iii, 464) where Euclid argues, for different lines $b$, $c$, and $d$:

Let the square on $b$ be equal to that by which the square on $c$ is greater than the square on $d$; therefore the square on $c$ is greater than the square on $d$ by the square on $b$. . . . And since the square on $c$ is five times the square on $d$, therefore the square on $c$ has to the square on $d$ the ratio which 5 has to one. Therefore, *convertendo*, the square on $c$ has to the square on $b$ the ratio which 5 has to 4, . . .

Here Euclid justifies the last step by his reference to conversion of ratios (V Definition 16 and Proposition 19):

$$\text{if } x:y::z:w \quad \text{with} \quad z<x \quad \text{and} \quad w<y \quad \text{then} \quad x:y::(x-z):(y-w).$$

It is not immediately clear how the argument is meant to go; for example, all the magnitudes $x$, $y$, $z$, and $w$ need to be homogeneous for V 19 to be applied whereas here, in the immediate context of XIII 11, they are not. (Incidentally, this passage is the only place I know of where the manuscripts of the *Elements* use numerals rather than words to denote numbers; see Chapter 7, below.) Yet despite the fluency in these apparently more complicated kinds of manipulations, the operations of doubling the antecedents,

$$\text{if } x:y::z:w \quad \text{then} \quad 2x:y::2z:w,$$

and asserting that

$$\text{if } x:y::z:w \quad \text{then} \quad 2x:y::z:\tfrac{1}{2}w$$

seem to require special and individual explanation, though their formal justification follows immediately from V 4. The intuition behind Euclid's treatment does not seem to be based on an understanding of the manipulation of common fractions, where these operations like doubling or halving the antecedent or consequent are straightforward; by contrast, these operations seem intractable for anthyphairetic and astronomical ratios (see B₈₅ and Section 9.3(a), below). Conversely, the operations *componendo* and *convertendo* (though not composition) are more straightforward and natural for anthyphairetic and astronomical ratios, but less so for common fractions.

There is nothing in the *Elements* that seems remotely connected with the general addition of two ratios. While composition can be conceived in terms of the addition of $x:y$ and $y:y$ to give $x+y:y$, there is no hint of this point of view, and the closely related manipulations of conversion and of V 12, that

$$\text{if } x:y::z:w \quad \text{then} \quad x:y::(x+z):(y+w),$$

(which is the *Parmenides* proposition in the case of equality) are very emphatically not directly related to the addition and subtraction of fractions, whatever the erroneous feeling among novices that they ought to be; also see the contrast between $(p\times q):(r\times s)$ and $(p+q):(r+s)$ described in A₇₅.

Let me now turn to the compounding of ratios—which is not to be confused with the composition of ratios though, as we shall see, they share the same Greek name. This does correspond to our multiplication of common fractions. The operation appears in *Elements* VI Definition 5 (which is almost certainly interpolated), VI 23, and VIII 5, all of which were quoted at the end of Section 1.2(d); they will also be quoted again below. The procedure, as it appears in VI 23 and VIII 5, but not VI Definition 5, can be described as follows. Given two ratios $b:c$ and $d:e$—and note yet once again, in passing, that Euclid gives no mathematically serviceable definition of ratio in the *Elements* adequate for a proof of VI 1, the *Topics* proposition, though he does, here and elsewhere, appeal to the idea of ratio—if we then write

$$b:c::x:y \quad \text{and} \quad d:e::y:z,$$

then $x : z$ is called the compound of $b : c$ and $d : e$. To bring out the connection with the multiplication of common fractions, we can carry out this operation by writing

$$b : c : : b \times d : c \times d \quad \text{and} \quad d : e : : c \times d : c \times e$$

(where the products $b \times d$, etc. are interpreted appropriately; this is a version of the *Topics* proposition). Then the compound of $b : c$ and $d : e$ is $b \times d : c \times e$. When $b$, $c$, $d$, and $e$ are lines, as in VI 23, then the lines $y$ and $z$ can be found for any line $x$, by the basic construction of a fourth proportional at VI 12; but when they are numbers, as in VIII 5, the operation requires the idea of continued proportion, the subject of VIII 4 and, indeed, of most of Book VIII. On the other hand, the garbled definition at VI Definition 5 seems to be striving towards an arithmetised operation, in which the two ratios are 'multiplied together' in some sense.

The role of compounding within the *Elements* is analysed in great detail in Mueller, *PMDSEE* and, to avoid an enormously long excursion through the evidence and its interpretations, I shall restrict myself to quoting some of his conclusions, with which I concur, and refer the reader to this book for the detailed analyses:

It is possible to say that compounding ratios is an analogue of multiplying fractions. The serious question for interpretation is whether compounding should be viewed as a representation of multiplying, i.e., as a device for representing the multiplication of fractions in the language of proportionality. I shall be arguing that compounding should not be viewed in this way [p. 88, on VI Definition 6].

[Euclid's] failure to exploit VIII 5 is a good indication that he does not construe compounding as a representation of multiplying. . . . Since there is no other candidate for the representation of the multiplication of fractions, it also seems unlikely that Euclid construes duplicating and triplicating as squaring and cubing [fractions] [pp. 92–3, on VIII 11 and 12].

It is not, of course, possible to show that Euclid does not have calculations or something like them in mind when he deals with ratios. But he certainly does not make any such concern explicit, and, as we shall see, the way he proceeds in applying ratios suggests anything but a calculational model. In particular, although he uses the compounding of ratios in later books, he shows no clear sense of its relation to duplicating and presumably, therefore, none of its relation to multiplication [pp. 135–6, on proportion, ratio, and calculation in Book V].

In general the geometric books confirm the impression gained from the arithmetic ones [VII–IX] that Euclid does not construe compounding as multiplication. VI 23 itself is, in a sense, evidence of this fact since the product of the lengths of two sides of a parallelogram does not produce a value of any mathematical significance [p. 154, on VI 23].

Euclid's failure to reduce VI 14 to 23 is one piece of evidence that he does not view compounding of ratios as some form of multiplication. Further evidence is provided by VI 19 and 20 . . . [p. 162, on VI 14, 19, and 20].

[Euclid's] failure [to see the notion of compounding in the solid Books XI and XII] is perhaps the strongest evidence that he does not construe compounding ratios as

multiplication [p. 221, on XI and XII; also see pp. 225–6 where the great simplifications that are made possible by using compounding of ratios are described].

Euclid's failure to prove this extension of XII 9 is perhaps some further confirmation of the view that the connections among compounding, multiplying, and volumes were not so immediately clear to him as they are to us [p. 229, on XII 9].

Also see the perceptive analyses in Saito, DRBS and CREA.

    Finally, let us look at the terminology used for some of these operations in the *Elements*. We find that the verb *suntithēmi*, 'to put together', its associated noun *sunthesis* and adjective *sunthetos*, the verb *sugkeimai* which serves as a passive, and the various participles, are used with an enormous range of meanings. Here are some examples of Heath's translations, with the exact Greek form given in each case. A word from this group can mean the sum, of two numbers:

Let two numbers AB, BC prime to one another be added (*sugkeisthōsan*). I say that the sum (*sunamphoteros*) is also . . . [VII 28],

or of two lines, of which examples abound in Book X; see, for example:

To find two lines incommensurable in square which made the sum (*sugkeimenon*) of the squares on them . . . [X 33].

It can also be used with overtones of multiplication, as in:

A composite (*sunthetos*) number is that which measured by some number [VII Definition 13].

It describes the operation on ratios called composition, or *componendo*:

Composition (*sunthesis*) of a ratio means taking the antecedent together with the consequent as one in relation to the consequent by itself [V Definition 14],

as well as that of compounding,

A ratio is said to be compounded (*sugkeisthai*) of ratios when the sizes of the ratios multiplied together make some (?ratio, or size) [VI Definition 4],

and

Equiangular parallelograms have the ratio compounded (*sugkeimenon*) of (the ratios of) their sides [VI 23, see also VIII 5].

It is used in connection with procedures of analysis and synthesis in Book XIII:

Synthesis (*sunthesis*) is an assumption of that which is admitted (and the arrival) by means of its consequences at something admitted to be true [see Euclid–Stamatis, *EE* iv, 198, and Heath, *TBEE* i, 137–40 and iii, 442–3].

In addition to these technical senses, it is used with the general sense of something being made up of, consisting of, constructed of, etc. We find these same words used elsewhere to describe logical and dialectical procedures; and the word *sunthesis* is employed in grammatical works and literary criticism in very much the same way as 'composition' is used in English.

    The precision of terminology that has been developed in parts of modern mathematics, by means of which different operations can be distinguished and

compared, is not found in the *Elements*. However, the spirit of such a study of these different operations can be discerned: there is ʿno sense of Euclid confusing addition, multiplication, composition, and compounding, even though they bear the same name. In this sense, my slaveboy's programme, described at $B_{87}$, is not anachronistic though, as Archytas objects at $A_{86}$, our use today of 'arithmetic' to describe operations of addition, subtraction, multiplication, and division with very general and different kinds of 'numbers' conveys a very different sense from that of the Greek '*arithmētikē*'.

Now let us turn to the objects and manipulations of Greek music theory.

### 4.5(d)  *The Sectio Canonis*

This is not the appropriate place to attempt a general survey of the role of music in early Greek scientific and moral theory. For this, I refer the reader to the articles in the *OCD*, s.v. music (by Mountford & Winnington-Ingram), with their references, and the articles and books by Barker, Bowen, and Burkert that will be cited here or listed in the Bibliography.

I wish here to illustrate the very limited comment about music theory set out in $A_{78}$–$A_{84}$: that it is conceivable that our understanding of mathematics might be helped by a study of the theory of music, in the way I have suggested that the study of cyclical phenomena in astronomy can help our understanding of *Elements* V, Definition 5. In the event, it will turn out that this contribution from music to mathematics is negligible or non-existent, but nevertheless the idea is worth trying. My illustration will be cast in the form of a description of the *Sectio Canonis*, a treatise on the construction of the musical scale.

The *Sectio Canonis* (Euclid–Heiberg, *Opera* viii (ed. Menge), 158–83, or Jan, *MSG*, 148–66; recent study in Barker, MAESC) is a short treatise on music theory. Its author is unknown and its date is uncertain, but it clearly belongs within a mathematical, as contrasted with empirical, study of music, and this places it broadly within the Pythagorean and Academic traditions, rather than those associated with Aristotle's Lyceum. The propositions that make up the bulk of the work are formulated in Euclidean style, and it is generally included within the Euclidean corpus. There is a complete translation in Barker, *GMW* ii, and I am grateful to Andrew Barker for providing the following literally rendered and annotated passages. We start with the introduction to the treatise:

If there were stillness and motionlessness, there would be silence; and if there were silence and nothing moved, nothing would be heard. Then if anything is going to be heard, impact (*plēgē*)[a] and movement must first occur. Then since all *phthongoi*[b] occur when some impact occurs, and since it is impossible for an impact to occur unless movement has occurred beforehand—and since of movements some are more closely packed (*pyknoterai*), others more sparse (*araioterai*), and the more closely packed ones make the *phthongoi* higher-pitched, the sparser ones lower-pitched—it follows that some *phthongoi* must be higher-pitched, since they are indeed composed of more closely packed and more numerous movements, while others are lower-pitched, since they are indeed composed of sparser and less numerous movements. Thus those that are higher-pitched than what is right[c] reach what is right [through] being slackened[d] by the

taking-away of movement, while those that are too low-pitched reach what is right [through] being tightened[e] by the addition of movement. We must therefore agree that *phthongoi* are composed of parts (*moriōn*), since they reach what is right by addition and taking-away. All things that are composed of parts are spoken of (*legetai*)[f] with respect to one another in a ratio (*logos*) of number, so that it is necessary that the *phthongoi*, too, be spoken of (*legesthai*) in a ratio of number with respect to one another. Of numbers, some are spoken of (*legontai*) in multiple ratio [to each other], some in epimoric, some in epimeric, so that it is necessary that the *phthongoi*, too, be spoken of (*legesthai*) in ratios of these sorts with respect to one another. Of these[g] the multiple and the epimoric are spoken of (*legontai*) with respect to one another under a single name.[h]

Again, of the *phthongoi* we grasp (*gignōskomen*) some as being concordant, others as discordant, and [we grasp] the concordant ones as making a single blend, the one [produced] out of both [*phthongoi*], the discordant ones as not [doing so]. Since these things are so, it is likely[i] that the concordant *phthongoi*, since they make for themselves a single blend of the *phōnē*,[j] the blend [produced] out of both, are among the numbers which are spoken of (*legomenōn*) with respect to one another under a single name, being either multiple or epimoric.[k]

*Notes:*
Round brackets indicate transliterations of the Greek, and square brackets contain words added to clarify the English version.

[a] *Plēgē*, literally 'a blow'. Also, instead of "impact and movement" one could read "an impact and a movement".

[b] *Phthongoi* are sounds, but in technical writings specifically pitched sounds, sounds each of which is at some one definite pitch: it is commonly used to mean 'note'. The term *psophos* is used for 'sound' in its most general sense, 'noise', whether or not it has definite pitch. The term *phōnē*, literally 'voice', and properly of the human voice, is often used in musical authors to mean 'musical sound', not necessarily of the voice, and not necessarily of a single definite pitch. See the occurrence indicated at note *j* below.

[c] *To deon*, what is right, correct, or proper, or what is required; here, evidently, the pitch that the musician requires.

[d] *Aniemenous*: this relates to the slackening of a string to lower pitch, but in standard jargon the verb is applied directly to the sound or note, in the sense 'to lower in pitch'.

[e] *Epiteinomenous*: the uses of this verb are parallel to those mentioned in the previous note.

[f] *Legetai*, from *legein*, 'to say': this is also the verb cognate with *logos* where *logos* means 'ratio', and this occurrence of it might therefore be translated 'are related', 'are ratioed'. But here the ambiguity, or perhaps conceptual assimilation, is part of the argument, since the fact that numbers standing to one another in multiple or epimoric ratios are 'spoken of with respect to one another under a single *name*' is taken to imply that things in these (but not other) kinds of ratios possess a real unity.

[g] That is, 'these numbers', not 'these ratios' as many commentators have assumed; see the similar passage at the end of the introduction: "the numbers which are said . . .". Also see Barker, MAESC, 2–3.

[h] The word for 'name' here (*onoma*) is not cognate with *logos*, *legein*, etc., and has no special mathematical usage.

[i] *Eikos*: 'A reasonable hypothesis', 'a fair assumption'.

[j] On the sense of *phōnē* see note *b* above. It seems to be carefully chosen here: the sound in question is a musical one, but it is not a sound of one definite pitch, a

'note', since it is precisely the sound constituted out of the blending of two notes. It is a phenomenon distinct from either of its components, or from the two considered merely as existing simultaneously: it is a 'third thing', but it is not a third note.

[k] The adjectives 'multiple' and 'epimoric' here qualify the *phthongoi*, not, grammatically speaking, the numbers.

The first half of the first paragraph seems to be related to another account of music given by Archytas in his Fragment B1 (Diels & Kranz, *FV* i, 431–5; recent detailed study in Bowen, FEPHS and Huffman, AAF1), except that while Archytas outlines what we might call a 'velocity hypothesis' in which sluggish and weak movements produce low-pitched sounds, quick and powerful movements produce high-pitched sounds, the *Sectio* sets out an 'impulse hypothesis'. The other apparent proponents of the impulse hypothesis seem later in time and associated with the Lyceum (see the pseudo-Aristotelian *Problems* XIX 39, 921a7–31, and *De Audibilibus* 803b27–804a9), though there are difficulties in the interpretation of these passages. For this reason, and in order to avoid overloading my dialogue yet further, I did not have my Archytas suggest developing a theory of the patterns generated by two interacting trains of impulses. These patterns are precisely what I have called astronomical ratios (see $E_{55}$ for an example) and their structure, in which near coincidences are generated and decay, can best be described in terms of the corresponding anthyphairetic ratios. The simplest patterns will be generated by the multiple ratios, and these are in some sense completely static, with a regularly occurring coincidence when the two trains of impulses are in phase. The patterns of the epimoric ratio (familiar from the vernier scale, which exploits the astronomical ratio $10:9$) will be almost static. But the pattern exhibited by the less simple ratios $[n_0, n_1, n_2, \ldots]$, with more than two non-zero terms, will have a shifting structure in which hierarchies of almost-periodic blocks of terms move in and out of prominence, rather as my Eudoxus described at $E_{67}$; a simple example of this phenomenon will occur with a ratio like $[1, 2, 100\,000\,000, \ldots]$. We could also apply this kind of analysis to the rhythmical structure of music, to explain why the rhythmical signatures described by multiple and epimoric ratios such as $3:2$ and $4:3$ have a structure that is immediately clear and easily understood, while signatures like $5:3, 8:3, 7:4$, etc. are much harder to comprehend by ear. There is some evidence of interest in these rhythmical ratios in Plato, Aristoxenus, and Aristides Quintilianus, but I shall not develop any of these rich ideas here.

In the second half of the introduction, we find what seems to be an argument that the study of *phthongoi*, pitched sounds, should be assimilated into mathematics. The argument is scarcely explicit, precise, or coherent, but it seems to go as follows, with my notes and additions in parentheses: pitched sounds consist of a succession of impacts, and since the pitch can be varied by adding or taking away movement and therefore impacts (and since each impact is an indivisible unit?), therefore the relation between two pitched sounds can be described as a ratio of two numbers. (Note, in passing, that the argument is

specious and the conclusion is false.) Any ratio (taken, as almost always in music theory, the greater to the less) will be either a multiple ratio $n:1$, or an epimoric ratio $(n+1):n$, or neither of these, an epimeric ratio. (On the use of the word *morion* or *meros*, plural *moria* or *merē*, 'part' and 'parts', see Sections 1.2(c) and 7.1(b). *Epimoric* means, literally, 'a part in addition', *epimeric* means 'parts in addition'.) The multiple and epimoric ratios are described by a single name. (This is enigmatic, and no completely satisfactory explanation presents itself; see the discussion and proposal in Barker, MAESC, 2–3. In $A_{80-2}$, I suggest a slight addition to Barker's argument, that this single name may be the single number $n$ which is needed to describe the simplest ratios $[n]$ and $[1, n]$, or, arithmetically, the simplest numerical and fractional quantities $n$ and $1 + n'$. These are precisely the multiple and epimoric ratios, and the Greek way of describing each of them will be given below.) Therefore the concordant intervals should then all be found among the multiple and epimoric ratios.

Some such conclusion does seem to have been held by Academic music theorists since later commentators remark, with some disbelief, that while musical practice and the theorists of the Lyceum regarded the octave plus a fourth ($2:1$ and $4:3$, making $8:3$) as a concordant interval, the Academic music theorists treated it as a dissonance or ignored it altogether; see Ptolemy–Düring, *HKP* 13. For example, no reference to this interval appears anywhere in the *Sectio Canonis*; also see my Archytas' evasion in $B_{83}$–$A_{84}$.

We now turn to the propositions of the *Sectio*. I shall first cite all of their enunciations and discuss briefly the proof of a few of them. The translation is again literal and the following details should be noted: The words *logos* (ratio) and *diastēma* (interval) are distinguished here; frequently they are conflated into 'ratio'. The propositions deal with combinations and decompositions of the multiple intervals (*pollaplasion diastēma*), for example the duple (*diplasion diastēma*), or triple (*triplasion diastēma*), and with the epimoric intervals (*epimorion diastēma*), the hemiolic (*hēmiolion diastēma*, 'half-whole', for $3:2$), epitritic (*epitriton diastēma*, 'third-in-addition', for $4:3$), and the epogdoic (*epogdoon diastēma*, 'eighth-in-addition', for $9:8$). The numbers of the propositions are set in square brackets because they are not found in the best manuscripts.

[1] If a multiple interval put together (*suntethen*) twice makes some interval, this interval too will be multiple.
[2] If an interval put together twice makes a whole that is multiple, then the interval will also be multiple.
[3] In the case of an epimoric interval, no mean number, neither one nor more than one, will fall within it proportionally.
[4] If an interval which is not multiple is put together twice, the whole will be neither multiple nor epimoric.
[5] If an interval put together twice does not make a whole that is multiple, that interval itself will not be multiple either.
[6] The duple interval is composed (*sunestēken*) of the two greatest epimoric [intervals], the hemiolic and epitritic.
[7] From the duple interval and the hemiolic, a triple interval is generated.

[8] If from a hemiolic interval an epitritic interval is subtracted (*aphairethēi*), the remainder left is epogdoic.

[9] Six epogdoic intervals are greater than one duple interval.

[10] The octave (*dia pasōn*) interval is multiple.

[11] The interval of the fourth (*dia tessarōn*) and that of the fifth (*dia pente*) are each epimoric.

[12] The octave interval is duple.

[13] It remains to consider the interval of a tone (*toniaion*) to show that it is an epogdoic.

[14] The octave is less than six tones.

[15] The fourth is less than two and a half tones, and the fifth is less than three and a half tones.

[16] The tone will not be divided into two or more equal intervals.

[17] The *paranētai* and *lichanoi* will be found by means of concords, as follows. . . .

[18] The *paranētai* and *tritai* do not divide the *pyknon* into equal intervals.

[19] To mark out the *kanōn* according to the so-called immutable *systēma*.

[20] It remains to find the moveable notes. . . .

Propositions 1–9 develop some purely arithmetical results about intervals, Propositions 10–18 translate these into statements about musical intervals, and Propositions 19 and 20, which some have argued are interpolations, describe how to mark out a particular kind of scale.

First, consider the terminology. The word *diastēma*, 'distance between' or 'interval', in a very general sense, is often treated by ancient and modern commentators as if it were here synonymous with *logos* 'ratio' (see A$_{80}$), but the author of the *Sectio* seems to maintain a careful distinction between the two. The mathematical Propositions 1–9 are permeated with the language of ratios, and they explicitly cite some results from some external mathematical theory of ratio or proportion, which are then translated into statements about intervals. See, for example, Proposition 2, which invokes *Elements* VIII 7:

If an interval put together (*sunthen*)[a] twice makes a whole that is multiple, then that interval will also be multiple.[b]
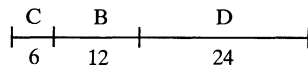


Let there be an interval BC, and let it be that as C is to B so B is to D (*hōs ho C pros ton B, houtōs ho B pros ton D*),[c] and let D be a multiple of C. I assert that B is also a multiple of C. For since D is a multiple of C, C therefore measures D. But we have learned (*emathomen de*)[d] that where there are numbers in [continued] proportion (*analogon*)—however many of them—and where the first measures the last, it will also measure those in between.[e] Therefore C measures B, and B is therefore a multiple of C.

*Notes:*

  [a] *Sunthen*, here and throughout the *Sectio*, has the mathematical sense of compounding, as in *Elements* VIII 5.

  [b] The figures show some variations between the manuscripts, and the numbers appended to them may be later additions.

  [c] This is exactly the same terminology for proportions that we find throughout Book VII.

*d* One of the very occasional explicit back-references in the *Elements*, in X 10, uses this same word, though with the different particle *gar*. See Section 1.2(b) and its Notes, above.

*e* This result follows immediately from *Elements* VIII 7. The argument will work only for continued proportion (*hexēs analogon* or, from *Elements* VIII 8 onwards, often *sunexēs analogon*): $a:b::b:c$, $b:c::c:d$, etc., the subject of *Elements* VIII. To see why, consider the simplest example of a non-continued proportion $2:3::4:6$. Also see Aristotle, *Nicomachean Ethics*, 1131a30–b16, and the note to *Elements* V Definition 8, in Heath, *TBEE* ii, 131.

We have here an enunciation phrased in terms of intervals between numbers. It is immediately translated into the terminology of proportions, and is established by means of a quoted result. Then, since the conclusion is to be related to the enunciation, there is an implied final translation back into the language of intervals.

Very similar remarks can be made about Proposition 3, in which the results of *Elements* VII 22 and VIII 8 are cited. The proposition is described by Boethius, who attributes it to Archytas, and it is also elsewhere translated by him (see Boethius–Friedlein, *IM* III 11 and IV 2). It has attracted an enormous commentary, and I refer the reader to the description in Heath, *TBEE* ii, 295, and the recent discussions in Burkert, *LS*, 442–7, and Knorr, *EEE*, 212–25 for further details.

Another example of the mathematical style can be found in Proposition 6:

The duple interval is composed (*sunestēken*) of the two greatest epimoric [intervals], that is from the hemiolic and the pitritic.
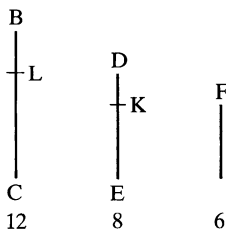


Let BC be hemiolic of DE, and let DE be epitritic of F. I say that BC is duple of F. I took away (*apheilon*) EK, equal to F, and CL, equal to DE. Then since BC is hemiolic of DE, BL is a third part (*triton meros*) of BC, and a half (*hēmisu*) of DE. Again, since DE is epitritic of F, DK is a fourth part of DE, and a third part of F. Then since DK is a fourth part of DE, and BL is a half of DE, DK will therefore be a half of BL. Now BL was a third part of BC: therefore DK is a sixth part of BC. But DK was a third part of F: therefore BC is a duple of F.

Alternatively:
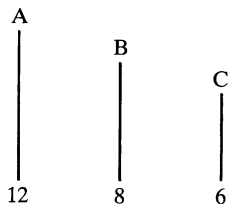
Let A be hemiolic of B, and let B be epitritic of C. I say that A is duple of C. Since A is hemiolic of B, A contains B and a half of B. Then two A's (*duo hoi* A) are equal to three B's. Again, since B is epitritic of C, B contains C and a third of C. Therefore three B's are equal to four C's. Now three B's are equal to two A's and therefore two A's are equal to four C's, and therefore A is equal to two C's: therefore A is the duple of C.

The arguments of both proofs are very basic and long-winded and both seem to avoid any explicit use of the words associated with ratio or proportion. Both styles of argument are found elsewhere in the treatise, so there is no suggestion that either proof is an interpolation. Consider the second proof first; it is an exploration of a particular case of the simple deduction that,

$$\text{if } n\text{A} = m\text{B} \quad \text{and} \quad m\text{B} = p\text{C}, \quad \text{then} \quad n\text{A} = p\text{C},$$

in which $m$ and $p$ are then reduced to their lowest terms and described. The proofs of Propositions 7 and 8 are similar, and perhaps that of Proposition 9 (cited below) can be viewed as an extension of this method. Now consider the first proof. This is an exploration of operations like: "I take away . . . CL, equal to DE [from BC]. Then, since BC is hemiolic of DE, BL is . . . a half of DE". This could be an almost explicit explanation of the statement that the anthyphairesis of BC : DE is [1, 2]; or it could also be an almost explicit statement of the arithmetised relation between BC and DE, that BC $= (1 + 2')$DE. Unfortunately for us, these kinds of descriptions are almost identical when applied to the anthyphairetic ratios $[1, n]$ and the fractional quantity $1 + n'$, and nowhere in the treatise is any more complicated kind of such a description explored. (See $A_{80}$–$B_{81}$; a similar example arose in Section 2.4(d), where the ratio $4673\,2' : 14\,688$ in the *Measurement of a Circle* was described as "greater than threefold and there are remaining $667\,2'$, which is less than the seventh part of $4672\,2'$". There, also, a little more information might have made it clear whether the underlying idea was anthyphairetic or fractional.)

The final citation of an external result occurs in Proposition 9:

Six epogdoic intervals are greater than one duple interval.
   Let A be one number. Let the epogdoic of A be B, let the epogdoic of B be C, let the epogdoic of C be D, let the epogdoic of D be E, let the epogdoic of E be F, and let the epogdoic of F be G. I say that G is more than the duple of A.
   Since we have learned how to find numbers of epogdoic of one another, the numbers A, B, C, D, E, F, G have been found. A is 262 144, B is 294 912, C is 331 776, D is 373 248, E is 419 904, F is 472 392, G is 531 441: and G is more than the duple of A.

Here A is $8^6$, B is $8^5 \cdot 9, \ldots$, and G is $9^6$; but, parenthetically, practically every numeral in the surviving manuscripts is incorrect. In particular, the sampi $\Upsilon$, for 900, appears variously as $\tau$, $\psi$, or $\Upsilon$; see Chapter 7, below. The reference in the proposition could be to *Elements* VIII 2:

To find numbers in continued proportion, as many as may be prescribed, and the least that are in a given ratio.

Indeed Euclid observes, in a porism, that the extreme terms of this continued proportion will be powers of the terms of the original ratio, just as in this application.

Let us now turn to the music theoretic manipulations in Propositions 10–20; but since these are described and analysed in detail in Barker, MAESC, I can be very brief. We find here that the terminology of ratios is confined to three carefully organised references to the mathematical theory. The first is right at the end of the long Proposition 12, to be quoted below; the second is in Proposition 16:

The tone will not be divided into two or more equal intervals.

It has been shown that it is epimoric. In the case of an epimoric interval neither several mean [numbers] nor one fall within it proportionally. Therefore the tone will not be divided into equal intervals.

and the third is a similar reference in Proposition 18.

The Introduction seems to provide the link between the mathematical theory of Propositions 1–9 and the musical theory of Propositions 10–20, and we can abstract that link into two principles:

The interval between two notes can be expressed in terms of numbers (*arithmoi*), and

All concords are either multiple or epimoric ratios.

As I remarked earlier, the second principle distinguishes Pythagorean and Academic musical studies from those of the Lyceum, and we can elaborate this, in the context of the *Sectio*, as follows. In Proposition 6, quoted above, the treatise investigates the relation

$$(4:3) \text{ and } (3:2) \text{ make } (2:1),$$

and, in Proposition 7, the relation

$$(2:1) \text{ and } (3:2) \text{ make } (3:1).$$

Then the music theoretic consequences are analysed in the comprehensive Proposition 12:

The octave interval is duple.

We have shown that it is multiple; it is thus either duple or greater than duple. But since we showed that the duple interval is composed of the two greatest epimorics, it follows that if the octave is greater than duple it will not be made up of just two epimorics, but of more. But it is made up of two concordant intervals, the fifth and the fourth. Therefore the octave will not be greater than duple: therefore it is duple.

But since the octave is duple, and the duple is made up of the two greatest epimorics, it follows that the octave is made up of the hemiolic and the epitritic, since these are the greatest. But it is made up of the fifth and the fourth, and these are epimoric. The fifth, therefore, since it is greater, must be hemiolic, and the fourth epitritic.

It is clear then that the fifth and octave is a triple. For we showed that the triple interval is generated from a duple and a hemiolic interval, so that the octave and a fifth is also a triple.

The double octave is a quadruple.

But nowhere does the *Sectio* discuss any mathematical or musical implication of the relation

$$(2:1) \text{ and } (4:3) \text{ make } (8:3).$$

Much more serious than this omission is the flaw to be found in the proof of Proposition 11:

> The interval of the fourth and that of the fifth are each epimoric.
>
> Let A be *nētē synēmmenōn*, let B be *mesē*, and let C be *hypatē mesōn* [three fixed notes on the scale, each a fourth apart]. The interval AC, being a double fourth, is therefore discordant [a musical fact]: it is therefore not multiple [!; see below]. Thus since the two equal intervals AB and BC when put together make a whole which is not multiple, neither is AB multiple [Proposition 1]. And it is concordant: therefore it is epimoric. The same demonstration applies also to the fifth.

The inference that a discordant interval is not multiple cannot follow from any argument or assumption stated or hinted anywhere in the treatise or elsewhere. Note that the second principle of the Introduction, described above, does *not* state that all multiple and epimoric intervals are concordant; this would ensure the truth of the proposition but it is clearly false, as the examples of $5:1$, $7:1$, and $9:8$ clearly illustrate. There seems to be no explanation of the flaw in this proof, nor any way of recovering the proposition, and since the remaining propositions depend on it, the remainder of the argument collapses.

It may be misplaced effort to look for a sophisticated mathematical explanation for a treatise so obscure and flawed as the *Sectio*, but nonetheless I have tried, in my dialogue. Just as the idea of ratio in an astronomical context might plausibly lead to the profound insight of *Elements* V Definition 5, as I proposed in Section 4, so an exploration of ratio in a musical context might yield some other mathematical insights. Moreover, the manipulations of music theory seem to depend fundamentally on the operation of compounding, an operation which seems to pose some serious problems for mathematicians. My purely speculative suggestion, set out in $A_{70-84}$, is that music theory might plausibly give some help with this problem. It might become apparent only later that the compounding of anthyphairetic ratios is an even more subtle operation, both conceptually and computationally, as I explained in Section 2, above. In the event, the proposal in $A_{84}$ that it may be easier to infer it from simple commensurable examples will turn out to be misguided, as I shall explain in Section 9.3(a).

Let me leave this brief discussion of the *Sectio Canonis* by emphasising a part of my opening description of the treatise: that its author is unknown and its date is uncertain.

### 4.5(e)  *Further problems*

The final speeches of my dialogue, $B_{89}–A_{96}$, introduce further questions about the connections between geometry and *arithmētikē* that I shall only discuss very briefly here. The description in $B_{89}–B_{91}$ of the geometrical operations is

included in preparation for the work of the next chapter, and to make the geometrical meaning of a statement like $\sqrt{2} \cdot \sqrt{3} = \sqrt{6} \cdot 1$ quite clear. The symbolism introduced is, of course, intended for the convenience of the modern reader, and there is no suggestion in Greek mathematics of any such developments; all of the symbolic manipulations should be read as shorthand descriptions of geometrical manipulations.

The straightforward construction in B93 of the lines involved in $\sqrt{2} \cdot \sqrt{3} = \sqrt{6} \cdot 1$ leads, in Fig. 4.2, to a configuration closely resembling that called the *arbēlos*, the shoemaker's knife, which is found in the Archimedean *Book of Lemmas*, Archimedes–Heiberg, *Opera* ii, 510–25 and Archimedes–Heath, *WA*, pp. xxxii ff. and 301 ff. This treatise is known only in an Arabic translation of what appears to be a later Greek compilation of heterogeneous results. The observation in B95 on the angles of an equilateral triangle is very suggestive. As a postulate:

That the angles of all equilateral triangles are equal to one another,

or a similar postulate for 'squares':

That the angles of all equiangular and equilateral rectangles are equal to one another,

it invites the explicit evaluation of Postulate 4 of *Elements* I:

That all right angles are equal to one another.

On the other hand, it is now well known that these postulates about similar and non-congruent figures are equivalent to Postulate 5 and that there are ways of excluding what is now called non-Euclidean geometry. Our evidence on this topic from early Greek geometry is very sparse, and I shall not enter the question of its interpretation here beyond observing that, although this postulate is generally described as if its origin lies in problems associated with parallel lines and indeed it is often referred to as the 'Parallel Postulate', the postulate itself is articulated without invoking the language of parallels:

That, if a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than two right angles.

Indeed it more resembles a converse of *Elements* I 17:

In any triangle two angles taken together in any manner are less than two right angles;

and it can be read as the assertion that it is possible to construct a triangle of any size with angles equal to those of a given triangle. Its formulation seems to be more closely concerned with the properties of similar figures than with parallel lines, and so with the considerations at the end of B95.

Our earliest treatises on Greek mathematics, *On the Moving Sphere* and *On Risings and Settings* by Autolycus of Pitane, who flourished at the end of the fourth century BC, are on spherical geometry, and we have later reports of work

by Eudoxus on subtle constructions involving spheres. However, we have no explicit evidence that Greek geometers ever considered the spherical model for geometry, as in B95, however attractive the idea may be. Spherical geometry is excluded from the *Elements* by Postulate 2:

To produce a finite straight line continuously in a straight line,

and also by Postulate 3:

To describe a circle with any centre and distance.

Postulate 1 can even be included in these considerations when it is interpreted as

To draw a [unique] line from any point to any point,

for this does not hold on a sphere: there are infinitely many lines joining antipodal points and two lines joining non-antipodal ones. Is it possible that the role of all of the postulates—and not just Postulate 5—was to exclude such exotic geometries? We cannot know; we have but little evidence on this topic. (For further discussions of spherical geometry, see Rosenfeld, *HNEG*, and Tóth, *PCA*.) My reason for introducing non-Euclidean geometry is to bring the little set of dialogues themselves in a full circle with my slaveboy's reflection at the end of B95, and so point to a further level of mathematical subtlety underlying Socrates' example of doubling a square in the *Meno*.

## 4.6 APPENDIX: THE WORDS *LOGISTIKĒ* AND *LOGISMOS* IN PLATO, ARCHYTAS, ARISTOTLE, AND THE PRE-SOCRATIC PHILOSOPHERS

### 4.6(a) *Plato*

The references to *logistikos* (the lemma form of *logistikē*) in Brandwood, *A Word Index to Plato*, divide between about twenty-five where the context is explicitly mathematical and about ten with the more general sense of 'intellectual principle' or 'reason'; while *logismos* (excluding ten instances found in the pseudo-Platonic *Definitions*, all of them non-mathematical) divide between about thirty explicitly mathematical references and forty usages with the more general sense of 'rational discussion' or 'reason'. It might also be interesting to check, in the same way, the 1500 or so entries for *logos*.

The relevant entries from the *Index* are reproduced below; they are divided into grammatical categories with lemmata underlined, as in the *Index*. Those items that seem to me to refer to specifically mathematical passages are set in bold-face type. An asterisk indicates some variation between the manuscripts or editorial emendation, details of which can be found in the *Index*. This list should cover all occurrences of these words in the Platonic and pseudo-Platonic

corpus. A similar index to *alogos* and *(ar)rhētos* can be found in the Appendix to Chapter 5.

λογιϲμόϲ *Timaeus* 34a.8; *Laws* I 644d.2.

λογιϲμόν *Euthyphro* **7b.10**; *Hippias Minor* **367a.9**, **c.5\***; *Republic* IV 440b.1, VII **522c.7, 524b.4**, IX **587e.5**; *Phaedrus* **274c.8**; *Timaeus* 30b.4, 33a.6; *Laws* VII 805a.3; *Epinomis* 974a.2; *Definitions* 412e.2, 416a.1,12.

λογιϲμοῦ *Meno* 100b.2; *Phaedo* 66a.1; *Symposium* 207b.7; *Republic* IV 439d.1, 441a.9, IX 586d.2; *Timaeus* 36e.6, 72e.2, 77b.5, 86c.3; *Critias* 121a.7; *Sophist* 248a.11; *Philebus* 21c.5, 52b.3; *Laws* I 645a.1,5; *Epinomis* 981c.4; *Definitions* 413c.3, 415e.6, 416a.18.

λογιϲμῷ *Meno* 98a.4; *Phaedo* 79a.3, 84a.7; *Republic* I **340d.4**, IV 431c.6, VI 496d.6, VIII 546b.1, X **603a.4**, 604d.5, 611c.3; *Parmenides* 130a.2; *Phaedrus* 249c.1; *Timaeus* 52b.2, 57e.1, 72a.1; *Laws* III 697e.2, VII 813d.1; *Epistle* VII 340a.2; *Definitions* 412b.4,5, 415d.10.

λογιϲμοί *Laws* VII **817e.6**, X **896c.9**.

λογιϲμούϲ *Hippias Minor* **367a.1**; *Protagoras* **318e.2**; *Republic* VI **510c.3**, VII **525d.1**; *Theaetetus* **145d.2**; *Statesman* **257a.7, b.7**; *Philebus* 11b.8; *Laws* VII **819b.2**.

λογιϲμῶν *Hippias Minor* **366c.6, 367b.7, c.2,3,5\***; *Hippias Major* **285c.5**; *Republic* VII **526d.8, 536d.5**; *Timaeus* **47c.2**; *Sophist* 254a.8; *Philebus* **57a.1**; *Laws* VII **809c.4**; *Definitions* 412a.7\*.

λογιϲμοῖϲ *Philebus* 52a.8; *Laws* X **897c.6**, XII **967b.3**.

λογιϲτέον (nt.) *Timaeus* 62a.1.

λογιϲτήϲ *Republic* I **340d.6**; *Statesman* **260a.5**.

λογιϲτικόϲ *Hippias Minor* **367c.4**; *Theaetetus* **145a.7**.

λογιϲτικόν *Republic* I **340d.3**.

λογιϲτικῷ *Republic* VII **525b.6**, IX **587d.11**.

λογιϲτικοί *Euthydemus* **290c.1**; *Republic* VII **526b.5**; *Laws* III **689c.9**.

λογιϲτική *Charmides* **166a.5**; *Gorgias* **450d.6, 451c.2,5**; *Republic* VII **525a.9**; *Statesman* **259e.1**; *Philebus* **56e.7**; *Definitions* 411e.10, 414c.3.

λογιϲτικήν *Gorgias* **451b.5**; *Republic* VII **525c.1**.

λογιϲτικῆϲ *Charmides* **165e.5, 166a.10**; *Hippias Minor* 366c.6; *Definitions* 412b.9.

λογιϲτικῇ *Statesman* **259e.5, 260a.10**.

λογιϲτικόν *Charmides* **174b.5**; *Republic* IV 439d.5, 440e.6,9, 442c.11, VIII 550b.1, 553d.1, IX 571c.4, d.7, 580d.4\*, X 605b.5; *Timaeus* 37c.1.

λογιϲτικοῦ *Republic* IV 440e.8, 441a.5, X 602e.1.

λογιϲτικῷ *Republic* IV 441a.3, e.4.

λογιϲτικά *Hippias Minor* 366d.5.


## 4.6(b) *Archytas*

Two fragments of Archytas, B3 and 4 in Diels & Kranz, *FV* ii, 436–8, are concerned with *logismos* and *logistikē*. They are cited here in full:

B3. *Logismos*, when discovered, stops strife and increases concord; when it occurs, there is no excess of gain, but there is equality; for by this we settle our disputes. By this, then, the poor take from the powerful, and the rich give to the needy, both sides trusting that through this they will get fair treatment. It is a rule and it prevents men from doing wrong: it stops those who know *logismos* before they do wrong, persuading them that they will not be able to escape notice, when they come to it: and it prevents those who do not know from doing wrong, showing by that very fact [i.e. that they do not know *logismos*] that they are doing wrong. [Translation adapted from Freeman, *APSP*, and Harvey, TKE; see Chapter 6, n. 11, below.]

B4. *Logistikē*, it seems, in regard to wisdom is far superior to all the other sciences, especially geometry, because *logistikē* is able to do more clearly any problem it will . . . and—a thing in which geometry fails—*logistikē* adds proofs and, at the same time, if the problem concerns forms, *logistikē* treats of the forms also. [Translation adapted from Freeman, *APSP*.]


### 4.6(c) *Aristotle*

The following entries are taken from Bonitz, *Index Aristotelicus*, which is probably far from complete. Aristotle almost always uses the words in the sense of the 'reasoning faculty'. Passages in bold-face type indicate a context which contains some additional reference, as when this 'reasoning faculty' is associated with the 'rational and irrational principles' of the soul, or, very occasionally, with some technical usage; but none of these passages seem to shed any light on the nature of mathematical *logistikē*. Again, there is a massive entry for *logos*, about fifteen times longer than these entries combined.

λογιϲμόϲ *Posterior Analytics* **88b12**, 100b7; *Topics* 145b2, b17; *Physics* **200a23**; *On the Soul* 415a8, 433a12, b29; *Metaphysics* 980b28, 1015a33; *Nicomachean Ethics* 1111a34, 1117a21, 1119b10, 1141b14, 1150b24, 1220a1, 1250a11, b13; *Politics* 1312b29, **1322b9**, **1334b24**, 1369b7; *Rhetoric to Alexander* 1429a17; *Fragment* 96 1493b32; *Fragment* 97 1493b37.

λογιϲταί *Politics* **1322b11**; *Fragment* 406 1546a4, a10, a21; *Fragment* 407 1546a26.

λογιϲτήριον *Fragment* 406 1564a4.

λογιϲτικοϲ *Topics* 126a8, a13, 128b39, 129a11 sqq, 134a34, 136b11, 138a34, b2, b13, 145a29, 147b32; *Physics* 210a30; *On the Soul* **432a25**, 433b29, 434a7; *Nicomachean Ethics* **1139a12**, 1182a18, a20, 1249a31, b30, 1250a3, a16; *Rhetoric* **1369a2**.


### 4.6(d) *Pre-Socratic philosophers*

The index to Diels & Kranz, *Die Fragmente der Vorsokratiker* contains nine columns of entries for *logos*, half a column for *logismos*, and one entry (Archytas B4; see above) for *logistikē*. Here are the entries for *logismos*, each with its reference number and volume, page, and line entries. Square brackets denote texts described as spurious; for further details of the classification, see *FV* i, 1. The bold type is as in Section 4.6(a), above.

| Anaxagoras | A66 | ii | 22.4 | |
|---|---|---|---|---|
| Archytas | **B3** | **i** | **437.7** | (see above) |
| Democritus | B187 | ii | 183.10 | |
| | B290 | ii | 205.19 | |
| | [B302 | ii | 222.19] | |
| | [B306 | ii | 223.32] | |
| Epicharmus | [B56 | i | 208.5 f.] | |
| | [B57 | i | 208.9] | |
| Hippias | **80A5** | **ii** | **256.2** | ($\equiv$ Plato, *Protagoras* 318e.2) |
| | **86A11** | **ii** | **328.32** | ($\equiv$ Plato, *Hippias Major* 285c.5) |
| Hippocrates of Cos | 64C2 | ii | 67.18 | (i.e. the physician) |
| Pythagoras | 9 | i | 101.2 | |
| Theodorus | **4** | **i** | **397.25** | ($\equiv$ Plato, *Theaetetus* 145d.2) |

# *ELEMENTS* IV, X, AND XIII:
## THE CIRCUMDIAMETER
## AND SIDE

## 5.1 THE CIRCUMDIAMETER AND SIDE,
## AND OTHER EXAMPLES

### 5.1(a) *The problem*

Let us now turn to the problem of the circumdiameter and side, introduced by my Socrates at $S_{41}$. It is very similar in spirit to the very fruitful problem of the diagonal and side, discussed in Section 2.2, out of which grew the remarkable discoveries of the dimension of squares in Chapter 3. We take a regular polygon with its circumscribed circle, and try to describe the relation between the circumdiameter and a side. Instead of the side we can take a diagonal; and instead of the diameter, we can take the radius, since doubling or halving a ratio has an effect on the anthyphairesis that seems difficult to elucidate. We can also consider the inscribed circle, or investigate the similar three-dimensional problems associated with an edge or diagonal of a regular polyhedron inscribed in or circumscribed around a sphere.

The geometrical constructions associated with several such problems are described in *Elements* IV and XIII. Most of them lead to lines that are either commensurable or related as the sides of squares, and they are described as such by Euclid; see *Elements* XIII 12 (equilateral triangle and hexagon), 13 (tetrahedron), 14 (octahedron), and 15 (cube). In these cases the anthyphairesis can be completely described by an additional short and now routine calculation. But three examples, the pentagon, icosahedron, and dodecahedron, give rise to more complicated relations than pure sides of squares. The worst kind of behaviour considered in the *Elements* occurs already with the pentagon, and I shall concentrate exclusively on this example. My geometrical descriptions throughout this chapter will be based on the treatment in Taisbak, *CQ*, to which I shall add an anthyphairetic motivation.*

### 5.1(b) *The pentagon*

Consider the circumradius and side or diagonal of a pentagon. There is no difficulty in evaluating and expressing the relations between these lines using

---

* I strongly encourage the reader to look at the Appendix, Section 10.1, before embarking on this chapter, which has been rewritten for the second edition, incorporating some of the material in my IRBT. All further notes will be put at the end and can be ignored by anyone whose main interest is to understand the material presented here; they contain background information and additional comments of a general nature.
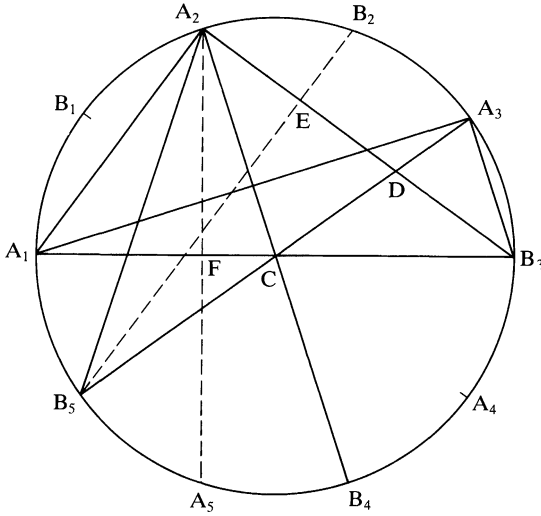
FIG. 5.1

the language and techniques at Euclid's disposal. Take a regular decagon $A_1B_1A_2B_2 \ldots A_5B_5$ inscribed in a circle (see Fig. 5.1), and note that Euclid's basic construction of the pentagon at *Elements* IV 11 does, in fact, start from the side of a decagon inscribed in a circle, though Euclid does not bring out this fact. Then the diameters, such as $A_1B_3$, $A_2B_4$, and $A_3B_5$, will intersect in the centre C of the circle. Now most of the angles of the figure can be expressed in terms of the angle subtended at the circumference by a side of the decagon, which is one-twentieth of the circle, or 18°. Hence many 36°–72°–72° isosceles triangles can be found, in particular $A_3B_3D$ (so $B_3D = B_3A_3 =$ the side of the decagon) and $A_2CD$ (so $A_2D = A_2C =$ the radius of the circumcircle = the side of the inscribed hexagon). Hence the line $A_2B_3$ is the sum of the sides of the inscribed hexagon and decagon, and so is divided in extreme and mean ratio; this result is proved at *Elements* XIII 9. Note that since $A_2B_5D$ is also a 36°–72°–72° triangle and $B_5EB_2$ is perpendicular to $A_2EB_3$, then $A_2E = ED = q$, where $2q = r$, the radius of the circle. Write $B_3E = p$; then, by XIII 3, one of the block of propositions on the metrical properties of the extreme and mean ratio (see Section 3.5(d), above),

$$B_3E^2 = 5A_2E^2, \quad \text{i.e.} \quad p^2 = 5q^2.$$

Hence

$$B_3A_3 = B_3D = B_3E - ED = p - q,$$

and

$$B_3A_2 = B_3E + EA_2 = p + q.$$

Now, from the right-angled triangle $A_1A_2B_3$,

$$A_1A_2^2 + A_2B_3^2 = A_1B_3^2,$$

so

$$A_1A_2^2 + (p + q)^2 = 16q^2,$$

and hence

$$A_1A_2^2 + p \cdot r = 10q^2 = 10(r/2)^2.$$

Similarly, from the right-angled triangle $A_1A_3B_3$,

$$A_1A_3^2 + A_3B_3^2 = A_1B_3^2,$$

or

$$A_1A_3^2 + (p - q)^2 = 16q^2$$

and hence

$$A_1A_3^2 = p \cdot r + 10(r/2)^2.$$

We, today, express these kinds of results in the arithmetised form:

$$\text{side of pentagon} = r/2\sqrt{(10 - 2\sqrt{5})}$$

$$\text{diagonal of pentagon} = r/2\sqrt{(10 + 2\sqrt{5})},$$

but we find nothing of this kind in the *Elements*. It is not that Euclid is unable or unwilling to articulate this kind of result. We can write these formulae as $\sqrt{(10(r/2)^2 \pm r \cdot \sqrt{5(r/2)^2})}$ and describe them, within the idiom of Books II or XIII 1–5, by:

The square on the side [or diagonal] of a pentagon is less [or greater] then ten squares on half the radius by the rectangle contained by the radius and the side of the square equal to five times the square on half the radius.

Nor is it that such a statement is difficult to understand. In Fig. 5.2(a), we start with "ten squares on half the radius" and then, in Fig. 5.2(b), double the base, halve the sides, and subtract "the [upper] rectangle contained by the radius $[r = 2q]$ and the side $[p]$ of the square equal to five times the square on half the radius [so $p^2 = 5(r/2)^2 = 5q^2$]". This leaves the lower, shaded rectangle equal to the square on the side of the pentagon; then the side itself can be constructed using *Elements* II 14. A similar procedure, in which $p \cdot r$ is added, not subtracted, will give the square on the diagonal. Nor is it that this kind of result is remote from the kinds of manipulations and evaluations that occur in the *Elements*: we find an exploration of the pentagon at XIII 11 which involves the vertices of the decagon, the projection of a side of the pentagon on an adjacent diameter (for example $A_1F$ in Fig. 5.1), and a line equal to a quarter of the radius, $q/2$ in my notation. Euclid's argument turns on the fact that
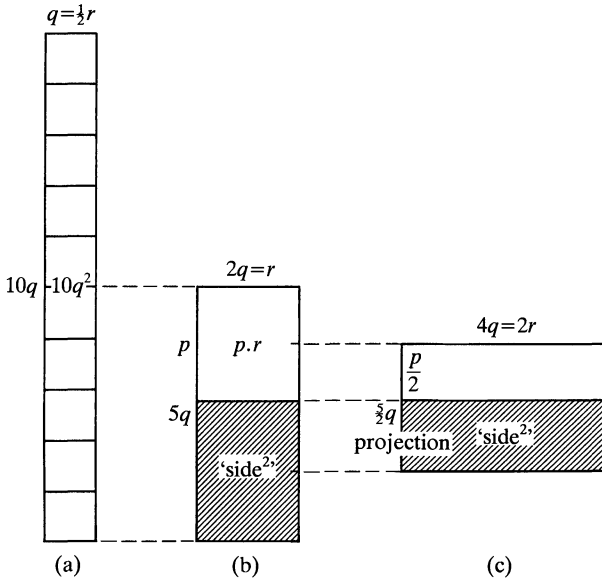
FIG. 5.2

$A_1A_2B_3$ and $A_1FA_2$ are similar right-angled triangles, so that

$$A_1A_2^2 = A_1F \cdot A_1B_3, \quad \text{i.e.} \quad \text{side}^2 = (\text{projection}) \cdot (4q).$$

This rectangle appears as the lower, shaded rectangle in Fig. 5.2(c), where it arises from a further doubling of the base and halving of the side of the rectangle of Fig. 5.2(b). While the likes of Fig. 5.2 are not generally drawn in the *Elements*, these kinds of figures do seem to follow the Euclidean style of argument much more closely than manipulations of the corresponding algebraic formulae.

For each of the cases of the circumdiameters and sides (or edges) of the pentagon, icosahedron, and dodecahedron, instead of providing something which corresponds to our kind of metrical description, Euclid suppresses all of the numerical constants that give essence to our evaluations and enunciates only qualitative results. For example, at XIII 11, we have:

If in a circle which has its diameter expressible (*rhētos*) an equilateral pentagon be inscribed, the side of the pentagon is the *alogos* ('without ratio') straight line called minor (*elassōn*).

This proposition means that the side of the pentagon can be expressed as the difference of the sides of two squares, where these squares satisfy certain further kinds of very elaborate qualitative, not quantitative, relations of a kind that have been developed, in the *Elements*, in the classification of Book X, and which I shall describe in Section 5.2, below.

### 5.1(c) *The extreme and mean ratio*

The extreme and mean ratio gives another illustration of Euclid's procedure. If a line of length 1 is cut in extreme and mean ratio, the longer segment has length $\frac{1}{2}(\sqrt{5}-1)$, the shorter $\frac{1}{2}(3-\sqrt{5})$, and their ratio is $\frac{1}{2}(\sqrt{5}-1):1$. While these arithmetised statements might seem to us today to contain some of the essential facts about the extreme and mean ratio, they do not occur in the *Elements*, though there is absolutely no doubt that Euclid could have articulated them, expressed geometrically, among the propositions that he does prove at *Elements* XIII 1–5. Instead this block of results finishes, at XIII 6, with:

> If an expressible (*rhētos*) straight line be cut in extreme and mean ratio, each of the segments is an *alogos* straight line, the one called apotome (*apotomē*).

This, again, is a qualitative statement that is, this time, sufficiently simple to be described without further terminology: each of the segments can be expressed as a difference of two lines which are incommensurable with each other, but such that the squares on these lines are commensurable with each other and with the square on the whole.

### 5.1(d) *Surd quantities*

Notwithstanding these reservations about the relevance of these kinds of arithmetised expressions, there is a very long and still active tradition of interpreting Book X in this style. Simon Stevin was far from the first to follow this path, but he is one of the best known; in 1585, he boasted that he possessed the key to Book X: the calculus of surd quantities.[1] Let us follow his suggestion and, without attempting to explain further what these 'surd quantities' are, explore a few more examples:

$$\sqrt{(2\pm\sqrt{3})} = \sqrt{\tfrac{3}{2}}\pm\sqrt{\tfrac{1}{2}}$$

$$\sqrt{(\sqrt{8}\pm\sqrt{6})} = \sqrt[4]{\tfrac{9}{2}}\pm\sqrt[4]{\tfrac{1}{2}}$$

$$\sqrt{(\sqrt{2}\pm 1)} = \sqrt{(\sqrt{\tfrac{1}{2}}+\tfrac{1}{2})}\pm\sqrt{(\sqrt{\tfrac{1}{2}}-\tfrac{1}{2})}$$

$$\sqrt{(\sqrt{5}\pm\sqrt{2})} = \sqrt{(\sqrt{\tfrac{5}{4}}+\sqrt{\tfrac{3}{4}})}\pm\sqrt{(\sqrt{\tfrac{5}{4}}-\sqrt{\tfrac{3}{4}})}$$

etc.

We get what seems, at first, to be a bewildering variety of behaviour. Since squares are easier to understand than square roots, let us also look at them:

$$(\sqrt{\xi}+\sqrt{\eta})^2 = (\xi+\eta+2\sqrt{\xi\eta}) = (\beta+\gamma) \text{ where } \beta = \xi+\eta \text{ and } \gamma = 2\sqrt{\xi\eta}.$$

Note that $\beta > \gamma$—the arithmetico-geometric inequality, X 59/60—and neither $\beta = \xi+\eta$ nor $\sqrt{(\beta^2-\gamma^2)} = \xi-\eta$ involves square root signs. A bit more explanation shows that these are sufficient conditions for $(\beta+\gamma)$ to be the square of one of these kinds of surds; and that explains the first example. Yet more exploration by anyone at ease with the mathematics that has developed

since the time of Stevin and which, I think, owes a not insignificant debt to him (see my FHYDF) will uncover what has been called "the quintessence of Book X":[2]

$$\sqrt{(\beta \pm \gamma)} = \sqrt{(\tfrac{1}{2}\beta + \tfrac{1}{2}\sqrt{(\beta^2 - \gamma^2)})} \pm \sqrt{(\tfrac{1}{2}\beta - \tfrac{1}{2}\sqrt{(\beta^2 - \gamma^2)})}. \qquad (*)$$

The first three examples given above satisfy conditions, the first of which was just described (namely $\beta$ the positive square root of a square; $\gamma$ the positive square root of a non-square; and $(\beta^2 - \gamma^2)$ a square) that lead to simplifications in this formula, while the fourth example represents the most general kind of case that can occur, in which none of the numbers involved are 'accidentally' rational, or are 'accidentally' rationally related to each other.[3]

One more example. Given $\beta$ and $\gamma$, we could choose $\xi$ at will, and write

$$\sqrt{(\beta \pm \gamma)} = \xi + (\sqrt{(\beta + \gamma)} - \xi);$$

then, for some special choice of $\xi$, we might get a pretty expression on the right-hand side. Formula $(*)$ is not doing this, at least not directly, for it requires that

$$\text{both} \quad \sqrt{(\beta + \gamma)} = \xi + \eta \quad \text{and} \quad \sqrt{(\beta - \gamma)} = \xi - \eta,$$

and it is almost irresistible for us nowadays to treat these two cases of addition and subtraction together. Euclid, however, gives completely separate but parallel treatments of his equivalents to these cases of addition and subtraction, only linking them together at the end; and therefore so shall I. Here then is the first way the formula misrepresents Book X; and Euclid must therefore specify just how the right-hand side is constituted, just what form the $\beta$ and $\gamma$ can have.

### 5.1(e)  *Anthyphairetic considerations*

Let us now try to explore these ratios anthyphairetically. We immediately face two problems: that the evaluation of the anthyphairesis is sometimes rather difficult, beset with theoretical and practical problems, and that the behaviour revealed by the heuristic search seems to be unpredictable.

The first problem in evaluating the terms of a ratio like $\tfrac{1}{2}\sqrt{(10 \pm 2\sqrt{5})} : (1$ or $2)$, using some algorithm such as those described in Section 2.3, has already manifested itself in Section 4.2, where the example of $(\sqrt{2} + \sqrt{3}) : 1$ was cited to illustrate the intractability of algorithms for anthyphairetic arithmetic. I shall illustrate the difficulties in more detail using this slightly simpler ratio. The algorithm of Section 2.3(b) depended on some kind of test to identify whether

$$p : q \text{ is } <, =, \text{ or } > \quad (\sqrt{2} + \sqrt{3}) : 1.$$

Suppose we manipulate boldly, ignoring problems of meaning and interpretation, and all historical and mathematical problems:

$$p : q \text{ is } <, \text{ etc.} \quad (\sqrt{2} + \sqrt{3}) : 1 \text{ according as } p \text{ is } <, \text{ etc.} \quad (\sqrt{2} + \sqrt{3})q,$$

$$\text{i.e. according as } p^2 \text{ is } <, \text{ etc.} \quad (5 + 2\sqrt{6})q^2,$$

provided we can justify these operations, including the arithmetical result that $\sqrt{2}\cdot\sqrt{3}=\sqrt{6}$, on which see B93,

$$\text{i.e. according as } (p^2 - 5q^2) \text{ is } <, \text{ etc. } 2\sqrt{6}q^2,$$

$$\text{i.e. according as } p^4 - 10p^2q^2 + 25q^2 \text{ is } <, \text{ etc. } 24q^4,$$

although if $p^2 < 5q^2$, some further investigation is needed,

$$\text{i.e. according as } p^4 + q^4 \text{ is } <, \text{ etc. } 10p^2q^2.$$

The trouble now, as anybody who applies the algorithm will find, is that the numbers grow far too quickly for the calculation to be carried very far; the first three terms are calculated in Table 5.1 and, at the next step, the calculation will overflow the capacity of the standard Greek numerals (see Chapter 7). No periodic or regular behaviour presents itself, though we notice incidentally that $(\sqrt{2}+\sqrt{3}):1$ is close to and greater than the ratio of the circumference to diameter of a circle. (Explanations of this can be seen in the perimeters of a square inscribed in, or of a hexagon circumscribed around, a circle, or in the areas of an octagon inscribed in and of a hexagon circumscribed around it; see Popper, *OSE* i, Chapter 6, n. 9.)

We attempt these calculations for the same motives that lay behind the calculations of Section 3.3 (see especially Table 3.1), in the hope that they may uncover some plausible hypotheses. After this, we hope that the calculations will fall away, to be replaced by deductive proofs, as happened with the problem of the dimension of squares; compare Sections 3.3–3.6. So let us continue to ignore any historical and mathematical scruples relating to any aspect of this calculation, and pass to the results, to see if

TABLE 5.1. Evaluation of $(\sqrt{2}+\sqrt{3}):1 = [3, 6, 1, 3 \text{ or more}, \ldots]$.
(Compare Table 2.1)

| Under-estimate | Over-estimate | New estimate $p:q$ | $p^4$ | $q^4$ | $p^4 + q^4$ | $10p^2q^2$ | Under/equal/over |
|---|---|---|---|---|---|---|---|
| 3:1 | 4:1 | 7:2 | 2401 | 16 | 2417 | 1960 | over |
| 3:1 | 7:2 | 10:3 | 10 000 | 81 | 10 081 | 9000 | over |
| 3:1 | 10:3 | 13:4 | 28 561 | 256 | 28 817 | 27 040 | over |
| 3:1 | 13:4 | 16:5 | 65 536 | 625 | 66 161 | 64 000 | over |
| 3:1 | 16:5 | 19:6 | 130 321 | 1296 | 131 617 | 129 960 | over |
| 3:1 | 19:6 | 22:7 | 234 256 | 2401 | 236 657 | 237 160 | under |
| 22:7 | 19:6 | 41:13 | 2 825 761 | 28 561 | 2 854 322 | 2 840 890 | over |
| 22:7 | 41:13 | 63:20 | 15 752 961 | 160 000 | 15 912 961 | 15 876 000 | over |
| 22:7 | 63:20 | 85:27 | 52 200 625 | 531 441 | 52 732 066 | 52 670 250 | over |

any plausible hypotheses would eventually present themselves. For the pentagon, we get:

$$\text{side}:\text{circumradius} = \sqrt{(10-2\sqrt{5})}:2$$

$$= [1,5,1,2,3,2,28,2,27,3,22,1,7,1,9,\ldots],$$

$$\text{side}:\text{circumdiameter} = \sqrt{(10-2\sqrt{5})}:4$$

$$= [0,1,1,2,2,1,6,1,56,1,54,1,1,1,10,\ldots],$$

$$\text{diagonal}:\text{circumradius} = \sqrt{(10+2\sqrt{5})}:2$$

$$= [1,1,9,4,1,1,1,2,1,1,2,3,7,1,2,\ldots], \quad \text{and}$$

$$\text{diagonal}:\text{circumdiameter} = \sqrt{(10+2\sqrt{5})}:4$$

$$= [0,1,19,2,3,6,5,1,1,1,3,2,1,\ldots].$$

We are again faced with behaviour similar to that encountered in the problem of the dimension of cubes (see Section 4.3), that no regularity seems to be evident. The purely anthyphairetic approach seems blocked.

We can try to investigate this problem from the other side, and attempt to evaluate the ratio of the side or diagonal to circumradius or diameter of the pentagon directly, by geometric arguments such as were given in Section 2.1 for the simpler cases of the diagonals and sides of regular polygons. We can indeed try, but I can see no possibility of progress with the example of the circumcircle of the pentagon. However the case of the extreme and mean ratio is amenable to this approach, for here we can both evaluate the ratio geometrically and explore the construction metrically, and then compare what we find. We then get

$$\tfrac{1}{2}(\sqrt{5}-1):\tfrac{1}{2}(3-\sqrt{5}) = [\overline{1}].$$

We also notice that this particular expansion is much easier to evaluate directly, since the test for whether

$$p:q \text{ is } <, =, \text{ or } > \tfrac{1}{2}(\sqrt{5}-1):\tfrac{1}{2}(3-\sqrt{5})$$

reduces to finding whether

$$p^2 \text{ is } <, =, \text{ or } > pq + q^2,$$

and the numbers involved in the arithmetical exploration only grow quadratically. But it is difficult to see, from just this one example, what may be happening in general, what it is that is so special about the relation $(\sqrt{5}-1):(3-\sqrt{5})$ that it yields a purely periodic anthyphairesis. Also the issue is complicated by another phenomenon: all examples of ratios of sides of squares seem to have the form $[n_0, n_1, \ldots, 2n_0]$, and so contain an even term somewhere in their anthyphairesis. Hence the extreme and mean ratio is an example of a periodic anthyphairetic expansion which cannot be expressed as a

ratio of sides of squares (see Section 3.5(b)). Can anything more be said about this new class of expressible ratios?

A final series of evaluations will illustrate the problem further. The following calculations, one of which requires a great deal more computational or analytical effort, yield

$$(\sqrt{2} + \sqrt{3}) : \sqrt{6} = [1, 3, 1, 1, 15, 1, 1, 1, 19, 2, 4, 1, 1, 3, 47, 1, 3, \ldots],$$

$$(\sqrt{2} + \sqrt{3}) : (\sqrt{6} - 1) = [2, 5, 1, 6, 4, 2, 2, 2, 106, 12, 2, 1, 1, 2, 2, \ldots],$$

$$(\sqrt{2} + \sqrt{3}) : (\sqrt{6} + 1) = [0, 1, 10, 2, 1, 1, 1, 15, 9, 5, 2, 6, 5, 4, 17, \ldots],$$

and

$$(\sqrt{2} + \sqrt{3}) : (\sqrt{6} - 2) = [6, 1, 2742, 1, 1, 1, 3, 1, 160, 1, 1, 2, 5, 3, 1, \ldots],$$

but

$$(\sqrt{2} + \sqrt{3}) : (\sqrt{6} + 2) = [0, 1, \overline{2}].$$

Let us explore to see why this last result may be so much simpler than the rest. I shall consider the reciprocal ratio $(\sqrt{6} + 2) : (\sqrt{3} + \sqrt{2})$, taken 'the greater to the less', and will work out in detail the geometrical manipulations which simplify this expression. We start from the identities

$$(\sqrt{3} + \sqrt{2}) \cdot (\sqrt{3} - \sqrt{2}) + \sqrt{2} \cdot \sqrt{2} = \sqrt{3} \cdot \sqrt{3},$$

so

$$(\sqrt{3} + \sqrt{2}) \cdot (\sqrt{3} - \sqrt{2}) = 1 \cdot 1,$$

and

$$(\sqrt{6} + 2) \cdot (\sqrt{3} - \sqrt{2}) = \sqrt{2} \cdot 1,$$

which correspond to geometrical configurations of rectangles which the reader is strongly encouraged to draw. (Note that the latter identity uses manipulations like $\sqrt{6} \cdot \sqrt{3} = \sqrt{18} \cdot 1 = 3 \cdot \sqrt{2}$, on which again see $A_{88}$–$B_{93}$.) Hence we have:

$$(\sqrt{6} + 2) : (\sqrt{3} + \sqrt{2}) = (\sqrt{6} + 2) \cdot (\sqrt{3} - \sqrt{2}) : (\sqrt{3} + \sqrt{2}) \cdot (\sqrt{3} - \sqrt{2})$$

$$\text{(by the } \textit{Topics} \text{ proposition)}$$

$$= \sqrt{2} \cdot 1 : 1 \cdot 1 \qquad \text{(by the evaluations above)}$$

$$= \sqrt{2} : 1 \qquad \text{(by the } \textit{Topics} \text{ proposition)}$$

$$= [1, \overline{2}].$$

However, this result seems to be very special, and were such a manipulation as this ever explored, it would perhaps be more likely to prompt an investigation of the behaviour of rectangles like $(\sqrt{p} \pm \sqrt{q}) \cdot (\sqrt{r} \pm \sqrt{s})$ than of the underlying properties of anthyphairesis.

These manipulations may seem to us to be very artificial and laboured. But when Euclid deals with a similar kind of result at *Elements* X 112–14, his

treatment is also notoriously complicated and, if viewed algebraically, difficult to follow. He never seems to recognise the relevance and simplification that would follow from using freely the basic identity $(b + c) . (b - c) = b^2 - c^2$. Further, there seems to be nothing corresponding to this identity in the configurations of Book II though it is easy to draw a variant figure of II 5 that deals with it. I shall return briefly to this question at the end of Section 5.2(f) below.

The material I have described so far has provided sufficient clues to permit an informed guess as to what might be happening, but I do not think that much of this information would be accessible in any such coherent form within the early Greeks' arithmetical techniques, while the underlying phenomenon itself, if it were ever guessed, would probably lie outside their techniques of deductive geometrical proof. The result is that the *only* known regular behaviour of anthyphairesis connected with the arithmetical operations of taking sums, products, and roots (of any order) is that ratios of the form $\pm(p \pm \sqrt{q}):r$ (or, in arithmetised terms, those numbers that arise as solutions of quadratic equations with integer coefficients) will have an anthyphairesis that is eventually periodic,

$$\pm(p \pm \sqrt{q}):r = [n_0, n_1, \ldots, n_k, \overline{m_0, m_1, \ldots, m_l}];$$

and, conversely, any such expression will give rise to a ratio that can be expressed in this way. (This converse can be demonstrated geometrically by the final generalisation of the procedure of synthesising ratios in Section 3.5 but, as I observed here, I feel that this also would be highly implausible as a reconstruction of Greek procedures since it involves an extended subsidiary quasi-algebraic computation.) Much more reasonable would be to propose that any such arithmetical heuristic explorations of the problems were abandoned at some early stage since they gave no insight and were computationally intractable.

## 5.2 *ELEMENTS* X: A CLASSIFICATION OF SOME INCOMMENSURABLE LINES

### 5.2(a) *Introduction*

An investigation such as I have described in Section 5.1 sets up a remarkable contrast. On the one hand, there is the spectacular success of the anthyphairetic exploration of the ratios of sides of squares, $\sqrt{n}:\sqrt{m}$, in which a wide range of arithmetical and geometrical techniques and problems are bound together into some apparently coherent whole. On the other hand, there is an equally spectacular lack of success in the search for any similar regular, predictable anthyphairetic behaviour of any other geometrical ratios, apart from the extreme and mean ratio. Such a contrast might lead to the growth of an attitude in which these sides of squares come to be considered as the basic underlying understandable geometrical objects in terms of which everything else should be described. I propose that Book X of the *Elements* may indeed be a product of such a point of view.

The book opens with four definitions:

1. Those magnitudes are said to be commensurable (*summetroi*) which are measured by the same measure, and those incommensurable (*asummetroi*) which cannot have any common measure.

2. Straight lines are commensurable-in-square (*summetroi dunamei*) when the squares on them are measured by the same area, and incommensurable[-in-square] when the squares on them cannot possibly have any area as a common measure.

3. With these hypotheses, it is proved that there exist straight lines infinite in multitude which are commensurable and incommensurable respectively, some in length only, and others in square also, with an assigned (*protetheisa*) straight line. Let the assigned straight line be called expressible (*rhētos*), and those straight lines which are commensurable with it whether in length and in square, or in square only, expressible, but those incommensurable with it *alogoi*.

4. And let the square on the assigned straight line be called expressible, and those areas which are commensurable with it expressible, but those which are incommensurable with it *alogoi*, and the straight lines that produce them *alogoi*, that is, in case the areas are squares, the sides themselves, but in case they are any other rectilineal figures, the straight lines on which are described squares equal to them.

I have rendered *rhētos* as 'expressible', but have left *alogos* in transliterated form. This latter word means 'without words' or, if the material of Book X is indeed not just about the lines and areas in themselves, but also their mutual relations, 'without ratio'. These Greek words are almost always rendered into English as 'rational' and 'irrational', respectively. But:

(i) According to Euclid's definitions, if we denote the assigned line by $a$, then both $3a : 2a$ and $\sqrt{3}a : \sqrt{2}a$ are ratios of expressible lines. However, in our arithmetised description of ratios today, the first corresponds to what is universally called a rational number, the second to an irrational number. So the usual translation can lead to mathematical confusions and misunderstandings.

(ii) The use of the words 'rational' and 'irrational' gives the impression that the definitions refer to a dichotomy in which every line is either rational or irrational; but this is not the case in the original Greek. In fact, Euclid makes no attempt in the classification of Book X to consider every possible kind of line; rather, he works outwards from the assigned line to generate new kinds of lines and areas that are described in terms of the lines and areas that have already been considered. He first describes the expressible lines in Definition 3, and then he goes on to describe thirteen different kinds of *alogoi* lines that arise from some very special constructions.

(iii) The word pairs *rhētos/arrhētos* and *logos/alogos*, which do describe dichotomies, are not uncommon from the fifth or fourth century BC onwards, when they are found in a variety of technical and non-technical contexts. The only surviving possibly mathematical pre-Socratic use of *alogos* is in the title of a book by Democritus, *Peri alogōn grammōn kai nastōn*, and of *arrhētos* is in a later and unreliable report of an aphorism of Lysis, *arithmos arrhētos theos*; both of these will be discussed in Section 8.3(b). An example of Plato's use of

these words occurs at *Theaetetus* 201e–202c, where Socrates describes a dream in a remarkable passage many of whose words have mathematical overtones, for example:

> . . . In that way, the elements (*stoicheia*) have no account (*aloga*) and are unknowable (*agnōsta*), but they are perceivable (*aisthēta*); and the complexes (*sullabai*) are knowable (*gnōstai*) and expressible (*rhētai*) and judgeable (*doxastai*) in true judgement (*alētheia doxa*). . . .

Given the many problems that are associated with these words it would seem best, in contexts that might involve unfamiliar nuances of specialised technical meanings, to make a special effort to keep track of the original terminology. So I have included an index to the occurrences of *alogos* and *(ar)rhetos* in Plato, Aristotle, and the pre-Socratic philosophers at the end of this chapter, and there is a general survey of the topic of incommensurability in the new Appendix, Section 10.1.

The suggestion of rendering *rhētos* as 'expressible' was made in van der Waerden, *SA*, 168–79; he also translates *alogos* as 'unreasonable'. Another approach is taken in Taisbak, *CQ*: he makes no attempt to translate the words but merely replaces them by neutral, easily remembered alternatives: 'red' for *rhētos*, 'amber' for *mēsos*, and 'obscure' for *alogos*, and this has the additional effect of enhancing his opinion that Book X may be nothing more than a virtuoso display of mathematical reasoning. This question of the motivation of Book X has openly perplexed commentators for the past five hundred years or more, and practically everybody now cites Stevin's description of Book X as the "cross of mathematicians":

> La difficulté du dixiesme Livre d'Euclide est à plusieurs devenue en horreur, voire jusque à l'appeler la croix des mathématiciens, matière trop dure à digérer, et en la quelle n'aperçoivent aucune utilité [quoted from Heath, *TBEE* iii 9].

More recent opinions about the motivation and role of the book are:

> We are prepared to face the possibility that there was no other point than to entertain [!] us with good logic [Taisbak, *CQ*, 58],

> The true merit of Book X, and I believe it is no small one, lies in its being a unique [?] specimen of a fully elaborated deductive system of the sort that ancient philosophies consistently prized [Knorr, *CM*, 60],

> I believe that a satisfactory account is provided by attributing to Euclid the desire to fit the edge of the inscribed regular icosahedron into a classification scheme. . . . One would, of course, prefer an explanation that invoked a clear mathematical goal intelligible to us in terms of our own notions of mathematics, and which, under analysis, would lead univocally [?] to the reasoning in Book X. Unfortunately Book X has never been explicated successfully in this way, nor does it appear amenable to explication of this sort. Rather, Book X appears to be an expedient for dealing with a particular problem and at the same time a mathematical blind alley [Mueller, *PMDSEE*, 270–1],

> [Euclid] probably plunged himself into this treatment of irrationality, forgetting every-thing else in his work—research for the sake of research, art for art's sake. Euclid's

behaviour reminds us of the immortal verses of Goethe's songster who refuses the king's ostentatious presents because he is content with his own song: "I sing as the bird sings up on high / Amidst the leaves and boughs. / The heartfelt song that comes as a cry / its own reward allows" [*Wilhelm Meisters Lehrjahre*, II 12. Tr. from Euclid–Frajese & Maccioni, *EE*, 568–9].

With more than 115 propositions, some very long, the book is massive (more than a quarter of the bulk of the *Elements*), monolithic, and forbidding:

Book X does not make easy reading. . . . Up to X 28 it goes fairly well, but when the existence proofs start with X 29 . . . ones does not see very well what purpose all of this is to serve. The author succeeded admirably [?] in hiding his line of thought [van der Waerden, *SA*, 172],

Simply put, Book X is a pedagogical disaster [Knorr, *CM*, 59],

Was the lot of the ancient mathematics student really such a pleasant one? We recall the frequent laments by Hellenistic writers over the harshness of the schoolmasters, their ready resort to corporal punishment. . . . "Learning comes with pain", they said. This meant real physical pain in the lower school; but did it mature into intellectual pain at university level? [Knorr, *RTCQ*, on Book X. As someone who has spent his life teaching mathematics, I would not like to think that I or any of my colleagues could behave like that now, but I must confess that I cannot be entirely sure!],

*Danger!* Ne s'aventurer dans la lecture de ce livre qu'après une solide préparation [The Introduction, by Itard, to the reprinting of Euclid–Peyrard, *OE*, xiv].

### 5.2(b)  *Preliminary comments*

No comprehensible description of Book X can avoid a massive reorganisation of Euclid's opaque and unhelpful order of presentation and I shall not dwell on this fact any further. I shall also set aside most of the textual issues: questions of whether the lemmas are interpolations or not, whether certain propositions are authentic, how some lacunae should be filled, and discussions of possible authorship, and will refer to Euclid as the author of Book X, although his role was possibly at most editorial; on Euclid and his/her/their *Elements*, see Sections 6.2 and 10.4. Such is the nature of the material that it imposes a more formal approach than I have adopted hitherto.

My account will be based on the interpretation given in Taisbak, *CQ*, Chapter 2, 26–61. This book sets out a purely geometrical account, and the author's quirky, idiosyncratic, witty style provides the perfect antidote to the relentless, repetitive monotony of the original. This treatment, like many of Taisbak's mathematical and historical opinions, parallels much in Knorr, *CM* (which is an elaboration of Knorr's earlier discussion in *EEE*, 278–85; also see Knorr, *ETB*), but while Knorr's descriptions and procedures can generally be translated into geometrical terms, the tendency of an incautious reader will be to interpret his notation arithmetically. A useful, short, clear, but highly arithmetised description is given in van der Waerden, *SA*, 168–72; and a long and detailed account, which develops simultaneously Book X and its application in Book XIII, can be found in Mueller, *PMDSEE*, 251–306, though additional complexity is introduced there both in the elaborate notation

that is used and by the way the side and circumdiameter of the icosahedron is made the locus for problems that already arise, in their most acute form, in the much simpler configuration of the side and circumdiameter of the pentagon that I described at the beginning of this chapter. I refer the reader to these accounts, especially to Taisbak and Knorr, for further details of the treatment that I shall present here.

The organisation of Book X differs from the other books of the *Elements* in several ways. It opens with four definitions which, like all of the definitions elsewhere in our best manuscript of the *Elements*, are unnumbered and, like many of the other sets of definitions, they do not have a heading 'Definitions'. (This kind of editorial tidying-up is clearly set in parentheses or given notes in the translations given in Mueller, *PMDSEE*, and discussed below at the beginning of Chapter 7.) But Book X is unique in having further definitions inset in the body of the text: six headed 'Second Definitions' between Propositions 47 and 48, and six more headed 'Third Definitions' between 84 and 85; it also splits between Propositions 72 and 73, as will be explained below. The book also contains some thirty other unnumbered items—corollaries, lemmas, porisms, and remarks, not all of them coherent or correct—interspersed in the text (for a discussion of this feature, see Knorr, *EEE*, 269, n. 40) which, purely for simplicity, I shall often refer to as 'Propositions', so that Proposition 1/2, for example, will be the remark between Propositions 1 and 2.

### 5.2(c)  *Commensurable, incommensurable, and expressible lines and areas*

The basic language of Book X is set out in its opening definitions (quoted above) and Propositions 1 to 26; then Propositions 27 to 35 are a series of general constructions of examples relevant to the theory which follows. Propositions 1 to 4 describe the anthyphairesis of two magnitudes; they have no apparent relevance to the material that follows but I am proposing that they provide the underlying motivation for the whole classification; Definitions 1 and 2 and Propositions 5 to 16 deal with commensurability and incommensurability (the first time that this topic has been introduced in the *Elements*; see Section 10.1), most of which I shall assume is familiar to the reader; and Propositions 17 and 18 deal with the standard Euclidean theory of application of areas, which I shall explain briefly later, when it is used.

I now go on to Definitions 3 & 4, the remark 18/19, and Propositions 19 to 26, and will make some comments about Propositions 27 to 35.

DEFINITIONS 3 & 4   We fix a line $a$, called ***the* assigned (or expressible) line**. Any other **line** $b$ such that $b^2$ is commensurable with $a^2$ will be called **expressible**. (For example, $2a$ and $\sqrt{2}a$ are expressible lines.) Hence expressible lines may be commensurable (i.e. in length, for example $\sqrt{2}a$ and $\sqrt{18}a$) or (incommensurable-in-length but) commensurable-in-square-only (for example $a$ and $\sqrt{2}a$). An **area** that is commensurable with $a^2$ is called **expressible**. A line or area that is not expressible is called ***alogos***.

REMARKS

(i) The definitions and propositions will be cross-referenced to the Euclidean definitions and propositions on which they are based; readers are strongly urged to read Euclid's own enunciations in every case, and complete the sketches of proofs, or supply one where none is given here—a geometrical, not an arithmetical, proof, of course—and then compare their proofs with Euclid's. The basic set of figures is collected together in Fig. 5.3. Remarks, parenthetic comments, and other such items will not be found in the *Elements*, unless otherwise indicated.

(ii) Euclid refers in Definition 3 to the assigned line and its square as 'the expressible'. I shall maintain the distinction between *an* expressible line (or area) and *the* expressible line (or square) by always calling the latter *the* assigned line (or square), and reserving the letter *a* (or $a^2$) for it.

(iii) Euclid's expression *hē to chōrion dunamenē* means 'the side of the square [equal to]', and is abbreviated in Heath's translation to 'side'; see his notes on Definition 4 and Proposition 54, *TBEE* iii, 13 and 19. I shall often refer to it as the 'side' or 'square side' and abbreviate it to '$\sqrt{}$'. I repeat: when it occurs here in the context of Greek mathematics, '$\sqrt{}$' should always be thought of as a line, not a number, the square *side*, not the square *root*. It can be constructed using *Elements* II 14.

(iv) I shall use the geometric shorthand that was introduced in Section 1.2(c), developed thereafter, and described in B$_{89}$ and B$_{91}$: lower-case italic Roman letters $a, b, c, \ldots$ will always denote *lines*; $b \cdot c$ the rectangle with sides $b$ and $c$; and $b^2$ the square with side $b$. Upper-case italic Roman letters will denote *areas*, $\sqrt{B}$ the square side of $B$, and if $\lambda$ is a positive integer, an *arithmos*, then $\lambda b$ or $\lambda B$ denote $\lambda$ concatenated copies of $b$ or $B$—there is no problem with the ambiguous notation $\lambda b \cdot c$ since, by II 1, $(\lambda b) \cdot c$ and $\lambda(b \cdot c)$ are equal; and $\sqrt{(\lambda(b \cdot b))}$ will be abbreviated to $\sqrt{\lambda b}$. Care should be exercised with this last notation since $\sqrt{\lambda b}$ has no arithmetical sense. Yet further, $\sqrt[4]{\lambda b}$ has neither arithmetical nor geometric sense and will never be used; instead, the notation $\sqrt{\sqrt{\lambda b}}$ will be used for $\sqrt{\lambda(\sqrt{\lambda b} \cdot b)}$, the square side of a rectangle with sides $\sqrt{\lambda b}$ and $b$. (Note how these definitions are formulations of the statements $\sqrt{\lambda b} \cdot \sqrt{\lambda b} = \lambda b \cdot b$ and $\sqrt{\sqrt{\lambda b}} \cdot \sqrt{\sqrt{\lambda b}} = \sqrt{\lambda b} \cdot b$; as an exercise, prove that $\sqrt{\lambda^2 \mu b} = \lambda \sqrt{\mu b}$.) The notation just introduced could be extended to include the general expressible area, where $\lambda$ is a fraction, but the proofs do not involve anything more than integral $\lambda$; the exploration of examples will, however, usually lead to fractions.

This notation is a pure shorthand for geometrical configurations, especially those involving rectangles. The algebraic-looking expressions that arise and the 'coefficients' like $\sqrt{2}$, $\sqrt{\sqrt{2}}$, etc., do not have any autonomous existence and they should never be manipulated without reference to some geometric figure, though this figure may often only be implied. Heath's translation of Book X occupies more than 250 pages of his [1926 iii] and many of its arguments stretch over almost two pages: after Proposition 17, all but a handful of them are about

manipulating rectangles and squares, and they can be accurately and briefly summarised in this geometric shorthand.

PROPOSITION 12   Any two expressible areas are commensurable, and anything commensurable with an expressible area is expressible.

PROPOSITION 15   The sum and difference of two expressible areas are expressible.

PROPOSITIONS 12 & 18/19   A line commensurable or commensurable-in-square-only with an expressible line is expressible.

PROPOSITIONS 9 & 10   Two lines are commensurable if and only if their squares have a ratio equal to some square number. Example: $\sqrt{2}a$ and $\sqrt{18}a$, again.

DEFINITION 3 & PROPOSITION 18/19   If two expressible lines are incommensurable, their squares will be commensurable.

PROPOSITION 15   The sum and difference of two commensurable expressible lines are expressible.

PROPOSITIONS 36 & 73   The sum (or difference) of two incommensurable expressible lines is *alogos*; it will be called a **binomial** (or an **apotome**) line. (As the numbering indicates, this is out of place here; it will be restated and proved later.)

PROPOSITION 19   The rectangle contained by commensurable expressible lines is expressible.
PROOF   See Fig. 5.3(a), where $h$ and $w$ are commensurable and expressible; by Proposition VI 1, $h . w : h^2 = w : h$, etc.

REMARKS
   (i) The role of VI 1, the *Topics* proposition, is analysed in Knorr, *EEE*, 259–60. His conclusion is that "there is not a single theorem within the theory of irrational lines proper which does not employ Aristotle's theorem [sc. VI 1] either explicitly or once-removed."
   (ii) All the figures here are laid out horizontally to emphasise their affinity with VI 1; the Euclidean figures start vertically and then, after Proposition 38, most of them become horizontal. Euclid usually refers to the height as "an expressible straight line" and the width as the *platos*, usually translated as "the breadth", which suggests the letter $b$, but I want to reserve that for another use. These letters—$a$, $b$, $c$, $d$, $e$, $h$, $q$, $u$, $v$, $w$, $x$, & $y$—have been chosen carefully to denote special kinds of lines, and will be used consistently.

PROPOSITION 20   If an expressible area $B$ is applied to an expressible line $h$, so $B = h . w$, then $w$ will be expressible and commensurable with $h$.

(a)

$h$ | $h.w$ | $h^2$

$w$     $h$

(b)

$h$ | $b^2 = h.v$ | $c^2 = h.w$

$v$     $w$

(c)

$b$ | $b^2$ | $b.c$

$c^2$

$b$     $c$

(d)

$h$ | $b^2 = h.u$ | $b.c = h.v$ | $c^2 = h.w$

$u$     $v$     $w$

(e)

$b$     $c$
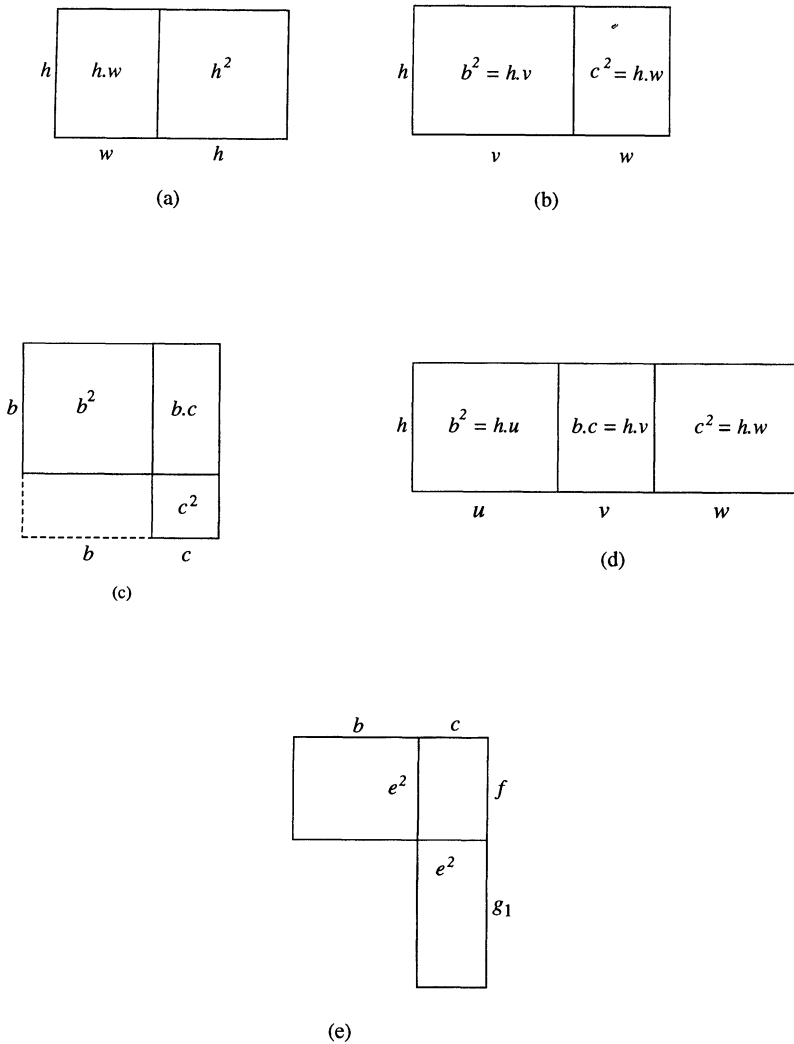
$e^2$ | $f$

$e^2$

$g_1$

FIG. 5.3

PROOF    Again see Fig. 5.3(a), etc.

Hereafter $h$ will be called **the height**; and $w$ will be called **the width of $B$ with respect to $h$**, these last four words usually being understood.

PROPOSITION 21    The rectangle contained by incommensurable expressible lines is *alogos*.
PROOF    See Fig. 5.3(a), yet again.

PROPOSITIONS 21 & 23/24    (Euclid gives no separate formal definition of this crucial idea.) An *alogos* **area** equal to the rectangle contained by incommensurable expressible lines is called **medial**. A **line** is called **medial** if it is the side of a medial area. For example: $\sqrt{2}a \cdot a$ is a medial area, $\sqrt{}\sqrt{2}a$ is a medial line.

PROPOSITION 22    If a medial area $b \cdot c$ (or, equivalently, the square on a medial line) is applied to an expressible height $h$, its width $w$ will be expressible and incommensurable with $h$.
PROOF    If $b \cdot c = h \cdot w$, then $h : c = b : w$ (VI 16), so $h^2 : c^2 = b^2 : w^2$ (VI 22). Since $h^2$ and $c^2$ are expressible, they are commensurable, so $b^2$ and $w^2$ are commensurable, etc. (Euclid's proof is slightly different, and uses VI 14, with its idea of reciprocal proportionality, associated with the operation of compounding; see Section 4.5(c).)

PROPOSITIONS 23 & 23/24    An area commensurable with a medial area is medial; and a line commensurable or commensurable-in-square-only with a medial line is medial.
PROOF    See Fig. 5.3(b): let $B = b^2$ be a medial area, and let $c^2$ be commensurable with $b^2$. Apply both to an expressible $h$, so $b^2 = h \cdot v$ and $c^2 = h \cdot w$. Then $v$ is expressible and incommensurable with $h$ (Proposition 22), so $w$ is expressible and incommensurable with $h$, so $h \cdot w = c^2$ is medial.

PROPOSITION 23/24    Two medial lines may be commensurable (e.g. $\sqrt{}\sqrt{2}a$ and $\sqrt{}\sqrt{32}a$), commensurable-in-square-only (e.g. $\sqrt{}\sqrt{2}a$ and $\sqrt{}\sqrt{8}a$), or incommensurable-in-square (e.g. $\sqrt{}\sqrt{2}a$ and $\sqrt{}\sqrt{3}a$).

Two medial areas may be commensurable or incommensurable. (Give examples!) It makes no *geometric* sense to say that incommensurable medial areas are commensurable-in-square, but we can give an *arithmetised* interpretation, in terms of the square and fourth roots of numbers: medial lines have length $\sqrt[4]{\lambda}$ and medial areas $\sqrt{\lambda}$, where $\lambda$ is not a perfect square, so any two medial areas are indeed arithmetically commensurable-in-square. It is even possible to articulate a version of this kind of result geometrically, using X 22: if two incommensurable medial areas are applied to an expressible height, then their breadths will be medial and commensurable-in-square-only. The fact that Book X contains no suggestion of this kind of result is yet another indication that its procedures are not being conceived in any arithmetical form.

PROPOSITION 24    The rectangle contained by commensurable medial lines is medial. For example: $\sqrt{}\sqrt{2}a$ and $\sqrt{}\sqrt{32}a$ contain the medial area $\sqrt{8}a \cdot a$.

PROPOSITION 53/54    A rectangle is the mean proportional between the squares on its sides.

PROOF   See Fig. 5.3(c): $b^2 : b \cdot c = b : c = b \cdot c : c^2$ by two applications of VI 1. This lemma duplicates a part of Euclid's proof of X 25.

PROPOSITIONS 25, 27, & 28   The rectangle contained by commensurable-in-square-only medial lines is either expressible or medial. Examples: $\sqrt{\sqrt{2}a}$ and $\sqrt{\sqrt{8}a}$ contain the expressible area $2a \cdot a$; and $\sqrt{\sqrt{6}a}$ and $\sqrt{\sqrt{24}a}$ contain the medial area $\sqrt{12}a \cdot a$.

PROOF   See Fig. 5.3(d): let $b$ and $c$ be medial lines, commensurable-in-square-only, and apply $b^2$, $b \cdot c$ and $c^2$ to the expressible height $h$, so $b^2 = h \cdot u$, $b \cdot c = h \cdot v$, and $c^2 = h \cdot w$. Then, by Proposition 22, $u$ and $w$ are expressible, and each is incommensurable with $h$, but since $u : w = (h \cdot u) : (h \cdot w) = b^2 : c^2$, they are commensurable with each other. Since, by the previous result, $b \cdot c$ is the mean proportional between $b^2$ and $c^2$, $v$ will be the mean proportional between $u$ and $w$. So, since $v^2 = u \cdot w$ where $u$ and $v$ are expressible and commensurable, $v$ will also be expressible. Finally, $v$ can be either commensurable or incommensurable with $h$, so $h \cdot v = b \cdot c$ can be either expressible or medial.

EXERCISES   Readers who wish to test their understanding of this material might now try to solve some of the problems posed in Propositions 27–35, but let me warn them that Euclid's examples are chosen strictly with a view to the theory that will follow and, until they have been put in their context, some of them (especially 33–5, which go beyond the corpus of lines so far described) will seem perverse and gratuitous. Moreover, while the enunciations seem to ask for specific examples, Euclid's proofs give general procedures for constructing all examples,[4] and this adds to their complexity. Since I have already given the two simplest and accessible results, Propositions 27 and 28, here are two more easy ones:

(i) If an expressible area is applied to a medial height, its width will be medial and commensurable-in-square-only with the height.

(ii) If a medial area is supplied to a medial height, its width will be medial and either commensurable or commensurable-in-square-only with the height.

Unlike the configurations described in these two exercises—which seem obvious manipulations to consider given the development of Book X so far— the lines in Book X that play the role of the height are always expressible (with the exceptions of Propositions 112–14, exceptions which do not detract from this general observation). This is one example (and I shall give another later, in note 6 which may indicate that Book X is not some purely formal and gratuitous exercise, as is suggested by Knorr and Taisbak (see Section 5.2(a), above), and which may seem the case from the description given so far).

The proof of one earlier result was postponed. Because this is important, indeed vital, to what follows, I give it here:

PROPOSITIONS 36 & 73   The sum (or difference) of two incommensurable expressible lines is *alogos*; it will be called a **binomial** (or **apotome**) line.

PROOF   Let $b$ and $c$ be incommensurable expressible lines; then, by II 4, $(b+c)^2 = b^2 + c^2 + 2b.c$. Now $b^2 + c^2$ is expressible and $b.c$ is medial; hence $2b.c$ is medial and so is incommensurable with $b^2 + c^2$. Hence $(b+c)^2$ is the sum of an expressible and a medial area, and so is not expressible, so is not commensurable with the expressible area $b^2 + c^2$; so it is *alogos*. The proof for $(b-c)$ is similar.

But I leave as final exercises:

PROPOSITION 26, generalised   Neither the sum, nor the difference of two incommensurable medial areas will be expressible or medial; and neither the sum, nor the difference of a medial and an expressible area will be expressible or medial. So the square sides of these areas will be *alogoi* lines. (The rest of Book X is, in effect, devoted to elucidating their properties; and the conclusion of these investigations is set out in Propositions 71 & 72, and 108 to 110.)

EXAMPLES   So far, only three kinds of lines have been named and proved to be *alogos*—medial, binomial, and apotome—but we are surrounded by other potential examples. For example:
- Higher order medials: $\sqrt{\surd\surd 2a}$, $\sqrt{\surd\surd\surd 2a}$, etc.; this is Proposition 115.
- 'Trinomials' and 'tri-apotomes' like $b \pm c \pm d$ for mutually incommensurable expressible lines; then 'quadrinomials', etc.
- 'Mixed binomials/trinomials/etc.', like $a \pm \sqrt{2}a \pm \sqrt{\surd 3a}$, etc.
- Lines of the form $\sqrt{b.c}$, $\sqrt{(b+c).d}$, etc., for any previously defined kinds of lines,

and there are completely different kinds of lines that are also almost certainly *alogoi*, for example:
- Edges of cubes, $\sqrt[3]{2a}$, etc.
- Cubic binomials, apotomes, trinomials, etc.
- Mixed 'edge–sides' like $\sqrt[3]{(\sqrt{2}a)}$ or $\sqrt{(\sqrt[3]{2}a)}$. (I cannot construct a purely geometrical proof that these lines are equal.)
- 'Circular irrationals' like the square sides of circles of expressible diameters.
- 'Elliptical irrationals', and perhaps also 'spiral irrationals',

and yet more combinations involving them.

Book X stakes off a little patch of coherent order in the middle of this jungle and, moreover, relates the otherwise unexplained lines that arise elsewhere in the *Elements* to this system. It is a remarkable achievement.

Readers who want to continue directly their reading of Book X should skip the next section.

### 5.2(d) *Interlude: surd numbers and alogoi magnitudes*

The word 'surd', as is well known, arises from a mistranslation of the Greek *alogos* into the Latin *surdus* (meaning silent, deaf, mute, etc.; cf. the French *sourd*) via the Arabic *aṣamm* (meaning deaf, etc.). (For more details see, for example, the *Oxford English Dictionary*, s.v. surd.) So, if we unscramble this

mistranslation, was Stevin's boast that the key to understanding Book X was to interpret it in terms of the *alogoi*? Surely Book X is nothing if it is not some kind of calculus of the *alogoi*.

Of course not. Stevin's surds are *numbers*, and he proposed to understand Book X in arithmetical terms. To this was soon to be added another dramatic development of his period in which he also participated and which was practically complete in its basics by the time of Descartes' *Géométrie* of 1637, namely the use of a symbolic algebraic notation. And I believe that the development of this algebraic notation owed no small debt to yet another contribution of Stevin's, his lucid introduction to the West of the idea of decimal fractions, in his pamphlet *De Thiende* of 1585.[5] For the first time there was a generally available, 'user-friendly', understandable, comprehensive way of doing arithmetic with the most general kinds of numbers. Mathematicians could now stop worrying about the details of each particular arithmetical manipulation and begin to abstract with confidence from the belief that any arithmetical calculation was possible. Of course the trick was as old as Babylon, and astronomers had been doing it to base 60 for more than three thousand years; but the Babylonians had a proper respect for the soft underbelly of this beguiling technique and avoided non-terminating expansions whenever they could; and sexagesimal arithmetic is still not user-friendly, even using pocket calculators which have it as a wired-in feature. There was for the first time, in the late sixteenth century, the general delusion that expressions like $\sqrt{2} \times \sqrt{3} = \sqrt{6}$ had some easily understood arithmetical sense. And it was a delusion that mathematicians would go on living with for 300 years of spectacular progress, but it remained unjustified until 24th November 1858 when Dedekind, as he tells us in his *Stetigkeit und irrationale Zahlen*, gave the first mathematically acceptable definition of the arithmetic involved and the first satisfactory explanation and proof of this identity; see Section 1.2(c).

These developments also affected geometry: the arithmetisation of geometry, which had been pottering along gently since the time of Heron and probably since the time of the merging of Greek geometrical and Babylonian arithmetical astronomy in the second century BC, exploded into action. This programme of arithmetisation of geometry starts by choosing, explicitly or implicitly, an assigned *unit* line; then all lines become endowed with a numerical length with respect to this line, with the essential feature that it is assumed that these lengths can be manipulated arithmetically. Similarly this unit line defines a *unit* square, whence all sufficiently simple plane regions (e.g. rectilinear, circular, etc.) are assigned numerical areas, and this definition is shown to be consistent with the definition of the area of a rectangle as the product of the lengths of its sides. Similarly for volumes. Lengths, areas, and volumes are then mixed and manipulated indiscriminately as dimensionless numbers, and we construct and solve expressions involving these numbers; and these expressions are abstracted into algebraic formulae which are also manipulated irrespective of their geometric sense. We then go on to explore the numbers in their own

right, and distinguish, for example, between rational, irrational, algebraic, transcendental, and other kinds of numbers.

Let us look at the effect the arithmetisation of geometry has on the basic language of Book X, such as I have described it so far. We have the assigned line *a*, so let us take that as the assigned *unit* line (and many commentators do indeed refer to it as such). Thereafter, instead of referring to lines $b, c, \ldots, x, y$, we refer to their lengths $\beta, \gamma, \ldots, \xi, \eta$, numbers; and instead of referring to geometrical areas $B, C, \ldots, X, Y$, we refer to their numerical areas, again numbers. So the essential geometrical difference between lines and areas that underlies Book X is not immediately captured by these formulae that are used to describe it. Worse, the length of expressible line can be either a rational number or the square root of a non-square rational number; an expressible area is a rational number; and a medial area is the square root of a non-square rational number. So we cannot immediately distinguish between expressible lines, expressible areas, and medial areas in these formulae.

If we have faith that Book X is doing something worth investigating and we want to try to understand just what it is, then the translation into the language we use to understand it should, at the very least, be able to represent clearly and without confusion the basic features of Euclid's text. In a choice between convenience for us and fidelity to Euclid, the second consideration should surely be paramount. So, while I think that an arithmetical interpretation can lead to insights into the mathematical problem that Euclid is exploring, I also feel that, if we are to understand Euclid's treatment, this arithmetised way of thinking may be a hindrance. And so also may be the complications caused by the translation of *rhetos* and *alogos* as rational and irrational, respectively.

### 5.2(e)  *The classification of Book X, and its use in Book XIII*

Choose and fix some expressible height *h*; and wherever there is a sum of two lines or two areas, *always write the larger one first.*

Propositions 36 to 72 of Book X describe properties of certain sums of pairs of lines or areas; then a parallel set of Propositions 73 to 110 describe precisely analogous properties of their difference. Proposition *n* corresponds to Proposition $(n + 37)$ for $36 \leqslant n \leqslant 70$; then Proposition 71, on an expressible area plus a medial area, splits into Propositions 108 (expressible minus medial) and 109 (medial minus expressible); and 72 (medial plus medial) corresponds to 110 (medial minus medial). Anyone who understands one of any of these pairs can immediately understand the other. Therefore I need only consider Propositions 36 to 72 here. In fact, the manuscripts of Book X have a break between Propositions 72 and 73, and heading: 'Second ordering of other ratios, those involving subtraction' (*deutera taxis heterōn logōn tōn kata aphairesin*).

Propositions 36 to 70 then split into five blocks of six and one block of five (Proposition 67 deals with two cases in one). If we understand the first proposition of each block properly, we can step through the rest almost

(+)

(−)

(a)

| $x.y$ | $y^2$ |
|---|---|
| $x^2$ | $x.y$ |

$\leftarrow x \rightarrow \leftarrow y \rightarrow$
$\leftarrow q_+ = x+y \rightarrow$

$e$

| $x.y$ | $y^2$ |
|---|---|
| $x^2$ | $x.y$ |

$\leftarrow x \rightarrow$
$\leftarrow q_- = x-y \rightarrow \leftarrow y \rightarrow$

(b)

$h.b = x^2 + y^2 = h.u + h.v$　　$h.c = 2x.y$　$h$

$\leftarrow b \rightarrow$
$\leftarrow w_+ = b+c \rightarrow$
$\leftarrow u \rightarrow \leftarrow v \rightarrow$

$h$　$h.b = x^2 + y^2 = h.u + h.v$
$h.c = 2x.y$

$\leftarrow b \rightarrow$
$\leftarrow w_- = b-c \rightarrow \leftarrow c \rightarrow$
$\leftarrow u \rightarrow \leftarrow v \rightarrow$

(c)

P　　$b/2$ $d/2$
S　R　Q　$c/2$
A　　C　　D B
$\leftarrow b/2 \rightarrow \leftarrow b/2 \rightarrow$
$\leftarrow d/2 \rightarrow$

P
S　R　Q
A　　C　　D B
$\leftarrow b/2 \rightarrow \leftarrow b/2 \rightarrow$
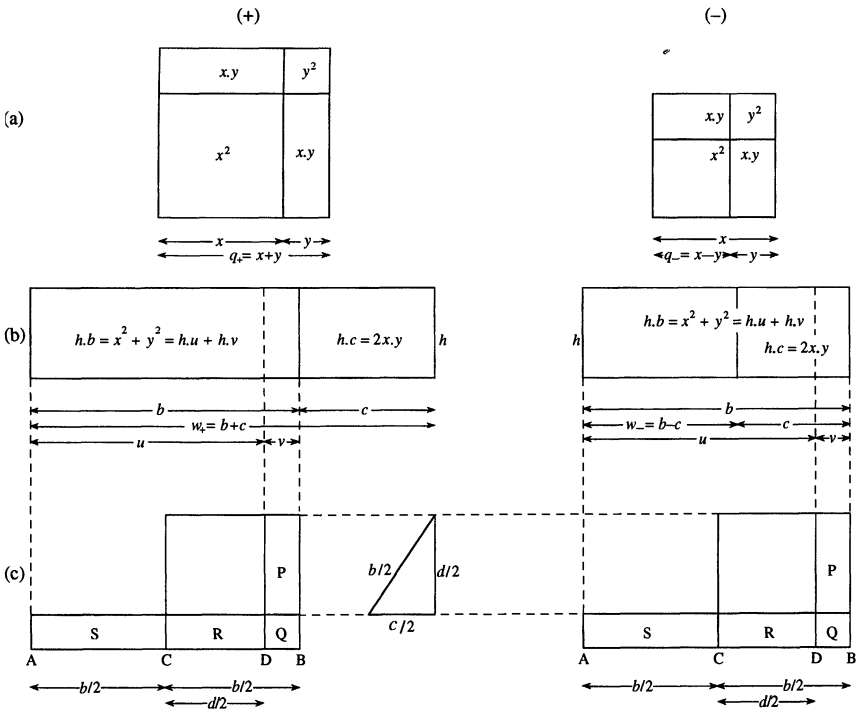$\leftarrow d/2 \rightarrow$

FIG. 5.4. To be read in conjunction with Table 5.2

TABLE 5.3. Typical examples of the six additive and subtractive lines, with respect to the height $h = a$. The letter $a$ has been omitted throughout for typographical convenience. For example, $(\sqrt{12} + 3) = (\sqrt{\sqrt{27}} + \sqrt{\sqrt{3}})^2$ is to be interpreted as 'the rectangle contained by $(\sqrt{12a} + 3a)$ and $a$ is equal to the square on $(\sqrt{\sqrt{27a}} + \sqrt{\sqrt{3a}})$'

| Class | Rectangle = Square | Name of the side of the square |
|---|---|---|
| 1 | $(3 \pm \sqrt{5}) = (\sqrt{\tfrac{5}{2}} \pm \sqrt{\tfrac{1}{2}})^2$ | binomial *or* apotome |
| 2 | $(\sqrt{12} \pm 3) = (\sqrt{\sqrt{27}} \pm \sqrt{\sqrt{3}})^2$ | first bimedial *or* first apotome of medial |
| 3 | $(\sqrt{8} \pm \sqrt{6}) = (\sqrt{\sqrt{\tfrac{9}{2}}} \pm \sqrt{\sqrt{\tfrac{1}{2}}})^2$ | second bimedial *or* second apotome of a medial |
| 4 | $(10 \pm \sqrt{20}) = (\sqrt{(5 + \sqrt{20})} \pm \sqrt{(5 - \sqrt{20})})^2$ | major *or* minor |
| 5 | $(\sqrt{5} \pm 1) = (\sqrt{(\sqrt{\tfrac{5}{4}} + 1)} \pm \sqrt{(\sqrt{\tfrac{5}{4}} - 1)})^2$ | side of an expressible plus a medial area *or* that, etc. |
| 6 | $(\sqrt{10} \pm \sqrt{8}) = (\sqrt{(\sqrt{\tfrac{5}{2}} + \sqrt{\tfrac{1}{2}})} \pm \sqrt{(\sqrt{\tfrac{5}{2}} - \sqrt{\tfrac{1}{2}})})^2$ | side of the sum of two medial areas *or* that, etc. |

TABLE 5.2, after Taisbak, *CQ*, 50. The six additive and six subtractive *alogoi* lines of Book X. To be read in conjunction with Fig. 5.4: $u+v=b$, $u.v=(c/2)^2$, and $b\,\text{Ȼ}\,c$; C = is commensurable with, Ȼ = is not commensurable with; exp = expressible, med = medial, and, in each row, terms with suffixes are commensurable if and only if the suffixes are equal

| Column | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Row | | | | | | | | | |
| Class of $w = b \pm c$ $b\,\text{Ȼ}\,c$ | $h\ \&\ b$ | $h\ \&\ c$ | $b\ \&\ d$ or $u\ \&\ v$ | $x^2 =$ $h.v$ | $y^2 =$ $h.v$ | $x^2+y^2 =$ $h.b$ | $2x.y =$ $h.c$ | Name of $q_+ = x+y$ | Name of $q_- = x-y$ |
| 1 | C | Ȼ | C | exp | exp | exp | med | binomial | apotome |
| 2 | Ȼ | C | C | $\text{med}_1$ | $\text{med}_1$ | $\text{med}_1$ | exp | first bimedial | first apotome of a medial |
| 3 | Ȼ | Ȼ | C | $\text{med}_1$ | $\text{med}_1$ | $\text{med}_1$ | $\text{med}_2$ | second bimedial | second apotome of a medial |
| 4 | C | Ȼ | Ȼ | | | exp | med | major | minor |
| 5 | Ȼ | C | Ȼ | | | med | exp | side of an expressible plus a medial area | that which produces with an expressible area a medial |
| 6 | Ȼ | Ȼ | Ȼ | | | $\text{med}_1$ | $\text{med}_2$ | side of the sum of two medial areas | that which produces with a medial area a medial whole |

automatically. So we need only understand five propositions—36, 42, 48, 54, and 60—and Definitions 47/48 to understand all of this material!

The geometrical configurations needed for all of these propositions, 36–110, are given in Fig. 5.4, on the left for addition, and on the right for subtraction. This composite figure, read in conjunction with Table 5.2, will describe how the procedure works; and together they comprise what can, I think, properly be called "the geometrical quintessence of Book X". They are based on Taisbak, *CQ*, 40, 42, and 50.

BASIC PROPOSITIONS ON ADDITION VI 28, X 16/17, 17, 18, & 59/60. Let $q = x + y$. If $w$ is the width of $q^2$ with respect to the height $h$, $q^2 = h \cdot w$, then $w$ can be written

$$w = b + c, \quad \text{where} \quad h \cdot b = x^2 + y^2 \quad \text{and} \quad h \cdot c = 2x \cdot y.$$

Conversely, given $w = b + c$, then $q = \sqrt{h \cdot w}$ can be written

$$q = x + y, \quad \text{where} \quad x^2 + y^2 = h \cdot b \quad \text{and} \quad 2x \cdot y = h \cdot c.$$

Note that, in these expressions, $b \geqslant c$ and $x \geqslant y$, by the convention stated at the beginning of this section. This detail is essential.

PROOF   If $q = x + y$, then $q^2 = x^2 + y^2 + 2x \cdot y$ (Fig. 5.4($a_+$) & II 4). Now construct the following rectangles and arrange them as in Fig. 5.4($b_+$): $h \cdot u = x^2$; $h \cdot v = y^2$, with $u + v = b$; and $h \cdot c = 2x \cdot y$. Since $x^2 + y^2 = (x - y)^2 + 2x \cdot y$ (Fig. 5.4($a_-$)), we have $x^2 + y^2 > 2x \cdot y$, i.e. $b > c$.

Conversely, given $w = b + c$ and $h \cdot w = q^2$, we first find how to write $b = u + v$ such that $h \cdot u = x^2$, $h \cdot v = y^2$, and $x \cdot y = \frac{1}{2}h \cdot c$.

Let us explore these relations. Since $x \cdot y$ is the mean proportion between $x^2$ and $y^2$ (see the discussions of Propositions 53/54 and 25, 27, & 28, above), $\frac{1}{2}h \cdot c$ will be the mean proportion between $h \cdot u$ and $h \cdot v$, so $\frac{1}{2}c$ will be the mean proportion between $u$ and $v$, so $u \cdot v = (\frac{1}{2}c)^2$. Hence we must solve the geometric problem $u + v = b$, $u \cdot v = (\frac{1}{2}c)^2$. This is precisely the problem of elliptical application of areas, introduced in Book X at Propositions 16/17, 17, and 18. Proposition 17 states, with my additions:

If there are two unequal straight lines [$b$ and $c$], and to the greater [$b$] there be applied a parallelogram [here a rectangle] equal to the fourth part of the square on the less [i.e. $u \cdot v = (\frac{1}{2}c)^2$] and deficient (*elleipon*) by a square figure [i.e. $u + v = b$] and . . . .

The solution of this problem is given at VI 28; it requires a condition (*diorismos*) which here becomes that $(\frac{1}{2}c)^2$ must not be greater than $(\frac{1}{2}b)^2$, precisely the hypothesis of our proposition. We therefore follow the procedure of VI 28, which is set out in Fig. 5.4($c_+$): bisect $AB = b$ at C, construct the square on BC, $(\frac{1}{2}b)^2$, and the square with side $CD = \frac{1}{2}d$ where $(\frac{1}{2}d)^2 = (\frac{1}{2}b)^2 - (\frac{1}{2}c)^2$, so that $\frac{1}{2}b$ is the hypotenuse of the right-angled triangle with shorter sides $\frac{1}{2}c$ and $\frac{1}{2}d$. But then

$$P + Q + R = \mathrm{BC}^2 - \mathrm{DC}^2 = (\tfrac{1}{2}b)^2 - (\tfrac{1}{2}d)^2 = (\tfrac{1}{2}c)^2,$$

while
$$P + Q + R = S + R \quad \text{since} \quad P + Q = S;$$
hence $S + R$ is our required elliptical application of $(\frac{1}{2}c)^2$ to AB. So AD = $u$ and DB = $v$ is the required splitting of AB = $b$.

Finally, we construct $x^2 = h \cdot u$ and $y^2 = h \cdot v$ on the same diagonal, as in Fig. 5.4(a$_+$), and verify immediately that $q = x + y$ gives the required decomposition of the square.                                    QED

COROLLARIES 17 & 18   With these notations, $u$ is commensurable with $v$ if and only if $b$ is commensurable with $d$.

PROOF    See Fig. 5.4(c$_{+ \text{ or } -}$): $u = \frac{1}{2}b + \frac{1}{2}d$, $v = \frac{1}{2}b - \frac{1}{2}d$ hence $b = u + v$ and $d = u - v$.                                    QED

Call $d = \sqrt{(b^2 - c^2)}$ the **square difference** of $b$ and $c$; Euclid's terminology is described in Heath, *TBEE* iii, 43; *he b ara tōs c meizon dunatai tē d*, 'the [square on] $b$ is greater than the [square on] $c$ by the [square on] $d$'.

This basic proposition describes the interplay of the following two fundamental moves around which Book X is constructed:

(r → s), rectangle → square: Given a line $w$, convert the rectangle $h \cdot w$ into a square $q^2$ and take its side: $q = \sqrt{h} \cdot w$, and

(s → r): Given a line $q$, take the width of $q^2$ with respect to the height $h$, $q^2 = h \cdot w$.

In Propositions 48 to 59, $w = b + c$ runs through a subclassification of binomial lines into six different kinds, and the operation (r → s), $q = x + y$, generates the six named additive *alogoi* lines. (There is a slightly confusing detail at the beginning, that the first kind of binomial $b + c$ will generate a binomial $x + y$; so we only get five new names of *alogoi* lines.) But Euclid does not set out the theory this way. Instead, he presents us, stone cold, with the ingredients $x$ and $y$ of the new *alogoi* in Propositions 31 to 35, defines these *alogoi* in 36 to 41, explores their properties in 42 to 47, introduces the subclassification of binomials in Definitions 47/48 and Propositions 48 to 53; and only then explains what he is doing when the operations (r → s) surface in 54 to 59 and (s → r) in 60 to 65.

We start this programme with a geometrical version of our opening exploration of $(\sqrt{\xi} + \sqrt{\eta})^2 = (\xi + \eta + 2\sqrt{\xi\eta}) = (\beta + \gamma)$ in Section 5.2. Compare the two approaches!

DEFINITION 47/48.1, AND PROPOSITIONS 54 & 60   If the square on a binomial $q = x + y$ is applied to an expressible height $h$, its width $w$ will be a binomial $w = b + c$ in which the expressible lines $b$ and $c$ satisfy: $b$ is commensurable with $h$; $c$ is incommensurable with $h$; and $d = \sqrt{(b^2 - c^2)}$ is

commensurable with $b$. Such a binomial will be called a **first binomial with respect to $h$**.

Conversely, if $w = b + c$ is a first binomial with respect to $h$, then the square side of $h \cdot (b + c)$ will be a binomial.

PROOF    See Fig. 5.4(+): we have $(x + y)^2 = h \cdot (b + c)$ where $h \cdot b = x^2 + y^2$, $h \cdot c = 2x \cdot y$, $x^2 = h \cdot u$, $y^2 = h \cdot v$, $b = u + v$, and $d = u - v$. Suppose, first, that $x + y$ is a binomial; then $x$ and $y$ are expressible lines, so $x^2$ and $y^2$ are expressible areas, so $h \cdot b = x^2 + y^2$ is expressible, so $b$ is expressible and commensurable with $h$. Similarly $h \cdot d = x^2 - y^2$ is expressible so $d$ is also expressible and commensurable with $h$. Finally, since $x$ and $y$ are incommensurable, $x \cdot y$ is medial, so $2x \cdot y = h \cdot c$ is medial, so $c$ is expressible and incommensurable with $h$. So $b + c$ is a first binomial with respect to $h$.

Conversely, if $b + c$ is a first binomial with respect to $h$, then $u = \frac{1}{2}(b + d)$ and $v = \frac{1}{2}(b - d)$ are also commensurable with $h$. So $x^2 = h \cdot u$ and $y^2 = h \cdot v$ are expressible areas, so $x$ and $y$ are expressible lines; and since $2x \cdot y = h \cdot c$ is medial, $x$ and $y$ are incommensurable. So $x + y$ is a binomial line.      QED

This proof is summarised in the column heading and top line of Table 5.2. The remaining five lines will correspond to the five possible other relations of (in)commensurability that can hold between $h$, $b$, and $d$, which will generate the remaining five classes of binomials with respect to $h$.

Abbreviate 'is commensurable with' as C, and 'is not commensurable with' as ¢, and organise the configurations as follows. (This notation and arrangement was introduced in Knorr, CM.)

SECOND    DEFINITIONS    47/48    There    are    the    following    six    mutually exclusive classes of binomials:

$$
\begin{array}{llllll}
\text{1st} & hCb & \& & h\text{¢}c & \& & bCd \\
\text{2nd} & h\text{¢}b & \& & hCc & \& & bCd \\
\text{3rd} & h\text{¢}b & \& & h\text{¢}c & \& & bCd \\
\text{4th} & hCb & \& & h\text{¢}c & \& & b\text{¢}d \\
\text{5th} & h\text{¢}b & \& & hCc & \& & b\text{¢}d \\
\text{6th} & h\text{¢}b & \& & h\text{¢}c & \& & b\text{¢}d
\end{array}
$$

If $w = b + c$ belongs to the $k$th class (or order), it will be called a $k$th **binomial with respect to $h$**.[6]

I shall call this **the subclassification of binomials with respect to $h$**; it is set out in columns 1 to 3 of Table 5.2. Can all cases be realised? (For example, we cannot have $hCb$, $hCc$, $bCd$, since $b + c$ is not then a binomial.) We must check that the following is possible:

PROPOSITIONS 48 TO 53    To find a $k$th binomial with respect to $h$.

In the narrow sense, these propositions are satisfied by a set of examples such as are listed in Table 5.3; but again Euclid gives a general method in each case, starting from the general constructions purportedly given in Lemma 28/29.[7]

Now look again at the caption of Table 5.2 and the headings of its columns. We have proved that:

(i) The entries in columns 1 and 6, and in columns 2 and 7, are equivalent, including the subscripts.

(ii) The entries in columns 3, 4, and 5 imply those in column 6, in rows 1 to 3.

It will be straightforward to read the table from left to right; this gives the process (r → s) which corresponds to Propositions 54 to 59, described below, and yields the five new *alogoi* lines. In order to read the table from right to left, (s → r), Propositions 60 to 65, we need the definitions of these new types of lines, which are given within Propositions 36 to 41. And Table 5.3 gives some examples of these new *alogoi* lines.

Here are the details. The process goes like clockwork.

PROPOSITIONS 37, 55, & 61    If $w = b + c$ is a 2nd binomial with respect to $h$, then $x$ and $y$ are medial, commensurable-in-square-only, and contain an expressible rectangle. Such a line $q = x + y$ is called a **first bimedial**, and it is *alogos*.

Conversely, if $q$ is a first bimedial, $w$ is a 2nd binomial with respect to $h$.

PROOF    See row 2: if $w$ is a 2nd binomial, $x^2 + y^2$ is medial (column 6); $u$ and $v$ are commensurable (column 3), so $x^2 = h . u$ and $y^2 = h . v$ are commensurable (columns 4 and 5); but neither $x^2$ nor $y^2$ can be expressible (since their sum is medial), so $x^2$, $y^2$, and $x^2 + y^2$ are medial and mutually commensurable. So $x$ and $y$ are medial, and commensurable-in-square-only, since $x . y$ is expressible. And $(x + y)^2 = x^2 + y^2 + 2x . y$ is the sum of a medial and an expressible area, so is *alogos*.

Conversely, if $q$ is a first bimedial, columns 4, 5, and 7 hold by definition, and column 6 follows immediately. But these imply the properties in columns 1–3.                                                                                    QED

PROPOSITIONS 38, 56, & 62    If $w$ is a 3rd binomial with respect to $h$, then $x$ and $y$ are medial, commensurable-in-square-only, and contain a medial rectangle. Such a line $q = x + y$ is called a **second bimedial**, and it is *alogos*.

Conversely, if $q$ is a second bimedial, $w$ is a 3rd binomial with respect to $h$.

PROOF    See row 3: as in the previous proposition, $x^2$, $y^2$, and $x^2 + y^2$ are medial and mutually commensurable. But this time $x . y = \frac{1}{2} h . c$ is medial and incommensurable with $x^2 + y^2 = h . b$. So $(x + y)^2 = x^2 + y^2 + 2x . y$ is the sum of incommensurable medial areas, so $q$ is *alogos* and not medial. So $x$ and $y$ are commensurable-in-square-only.

Conversely, again, columns 6 and 7 imply columns 1 and 2. Also, since the areas in columns 4 and 5 are commensurable, this implies column 3.    QED

PROPOSITIONS 39, 57, & 63   If $w$ is a 4th binomial with respect to $h$, then $x$ and $y$ are incommensurable-in-square and such that $x^2 + y^2$ is expressible and $x \cdot y$ is medial. Such a line $q = x + y$ is called a **major**, and it is *alogos*.
   Conversely, etc.
PROOF   Surely clear from row 4!

PROPOSITIONS 40, 58, & 64   If $w$ is a 5th binomial with respect to $h$, then $x$ and $y$ are incommensurable-in-square and such that $x^2 + y^2$ is medial and $x \cdot y$ is expressible. Such a line $q = x + y$ is called the **side of an expressible plus a medial area**, and it is *alogos*.
   Conversely, etc.
PROOF   Clear, even to the description of the names, from row 5.

PROPOSITIONS 41, 59, & 65   If $w$ is a 6th binomial with respect to $h$, then $x$ and $y$ are incommensurable-in-square and such that $x^2 + y^2$ and $x \cdot y$ are incommensurable and both medial. Such a line $q = x + y$ is called the **side of a sum of two [incommensurable] medial areas**, and it is *alogos*.
   Conversely, etc.
PROOF   See row 6: we need only verify that $x^2 + y^2 = h \cdot b$ and $2x \cdot y = h \cdot c$ are incommensurable. But $b$ is, as always, incommensurable with $c$.      QED


   We now prove the following:

PROPOSITIONS 42 TO 47   Each of these *alogos* lines $q$ can be expressed in only one way as a sum $x + y$ satisfying the descriptions in columns 6 and 7 of Table 5.2.


We use the following:

LEMMA 41/42   If $x + y = x_1 + y_1$ with $x > x_1$, then $x^2 + y^2 > x_1^2 + y_1^2$.

PROOF   Write $x + y = 2s$, $x - s = d$, and $x_1 - s = d_1$; then $d > d_1$. We now have exactly the configuration of II 5, so $x \cdot y + d^2 = s^2 = x_1 \cdot y_1 + d_1^2$. Hence $x \cdot y < x_1 \cdot y_1$, and   since   $x^2 + y^2 + 2x \cdot y = x_1^2 + y_1^2 + 2x_1 \cdot y_1$,   we   see   that $x^2 + y^2 > x_1^2 + y_1^2$.      QED

PROOF OF PROPOSITIONS 42 TO 47   Consider first the case of the binomial. Suppose $q = x + y = x_1 + y_1$, both binomials. If $x$ and $x_1$ are not equal, suppose $x > x_1$. Then $x^2 + y^2 + 2x \cdot y = x_1^2 + y_1^2 + 2x \cdot y_1$, and, by the lemma, $(x^2 + y^2) > (x_1^2 + y_1^2)$. Hence $(x^2 + y^2) - (x_1^2 + y_1^2) = 2x_1 \cdot y_1 - 2x \cdot y$. But this left-hand side is expressible, while the right-hand side is not, which is not possible.
   A similar argument is given by Euclid for any $q$ in rows 2, 4, and 5: look at columns 6 and 7. However the following argument, given by Euclid for rows 3 and 6, will, in fact, deal with all of the rows 2 to 6:

Let $q = x + y$ be any of the *alogoi* lines from rows 2 to 6 and suppose, as above, that $q = x + y = x_1 + y_1$ with $x > x_1$. Then, by the operation (s → r) of Propositions 60 to 65, $q^2 = h \cdot w$, where $w$ is a binomial, which we can express as $w = b + c$ with $h \cdot b = x^2 + y^2$ and $h \cdot c = 2x \cdot y$, or $w = b_1 + c_1$ with $h \cdot b_1 = x_1^2 + y_1^2$. Hence $b > b_1$, by the lemma just proved; and thus $w$ has been written as a binomial in two different ways, which contradicts the result already proved for binomials.                                                                    QED

It is a basic feature of expressible and medial lines that a line commensurable with an expressible line is also expressible, and a line commensurable with a medial is also medial. Euclid proves, in Propositions 66 to 70, that this applies to the other additive *alogoi* lines, also including, in 66, that a line commensurable with a $k$th binomial is also a $k$th binomial. It is also a basic feature of expressible and medial lines that the same obtains for commensurability-in-square, and Euclid ought to have proved that this applies to the other *alogoi* lines also, for he appears to use it in XIII 18, as I shall explain later. (But this property does not apply to the subclassification of binomials.). And, at the same time, we can prove that the lines are divided in the same ratio.[8] So we have:

PROPOSITION 66, strengthened   A line commensurable with a $k$th binomial with respect to $h$ is also a $k$th binomial with respect to $h$, and is divided in the same ratio.

(66 to 70, strengthened)   A line (commensurable or) commensurable-in-square with any of the additive *alogos* lines of Table 5.2, column 8, is the same kind of *alogos* line, and is divided in the same ratio.

PROOF   It is straightforward to see that a line $w_1$ commensurable with a $k$th binomial $w = b + c$ is also a $k$th binomial divided in the same ratio (i.e. $w_1 = b_1 + c_1$ with $b:c = b_1:c_1$; so $b:b_1 = c:c_1 = (b+c):(b_1+c_1)$). And it is also obvious why this does not apply to commensurability-in-square, since the width of the square on any binomial applied to $h$ is a 1st binomial with respect to $h$.

Now suppose $q_1$ is commensurable-in-square with some *alogos* line $q = x + y$ of column 8, so $q_1^2 = h \cdot w_1$ is commensurable with $q^2 = h \cdot (b+c)$; so $b + c$ is a binomial and $w_1$ is commensurable with $b + c$; and so, by what we have just proved, $w_1$ is also a binomial, $w_1 = b_1 + c_1$, in the same class and divided in the same ratio. So $q_1 = x_1 + y_1$ will be the same kind of *alogos* line as $q$. Moreover $(x + y)^2 : (x_1 + y_1)^2 = (b+c):(b_1+c_1) = b:b_1 = c:c_1$. We also have $(x^2 + y^2) : (x_1^2 + y_1^2) = h \cdot b : h \cdot b_1$ and $2x \cdot y : 2x_1 \cdot y_1 = h \cdot c : h \cdot c_1$, from which we easily deduce that $x:y = x_1:y_1$.                                    QED

Propositions 71 and 72 give a complete answer to the part of the final exercise of Section 4 above which deals with addition. I give 72 first, for its enunciation is a little startling:

PROPOSITION 72   The side of the sum of two incommensurable medial areas is either the side of the sum of two medial areas, or is a second bimedial! (The first alternative is, of course, the name of the additive *alogos* line in row 6 of Table 5.3.)

PROOF   Let $B$ and $C$ be incommensurable medial areas: apply both to $h$, $B = h.b$, $C = h.c$ in the configuration of Fig. 5.4($b_+$). Then $h\mathbb{C}b$, $h\mathbb{C}c$, and $b\mathbb{C}c$. Either $b\mathbb{C}d$, and $b + c$ is a 6th binomial, so $q = \sqrt{h}.(b + c) = \sqrt{(B + C)}$ is the side of the sum of two medial areas; or $bCd$, so $b + c$ is a 3rd binomial, and $\sqrt{(B + C)}$ is a second bimedial.                      QED

We see that the second case is a special case of the first. Nowadays we might be inclined to explain what is going on: "The side of the sum . . . is, in general, the line called 'the side of the sum of two medial areas'; in particular cases, it is the line called a second binomial".[9] There is also:

PROPOSITION 71   The side of the sum of an expressible and medial area is either . . . .

We have, I think, now covered all propositions up to 72; hence, by applying our knowledge to the subtractive cases, we have also covered 73 to 110, and so have described all of the additive and subtractive *alogoi* lines of columns 8 and 9 of Table 5.2; and also the subdivision of the binomials and apotomes into six classes each; and the medial line. We need once again to check that they are all different. We start with:

PROPOSITION 111   No binomial is commensurable with any apotome.

PROOF by contradiction.   Suppose $q = x + y = x_1 - y_1$ where $x$ and $y$, and $x_1$ and $y_1$ are pairs of incommensurable expressible lines. Write $q^2 = h.w$, then $w = b + c = b_1 - c_1$, where $b + c$ is a 1st binomial and $b_1 - c_1$ a 1st apotome; so $b$ and $b_1$ are both commensurable with $h$, and $c$ and $c_1$ are both incommensurable with $h$. Suppose $b < b_1$; then $c = (b_1 - b) - c_1$, so the expressible line $c$ has been written as an apotome, which is impossible. The case $b > b_1$ is identical.                      QED

We see immediately that this result also holds for commensurability-in-square. And we then use this result as a bridge to prove:

PROPOSITIONS 72/73 & 111/112, strengthened.   The thirteen classes of *alogoi* lines: medial, binomial, apotome, first bimedial, first apotome of a medial, . . . are all disjoint; and any line that is commensurable or commensurable-in-square with a line in a given class is also in that class.

There is some scholarly discussion over whether Propositions 112 to 115 are later interpolations into the manuscripts or not. However since I am leaving aside all but the most blatant textual questions, I now go on to consider them.

PROPOSITIONS 112 & 113   If an expressible area $e^2$ is applied to a binomial (or apotome) height, its width will be an apotome (or binomial) the terms of which are commensurable with the corresponding terms of the binomial (apotome), and in the same ratio, and this apotome (binomial) will be of the same order as the binomial (apotome) with respect to any given expressible line $h$.

PROPOSITION 114   If $(b + c)$ is a binomial, $(f - g)$ an apotome, $b$ and $f$, and $c$ and $g$ are commensurable, and $b : c = f : g$, then $(b + c) \cdot (f - g)$ is expressible.

The arithmetical versions of these results are simple variations on the identity $(\beta + \gamma) \cdot (\beta - \gamma) = (\beta^2 - \gamma^2)$, but the intricacies of Euclid's proofs illustrate the increasing divergence between the arithmetical and geometrical points of view.[10] I explain here the proof of 112 in detail:

PROOF OF 112   See Fig. 5.3(e); suppose the expressible area $e^2$ is applied to the binomial $(b + c)$ as height, so $e^2 = (b + c) \cdot f$, then its width $f$ will be *alogos*. We shall try to find $g$ so that $f = (f + g) - g$ displays that $f$ is an apotome. (This procedure is familiar from Euclid's treatment of apotomes, where $g$ is called the **annex** of $f$; see, for example, Propositions 79 to 84.) So we need to find the appropriate expressible line $(f + g)$ bigger than $f$. Various possibilities suggest themselves: to apply $e^2$ to $b$ or to $c$, or to $b - c$, and take its width. I leave the reader to explore the first and third alternatives, which appear to me to lead nowhere; here, following Euclid, we pursue the second.

If $e^2$ is applied to $c$, its width will be expressible, commensurable with $c$, and bigger than $f$; so write it as $f + g_1$. Then

$$e^2 = c \cdot (f + g_1) = (b + c) \cdot f.$$

Hence

$$(b + c) : c = (f + g_1) : f$$

and so, *separando*,

$$b : c = g_1 : f.$$

Unfortunately, in the resulting expression, $f = (f + g_1) - g_1$,

(i) $g_1 = (f + g_1) - f$ is the difference between an expressible and an *alogos* line, so is *alogos*, and

(ii) we do not have $(f + g_1) : g_1 = b : c$.

So try again. We first deal with the second condition by finding $g_2$ such that

$$(f + g_2) : g_2 = b : c,$$

$$\text{so also} = g_1 : f,$$

$$\text{and so too} = (f + g_1 + g_2) : (f + g_2).$$

[We are now at Heath, *TBEE* iii, 244.11–18.] Rearranging the order of this,

$$(f + g_1 + g_2) : (f + g_2) = (f + g_2) : g_2 = b : \check{c} = g_1 : f,$$

and we see a continued proportion in the first two terms, so (by VI 19, for squares),

$$(f + g_1 + g_2)^2 : (f + g_2)^2 = (f + g_1 + g_2) : g_2$$

and also $= (f + g_2)^2 : g_2^2$

and also $= b^2 : c^2$, which are commensurable.

Hence $(f + g_1 + g_2)Cg_2$, so $(f + g_1)Cg_2$, and since $(f + g_1)Cc$, we see that $g_2Cc$; hence also $g_2$ is expressible. Further, since $(f + g_2)^2Cg_2^2$, $(f + g_2)$ is also expressible; and since $b\cancel{C}c$, $(f + g_2)\cancel{C}g_2$. Hence $f = (f + g_2) - g_2$ is indeed an apotome with annex $g_2$, divided into the same ratio as the binomial $b + c$. Moreover $(f + g_2) : b = g_2 : c$, where $g_2Cc$, so the terms of $(f + g_2) - g_2$ are commensurable with the terms of $b + c$. Finally, Proposition 14 completes the verification that $b + c$ and $(g + g_2) - g_2$ belong to the same class with respect to any expressible line $h$.    QED

The proof of Proposition 113 runs exactly parallel to this.[11] And I recommend readers to construct their own proof of Proposition 114 before working carefully through Euclid's proof, and then to ponder on any discrepancies between the two lines of thought. Euclid concludes Proposition 114 with:

PROPOSITION 114/115    An expressible area can be contained by *alogoi* lines.

But we have already seen, in Proposition 25, that this happens for two appropriate medial lines.

After Proposition 115 (on higher order medials; see the end of Section 5.2(c)), the manuscripts of Book X end with a rag-bag of unnumbered material, almost certainly interpolated: alternative proofs of Propositions 115, 106, and 107; two long-winded proofs that the side of a square is incommensurable with its diagonal;[12] and constructions of incommensurable plane and solid figures. All of this is relegated into an appendix by Heiberg (see Euclid–Stamatis, *EE*, 228–36) and most of it is not mentioned in Heath, *TBEE*.

Consider now the applications of Book X in Book XIII, in Propositions 6, 11, 16, 17, and 18. Here I shall proceed by taking skeleton *quotations*, slightly paraphrased, from Heath's translation of the proofs of these propositions, reading, as always, 'rational' and 'irrational' as 'expressible' and '*alogos*' respectively, and altering the labels of the lines to which these proofs refer, replacing them by the letters that have been used uniformly throughout this section and in Figures 5.3 & 5.4 and Table 5.2. Further comments, and page and line references to Heath, *TBEE* iii, are placed in square brackets.

PROPOSITION XIII 6    If an expressible straight line be cut in extreme and mean ratio, each of the segments is the *alogos* straight line called apotome.

[PROOF]    Let $h$ be an expressible straight line, let it be cut in extreme and mean ratio, and let $g$ be the greater segment; I say that each of $g$, $h - g$ is the *alogos* straight line called apotome. For . . . $\left(\frac{1}{2}h + g\right)^2 = 5\left(\frac{1}{2}h\right)^2$ [by XIII 1]. Therefore . . . $\left(\frac{1}{2}h + g\right)$ is expressible . . . and incommensurable in length with $\frac{1}{2}h$; therefore . . . $g\left[=\left(\frac{1}{2}h + g\right) - \frac{1}{2}h\right]$ is an apotome. Again . . . $h \cdot (h - g) = g^2$ [this is the basic property of extreme and mean ratio]. Therefore the square on the apotome $g$, if applied to the expressible straight line $h$, produces $(h - g)$ as breadth. But the square on an apotome, if applied to an expressible straight line, produces as breadth [i.e. width] a first apotome [with respect to that line]; therefore $(h - g)$ is a first apotome [with respect to $h$] . . .                   QED

Next, we consider the side of a pentagon inscribed in a circle.

PROPOSITION XIII 11    If in a circle which has its diameter expressible an equilateral pentagon be inscribed, the side of the pentagon is the *alogos* straight line called minor.

In the first half of the proof of this proposition, a configuration is set up in which, in Fig. 5.5, B is a vertex of the pentagon inscribed in a circle of diameter BH, F is the centre of the circle, $FK = \frac{1}{8}BH$, and M is the intersection with the diagonal of the pentagon perpendicular to BH. The geometry of the pentagon is then explored quantitatively; but it is not pointed out before the final lines of the proof that this exploration implies that the square on the side of the pentagon is equal to the rectangle BH . BM. Euclid therefore proves that $BM = BK - KM$ is a fourth apotome with respect to BH. So let $BH = h$, $BK = b$, $KM = c$, so $BM = b - c$, and now pick up the proof at 463.21.

Therefore $c^2 = 5(h/8)^2$, . . . therefore $c$ is expressible. And . . . $b = 5(h/8)$ therefore . . . $b$ is incommensurable in length with $c$. And each of them is expressible. Therefore $b$, $c$ are expressible straight lines commensurable in square only,    [464]    . . . therefore $b - c$ is an apotome and $c$ is the annex to it. I say next that $b - c$ is also a fourth apotome [with respect to $h$]. Let . . . $d^2 = b^2 - c^2$; therefore . . . [there follows a proof that BK is commensurable with BH; but this was already proved, in effect, at the beginning] $b^2$ has to $c^2$ the ratio which 5 has to 1. Therefore *convertendo*, $b^2$ has to $d^2$ the ratio which 5 has to 4, and this is not the ratio which a square number has to a square number; therefore $b$ is incommensurable with $d$ . . . therefore $b - c$ is a fourth apotome. But the rectangle contained by a rational straight line and a fourth apotome is *alogos*, and its square side[13] is alogos, and called minor.

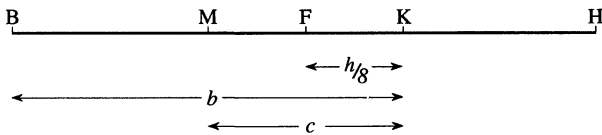The classification of the edge of the icosahedron appeals to this result for the pentagon:



FIG. 5.5

PROPOSITION XIII 16    To construct an icosahedron and comprehend it in a sphere, like the aforesaid figures; and to prove that the side of the icosahedron is the *alogos* straight line called minor.

Let the diameter . . . of the given sphere be set out . . . [five pages later] . . . therefore the icosahedron has been comprehended in the given sphere.    [486.1]    I say next that the side of the icosahedron is the *alogos* straight line called minor. For, since the diameter of the sphere is expressible, and the square on it is five times the square on the radius of the circle EFGHK, therefore the radius of the circle EFGHK is also expressible; hence its diameter is expressible. But, if an equilateral pentagon be inscribed in a circle which has its diameter expressible, the side of the pentagon is the *alogos* straight line called minor. And the side of the pentagon EFGHK is the side of the icosahedron. Therefore the side of the icosahedron is the *alogos* straight line called minor.

Similarly, the classification of the dodecahedron appeals to the result for the extreme and mean ratio; I leave the reader to work out the details in XIII 17.

Finally we have:

PROPOSITION XIII 18    To set out the sides of the five figures and to compare them to one another.

. . . [506.22]    Therefore the square on the side of the pyramid is four-thirds of the square on the side of the octahedron, and double of the square on the side of the cube; and the square on the side of the octahedron is one and a half times the square on the side of the cube. The said sides, therefore, of the three figures, I mean the pyramid, the octahedron and the cube, are to one another in expressible ratios. But the remaining two, I mean the sides of the icosahedron and the side of the dodecahedron, are not in expressible ratios either to one another or to the aforesaid sides; for they are *alogos*, the one being minor and the other apotome.

As was pointed out in this discussion of Propositions X 66 to 70, the result of this final sentence is not proved in Book X, but it is easy to extend the proofs of Propositions 103, 105, and 111/112 to cover it.

In addition to these explicit references to Book X in Book XIII, there are several other propositions which may well be implicitly related to the classification. For a typical example:

PROPOSITION XIII 12, supplemented    If an equilateral triangle be inscribed in a circle, the square on the side of the triangle is triple of the square on the radius of the circle [and so, if the radius is expressible, then also is the side of the triangle].

Similar inferences follow immediately from Propositions 1 to 4 and 12 to 15; and most of the remaining results of Book XIII establish intermediate steps in the proofs of all of these results.

This completes our reading of the text of Books X and XIII. The only other surviving use that is made of the Book X classification is found in Pappus, *Collection* IV; it is described in Heath, *TBEE* iii, 9–10. And both Proclus and Pappus report that Apollonius worked with the so-called 'unordered *alogoi*'; for a discussion with quotations, see Heath, *TBEE* iii, 10 and 255–9.

### 5.2(f) *Euclid's presentation of the classification*

As the quotations in Section 5.2(a) indicated, few people today can find a kind word to say about Euclid's exposition, no matter how much they may elsewhere praise the austere wonder of the mature synthetic style that we find in Archimedes and Apollonius. So let me try to offer two small crumbs of justification on his behalf.

The first feature I want to discuss is the way the first explanation of the additive *alogoi* lines is deferred until Propositions 54 to 59. The complications this causes are particularly acute in the case of the lines of rows 4 to 6; we get arcane examples constructed in Propositions 33 to 35 (and the relation of these lines to the classification of Book X, when it is eventually developed, will never be explored by Euclid); then these lines are used in the bizarre Propositions cum Definitions 39 to 41, explored further in Propositions 45 to 47, then left on one side. Some completely new distinctions are introduced in Definitions 47/48 and explored in Propositions 48 to 53. Only then, as I say, is the mechanism revealed. The exposition would be greatly simplified if the binomial and apotome lines were defined along with the medial somewhere around Proposition 20; the subclassification of binomials of Definitions 41/42 given then; and the process (r → s) of Propositions 54 to 59, and (s → r) of Propositions 60 to 65, given with its associated propositions on construction, *alogos*-ness, and uniqueness. Or if not this radical reorganisation, then at least some gesture in this direction.

The whole theory starts up with the choice of the assigned line $a$; but observe how this more amenable description I have just proposed requires the second choice with which I opened Section 5.2(c): the choice of the express-ible height $h$. *But also observe that this height $h$ does not enter Euclid's construction and description of the alogoi lines in Propositions 33 to 47.* The choice of $h$ determines the subclassifications of binomials and apotomes, and that determines the order in which the *alogoi* lines are generated in Proposi-tions 54 to 59; if we replace $h$ by $h_1$, incommensurable with $h$, this subclassification will be changed. The following example will illustrate this clearly.

The binomial $w = 2a + \sqrt{3}a$ is a 1st binomial with respect to $h = a$, and

$$q = \sqrt{h} \cdot w = \sqrt{\tfrac{3}{2}}a + \sqrt{\tfrac{1}{2}}a$$

is again a binomial; but it is a 3rd binomial with respect to $h_1 = \sqrt{2}a$, and

$$q_1 = \sqrt{h_1} \cdot w = \sqrt{(\sqrt{2}a \cdot (2a + \sqrt{3}a))} = \sqrt{(a \cdot (\sqrt{8}a + \sqrt{6}a))} = \sqrt{\sqrt{\tfrac{9}{2}}}a + \sqrt{\sqrt{\tfrac{1}{2}}}a$$

is now a second bimedial. So this way of generating the *alogoi* lines depends on an arbitrary choice of height $h$, which then drops out of the procedure. And Book X, in Euclid's procedure, is presented in such a way as to deal first with everything that does not depend on choosing $h$, first for the additive lines, then for the subtractive. Note also that in the application of Book X to the

classification of the edge of the icosahedron in XIII 16, the lines that play the roles of *a* and *h* are different and incommensurable.

It would be easier to argue vigorously for this justification of the presentation of Book X if its author had been a little more punctilious over some details! For example, casual readers, searching for some information about the *alogoi* lines, will come across the blocks of formal definitions, indeed headed 'Definitions', at 47/48 and 84/85; but they will have to search more closely and read more carefully to find the essential definitions buried in Propositions 21, 36 to 41, and 73 to 78. So one might be inclined to think that Book X contains 23 different kinds of *alogoi* lines: medial, six binomial, six apotome, two bimedial, etc. But of these, the subclassification of the binomials and apotomes is, or should be, internal to the mechanism of Book X, while the subclassification of the bimedials and the apotomes-of-a-medial is part of the classification proper. And only when one reaches the end of 111/112 does Euclid make it explicit that the classification embraces only thirteen kinds of lines, none of which are described in the three blocks of formal definitions.[14]

Yet more, Euclid is very casual in his references to this line *h* which I have consistently referred to here as 'the height', and which plays this determining role from the Second Definitions 47/48 onwards. Euclid introduces it as "Given an expressible line . . ." (as in Definitions 47/48); or gives no quantification, as in Propositions 54 to 59: "If an area is contained by an expressible line and a *k*th binomial . . .", and similarly, in Propositions 60 to 65; or he does not refer to it at all, as in Propositions 48 to 53: "To find a *k*th binomial". (All of these quotations are taken from the enunciations of these propositions, but this kind of treatment is also followed through consistently in the proofs.) Yet more, when Euclid applies the theory in XIII 6 and 11, outside the context of the development of the theory in Book X, he refers to the 1st and 4th apotomes respectively, but fails to say to which heights these refer. I have added these missing details to my discussions of all of these examples.

For my second crumb of justification of Euclid's treatment, consider the spectacular absence of specific examples and illustrations from Book X. However useful and illuminating these examples might be, the fact remains that it is not Euclid's practice to offer this kind of assistance anywhere in the *Elements*. But still, the emphasis on the qualitative aspect of a classification, which we now see as essentially quantitative, is striking: contrast Euclid's statement that the side of a pentagon inscribed in a circle of expressible diameter is minor with our nearest common equivalent, that its length is $\sqrt{(10 - 2\sqrt{5})}d/4$. Euclid may have a further justification, however, that in any specific example in rows 3 or 6, for instance

$$(\sqrt{2a} + \sqrt{3a})^2 = (5a + \sqrt{24a}) . a,$$

he must either state and prove the lemma that $\sqrt{2a} . \sqrt{3a} = \sqrt{6a} . a$, or state this as a general hypothesis, that $\sqrt{\lambda a} . \sqrt{\mu a} = \sqrt{\lambda \mu a} . a$, or produce a truly general proof; but such a proof may have been way beyond his means. (See note [6],

above.) By avoiding examples, he avoids both the problem of dealing with this issue, and the pretence (that almost all—if not all—modern commentators seem to adopt) that this problem does not exist. Given the constraints imposed on his expository style, I think his tactic is justified.

## 5.3 THE SCOPE AND MOTIVATION OF BOOK X

I shall set out my proposals in some theses and antitheses which culminate in my principal thesis.

(i) Euclid's treatment throughout the *Elements* is purely geometrical, with no undertones of arithmetised mathematics.

I have already illustrated this thesis in many places throughout this book, in particular at several places in the discussion of *Elements* X, and it should be unnecessary to dwell on it further here.

(ii) Euclid's treatment in *Elements* II and X is two-dimensional, based on squares and rectangles, and he nowhere shows any interest in giving any similar results for cubes or other three-dimensional figures.

This thesis is a truism, so I will bring out its implications more prominently. The basic results of a three-dimensional exploration would be straightforward: the *Topics* proposition in three dimensions, which is, in fact, proved as XI 25; the cubic version of X 9, which would follow from VIII 27 in the same way that the square version is apparently based on VIII 26; the three-dimensional analogue of the results of Book II, such as the geometrical identity

$$(x + y)^3 = x^3 + 3x^2 . y + 3x . y^2 + y^3$$

illustrated in Fig. 5.6; etc. We could then define a new class of 'cubic expressibles' by taking the edges of cubes commensurable with the assigned cube, then define new kinds of 'cubic' medials, binomials, and apotomes, and begin to explore their properties. Such manipulations suggest themselves immediately, and the early results are attractive exercises in the kinds of
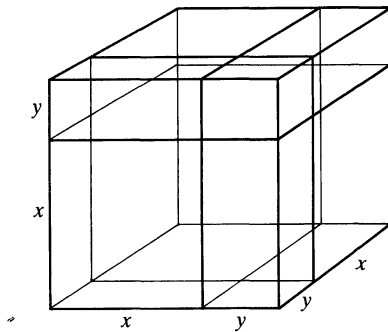


FIG. 5.6. $(x + y)^3 = x^3 + 3x^2 . y + 3x . y^2 + y^3$

reasoning found throughout the plane *and solid* geometry of the *Elements* and in our other fragmentary evidence of early Greek mathematics. Yet, with the sole exception of the three-dimensional *Topics* proposition at XI 25, none of these results is found in the *Elements*. For example, while the general idea of II 1 is repeated in the particular cases of II 2 and 3, and has indeed already been silently assumed in the proof of I 47, nowhere do we find its three-dimensional analogue articulated; nowhere do we find the result of Fig. 5.6 set out. The contrast is striking: the explorations of Books II and X are strictly restricted to two-dimensional phenomena.

(iii) The subject of Book X is *not* the manipulation or classification of quadratic surds, nor of the roots of quadratic or biquadratic equations with integer or fractional coefficients.

This statement is a reformulation of thesis (i), and has been discussed at lengths in Section 5.1, above. The notes in Heath, *TBEE* iii, provide very interesting and useful discussions of textual and terminological matters concerning Book X, but his mathematical analysis is, I believe, irrelevant. The same comment goes for the analyses in Pappus–Thomson & Junge, *CPBXEE*, 17–32 and Euclid–Stamatis, *EE* iii, pp. xi–xv. The perplexities that follow from such kinds of arithmetised and algebrised descriptions are raised several times in Mueller, *PMDSEE*, Chapter 7.

(iv) The implied underlying theme of the book is some qualitative description of the ratios of certain kinds of lines, and words *rhētos* and *alogos* are indeed referring specifically to the *ratios* of *lines*.

The best way of establishing this thesis is to appeal to Euclid's text, and look at the role played by Book X in Book XIII, where it is used to classify other lines arising in the extreme and mean ratio (XIII 6), the side and circumdiameter of the regular pentagon (XIII 11), and the edge circumdiameter of the icosahedron (XIII 16) and dodecahedron (XIII 17). This cycle of results on the five regular solids is then summarised in XIII 18 and, in the course of this proof, we read:

> The said sides, therefore, of the three figures, I mean the pyramid, the octahedron, and the cube, are to one another in expressible ratios (*logois rhētois*). But the remaining two, I mean the side of the icosahedron and the side of the dodecahedron, are not in expressible ratios either to one another or to the aforesaid sides; for they are *alogoi*, the one being minor and the other an apotome.

Also, Euclid makes no attempt to build up a systematic classification of *areas*; see, for example, his offhand treatment of the medial area in X 21–3. The areas that are named are introduced only with a view to developing the classification of lines.

(v) The idea of ratio that underlies Book X is not arithmetised but anthyphairetic.

Consider first arithmetised mathematics, mathematics in which the idea of number is extended beyond the *arithmoi* to include fractional quantities, and

then further to include more and more general kinds of numbers, and where this general kind of number is then used to define and manipulate ratios. The striking dichotomy here is between fractions, the 'rational numbers' of modern mathematics, which describe ratios of *arithmoi*, and the rest, the 'irrational numbers', which cannot be described in finite terms by ratios of *arithmoi*. This dichotomy pervades arithmetised mathematics and so now dominates our thinking about ratios; for example, within basic arithmetised mathematics, we can perceive little qualitative difference between irrational square roots and irrational cube roots. This dichotomy is not suggested anywhere by any of the terminology of Book X and does not correspond to any of the classifications found there.

Now consider anthyphairetic mathematics. Those ratios that can be now completely understood and described in finite terms by the *arithmoi* include the ratios of the sides of commensurable squares, that is the ratios of expressible lines $\sqrt{m} : \sqrt{n}$, and some sporadic examples like the extreme and mean ratio. If we then wish to extend outwards from these accessible, expressible lines, it would be natural to start with those lines that arise from the simplest operations of adding, subtracting, and squaring pairs of expressibles. This seems to correspond closely to what we find in Book X, which starts from the expressible lines and goes on to construct medials, apotomes, and binomials, which are truly baffling from an anthyphairetic viewpoint. These new lines are profoundly *alogoi*, without reason, and Book X is then one of the very few successful attempts to find order in a corner of this mathematical jungle. Sections 9.2(d) and 9.3(b) will describe briefly the modern and largely unsuccessful further attempts at this kind of exploration.

### 5.4 APPENDIX: THE WORDS *ALOGOS* AND (*AR*)*RHĒTOS* IN PLATO, ARISTOTLE, AND THE PRE-SOCRATIC PHILOSOPHERS

For an explanation of this list, see the Appendix to Chapter 4.

#### 5.4(a) *Plato*

ἄλογον (m.) *Timaeus* 42d.1.
ἀλόγους *Republic* VII **534d.5**.
ἄλογος (f.) *Definitions* 414c.7, 416a.23.
ἄλογον *Theaetetus* **201d.1**; *Timaeus* 47d.4.
ἀλόγου *Timaeus* 28a.3; *Laws* III 696e.1.
ἀλόγῳ *Republic* IX 591c.6; *Timaeus* 69d.4.
ἄλογον (nt.) *Gorgias* 465a.6, 496b.1, 519e.3; *Phaedo* 62b.2, c.6, 68d.12*; *Symposium* 202a.6; *Republic* X 609d.9, 11; *Parmenides* **131d.2**, **144b.3**; *Theaetetus* **199a.3**, **203d.6**, **205c.9**, **e.3**; *Timaeus* 51e.4; *Sophist* 219e.4, 238c.10, e.6, 239a.5, 259a.1; *Philebus* 55b.1; *Axiochus* 365e.5.
ἀλόγου *Philebus* 28d.6.

ἄλογα *Protagoras* 321c.1; *Theaetetus* **202b.6**, **203a.4**, **b.6**; *Sophist* 241a.5*, 249b.1.
ἀλογώτερον (nt.) *Charmides* 175c.7; *Gorgias* 519d.1.
ἀλογώτατον (nt.) *Apology* 18c.8.
ἀλογώτατα *Philebus* 55c.3.
ἀλόγως *Gorgias* 501a.6; *Republic* IV 439d.4; *Phaedrus* 238a.1; *Timaeus* 43b.2, e.3, 53a.8; *Laws* II 669d.4, IX 875b.8.

ἄρρητον (f.) *Laws* VII 788a.3.
ἀρρήτων *Republic* VIII **546c.5**.
ἄρρητον (nt.) *Sophist* 238c.10, e.6, 239a.5; *Laws* VI 754a.4; *Alcibiades* I 122d.2.
ἄρρητα *Hippias Major* **303c.1**; *Symposium* 189b.4; *Sophist* 241a.5*; *Laws* VII 793b.3, 822e.2.
ἀρρήτων *Hippias Major* **303b.7**.

ῥηταί *Theaetetus* 205d.9.
ῥητάς *Theaetetus* 202b.7.
ῥητῶν *Republic* VIII **546c.5**.
ῥητόν *Theaetetus* 205e.7; *Epistle* VII 341c.5.
ῥητά *Hippias Major* **303b.8**; *Republic* VIII **546c.1**, *Laws* VII 817d.3; *Epistle* VII 341d.5.
ῥητοῖς *Symposium* 213a.2; *Laws* VIII 850a.7.

5.4(b)  *Aristotle*

ἄλογος *Posterior Analytics* **76b9**; *Physics* 188a5, 212b30, b34, 252a24; *On the Heavens* 289a6, b34, 291b13; *Meteorology* 355a21, a35, 362a14, 366a9; *On the Soul* **432a26**; *Prophesying by Dreams* 463a23; *Generation of Animals* 722b30, 734a7; *Problems* 954b35; *On Indivisible Lines* **968b18**; *Metaphysics* 999b23, 1002a29, 1046b2, **1083a8**; *Nicomachean Ethics* 1095b15, 1102a18–1103a10, 1111b1, b13, 1117b24, 1120b18, 1172b10; *Magna Moralia* 1198a17; *Eudeman Ethics* 1218a29; *Politics* 1334b18, b21; *Oeconomica* 1343b13; *Rhetoric* 1369a4, 1370a18; *Poetics* 1460a13, 1461b14.

ἀρρήτως *Fragment* 40 1481b6.

ῥητός *On Indivisible Lines* **968b15**, **b18**; *Metaphysics* 1017b1, b3; *Nicomachean Ethics* 1162b26, b31, 1163a5.

5.4(c)  *Pre-Socratic philosphers*
ἄλογος

| Democritus | | | | |
|---|---|---|---|---|
| { | **A33** | ii | **91.20** } | (See Sections 5.2(a) and |
| { | **B11p** | ii | **141.24** } | 8.3(b)) |
| | A105 | ii | 109.35 | |
| | A116 | ii | 111.25 | |
| | B164 | ii | 177.1 | |
| | B292 | ii | 206.6 | |

| Empedocles | B136 | ii | 367.4 | |
| Heraclitus | A16 | i | 147.35, 148.19 and 35 | |
| Leucippus | A22 | ii | 77.1 | |
| Parmenides | A45 | i | 226.2 | |
| Philolaus | B11 | i | 412.11 | |
| Protagoras | C1 | ii | 269.23 | ($\equiv$ Plato, *Protagoras*, 321*c*) |

ἄρρητος
| Lysis | [4 | i | 421.5] | (See Sections 5.2(a) and 8.3(b)) |

## NOTES

1, p. 156: Stevin described Book X as "the cross of mathematicians"—"*La difficulté du dixiesme Livre d'Euclide est à plusieurs devenue en horreur, voire jusque à l'appeler la croix des mathématiciens, matière trop dure à digérer, et en la quelle n'aperçoivent aucune utilité*". (This particular quotation is in G. Loria, *Le scienze esatte nell' antica Grecia* (1914); it then was passed on by Heath, *TBEE* iii, 8–9, and thereafter by most modern commentators on Book X.) Knorr, CM, opens with a different but similar quotation, and a nice general discussion which I wish I could include in its entirety here, along with its three notes—better, I wish I had written it!—for I can conceive of no better introduction to my own treatment. (I would change only one word: for 'algebraic' in lines 23 and 26, read 'arithmetical'.) However the references to Stevin need some clarification and correction. Stevin's *L'Arithmétique* (1585) is an octavo volume consisting of *Le Premier* and *Le Second Livre d'Arithmétique*, followed by *Les Quatres Premiers Livres d'Algèbre de Diophante d'Alexandrie*, paginated together. This was issued bound with *La Pratique d'Arithmétique*, a second collection of tracts which include *La Pratique* itself, *La Reigle d'Interest avec ses Tables* (originally published in Dutch in 1582, and subsequently reissued in a corrected Dutch version in 1590), *La Disme* (the French translation of *De Thiende*, also of 1585), and the *Traicte des Incommensurable Grandeurs*, all paginated together in a second sequence. Loria's passage is from *Le Premier Livre d'Arithmétique*, pp. 36–7 and Knorr's from the *Traicte des Incommensurables Grandeurs*, p. 162 of the second sequence. Stevin's mathematical works were then reissued by Albert Girard in 1634 as the *Œuvres Mathématiques* in six parts, paginated separately but bound together, and the first part contains *L'Arithmétique* and *La Pratique d'Arithmétique* with an added *Appendice Algébrique*, now paginated in one sequence; and these two quotations occur on pages 10 and 213 respectively. Finally, excerpts are again reprinted in Stevin–Struik, *The Principle Works of Simon Stevin*, and these quotations can be found in vol. ii(b), 535–6 and 713 respectively. *La Pratique d'Arithmétique* also contains the *Appendice des Incommensurables Grandeurs* mentioned by Heath; this is pages 187–201 of the 1585 edition, second sequence, and pages 218–22 of the 1634 *Œuvres*, but it is not included in Stevin–Struik, *PW*, and neither quotation occurs therein. Knorr notes the modern German translation of *De Thiende*; it is worth adding that the Dutch original and Richard Norton's English translation of 1608 are in *The Principal Works* iia, on pp. 386–455, and this English translation has often been reprinted.

2, p. 157: This algebraic interpretation was apparently first formulated by Chasles, its description as "the quintessence of Book X" is due to Junge, and van der Waerden and Knorr cocnur with it; see Pappus–Thomson & Junge, *CPBXEE*, 23; Knorr, CM, 64 n. 32; and van der Waerden, *SA*, 168–72. My own "anyone at ease . . . will uncover" is, of course, typical crass mathematician's arrogance and insensitivity to the historical past, and should be ignored; it is not meant to be taken seriously.

3, p. 157: In my IRBT, 255–8, I give examples of how this formula both illuminates and distorts the interpretation of Book X and why, ultimately, it is conﬁcal to think of it as the quintessence of anything, for it purports to explain something reasonably simple, $\sqrt{(\beta \pm \gamma)}$, in terms of two much more complicated expressions!

4, p. 170: Or which purport to generality. But Knorr, CM 57–8, points out, apparently for the first time, that the second Lemma 28/29 is grossly defective when viewed, as apparently presented, as a general construction. He concludes: "Euclid's handling of the second lemma is thus yet another unsettling reminder that the proofs of Book X are not always as well conceived as they commonly are purported to be."

5, p. 172: I do not wish to imply that Descartes' geometry is arithmetised; his algebra is as impeccably geometrical as my approach in Section 5.2 above, with the mild exception that he does not seem to explain what he means by $\sqrt{2}$, $\sqrt[3]{2}$, etc. The key step underlying his algebra is to fix an assigned line $a$, and then define a new kind of multiplication of $x$ and $y$ to be the width of the rectangle $x \cdot y$ when applied to $a$; so this product is again a line. In fact, this same kind of operation is used repetitively throughout Book X in a different context as we shall soon see, but here without any overtones of any analogy with the multiplication of two numbers.

However, I think it can be argued that Descartes encouraged his readers to think in arithmetical terms and build on their arithmetical understanding of geometry; and only to resort to his geometrical definitions for a formal justification. A similar schizophrenia is common among mathematicians today. To return to the example of Sections 1.2(a) and (e), for most mathematicians $\sqrt{2} \times \sqrt{3} = \sqrt{6}$ might conjure up something closer to $1.4142\ldots \times 1.7321\ldots = 2.4495\ldots$ than a complicated statement about equivalence classes of Cauchy sequences of rational numbers, or Dedekind cuts, even though this arithmetical approach is defective in the essential detail that the multiplication is not effectively defined.

6, p. 178: Compare this table with the text itself of Euclid's Second Definitions 47/48 and Plato's *Theaetetus* 192a1–194b6. I cannot believe that the mathematician(s) who developed Book X did not have some schematic way of setting out these descriptions. A formally minded person will notice that the 4th, 5th, and 6th classes can be further subdivided according as to whether $cCd$ or not, and the 5th and 6th classes subdivide differently according as to whether $hCd$ or not. But all these further subdivisions of the $k$th binomial still generate the same *alogos* line in Euclid's classification. This may be another indication that Euclid is concerned with something more than a fully worked-out formal system.

7, p. 179: See note 4, above, and its text.

8, p. 181: Junge noticed the first problem with commensurability-in-square; see Pappus–Thomson & Junge, CPBXEE 175 n. 131, with thanks to Wilbur Knorr for this reference. The proofs strengthening Propositions 66 to 70 are due to Mueller; see his *PMDSEE* 283 and 199. This observation about the ratio of the two lines of the binomial is due to Knorr; see his CM 65 n. 38.

9, p. 182: Mathematicians today refer to these special cases using pejorative words like 'degenerate' or 'non-generic' and, of late, the interest has been in developing a precise language for exploring the typical or generic behaviour; for example: 'almost all numbers are irrational, indeed transcendental'; 'almost all continuous functions are nowhere differentiable'; 'almost all orbits in the three-body problem are bounded', etc. Euclidean geometry is, *par excellence*, the study of the degenerate case, the specific example, the very particular construction. The fact that some behaviour happens to manifest itself is sufficient to establish its importance. So the fact that the fourth additive and subtractive lines happen to arise in the construction of the pentagon—see below—is sufficient to justify the special names they are given of 'major' and 'minor'. Knorr (in *EEE*, CM, and ETB), Mueller (in *PMDSEE*), and

Taisbak (in *CQ*) speculate on the origin of these names. I agree that this is worth investigating, but I feel it may be more pertinent to speculate first on the reason for the basic terminology of the *rhētos* and the *alogos* that underlies the programme of Book X. At the risk of becoming tedious, I say again that, at the very least, one should not obliterate these features by the misleading mistranslations as 'rational' and 'irrational'.

10, p. 183: I cannot see how to construct a geometric proof around the identity $\epsilon^2/(\beta + \gamma) = (\epsilon^2/(\beta^2 - \gamma^2))\beta - (\epsilon^2/(\beta^2 - \gamma^2))\gamma$, which is an arithmetical version of Proposition 112, provided an appropriate litany is added about square roots of rational numbers which leads to the conclusion that the coefficients $\epsilon^2/(\beta^2 - \gamma^2)$ are rational.

11, p. 184: In fact the close correspondence between 112 and 113 causes Euclid to abandon completely one of his rules of style. When Euclid uses letters in setting out a proposition, to refer to the details of his figure, he normally introduces them in strict alphabetical order: $A$, $B$, $\Gamma$, $\Delta$, $E$, . . . . But the letters appear in Prop. 113 in the order determined by the previous proposition: $A$, $B$, $\Delta$, $K$, $\Theta$, $\Gamma$, $H$. Earlier propositions of Book X show minor infringements of this stylistic principle—see Propositions 48 to 53, and 108—but nothing as spectacular as this.

12, p. 184: These are two versions of the so-called Pythagorean proof that $m^2 = 2n^2$ leads to a contradiction. Knorr proposes, very plausibly, that this material was interpolated in Book X after the time of Alexander of Aphrodisias (early third century AD), as a result of the continuing activity of commentators who wished to deal with the remarks of Aristotle at *Prior Analytics* 41a23–30 and 50a35–38 about "e.g. the falsity of what follows from the assumption that the diagonal is commensurate, viz. that then odd numbers are equal to evens". See Knorr, *EEE*, 228–31, and Sections 8.3 and 10.1 below.

13, p. 185: Heath translates this as 'square root' (*TBEE*, 464, 5 lines up), which may indicate that, after all, he is thinking arithmetically.

14, p. 188: I maintain the distinction between the internal mechanism and the end-product of the classification by referring to the 1st, 2nd, . . . , 6th binomials and apotomes, but the first and second bimedials and apotomes of a medial, etc. And were the explanation not to be found at Proposition 111/112, I am certain that there would be scholarly arguments about how many different types of *alogoi* lines Euclid described: 23, 13, or 11. See, for example, Proposition 72/73 and its note in Heath, *TBEE* iii, 158.

# PART TWO

---

# EVIDENCE

> The papyrologist works with an ancient script which he trans-
> forms into a modern script, with the language denoted by the
> script, and with the meaning conveyed by the language. This
> linguistic trinity—script, language, and meaning—is fixed
> irremovably at the centre of his activity, and it is an indivisible
> trinity.
>
> H. C. Youtie, *The Textual Criticism of*
> *Documentary Papyri*, p. 16.

Chapter 6 speaks for itself, and needs no further introduction. Chapter 7
describes a wide range of different kinds of texts—scientific and commercial,
advanced and elementary, abstruse and pedagogic—whose common element,
examined in great detail here, is their treatment of numbers and, especially,
fractions. I wish to illustrate how the implications of a kind of arithmetical
manipulation might change as a text is recopied with different conventions of
writing or printing, or translated, or interpreted in a different mathematical
context. All the examples chosen relate to the question whether the Greek
mathematicians of the fifth and fourth centuries BC had at their disposal the
manipulations of common fractions such as $p/q + r/s = (ps + qr)/qs$,
$p/q \times r/s = pr/qs$. I believe that we have no good evidence on which to
argue that they did, although reservations about this point of view can be
found in Knorr, TFAEG and WEM. Of course, insofar as fractional quantities
were manipulated correctly, they did use procedures that were equivalent to our
operations with common fractions, but their different conceptions and opera-
tions may have given rise to different intuitions and abstractions about the
underlying mathematics.

Another purpose of Chapter 7 is to set in higher relief the non-arithmetised
nature of early Greek mathematics; were it to have been arithmetised, then we
shall see just what kind of arithmetic would have been used. (There is a brief
summary of the way this aspect of the overall argument fits together in Fowler,
LFEGM.) However, I nowhere consider the important and vast subject of
Ptolemaic astronomy, which is a blend of Babylonian arithmetical and Greek
geometrical methods, but only note, in passing, that we have no Greek evidence
of any use of or interest in Babylonian sexagesimal arithmetic before the second
century BC.

# THE NATURE OF OUR EVIDENCE

## 6.1 *ΑΓΕΩΜΕΤΡΗΤΟΣ ΜΗΔΕΙΣ ΕΙΣΙΤΩ*

Plato, it is said, placed an inscription over the door to the Academy: "Let no one unskilled in geometry enter". Since my aim is to propose a new and central role for Plato in the development of early Greek mathematics, this well-known and often-cited story provides an epitome of the main theme of this book.

Let us, however, pause a moment and enquire about the evidence for the story. Plato (*c.* 429–347 BC) founded his school sometime around 385 BC, and the 'early Academy',[1] the fourth-century period of existence under its first three directors or scholarchs, Plato, his nephew Speusippus, and Xenocrates, forms the shadowy background to the mathematical developments discussed here. The Academy is then commonly believed to have continued in existence for almost a millennium, directed by a 'golden chain' of scholarchs who kept alive and developed the Platonic tradition, until the closure of the Athenian schools by an order promulgated by the emperor Justinian in AD 529, the same year that St Benedict founded the monastery of Monte Cassino (thus, coincidentally, providing historians with a clear and symbolic break between pagan antiquity and medieval Christianity). Such is the almost universal impression to be gained from both scholarly and popular accounts alike, yet the evidence for every remark of substance in the previous sentence is found, under scrutiny, to be either lacking or distorted, and this whole line of interpretation is unfounded. This is not the place to attempt a summary of this scrutiny and details of alternative interpretations,[2] beyond the following

---

[1] See Cherniss, *REA*. This short and erudite book is essential reading for anybody interested in the scanty evidence concerning the early Academy and what has been made of it by modern scholars.

[2] See Cameron, LDAA, Glucker, *ALA*, and Lynch, *AS*. These works are each fascinating and authoritative syntheses of a wide range of different kinds of evidence. Cameron poses and discusses the following questions: "What was Justinian's motive? Did he give the last push to a tottering edifice, or destroy a thriving intellectual centre? Indeed, did he actually succeed in destroying anything at all? What did the philosophers do on their return?" (p. 7). Lynch's book assembles the literary, archaeological, epigraphical, and legal evidence concerning the Athenian schools in general, and so has much of importance to say about the Academy. Glucker's massive re-evaluation of the context of later Platonism must, if it is accepted, modify or put into question practically everything previously written on the subject. His work is set out in an equally massive book of some 450 densely argued pages, with a second instalment promised, directed at a readership so different from what I envisage here that I dare to suggest the following page references to his main conclusions for the interested but non-specialist reader: pp. 88–97 (Philo and Antiochus); 138–9, 145, and 153–8 (on the revival of Platonic studies in Asia Minor and Syria in the second century AD, but its absence from

general conclusions. It now appears that the Academy fell into disuse after the directorship of Philo, who went to Rome,[3] a refugee from the Mithridatic wars, sometimes around 85 BC, and soon after this the tradition of philosophical schools in Athens languished. Then there was a general revival of interest in Plato's philosophy and mathematics in the second century AD, in Syria and Asia Minor, at which time a few philosophers may have migrated back to Athens. However, Athens does not seem to have become once more celebrated for its philosophical activity before the fifth century,[4] when the study of Plato's work again flourished in a prosperous and active school directed successively by Plutarch of Athens, Syrianus, Proclus, Marinus, Hegias, and Damascius; this school appears to have survived Justinian's edict; and its next scholarch, Simplicius, appears to have continued to write, have students, draw his emoluments, and even practise his pagan religion—albeit in reduced circumstances and more circumspectly—into the second half of the sixth century. But

Athens until the revival of the neo-Platonic school of Athens in the fifth century AD); 226–55 (the school's property, including a summary of recent archaeological investigations); 296, 306–15, and 321–2 (the 'Golden Chain' of scholarchs and Proclus' mystical attitude to Plato's 'divine philosophy'); 322–9 (a summary and extension of Cameron, LDAA, on the 'end' of the Athenian schools); and 364–79 (the succession of the Peripatetic, Stoic, and Epicurean schools and the attitude of the Athenian citizens to the philosophical activity within their city). My only unease about these investigations concerns their neglect of the role of mathematics within philosophy throughout the period with which they are concerned. Compare, for instance, the conclusions of Cherniss, quoted at the beginning of Chapter 4, with the only three allusions to mathematics I could find in Glucker's book: two passing references to Theon of Smyrna on pp. 136 and 212, and the following passage on p. 260: "From our passage [Plutarch, *De E apud Delphos*, 387f.] we learn that he [Plutarch] had already studied some mathematics—a subject with which, at the dramatic date of the dialogue [AD 66/7] he was somewhat too passionately involved". This disdain, or distaste, for mathematics is general. See, for example, the review of Cherniss, *REA* by Solmsten (*Classical Weekly* 40 (1946), 164–8, on 168): "This [mathematics] was the kind of work to which Plato kept the members of the Academy with inexorable strictness, refusing to discuss with them, especially with the younger men, the *more serious philosophical problems* connected with his theory of ideas" (my emphasis); and it would be easy and profitless to point to scores of similar instances.

According to Cameron, Monte Cassino is now believed to have been founded in AD 530!

[3] Glucker's stronger conclusion, that Philo died in Rome, is queried in Sedley, EA; see especially n. 2.

[4] Here, for example, is one anecdotal piece of evidence for the dearth of intellectual activity in Athens, in a letter from the neo-Platonic philosopher and student of Hypatia, Synesius of Cyrene, to his brother, written around AD 396: ". . . And as for Athens: A curse on the accursed boatman who brought me here! Why, present-day Athens has nothing worth venerating except for the famous names of its places. It's just like the skin that is left from a sacrificed victim: a remainder of a life that once was. That's the condition of things now that philosophy has gone from here. All that remains for us to to travel around and wonder at the Academy, the Lyceum, and—by Zeus—the painted Stoa which gave its name to the philosophy of Chrysippos, though it's now no longer painted since the proconsul took away the panels on which Polygnotes of Thasos displayed his art. In our time Egypt has received and continues to cherish the creative works of Hypatia, while Athens—which was formerly the dwelling place of the wise—now has only bee-keepers to make her famous" (*Epistle* 136, quoted in Lynch, *AS*, 195f.).

though this Athenian neo-Platonic school consciously revived and emulated the tradition of Plato's Academy, it never seems to have been called 'The Academy' by its contemporaries. In fact this name, with its modern connotations, is an obstacle to the solution of what is characterised, in Cherniss, *The Riddle of the Early Academy*, as the more genuine problem of understanding Plato's Academy:[5]

What, then, did Plato really do in his Academy? . . . 'Academy' and 'Academic' are terms which men of formal training, who speak as their own the modern European tongues, have been pleased to apply to themselves and their organisations. Is it not wonderful, therefore, that by a more or less unconscious retrojection modern scholars have attached the particular significance which 'Academy' has in their own milieu to the garden of Plato's which was situated in the suburb northwest of Athens called 'Academia' after a mythical hero. . . . The external evidence for the nature of the Academy in Plato's time is extremely slight . . . [*REA*, 61 f.].

The same point is made in Glucker, *Antiochus and the Late Academy*:

For the ordinary educated Roman of [Cicero's] age, the place is just what it appears to be to a visitor to Athens—a gymnasium. To us, influenced as we are by the Platonic tradition, the word 'Academy' has come to mean an institution of learning, a learned society, or at least a place of theoretical ('academic') education. In ancient Athens, the Academy was first and foremost a public park dominated by its gymnasium, and the connection between it and Plato's school was only one of the numerous historical reminiscences in an area rich in history. (In order to obtain the right perspective, one can add that the excavations in the area of the Academy have not proved to be a tourist attraction in our times, although they are mentioned in most guidebooks to Greece.) It is in such a light that Pausanias, in what is essentially a tourist guide [*Description of Greece*, I xxix–xxx], describes the area.[6] He proceeds to describe in detail monuments like the enclosure to Artemis nearby and the small shrine of Dionysus Eleutherus . . . . He lists the graves of the various war heroes and of distinguished statesmen and philosophers . . . . He [describes] the altar to Eros built by Charmus . . . ; the altar of Prometheus in the Academy itself; the altar to the Muses—most probably the one dedicated by Plato and decorated by Speusippus (Diogenes Laertius, *Lives of the Eminent Philosophers* IV, 1); an altar to Athena and an olive tree. It is only then that he mentions that not far from the Academy is a memorial to Plato, in a place where a god indicated to Socrates in a dream that Plato was to be the greatest of philosophers. Pausanias narrates the dream in great detail, but says nothing to the effect that Plato was buried on what was his own estate near the Academy (if that was the case), or of the very existence of such an estate. He passes on immediately to a description of the tower of Timon the misanthrope, whose property, as we know from other sources, was near Plato's school. . . . The traveller in Attica was shown Plato's grave, but he was not told of the connection between the Academy and Plato himself or his school [*ALA*, 244 f.].

---

[5] The 'riddle' of Cherniss' title is to understand why some modern scholars need to propose an 'unwritten doctrine' to account for the discrepancy between Aristotle's account of Plato's theory of idea-numbers and what is found in Plato's own writings.

[6] Pausanias lived during the second century AD. The *OCD* observes: "He loves all religious and historical remains .ꞏ. . on which he writes plainly and honestly. His accuracy herein is confirmed by existing remains".

To return, then, to the inscription. It is most unlikely that the archaeological record could now be of any help in verifying or refuting the proposal that Plato himself had any part in putting it up[7] so we have, instead, to rely on references to it by ancient writers;[8] but these cannot be traced back any earlier than the fourth century AD. The earliest such reference occurs in an oration written in AD 362 by the emperor Julian the Apostate, who mentions an inscription over the entrance to Aristotle's classroom and alludes to another over Plato's, without giving any details of their contents. No other reference to Aristotle's inscription is known, and the first details of the wording of Plato's inscription come from an anonymous scholiast[9] who has been identified as probably the fourth-century orator Sopatros, in an annotation of a manuscript of Aelius Aristides. Sacred places sometimes had inscriptions such as "Let no unfair or unjust person enter"; and this scholium implies that the author of Plato's inscription

[7] The site of the Academy has been buried under the sprawl of twentieth-century Athens without ever being properly excavated, and the scattered reports of the partial excavations have never been brought together into any kind of comprehensive work; see Whycherly, PAPS, especially Part II, 2–10. A full discussion of the archaeological evidence, with references to some fifty archaeological reports, can be found in Glucker, ALA, 237–46; he concludes: "The one subject on which the excavations have not so far shed any new light is precisely the issue which Mr. Aristophron hoped to clear up with the aid of these excavations: the connection between the Academy and the school of philosophy established in that area by Plato. . . . As to the history of Plato's school, we are at present still largely at the mercy of our literary sources. And the literary sources hold their peace. Or rather, they behave as though any connection between the Academy and the philosophical school founded by Plato was restricted to the 'Classical' period of the school, between Plato and the pupils of Carneades and Clitomachus [i.e. up to the end of the second century BC]". Lynch, AS, 17–31, also deals with the archaeological evidence, but mainly with reference to the Lyceum, which was to the east of Athens.

[8] My summary of these sources is based on Saffrey, AME (= $A\Gamma E\Omega METPHTO\Sigma$ $MH\Delta EI\Sigma$ $EI\Sigma IT\Omega$, Une inscription légendaire). Also see Riginois, Platonica, The Anecdotes Concerning the Life and Writings of Plato, in which the sources of 148 anecdotes (including this) are examined; very few can be traced back to the first century AD or before.

[9] A scholium is a marginal note in a manuscript (for some examples, see Plates 5 & 6) and its author is the scholiast. The scholium reads:

εἰ δὲ ἡ γεωμετρία ἐπεγέγραπτο δὲ ἔμπροσθεν τῆς διατριβῆς τοῦ Πλάτωνος ὅτι $A\Gamma E\Omega METPHTO\Sigma$ $MH\Delta EI\Sigma$ $EI\Sigma IT\Omega$. ἀντὶ ἀνίσος καὶ ἄδικος. ἡ γὰρ γεωμετρία τὴν ἰσότητα καὶ τὴν δικαιοσύνην ζητεῖ.

(There had been inscribed at the front of the school of Plato. "Let no one who is not a geometer enter'. [That is] in place of 'unfair' or 'unjust': for geometry pursues fairness and justice.)

Andrew Barker, who has provided this and the following translation, notes: "The piece is riddled with ambiguities. 'No one who is not a geometer' might be paraphrased as 'no one who is not geometrically minded'. 'No one unskilled in geometry' doesn't quite work here, since the sequel suggests that ageometrētos indicates a disposition as much as a skill. Anisos 'unfair', literally 'unequal', is used with something of the sense of 'antiegalitarian'. In this usage, someone who is isos pursues equality as an end or ideal, while someone who is anisos rejects it—see Aristotle, Nicomachean Ethics 1129a32–b1. 'Unfair' or 'unjust': the scholiast may possibly mean 'unfair or unjust', or again 'unfair and unjust'. In such a context kai may mean either 'and' or 'or', and in the latter case there is no way of telling which arrangement of inverted commas best expresses the scholiast's intention."

has substituted *ageōmetrētos*, 'ungeometrical', for *anisos kai adikos*, 'unfair or unjust', in the normal formula; also the scholium seems to indicate that the inscription was not put up by Plato. The story is repeated and used by the sixth-century Alexandrian neo-Platonic philosophers Philoponus, Olympiodorus, Elias, and David to impute a variety of different motives to Plato in having the inscription put up in the first place. Finally, the most commonly used standard source for the story is the twelfth-century Byzantine Johannes Tzetzes (described in the *OCD* as a "copious, careless, quarrelsome, Byzantine poly-math") in his *Book of Histories* (or *Chiliades*), a review of Greek literature and learning in 12 674 verses.[10] Thus the evidence for the story is late and of doubtful provenance.

In support of the evidence, it can be added that the word *ageōmetrētos* is found in technical use during the fourth century BC (Aristotle, *Posterior Analytics* 77b12–13: "One should therefore not discuss geometry among those who are *ageōmetrētoi*, for in such a company an unsound argument will pass unnoticed."), so the inscription is philologically acceptable; Julian's account of the inscription seems to allude to a well-known story, so the anecdote could have been in circulation before his time (but how much earlier?); and there can be no doubt that Plato regarded mathematics as being extremely important. However, that the inscription corresponds so well with Plato's own opinions can be used as an argument against its authenticity: it could be a later elaboration of Plato's words, for example at *Gorgias* 508a: "Geometric equality is of great importance among gods and men".[11] The main

---

[10]  Tzetzes, *Chiliades VIII*, 974–7:

Πρὸ τῶν προθύρων τῶν αὐτοῦ γράψας ὑπῆρχε Πλάτων·
    Μηδεὶς ἀγεωμέτρητος εἰσίτω μου τὴν στέγην·
Τουτέστιν, ἄδικος μηδεὶς παρεισερχέσθω τῆδε·
    Ἰσότης γὰρ καὶ δίκαιόν ἐστι γεωμετρία.

(On the front of his doorway Plato had written 'Let no one who is not a geometer enter my house.' That is, 'Let no one who is unjust come in here', for geometry is equality and justice.) Andrew Barker notes: "Similar comments to those on Sopatros' scholium apply here. In both passages I would take *adikos* to mean 'unjust' rather than the more general 'unrighteous', since this facilitates the parallel with geometry, which presumably has to do with 'due proportions' and the like; but the more general senses can't be ruled out if 'righteousness' is being understood, for instance, within a broad conception of virtue as a mean. In any case, writers were not always conscious of the ambiguity, or sometimes appear to play on it deliberately—as (arguably) does Plato himself."

[11]A bizarre, though possible, explanation of the background of 'geometric equality', at *Gorgias* 508a, has been proposed in Harvey, TKE: that this was part of an intellectual argument developed by the aristocratic Archytas and Plato to discredit democracy. In place of the simple political formulation that all citizens are equal, which loosely corresponds to the application of an *arithmetic* proportion in which each term stands at an equal distance from its neighbour, one should rather reward each according to his value, and this is better described by a *geometric* proportion, in which each term bears the same relationship to its neighbour. Harvey even proposes that *logismos*, in Archytas' fragment B3, might mean geometric proportion, but this surely cannot be the case for Plato's *logismos*; see Section 4.2. Also see de Ste. Croix, *CSAGW*, 413–14.

argument against the authenticity of the story is the fact that no direct evidence can be found earlier than the fourth century AD, and 747 years separate the tentative date of the Academy's foundation and the date of Julian's oration. There can also be added a further kind of negative evidence *ex silentio*: Aristotle (384–322 BC), who spent from his eighteenth to thirty-seventh year in the Academy, leaving on Plato's death in 348/7 BC, was very critical of the mathematical bias of Academic philosophy. For instance, in a discussion of Plato's theory of ideas (*Metaphysics* 992a32–b1) he wrote: "Mathematics has come to be identical with philosophy for modern thinkers, though they say it should be studied for the sake of other things", and a little later, he adds that, in this approach, "The whole study of nature has been annihilated".[12] It is therefore curious that Aristotle nowhere points to the inscription as a flagrant and defiant advertisement by Plato of the very tendency of which he, Aristotle, is so unequivocally critical.

I have dwelt on this story of Plato's inscription at such length because it provides a convenient way of introducing what we find, time and time again, when we study the spectacular developments in mathematics that took place in the period which culminated in the compilation of Euclid's *Elements* and *Data* at the end of the fourth century BC: the evidence for many stories cannot be traced back any earlier than a period five, six, seven, or even eight hundred years after the event, and we often have little indication of whether the stories are authentic, plausible, or misleading. It is now worth pursuing this question even further and exploring the form in which the surviving evidence, early or late, reaches us.

## 6.2 EARLY WRITTEN EVIDENCE

Suppose we take a substantial text relating to the early developments of Greek mathematics—Euclid's *Elements*, for example. What, precisely, are we considering?

Euclid himself is an elusive figure.[13] We know nothing of him from contemporary references, and the only thing that can be hazarded with any degree of confidence is that he may have lived sometime around 300 BC, since the general theory of proportion to which Aristotle appears to allude, at *Posterior Analytics* I.5, 74a17–25, appears in polished and complete form in Book V of the *Elements*. He most probably lived before Archimedes, since Archimedes refers to earlier, possibly Euclidean, works,[14] though the only

---

[12] Also see Cherniss, *REA*, *passim*, and especially p. 68.

[13] See the *DSB* for articles on Euclid (by I. Bulmer-Thomas) and the transmission of the *Elements* (by J. Murdoch; this deals mainly with the Arabic and Latin traditions that I shall not consider here).

[14] In *Sphere and Cylinder* I 2 and 6; *Quadrature of the Parabola* 3, *Conoids and Spheres* 3; and *Method* I and II; these passages are cited in the testimonia in Euclid–Stamatis, *EE* i, pp. xii–xiii and some are excerpted in Thomas, *SIHGM* ii, 50 f. (for *Sphere and Cylinder* I 2)

explicit citation, in *Sphere and Cylinder* I, 3: ". . . by the second [proposition] of the first of the [books] of Euclid", is uncharacteristic of Archimedes' style, and so is very probably a later addition to the manuscript. A stronger argument for dating some kind of Euclidean text to around 300 BC will be described below, but the evidence for this will also not be without its complications.[15]

Nine or ten works are attributed to Euclid, of which the *Elements*, *Data*, *Division of Figures*, *Phaenomena*, *Optics*, *Catoptrica*, and *Sectio Canonis* survive in one form or another. Of these the most varied and important is the *Elements*, which comprises thirteen books containing, in all, more than 133 definitions and 465 propositions, occupying more than 20 000 lines of text.[16] It deals with a variety of topics in a recognisably varied range of mathematical and literary styles. Euclid is the compiler, not the author, of the work: he is believed to have taken source works by other mathematicians and edited them, adapting and rearranging the material, perhaps even inserting new material of his own, to make the complete treatises. The original source works cannot have been written very long after the change in Athenian intellectual and cultural activity from an oral to a written tradition; this appears to have taken place in the sixth, fifth, and fourth centuries, BC, and a range of explanations has been proposed to account for this development.[17] Let us now consider what actual writings

and 224 f. (for the *Method*). However, there are indications that Archimedes may have learned some mathematics from pre-Euclidean sources; see Knorr, APEPT.

[15] The suggestion that Euclid was the pen-name of a group of mathematicians or the name of a teacher, used collectively by his students, was put forward in Itard, *LAE*, 11. The fact remains that we have no facts on this subject. For a recent discussion of the difficulties posed by Proclus' attempt to locate Euclid (*CFBEE* 68, lines 10 ff.) see Fraser, *PA* ii, 563, n. 82; other ancient references are discussed in notes 84–91 and the text to these notes.

[16] A 'line of text' means a line of the critical edition, Euclid–Heiberg, *Opera*. The text and scholia of the *Elements*, vols. i–v, without Latin translation and with added prefatory material, has been reissued in Euclid–Stamatis, *EE*; but see notes 34, 40, 58, and Chapter 7 n. 1, below.

[17] See, for example, for one of the most popular of the explanations, Hall, *CCT*, 26–7: "Up to the end of the sixth century BC Greek literature is in this state of ceaseless flux, and is exposed to all the dangers of a tradition that is practically oral. And then the change comes swiftly and suddenly with the birth of a new form of literature, not local nor occasional nor professional as the older forms had been, but Pan-Hellenic in its appeal, although it sprang from a single city-state. This new form was Attic Tragedy, which never lost the hold which it rapidly obtained over the Greek race in all quarters of the ancient world. The enthusiasm for Tragedy created a reading public, since but few Greeks could hope to see the masterpieces of the great dramatists performed in Athens. Thus an impulse was given to the production of books which ends in the growth towards the end of the fifth century of an organised book trade with its centre in Athens." This from a book published in 1913; nobody today would speak with such confidence. Here, for example, is Turner, *ABFFCBC*, 16: "Let us . . . attempt to trace the steps by which books came into common use in Athens. The target before such an inquiry is a correct appreciation of the part played by the written word in the revolution which took place in the techniques of thought during the fifth century BC. It is a target against which Wilamowitz scored a brilliant and authoritative miss [in *Einleitung in die griechischen Tragödie* (1907), Chapter 3]. He laid it down that the texts of the Greek tragedies were the first books. This positive statement I hope to demonstrate as false; while his rigorous definition of a book as 'what is published by its author through the medium of an organised book trade for

survive, and in what form. Here I shall only lead up to consideration of the Greek Euclidean tradition; there is also, up to the tenth century AD, a debased Latin tradition; and also a rich but largely unexplored Arabic tradition which plays an important role in the transmission of Greek mathematical ideas.[18]

Inscribed, as opposed to written, texts will scarcely enter into our discussion. The most common materials on which writing is found are pieces of pottery or stone (called 'ostraca') or wooden writing boards or wax tablets, for ephemeral work,[19] and, for more enduring records, papyrus rolls: the pith of the papyrus plant was fashioned into oblong sheets which were then immediately joined by the maker to form rolls about thirty feet long. The height of the rolls varied greatly, and might be altered by cutting, but a typical range of heights for literary texts would be eight to twelve inches.[20] An exceptionally narrow roll, P. mich. iii 146, only four inches high, will be cited in the next chapter; it is a pocket-sized ready reckoner. The writing was set out, column by column, side by side on the roll, in capital, unaccented letters without gaps between words, but with the words occurring at the end of a line divided between lines according to strictly observed rules.[21]

the benefit of an expectant public' is to import from the nineteenth century a criterion which obscures understanding of the fifth." Also see the mythical story told by Plato of the invention of "*arithmos te kai logismos*, geometry and astronomy, not to speak of draughts and dice, and above all writing" at *Phaedrus* 274d–275b, quoted in Section 1.2(e), and the notes to that section. For a further discussion, which includes the oriental background against which Greek culture is set, see Pfeiffer, *HCS*, Chapter 2. The earliest evidence for the development of the Greek alphabet out of the Phoenician is found in the eighth and seventh centuries BC and comes from graffiti on pottery; for a summary and references, see Coldstream, *GG*, 295–302 and 382.

[18] For references concerning the Arabic and Latin traditions, see note 13 and Toomer, *LGMWAT*. For further details of the role of writing in ancient Greek culture, see Turner, *GP*; I shall rely heavily on this and other works by Turner in what follows. Wherever possible, I shall refer to papyrus texts by the standard method adopted by papyrologists: an abbreviation for the published collection or publication, followed by the publication number; the abbreviations are expanded in *GP*, Chapter IX or the Introduction to Liddell, Scott, & Jones, *GEL*. If a text has no such publication number, but appears in the second edition of Pack, *GLLTGRE*, its reference number there will be given and, for anything particularly important to our topic here, further details of its publication. In this way, I hope that every relevant cited text can be easily located, even by somebody who has never previously ventured into the papyrological literature. The Euclidean texts will also eventually be included in *Corpus dei papiri filosofici greci e latini*.

[19] See Diogenes Laertius-Hicks, *LEP*, III 37: "Some say that Phillip of Opus copied out the *Laws* [of Plato], which were left on waxen tablets." For illustrations of tablets, complete with the schoolboys' writing exercises, see Turner, *GMAW*, Plate 4, and *PW*, Plate 6; other examples will be cited in the next chapter. Tax-collectors' receipts and calculations were often made on ostraca; see Plate 8 for an example.

[20] I am here only discussing Greek texts (i.e. texts written in Greek). Papyrus as a writing material is much older: around 3000 BC, at Saqqara, a blank roll was placed in the tomb of the Vizier Hemaka for the use of the dead man. See Černy, *PBAE*; Cockle, *RCP*; Lucas, *AEMI*; and Lewis, *PCA*.

[21] See Turner, *GMAW*, 19–20. One rule, about division between doubled consonants, may

SURVIVINGEARLYGREEKPAPYRIARESETOUTJUSTLIKETHISAND
INSCRIPTIONSARESOMETIMESCARVEDBOUSTROPHEDONICALLY
BACKWARDSANDFORWARDSONALTERNATELINESINDIVERSE
DIALECTSINALIVINGCHANGINGVARIETYOFSTYLESASSOCIATED
WITHLEGALADMINISTRATIVELITERARYSOCIALANDTECHNICAL
SUBJECTSANDINANEVOLVINGANDVARYINGSTYLEOFHANDS
WITHTILDASFILLINGTHEENDOFTHELINE~~~~~~~~~~~~~~~~

Accents, marks to resolve ambiguous word divisions, and even indications of change of speaker in a play, were only slowly and irregularly introduced during the Alexandrine period, and it is not before the ninth-century Byzantine manuscript tradition that written Greek approached its now standard form, with upper and lower case letters, spaces separating words, and an elaborate array of accents and breathings; this development will be described in the next section. The most common type now used for printed Greek (it is used in this book) is based on the beautiful Greek handwriting of the Cambridge scholar Richard Porson (1759–1808) which itself was based on this Byzantine script; it was designed and first used by Cambridge University Press for his posthumously published editions of Euripides.[22] For an example of a font designed to correspond more closely to early Greek script see Plate 2, a reproduction from volume i of *The Oxyrhynchus Papyri*; this experiment was gradually abandoned after a few years, and the editors of the series now use Porson's Greek types, with a recently introduced and increasingly popular variation, which I have adopted, of a sigma in the form of a ϲ, without any distinction between medial and final forms.[23] (In fragmentary unspaced texts, such a distinction may be problematic or impossible to determine.)

Papyrus rolls are fragile and will tear, if handled too roughly; they are inconvenient, as any microfilm user will appreciate: Athenian vase-paintings show readers getting into difficulties with a twisted roll, and the aged Verginius Rufus broke a hip while trying to collect up one he had dropped;[24] but they continued to be the main vehicle for permanent records until between the second and sixth centuries AD, when a gradual process of 'codification', of transcription of manuscripts into codices, was carried out. (A codex is a forerunner of the modern book: flat sheets of material are gathered, folded,

have forced a scribe in an example which follows to cram in an extra $\mu$ above the line when splitting the word $\gamma\rho\alpha\mu\,|\,\mu\eta$ see Turner, Fowler, Koenen, & Youtie, EEI, on lines 27–8 of P. Mich. iii 143.

[22] See Sandys, *HCS* ii, 424–30. There is a facsimile of Porson's handwriting in Watson, *The Life of Richard Porson, M.A.*, opposite p. 260. (Porson was celebrated for his witticisms and epigrams; this is a Greek epigram boldly rendered into English: "The Germans in Greek / Are sadly to seek; / Not five in fivescore, / But ninety five more: / All save only Herman, / And Herman's a German." He was also celebrated for his drinking.)

[23] In fact, on the insubstantial evidence cited in the previous note, Porson mainly used the ϲ sigma; but, curiously, there is one σ, in Τεύτονοσ, where we now would use ς.

[24] See Turner, *GP*, 7.

and stitched. The word derives from the Latin name of a hinged set of wooden tables. Early codices were made from either papyrus or parchment.[25])

Do we then have a papyrus roll containing an early manuscript of the *Elements*? That would be far too much to expect. Papyrus will rot unless kept perfectly dry, though it can be preserved in a damp climate if it is baked at a high temperature until it is carbonised, and then kept undisturbed—like burned paper, it is then useless for everyday purposes, and the writing is frequently illegible, but it is in a stable and relatively indestructible state. This is certainly the main reason why only two pieces of papyrus have been found in Greece, both of them in tombs. The oldest dates from the second half of the fifth century, and so is perhaps the earliest surviving Greek papyrus; it comes from a recently uncovered tomb between Athens and Sounion which contained a male skeleton, a papyrus roll, a bronze pen with a split nib, the remains of a tortoise shell lyre, and two oil flasks; "The papyrus, a shapeless, flattened mass, will need a miracle to be preserved and unrolled, as it did to survive 2500 years".[26] The second piece, a line-by-line commentary on an Orphic cosmological poem, was found in 1963, in Derveni, north of Thessaloniki; it has been preserved through being baked during the burial rites.[27] These two texts, together with a cache found in Herculaneum, to be described below, may be the only written (again, not inscribed) Greek texts that have not passed through Egypt in the course of their transmission down to us.

No surviving Greek papyrus found in Egypt can be confidently dated by handwriting, context, or other internal evidence to before Alexander's conquest of Egypt, though a roll of *Persae* by Timotheus, Pack no. 1537, was found, in 1902, at Abusir, in a sarcophagus that has been dated by some, on archaeological evidence, to the middle of the fourth century, but this is disputed.[28] For a long time, the oldest document that could be securely dated on internal evidence was P. Eleph. 1, a marriage contract of 311 BC from Elephantine Island, Assuan.[29] This may now have been supplanted by a boldly written notice found

---

[25] The origins and development of the Greek codex are still far from understood, and literary texts on papyrus rolls are still occasionally found in the ninth century. See Roberts & Skeat, *BC*, who also discuss the transition from papyrus to parchment.

[26] Report of *The Times*, 25 May 1981, supplemented by Cockle, RCP, 147.

[27] There is a very interesting preliminary account of the discovery of the Derveni papyrus, followed by a most illuminating transcript of a discussion, in Kapsomenos, OPRT; for a photograph, brief discussion, and bibliography, see Turner, *GMAW*, no. 51. The text still has not had its definitive publication; see Turner, Tsantsanoglou, & Parássoglou, ODP.

[28] For a bibliography, see Pack no. 1537.

[29] The text and translation is in Hunt & Edgar, *SP* i, 2–5, and it is worth quoting in full: "In the seventh year of the reign of Alexander, son of Alexander, the fourteenth year of the satrapship of Ptolemy, in the month Dius. Marriage contract of Heraclides and Demetria. Heraclides takes as his lawful wife Demetria, Coan, both being freeborn, from her father Leptines, Coan, and her mother Philotis, bringing clothing and ornaments to the value of 1000 drachmae, and Heraclides shall supply to Demetria all that is proper for a freeborn wife, and we shall live together wherever it seems best to Leptines and Heraclides consulting in common. If Demetria is discovered doing any evil to the shame of her husband Heraclides,

in the Sacred Animal Necropolis at Saqqara which reads: "Of Peukestas. No one is to pass. The chamber is that of a priest. (*ΠΕΥΚΕΣΤΟΥ / ΜΗΠΑΡΑΠΟΡΕΥΕΣΘΑΙΜΗ / ΔΕΝΑΙΕΡΕΙΩΣΤΟΟΙΚΗΜΑ*)", or, more colloqually, "By order of the commander-in-chief Peukestas. Out of bounds to troops. Ritual areas." This notice can perhaps be dated, on the evidence of the name of Peukestas, to around 330 BC.[30]

Let us now turn to technical and Euclidean texts. Perhaps our earliest scientific text is P. Hib. i 27, a parapegma related to Eudoxus' calendar; this gives a very clear and important illustration of early numerical practice, and it will be described and reproduced in the next chapter (see Section 7.1(c) and Plate 7). Our earliest glimpse of Euclidean material will be the most remarkable for a thousand years: six fragmentary ostraca containing text and a figure, Pack no. 2323, found on Elephantine Island in 1906/7 and 1907/8 and published, with photograph and commentary, in Mau & Müller, MOBS. These texts are early, though still more than 100 years after the death of Plato (they are dated on palaeographic grounds to the third quarter of the third century BC); advanced (they deal with the results found in *Elements* XIII, 10 and 16, on the pentagon, hexagon, decagon, and icosahedron); and they do not follow the text of the *Elements* (though they are clearly connected with the material of these propositions). So they give evidence of someone in the third century BC, located more than 500 miles south of Alexandria, working through this difficult material; and the fact that the texts are on potsherds and are not the received text of the *Elements* suggests that this may be an attempt to understand the mathematics, and not a slavish copying or learning of the material. For example, the text jumps straight from the middle of Proposition 16 (Heath, *TBEE* iii, on p. 483) to the porism found at the end, and it contains nothing on the classification of the side as a minor line. These fragments are the evidence, mentioned earlier, which corroborates assigning Euclid to *c.* 300 BC,[31] but one

she shall be deprived of all that she brought, but Heraclides shall prove whatever he alleges against Demetria before three men whom they both accept. It shall not be lawful for Heraclides to bring home another wife in insult of Demetria nor to have children by another woman nor to do any evil against Demetria on any pretext. If Heraclides is discovered doing any of these things and Demetria proves it before three men whom they both accept, Heraclides shall give back to Demetria the dowry of 1000 drachmae which she brought and shall moreover forfeit 1000 drachmae of the silver coinage of Alexander. Demetria and those aiding Demetria to exact payment shall have the right of execution, as if derived from a legally decided action, upon the person of Heraclides and upon all the property of Heraclides both on land and on water. The contract shall be valid in every respect, wherever Heraclides may produce it against Demetria, or Demetria and those aiding Demetria to exact payment may produce it against Heraclides, as if the agreement had been made in that place. Heraclides and Demetria shall have the right to keep the contracts severally in their own custody and to produce them against each other. Witnesses: Cleon, Gelan; Anticrates, Temnian; Lysis, Temnian; Dionysius, Temnian; Aristomachus, Cyrenaean; Aristodicus, Coan."

[30] See Turner, CCOS.
[31] The first reference to these ostraca in connection with the problem of locating Euclid appears to be Fraser, *PA* ii, 558, n. 43.

complication in this proposal is that we cannot be sure of their relationship to the received text of the *Elements*: it is not inconceivable that the fragments represent a preliminary version of the final text, not the other way round, though they provide, in themselves, slender evidence for such a proposal.

One of the first and most spectacular discoveries of papyrus was at Herculaneum: a library of some 1800 rolls that had been buried in boiling volcanic ash and rain during the eruption of Vesuvius in AD 79. The town was rediscovered when a well was being dug in the gardens of a house on the lower slopes of Mount Vesuvius in 1752, and it has subsequently been excavated. The publication of these texts became a prestige project commissioned by the Bourbon Kings of the Two Sicilies and, between 1802 and 1806, the Prince of Wales (later King George IV) sent his chaplain John Hayter to supervise the unrolling of the papyri; then, in 1819 and 1820, he sent Sir Humphry Davy (later President of the Royal Society) and the Reverend Peter Elmsley (later Camden Professor of Ancient History at Oxford) to make chemical experiments on the papyri and to decipher them.[32] Initially great expectations were aroused that their contents might unlock the secrets of antiquity, but unfortunately very little of the material proved to be legible, and the books in the library are mostly restricted to Epicurean philosophy. One of the papyrus rolls (P. Herc. 1061) does provide us, however, with a second glimpse of the *Elements*, in an essay by Demetrius Lacon. The title page of the roll, apart from the author's name and a few letters at the end, has broken away and much of the text that follows is either missing or illegible. Generally the state of the texts is such that photographs show little of use (see Plate 1); but as the various papyrus rolls from Herculaneum have been unrolled or broken open, layer by layer, sometimes inevitably being destroyed in the process, they have been drawn, and some drawings and a transcription of column 9, from Heiberg, QPTM, are also given in Plate 1.[33] Column 8 lines 9–17 cites *Elements* I Definition 15, the definition of a circle; column 9 cites *Elements* I 9, the proposition on bisecting an angle, and its figure and a summary of the proof are given in column 10; column 9, lines 12–15 cites *Elements* I 3, a proposition that is used in the proof of I 9; and finally column 11 gives the enunciation and figure of I 10. The text itself appears to be a defence of an Epicurean doctrine that allowed the possibility of an infinite divisibility of a line, against criticism by Polyaenus, and is believed to date from around 100 BC.

Our next Euclidean fragment[34] comes from a dramatic and extensive find of papyri at Oxyrhynchus, in January 1897, by B. P. Grenfell and A. S. Hunt, in

---

[32] See Davy, SOEPFRH. A survey of recent restoration work on the papyri, especially by A. Fackelmann, is given in Cockle, RCP, 158. There is a recent bibliography of the Herculaneum papyri, Gigante, *CPE*; and Cavallo, LSSE, contains sixty-four remarkable plates of photographs, by T. Dorandi, of fragments of the collection.

[33] Heiberg's conclusions have not all been accepted by later scholars; see, for example, De Falco, *EDL*. For further work on this fragment, see Angeli & Dorandi, PMDL.

[34] The next three texts are unreliably reprinted in Euclid–Stamatis, *EE* i, Appendix II.

one of a series of expeditions for papyri sponsored by the Egypt Exploration Fund (later to become the Egypt Exploration Society). Oxyrhynchus was a large Graeco-Roman town in the desert to the west of and above the Nile valley, adjacent to the modern Arab village Behnes, about 120 miles south of Cairo. A series of rubbish mounds yielded an enormous quantity of papyrus: "The flow of papyri soon became a torrent it was difficult to keep pace with",[35] and indeed about half of the papyri found there still remains to be transcribed and published. Publication is still taking place in the series *The Oxyrhynchus Papyri*, 65 volumes to 1988, the centenary year; and volume i, edited by Grenfell and Hunt and published in 1898, included a fragment containing the enunciation of *Elements* II 5, together with a roughly drawn figure (see Plate 2). This fragment is dated by Grenfell and Hunt to the third or fourth century AD, but Sir Eric Turner has revised this dating in some notes that he very kindly provided me on the published Euclidean fragments from Egypt. Of this, he wrote[36] (with my additions in square brackets):

P. Oxy. i 29. Height 8.5 cm, width 15.2 cm.

A *kollesis* (a pasted join between two sheets of papyrus, the join being made by the manufacturer at the time the roll was put together) runs through the middle of the column of writing. Thick and poor quality papyrus. Writing in a single broad column. Lower margin about 3 cm, broken above. The *kollesis* and the fact that this is a II 5 show that the fragment is almost certainly from a longer (? complete) manuscript of Euclid [to which can also be added the evidence of the broken letters in the first line: these probably come from the end of either the enunciation, or the conclusion of II 4]. An empty space at the left (with no trace of the preceding column) of at least 3 cm shows that this was a generously laid out manuscript. But the figure is drawn without rule, free-hand, and carelessly [and is unlabelled].

The handwriting is a medium to large rounded capital, roughly bilinear, fairly quickly written, clear to read but with no claim to be regarded as calligraphic. I should date it confidently to the end of century i AD/early century ii, say AD 75–125. . . . [Here follows a discussion of palaeographical features, supporting this date; see note 37.]

Note that this text was one of the first finds of Grenfell and Hunt at Behnesa (Oxyrhynchus). They published in a great hurry, and no doubt included this piece to illustrate the wide range of subjects on which papyri could throw light. Their palaeographical (especially dating) framework was still to be worked out. By 1902 I do not think that the dating offered would have been 'third or fourth century'.

[35] See Turner, *GP*, Chapter 3 and James, ed., *EEEES*, for full and exciting accounts of this and other discoveries. This quotation is from Grenfell and Hunt's own account, cited in *GP*, 29.

[36] Personal communication. Turner's query about the dating was noted in Pack, *GLLTGRE*, no. 368.

[37] Turner's notes continue: "The verticals of $\iota$ $\tau$ $\eta$ $\gamma$ $\nu$ $\pi$ often terminate at the foot with left-pointing serifs, $\nu$ is widely spread. The hand is rounded and fairly upright, but the writing takes an upward slope to the right. Note the $\tau$ in $\mu\epsilon\tau o\xi\nu$ (and elsewhere), the frequent cursive $\epsilon$, and the enlarged initial $\epsilon$ in $\epsilon\alpha\nu$. Some parallels for the handwriting, in general, in dated documents are: P. Oxy. 378 (BM 809) col. ii (Domitian); P. Ryl. 597; P. Amh. 66 (AD 124); P. Lond. 298 (AD 124); and among undated documents: P. Oxy. 1793 (Callimachus); P. Ryl. 484 (assigned to this period); P. Lit. Lond. 132 (Hyperides). There are tremata on initial $\ddot{\iota}$, $\ddot{\nu}$. One final $\nu$ is extended to fill the line (this is not significant for the date)."

Their transcription is very good. Of line 1 only $\pi\epsilon\rho$ is certainly read. The traces are compatible with the restoration [of the enunciation or conclusion of II 4], but a guess of a different word could be verified, i.e. stated to be acceptable or not. Between lines 1 and 2, a paragraphus indicated that a new section begins: also see the cursive $\epsilon' = $ number 5 in the left-hand margin and the enlarged initial letter of $\epsilon\alpha\nu$.

I doubt whether the restoration of the figure at the end of lines 4 and 5 is correct [also see the discussion below], because on this restoration there is no reason why in line 2 the writing after $\gamma\rho\alpha\mu\mu\alpha$ should not continue to the right, instead of resuming at the beginning of line 3. Perhaps the figure was something quite different. All there is to go on is the short portion of a horizontal line visible more or less accidentally, since the papyrus is broken off above line 4, $\pi\epsilon\rho\iota\epsilon\chi o\mu\epsilon\nu o\nu$.

The two 'errors' in Greek are superficial: $\tau\epsilon\tau\rho\alpha\gamma\omega\nu o\nu$ is probably attracted into the genitive by the preceding $\eta\mu\iota c\epsilon\iota\alpha c$. In $\mu\epsilon\tau o\xi\upsilon$ the $o$ is not an error but the use of an alternative form of the word common in the later koine.

I have the impression that there *may* be ink (one or two letters?) at the top left of the figure level with line 9 and at the lower right. But my photograph has gone brown, and these may not really be ink.

The text of II 5 reads, in Heath's translation:

If a straight line be cut into equal and unequal segments, the rectangle contained by the unequal segments of the whole together with the square on the straight line between the points of section is equal to the square on the half.

There is a space to the right of lines 1 to 4, and traces of a horizontal line on line 4, invisible on the photograph in Plate 2: these led Grenfell and Hunt to propose that the space was taken up with a figure showing two lines: 'a straight line cut into equal segments' and 'a straight line cut into unequal segments'. As Turner points out above, and because such a figure is quite unprecedented in our early evidence of the text, this is implausible. The question of the figure is discussed by Heiberg who also disagrees with this restoration, but I cannot understand fully his alternative proposal.[38] An alternative reconstruction, which I have not seen discussed in the literature, is as follows. As was noted above, the broken words in line 1 could come either from the enunciation (*protasis*) or the conclusion (*sumperasma*) of II 4, since the conclusion is a

---

[38] Heiberg, QPTM, 148: "La figure, elle aussi, est identique à celle de nos manuscrits; il est étrange qu'elle ne paraisse pas porter trace de lettres et qu'elle soit placée immédiatement après la protase: dans nos manuscrits mathématiques, la figure se trouve régulièrement à la fin du théorème. Peut-être cette place insolite fournit-elle la raison pour que l'$\epsilon$ ajouté désigne la figure comme appartenant à ce théorème. Les éditeurs expliquent le peu de longueur des lignes 2–4 en admettant qu'il y a eu à leur droite une figure, soit deux lignes droites divisées pour exemplifier les mots $\epsilon\iota c$ $\iota c\alpha$ $\kappa\alpha\iota$ $\alpha\nu\iota c\alpha$; ils ajoutent qu'on ne voit les traces que de la ligne inférieure. Mais une figure de ce genre est tout à fait inouïe dans nos manuscrits mathématiques et on ne peut plus inutile, en sorte que cette explication est sujette à caution. On devrait plutôt penser que la fin nécessaire du théorème précédent (II 4) $\ddot o\pi\epsilon\rho$ $\ddot\epsilon\delta\epsilon\iota$ $\delta\epsilon\hat\iota\xi\alpha\iota$, qui, selon les éditeurs, a rempli, sous quelque forme abrégée, le reste de la ligne 1, a eu la distribution que voici: $\ddot o\pi\epsilon\rho$ $\ddot\epsilon\delta\epsilon\iota$ a eu sa place à la ligne 1, tandis que $\delta\epsilon\hat\iota\xi\alpha\iota$ s'est trouvé au-dessous, devant les ligne 2–3; alors la barre qui suit la ligne 4 a pu être mise comme signe de séparation." But what is meant by "$\delta\epsilon\hat\iota\xi\alpha\iota$ s'est trouvé au-dessous, *devant* les lignes 2–3"? There is an ample margin to the left which contains no trace of writing.

word-for-word repetition of the enunciation.[39] This fragment does not appear to come from a scribe's copy of the work; with its quickly written hand on poor quality papyrus, it is more likely to be working notes by somebody who understood the material. This opinion is corroborated by the roughly drawn figure and, in particular, by the way in which the figure is unlabelled and the diagonal line is undoubled: the diagonal enters only in the construction of the figure and does not play any part in the proof, and an unlabelled figure would provide poor help for understanding the text of Euclid's proof. The text could have been written by somebody who wrote down the enunciation, drew the figure, and then worked through the proof doubling the lines on the figure as he did so, but without writing down any of the details of the proof or labelling the figure; observe the generous space immediately below the last line and the second paragraphus at the end of the fragment. (The proof is a very long and elementary verification of the associated geometrical construction; see the discussion of Book II in Chapter 3.) This would explain why the figure came immediately after the enunciation, rather than at the end of the proof, why the figure is unlabelled and the diagonal is undoubled, and it would also explain line 1 and the gap to the right of lines 1 to 4: these could have contained, in similar style, the last line of the *enunciation* of II 4 and the associated figure for this proposition; and the fragment could have been part of a manuscript of notes by someone working through Book II of the *Elements*.

The next text, P. Fay. 9, is indeed fragmentary; its photograph and transcription are reproduced in Plate 3. It was edited and published in 1900 in Grenfell, Hunt, & Hogarth, *FTP*, along with other papyri found in the Fayûm, a region some fifty miles north of Oxyrhynchus.[40] It is assigned to the latter half of the second century AD and it contains part of *Elements* I 39 and 41. The difficulties of this fragment arise from the fact that no single complete line of text survives, and when the surviving letters are compared with the received text, the lines of the fragment seem to contain a very irregular number of letters: between nearly corresponding points on the first ten lines of the fragment, there are 32, 39, 81, 26, 41, 38, 36, 40, and 23 letters in the received text, to be described in the next section.[41] It is worth quoting part of the editors' discussion:

The general tendency of the fragment is towards compression; and some agreements with the manuscript called p [= Paris gr. 2466] are noticeable. The irregularities of the

---

[39] For this reason the conclusion is often omitted from editions and translations of the *Elements*; see how Heath generally finishes his translations with "Therefore etc." from I 4 onwards, as explained in his note to that proposition. Ancient authors, editors (perhaps even Euclid), or scribes did the same: the conclusion is often omitted in Books VIII and IX, or cut short, from Book X onwards, with the words καὶ τά ἑξῆc (and the rest); see Heath, *TBEE* i, 58.

[40] The text is reprinted in Euclid–Stamatis, *EE* i, 188 f., but there is a displacement of its vertical alignment and some of the spacing is incorrectly reproduced, and the final two lines of fragment (b) should read: ] ταυταιϲ α] | πα]ραλληλ[ο]ϲ [.

[41] These figures, which are based on my own counts, differ slightly from those given by the editors.

text followed by the papyrus extended to the order of propositions. Proposition 39 is immediately succeeded by Proposition 41. Proposition 40 was either omitted or else placed in some other position. It is noticeable that the diagram of Proposition 39 is drawn at the end of the demonstration, instead of, as is usual, at the beginning. [Incorrect: the diagrams in the manuscript usually come at the end of the demonstrations. See note 38 and Plate 5.]

An easy explanation of the eccentricities of this fragment would be obtained if it could be supposed that it did not form part of a regular book, but was merely an imperfectly remembered exercise. But this is not a satisfactory view. The words are correctly spelled, and the handwriting, though not of the regular literary type, is by no means ill formed. Its date is apparently the latter half of the second century AD. The papyrus was found with a number of documents belonging to the reigns of Antoninus, Marcus, and Commodus.

Sir Eric wrote of this fragment:

I am inclined not to dismiss this text as yet. At least it deserves further scrutiny, but it is going to be difficult to do it. The photographs from Mount Holyoke show that it has come to pieces (of course, perhaps it was always in pieces[42]), and it has been put into the frames just any how. The placing of the pieces doesn't agree with Grenfell and Hunt's publication, and they are obviously wrong in their relation to each other.

Grenfell and Hunt's date is probably correct enough. On page 97 they point out that the papyrus does not seem to correspond with the manuscripts, and leave it at that; I wonder if this is sufficient.

The reason why I think it would be interesting to know more is in the drawing of the figure, and its two surviving letters. It seems to be that these letters are written by the original scribe, but with a different, sharper pen (perhaps with the one with which he ruled the lines of the figure). In that case he was taking some trouble (using a rule), and perhaps these indications are enough to suggest that we have part of a real manuscript for the *Elements* here, and its difficulties and divergencies deserve further investigation.

The challenge of matching the fragment to the text had already been taken up by Heiberg, who also gave an extended discussion of the question of the missing Proposition 40. Unfortunately his proposals are too long and detailed to quote or summarise here, but his overall conclusion is also that this fragment, as he reconstructs it, diverges greatly from our text as received from other sources; but that this in no way justifies rejecting the text, in view of the character of the writing.[43] Therefore the fragment remains as a tantalising

---

[42] Miss Anne Edmonds, Librarian of Mount Holyoke College, where the papyrus is now kept, has kindly sent me copies of correspondence with organisers of the Egypt Exploration Fund that indicate that the fragment was already badly broken up when it was received by the college in 1903/4: "The last [papyrus] may be a puzzle or lesson to some mathematicians who can read Greek. I fear that it is very much more mutilated now than it was when the account of it was written" (Letter from W. W. Goodwin, 12 June 1904).

[43] Heiberg, PE. After a brief discussion of P. Herc. 1061 and P. Oxy. i 29 on pp. 47–8, his discussion of this text and his proposed reconstruction follows on pp. 48–53. Here is his general conclusion: "Bei dieser Reconstruction bin ich so wenig wie möglich von unseren Hss. abgegangen, und wenn auch das positive hier und da unsicher bleibt, steht das negative Ergebniss, wozu schon die Herausgeber gelangt sind, unumstösslich fest, dass der Papyrus sehr stark von unserem Text abweicht, der allgemein als vortrefflich gilt. Man könnte daher

unresolved puzzle: that the surviving papyrus fragment that most closely resembles a professional scribe's copy differs so much from the received text.

Next, a fragment containing the first ten definitions of Book I of the *Elements*, P. Mich. iii 143, published by F. E. Robbins in Winter, Robbins, & others, *PUMC* iii, 26–8. Sir Eric's notes began: "I have compared the photograph and text as published. At a number of points I would quarrel with the transcript (all would need to be checked with the original). . . ." The subsequent re-examination has led to a republication of the text in which the original editor's assessment that it "might well be the attempt of an older youth to reproduce from memory definitions imperfectly learned" has now been reversed to "One might guess that the piece was written by a schoolmaster, a grammaticus, making sure he has the text which he wanted to dictate to his students or on which he intended to base his next lesson"! For further details and a photograph, see the re-publication, Turner, Fowler, Koenen, & Youtie, EEI.

One more Euclidean fragment has been published recently: P. [inv.] Berol. 17469, in Brashear, VNTAB (see Plate 4). This is dated on palaeographic grounds to the second century AD, and it gives the figures and enunciations of *Elements* I 8 (fragmentary), 9, and 10 (fragmentary). (The editor's restoration of the figure to Proposition 8 omits the second half of Euclid's text figure; one is tempted to offer a different restoration to include this, but there is no evidence of it in the traces on the fragment, even where one would expect them.) This is the only Greek text I know of where the figures are placed before the enunciations, something that seems to us to be a logical place for them; we find them alongside the text (see Plates 1 and 2), in the middle (Plate 3) and, universally in our minuscule Byzantine manuscripts, at the end (see Plate 6 and note 38, above). There are several possibilities for its origin: it could be a teacher's or student's notes on the text, an *aide-mémoire*, though it is very well produced for such a use: the figures are drawn with a ruler, and the text is carefully written. It could be a thematic index to some or all of the *Elements* (such as we find in Mueller, *PMDSEE*, though there restricted to the enunciations); this is often done for volumes of music, which give the first two or three bars of each piece, and modern catalogues of manuscripts, which give the incipits of each item. We cannot know, though there is one thing that it is very unlikely to be: in modern texts, the figures or plates are often separated from the text and published as a separate volume (this is done by Neugebauer and his associates; see, for example, his *HAMA*); but the figures here are unlettered, which would render them useless for this purpose. However, whatever its

versucht sein, wie es auch geschehen ist, das ganze wegzuwerfen als eine verwilderte Ueberlieferung. Die Herausgeber haben an ein 'imperfectly remembered exercise' gedacht, verwerfen aber mit vollem Recht diesen Gedanken angesichts der correcten Buchstabierung und des ganzen Schriftcharakters. Dass wir ein wirkliches Buch vor uns haben, bestätigt die Normalzeile, und es kommen noch andere Umstände hinzu, wodurch die einfache Verwerfung dieser unserer ältesten handschriftlichen Quelle ganz unmöglich wird" (p. 50).

purpose was, it does lend support to my proposed reconstruction above of P. Oxy. i 29 as part of a similar text of enunciations and unlettered figures, though that text was much more casually written and its figure more crudely drawn.

Such is the extent of the published Euclidean papyrus fragments: material corresponding to about sixty complete lines of Heiberg's text, plus fragmentary information about a further sixty lines.[44] Moreover, neither the earliest evidence in the Elephantine ostraca, nor that in the fragment that most resembles a professional scribe's copy, P. Fay. 9, follows the received text. In addition to these fragments on papyrus, there is a parchment palimpsest from the seventh or eighth century, British Library Add. gr. 17211, which consists of five partially legible parchment leaves containing parts of Book X and Book XIII Proposition 14 (but there numbered 19), which was collated by Heiberg in his edition of the Elements.[45]

There is a persistent and widespread belief that the lack of surviving texts from this early period is due to the burning of the library of Alexandria by the Arabs after their conquest in AD 642. The evidence connected with this story is very well and authoritatively presented in Butler, *The Arab Conquest of Egypt* (1902), 401–26, and the supplemented second edition by P. M. Fraser (1978) adds a summary of subsequent work, which does not significantly alter Butler's conclusions.[46] I can do no better than to quote two passages, and refer the reader to this book for a full discussion.

The story as it stands in Abû 'l Faraj is well known and runs as follows. There was at this time a man who won high renown among the Muslims, named John the Grammarian. He was an Alexandrian, and apparently had been a Coptic priest, but was deprived of his office owing to some heresy by a council of bishops held at Babylon. He lived to see the capture of Alexandria by the Arabs, and made the acquaintance of 'Amr, whose clear and active mind was no less astonished than delighted with John's intellectual acuteness and great learning. Emboldened by 'Amr's favour, John one day remarked, 'You have examined the whole city, and have set your seal on every kind of valuable: I make no claim for aught that is useful to you, but things useless to you may be of service to us.' 'What are you thinking of?' said 'Amr. 'The books of wisdom,' said John, 'which are in the imperial treasuries.' 'That', replied 'Amr, 'is a matter on which I can give no order without the authority of the Caliph.' A letter accordingly was written, putting the question of Omar, who answered: 'Touching the books you mention, if what is written in them agrees with the Book of God, they are not required: if it disagrees, they are not desired. Destroy them therefore.' On receipt of this judgement, 'Amr accordingly ordered the books to be distributed among the baths of Alexandria and used as fuel

---

[44] At least one more Euclidean text can be expected from the Oxyrhynchus collection.

[45] A palimpsest is a parchment (as here) or other medium that has been washed or rubbed clean to make way for another text; sometimes parts of the lower text can still be read. Here the *Elements* was the earlier text, replaced by a text of Homer; the legible parts of the original are published in Heiberg, PEE. There is a brief general account of palimpsest discoveries in Reynolds & Wilson, *SS*, 174 ff.

[46] Additional references, for which I am indebted to P. M. Fraser, are Furlani, GFIBA and SIBA, and Meyerhoff, JGPAAM.

for heating: it took six months to consume them. 'Listen and wonder', adds the writer [*ACE*, 401–2].

After twenty pages of discussion, Butler concludes:

It may not be amiss to briefly recapitulate the argument. The problem being to discover the truth or falsehood of the story which charges the Arabs with burning the Alexandrian Library, I have shown—
(i) that the story makes its first appearance more than five hundred years after the event to which it relates;
(ii) that on analysis the details of the story resolve into absurdities;
(iii) that the principal actor in the story, viz. John Philoponus, was dead long before the Saracens invaded Egypt;
(iv) that of the two great public Libraries to which the story could refer, (a) the Museum Library perished in the conflagration caused by Julius Caesar, or, if not, then at a date not less than four hundred years anterior to the Arab conquest; while (b) the Serapeum Library either was removed prior to the year 391, or was then dispersed or destroyed, so that in any case it disappeared two and a half centuries before the conquest;
(v) that fifth, sixth, and early seventh century literature contains no mention of the existence of any such Library;
(vi) that if, nevertheless, it had existed when Cyrus set his hand to the treaty surrendering Alexandria, yet the books would almost certainly have been removed— under the clause permitting the removal of valuables—during the eleven months' armistice which intervened between the signature of the convention and the actual entry of the Arabs into the city; and
(vii) that if the Library had been removed, or if it had been destroyed, the almost contemporary historian and man of letters, John of Nikiou, could not have passed over its disappearance in total silence.
  The conclusion of the whole matter can be no longer doubtful. . . . One must pronounce that Abû 'l Faraj's story is a mere fable, totally destitute of historical foundation.
  My only concern in this matter has been to establish the truth, not to defend the Arabs. No defence is necessary: were it needful, it would not be difficult to find something in the nature of an apology. For the Arabs in later times certainly set great store by all the classical and other books which fell into their hands and had them carefully preserved and in many cases translated. Indeed they set an example which modern conquerors might well have followed. . . . [*ACE* 424–6].

For the answer to the question posed at the beginning of this section, let us now turn to our most important source of information about Greek literature and science: Byzantine and medieval manuscripts from the ninth century AD onwards.

## 6.3 THE INTRODUCTION OF MINUSCULE SCRIPT

Two developments in writing appear to have had an important effect on the survival and transmission of Greek manuscripts. The first, the change from the papyrus roll to the codex form, was described briefly in the last section;[47] the

---

[47] See, especially, n. 25.

second, to be described here, was the development of a new script.[48] Although the further developments in the substrate of writing—the change from papyrus to parchment and the introduction of paper by the Arabs, who learnt the process from Chinese prisoners of war taken at Samarkand, perhaps in AD 751—were ultimately to prove very important in the spread of written and, later, printed books, in the short run they were less influential.

Up to the beginning of the ninth century AD, almost all Greek literary works continued to be written in variants of the unspaced capital script illustrated in the preceding section. This seemed to be slow to write, to be difficult to read, and to require a generous area of papyrus or parchment. Early administrative and financial documents used a wide range of abbreviated, cursive scripts that can be very difficult to decipher; for an example, see Plate 8. Various uses of cursive scripts for literary texts had not previously led to the emergence of a distinctive, stable, and widely employed script; but then, sometime around AD 800, in circumstances that are still far from clear, a new kind of compact, cursive script took over and thereafter came to dominate the production of literary texts. This script soon came to employ majuscule and minuscule letters, ligatured groups of letters, spaced words, and a full range of accents and breathings; and very much more text could be fitted onto the page. It is now generally referred to as 'minuscule' script, and almost all of the Greek texts that survive have been through a process of minusculisation; indeed this was, in most cases, necessary for their preservation, as very little survives in the old script.

The first precisely dated manuscript in minuscule, Leningrad gr. 219, copied in AD 835, is a Gospel named after the Archimandrite Porphoryrij Uspenskij; another text which contains a few leaves of minuscule, and may date from 813–20, is Leiden B.P.G. 78, a copy of Ptolemy's *Handy Tables* with a commentary by Theon. Since the minuscule script in both of these examples does not look primitive or experimental, it may have been developed as much as fifty years earlier. The other scientific manuscripts that survive from this early period illustrate further how the old capital script still continued in use (Ptolemy's *Almagest* survives complete in the older style in Paris gr. 2389 and in minuscule in Vatican gr. 1594, and the old capital form is often retained for the headings of books and chapters that are otherwise written in minuscule), and how the earliest minuscle versions might not necessarily be the best (Euclid's *Elements* is found first in a beautiful minuscule manuscript dated to AD 888, Bodleian D'Orville 301 (see Plate 5), then a better text is found in the slightly later minuscule Vatican gr. 190 (see Plate 6); these texts will be referred to by the letters B and P respectively[49]). Another ninth-century manuscript, Vatican gr. 204, contains, in the following order, an important collection of scientific

---

[48] My account will be very heavily based on Lemerle, *PHB*, Reynolds & Wilson, *SS*, and Wilson, *SB*.

[49] It is editoral practice to associate letters, called sigla, with the manuscripts used in a critical edition of a text. The *conspectus siglorum* for Euclid–Heiberg, *Opera*, is P = Vatican gr. 190 (first used by Peyrard for his Greek, Latin and French edition of 1814–18);

texts:[50] Theodosius, *Sphaerica*; Autolycus, *On the Moving Sphere*; Theon's recension of Euclid's *Optics* with *Prolegomena*; Euclid, *Phaenomena*;[51] Theodosius, *On Habitations, On Nights and Days*, and *On Days and Nights* II; Aristarchus, *On the Sizes and Distances of the Sun and Moon*; Autolycus, *On Risings and Settings* I and II; Hypsicles, *On Rising-times*;[52] Euclid, *Catoptrica*; Eutocius, *Commentary on Books I–III of Apollonius' Conics*; Euclid, *Data*; Marinus, *Commentary on Euclid's Data*; and *Scholia to Euclid's Elements*, incomplete at the end.

Two people in the ninth and early tenth century were important in the preservation and transmission of Greek scientific texts: Leo the Philosopher (*c.* 790 – after 869), Archbishop of Thessaloniki 840–3 and later head of the school established under the emperor Bardas; and Arethas (*c.* 850–?), Archbishop of Caesarea in Cappadocia.[53] Leo's interest in mathematics and astronomy led him to acquire a copy of some works of Archimedes, Apollonius' *On Conic Sections*, some works on mechanics by the unknown authors Marcellus and Quirinus, a geometrical treatise by Proclus, and a work by Theon; we know this[54] from some obscure epigrams he wrote, preserved in the *Greek Anthology* IX 200, 202, and 578, which do not permit closer identification. He certainly must have possessed a copy of Ptolemy's *Almagest* though not, as is commonly supposed,[55] the manuscript Vatican gr. 1594, and he wrote a note on *Elements* VI 5 that is found, in the later scribe's hand, in the Euclid manuscript B;[56] but none of the books owned by Leo are known to survive. Eight volumes of the library of Arethas have survived and been identified through his marginal notes. The earliest is our earliest surviving complete manuscript of the *Elements* (see Plate 5); it was written by the scribe Stephanus, possibly one of the most accomplished by Byzantine calligraphers,

B = Bodleian D'Orville gr. 301; F = Florence Laurenziana gr. 28.3; V = Vienna phil. gr. 103; b = Bologna gr. 18 and 19; p = Paris Bibliothèque Nationale gr. 2466; q = Paris gr. 2344; and L = London Add. gr. 17211. For more details of these manuscripts, see vol. v. of *Opera*, reprinted in Euclid–Stamatis, *EE*, $v_1$ (in Latin) or Heath, *TBEE* i, Chapter 5 (in English). Further Euclidean manuscripts are described in Heiberg, PE.

[50] This is taken from the description, due to Menge, given in Aristarchus–Heath, *AS*, 325 f.

[51] There is a photograph of an opening of this text in Neugebauer, *HAMA* iii, Plate VIII, and a very interesting discussion of text figures in general, and those in this manuscript in particular, in ii, 751–5.

[52] This is the earliest surviving *Greek* text to divide the circle into 360 degrees, or to show any other influence of Babylonian sexagesimal arithmetic; see Chapter 7, below.

[53] For full biographies, see Wilson, *SB*, Chapters 4 (Leo) and 6 (Arethas), and the *DSB* (Leo).

[54] Except for information about the Archimedes manuscript. Leo may have owned the lost archetype of most of our surviving manuscripts of Archimedes' works; see Chapter 7 n. 45 and text.

[55] See Wilson, TBS, where it is pointed out that the identification is based on a note on folio 263[v], but this is written in a much later hand, of the thirteenth century or afterwards, and is of no particular significance.

[56] See Euclid–Heiberg, *Opera* v, 714 (= Euclid–Stamatis, *EE* $v_2$, 341): *Hupomnēma scholion eis tas tōn logōn sunthesis te kai aphairesin Leontos.*

in AD 888, and it cost fourteen gold pieces, at a time when salaries of civil servants started at seventy-two gold pieces a year, though they could rise in exceptional circumstances to 3500. Arethas also commissioned an important copy of Plato which contains all the main dialogues except the *Republic, Laws*, and *Timaeus*: Bodleian E. D. Clarke 39, written by John the Calligrapher in 895 and costing twenty-one gold pieces. These are the only two manuscripts of his that bear directly on our interests here, though it has been suggested that a copy of Pappus' *Collection* (Vatican gr. 218) might have belonged to his library, on the flimsy grounds that its minuscule script hints at an origin in Constantinople and the subject matter coincides with one of Arethas' interests.[57]

It is evident that most of our texts of the *Elements* derive from an edition made by Theon of Alexandria in the middle of the fourth century AD. The scribe of the manuscript P appears also to have had a pre-Theonine manuscript to which he also referred, and this enables us to get a glimpse of an earlier recension of the text; it forms the most important text for Heiberg's critical edition, in his *Euclidis Opera Omnia* (1883–1916). The other texts he used and the textual procedures he followed are described, in Latin, in his *Prolegomena Critica* (Euclid–Stamatis, *EE* v₁, pp. xvi–lxxxix), and the evidence from some of the papyrus fragments and further Byzantine manuscripts is discussed in his later article 'Paralipomena zu Euklid'.[58] All of this material is described in Heath's admirable translation, *The Thirteen Books of Euclid's Elements* (= *TBEE*), Chapters 5 and 6.

Plate 6 gives a photograph of a folio of the Vatican Manuscript P; the text starts in the middle of Definition 19 of *Elements* I and continues through to the middle of Common Notion 3. Plate 5 is an opening of the Bodlein manuscript B, starting in the middle of the enunciation of *Elements* II, Proposition 5 (compare this with the same text in Plate 2) and continuing into Proposition 6. Observe the wealth of scholia, some copied by the scribe from his exemplar, some added later, the interlinear annotations, and the many figures. Another facsimile from B, of folio 45 verso, *Elements* III 4 and the beginning of 5, can be found as a frontispiece to Heath, *TBEE* i, with brief description on p. xi.[59]

To summarise and answer the question which opened Section 6.2: our sources before AD 888 tell us about approximately 1 per cent of the eventual

---

[57] See Wilson, *SB*, 129–30: "We must not imagine Arethas as being rich enough to commission or acquire all the important or calligraphic manuscripts produced in his lifetime, nor should we think so ill of the culture of the capital as to suppose him the only man of his generation willing to explore the abstruse areas of mathematics". There is a detailed description of Vatican gr. 218, with new critical edition and translation of Pappus, *Collection* VII, in Pappus–Jones, *BSC*.

[58] Beware; the bibliography of Euclid–Stamatis, *EE* i, pp. xxxvii–xxxix does not contain Heiberg's very important article PE, and still less does the reprinted text incorporate any of the material therein; the existence of this article is only noted on *EE* v₁, p. xv. (Concerning this page, also see Chapter 7, n. 1, below.)

[59] For a photographic reproduction of a complete mathematical manuscript, see Heron–Bruins, *CC* i; this manuscript is our only surviving copy of Heron's *Metrica*.

text of the *Elements*, and possibly a similar amount of other scientific and literary texts; almost all of the surviving Greek material comes in the form of Byzantine or medieval manuscripts. On the other hand, the wealth of documentary papyri, dealing with all manner of financial, administrative, legal, and personal matters, have vastly increased our understanding of everyday aspects of life in Graeco-Roman Egypt.

In the next chapter I shall examine in detail part of one small family of Byzantine manuscripts of a short mathematical work, Archimedes' *Measurement of a Circle*, and describe some other kinds of texts, both scientific and commercial, found on papyrus.

# NUMBERS AND FRACTIONS

## 7.1 INTRODUCTION

### 7.1(a) *Numerals*

Almost the only notation for numbers (i.e. numerals) found in the earliest and best manuscripts of Euclid's works (apart from the *Sectio Canonis*) is used to label the propositions;[1] with the exception of three numerals in *Elements* XIII 11,[2] those few numbers that are used in the text are written out as words. This is an extreme case; most mathematical texts, and all manner of financial, administrative, and legal documents, employ a notation that will underlie the main topic of this chapter. But first I shall start by severely delimiting and emphasising the scope of my discussion here:

(i) The Greeks employed two different kinds of numerals, the so-called acrophonic and alphabetic systems. There are wide regional variations in the acrophonic system; the earliest and best-represented version in our evidence, that found in Attica, is based on a variety of different forms of the six elementary symbols: $I = 1$, $\Pi$ (for $\pi\acute{\epsilon}\nu\tau\epsilon$; the right-hand leg of the pi is shorter than the left) $= 5$, $\Delta$ (for $\delta\acute{\epsilon}\kappa\alpha$) $= 10$, $H$ (for $\acute{\epsilon}\kappa\alpha\tau\acute{o}\nu$) $= 100$, $X$ (for $\chi\acute{\iota}\lambda\iota\sigma\iota$) $= 1000$, and $M$ (for $\mu\acute{\upsilon}\rho\iota\sigma\iota$) $= 1\,0000$, together with four compound symbols, also found in a variety of forms, which denote 5, 500, 5000, and 5 0000. So much is well known

---

[1] The definitions are not numbered in the best manuscripts of the *Elements*, and they are grouped together in blocks connected by $\delta\acute{\epsilon}$; see Plate 6. For example, in Book I, Definitions 1–3, 4–6, 7–10, 11–12, 15–18, and 19–22 are so connected; for a discussion, see Turner, Fowler, Koenen, & Youtie, EEI, 19–20. The numbers that occur in the manuscripts of the (possibly pseudo-Euclidean) *Sectio Canonis*, Proposition [9] (here the propositions are unnumbered) require some emendation for them to make sense; see Section 4.5(d). This, we shall see, is typical. There are abundant numerical illustrations in the later scholia to the *Elements*, one of which has entered the text of some manuscripts; see Euclid–Stamatis, *EE* ii, 237–8. In particular, there is a frequently cited later annotation on folio 32$^v$ of B which consists of the Arabic numerals for 1 to 9 (see Vogel, BIZB, which reproduces facsimiles of the text, and Wilson, MP, 401–4 for an informed discussion); unfortunately the description of this in Euclid–Heiberg, *Opera* v, p. xix has been omitted from Euclid–Stamatis, *EE*. This missing text reads, insofar as it can be reproduced here: "*De notis numeralibus arabicis, quae in scholiis Vindobonensibus maxime in libro X occurrunt, hoc tantum commemorabo, scholia illa manu $V^b$, h.e. sine dubio saec XII, exarata esse. pro numero 5 usurpatur 0, nostrum vero 0 punctum est vel $^\circ$; prorsus similes sunt series numerorum in B fol. 32$^V$ (ad initium libri 11) m. rec. [see below] et in b ad II, 1 m. rec. $\psi\hat{\eta}\phi\sigma\varsigma$ $\acute{\iota}\nu\delta\iota\kappa\acute{\eta}$ [see below].*" The omitted passages reproduce the form of these Arabic numerals; examples of these can also be seen in many of the scholia in Book X, for example nos. 152, 155, 170, 183, 184, and many from no. 356 onward.

[2] This passage was quoted in Section 4.5(c), above.

and frequently described,[3] but what is less frequently pointed out is that, with the exception of a few early papyrus accounts recently found at Saqqara, of which only one has so far been published,[4] and occasional archaic usages in literary stichometry, the acrophonic system is only found on inscriptions, and has little relevance to mathematics.[5] I shall consider exclusively the alphabetic system.

(ii) We must differentiate clearly between texts which are copies of earlier texts, such as Byzantine manuscripts, and texts which may be originals or which, if copies, more closely reflect early practice, such as some material found on papyrus. One objective of this chapter will be to emphasise the problem of assessing the effect of the later modification of the conceptions and notations, especially in the descriptions of fractional quantities, during the ninth to sixteenth centuries, the period during which most of the surviving texts were copied. But note that I shall follow here the almost universal practice of presenting any Greek words or texts in modern printed form, as they have been transcribed by their editors,[6] and will only occasionally refer to the wide diversity of different styles which are found in the original manuscripts.

(iii) In scientific computations from the time of Hipparchus and Hypsicles (*c.* 150 BC), and especially in Ptolemy's *Almagest* and its commentaries, the Greeks employed an alphabetic version of the sexagesimal system found earlier in Babylonian cuneiform clay tablets that date back almost to 2000 BC. But there is no trace whatsoever of sexagesimal numbers in any Greek text from before the second century BC, and I shall not consider them further here.[7]

---

[3] See, for example, Heath, *HGM* i, 29 ff., Thomas, *SIHGM* i, 41 ff., etc.

[4] Turner, FODMS. Also see Section 6.2, n. 30 and its text.

[5] For descriptions of the acrophonic system, based on a thorough familiarity with the epigraphical evidence, see Guarducci, *EG* i, 417–28; Tod, *AGNS*; and the more specialised Johnston, *TGV*, Chapter 6, and Lang, NNGV. Concerning stichometry—the labelling or counting of lines or verses in a text—see Tod, *AGNS*, 32 f. (= GNN 130 f.), Turner, *GP* 94 f., McNamee, *AGLPO*, 122, and, for examples, P. Ryl. iii 540, P. Oxy. x 1231 and xlii 3000, and P. Herc. 1151. These uses of acrophonic numerals are not unlike the way we have retained Roman numerals in stylised contexts such as labelling the books of the Bible (as in 'I Kings' for 'the first book of Kings') or on clock dials (which consistently employ the form IIII, now elsewhere very unusual).

[6] So one text, P. Lond. ii 265, in Section 7.3(c) below, will be given without diacritical marks, while others have them added or corrected by their editors. For a not entirely successful attempt to use characters that more closely resemble early script, see Fowler & Turner, HP. The editors of *The Oxyrhynchus Papyri* gradually abandoned their original practice of employing the specially created font that can be seen in Plate 1 for their diplomatic transcriptions (i.e. transcriptions as close as possible to the written text).

[7] The earliest surviving text which contains anything resembling the sexagesimal system is Hypsicles, *Anaphorikos (On Rising-times)*; see Hypsicles–De Falco, Krause, & Neugebauer, AG. (This manuscript is in Vatican gr. 204; see Chapter 6, n. 52 and text.) Also there are later secondary reports that Eratosthenes (*c.* 250 BC) divided the circumference of a circle into sixty parts; the Keskinto inscription in Rhodes, *IG* xii(1) 913, has a division of the circle into 360°; and the Antikythera Mechanism of 80 BC has brass circular scales divided into 360° (see Price, *GG*). Numerical practice is rarely consistent, and different systems are often mixed together, even within a single context. For discussion and references, see Neugebauer, *HAMA* ii, 590 ff. and 698 ff.

(iv) Any convenient and easily remembered sequence of labels or marks can be used to stand for the sequence

first, second, third, fourth, fifth, ... ;

for example we can, and often still do, use the letters of the Greek alphabet, $\alpha$, $\beta$, $\gamma$, $\delta$, $\epsilon$, .... Sometimes, as for example in labelling the books of the *Iliad* and *Odyssey*, the next letter in this sequence is $\zeta$, as we might expect; in this case there will be no numerical overtones, and we have a pure letter labelling system. Our earliest evidence about numerical notations is of such kinds of usage, and this also will not concern us here;[8] ordinal arithmetic will not become a part of mathematics before the nineteenth century.[9]

The main topic of this chapter will be the use of the alphabetic numerals and the associated system for treating fractional quantities, in school, commercial, and scientific texts. I shall regard it as very important to pin down as many as possible of the various points of the discussion to illustrative examples which, in many cases, will then be reproduced and transcribed in the plates, and translated and described in the text.

The standard Greek alphabet consists of twenty-four letters. In the alphabetic numeral system,[10] this is supplemented by three letters:[11]

- *vau* or *digamma*, originally written $\subset$ or $F$, which later, in minuscule, is called *stigma* and is transcribed as $\varsigma$, a letter not unlike, but not identical with, the final form of sigma in Porson's Greek types;
- *koppa* $\varsigma$, or $\varphi$, and
- *san* or *ssade* $T$, later, in minuscule, called *sampi* and written $\lambda$.

These twenty-seven letters were used in three groups of nine to represent the unit digits:

[8] See Tod, LLGI. For examples of confusion between $\zeta$ as the sixth letter of the Greek alphabet and the numeral for seven, see Keaney, ETTHP, 296 n. 2, and Sharples, NETDP, 144.

[9] Landau's joke at the beginning of his *Grundlagen der Analysis* (1930) could scarcely have been understood before the twentieth century: " 'Theorem 1', 'Theorem 2', ... , 'Theorem 301' are simply labels for distinguishing the various theorems, ... and are more convenient for the purposes of reference than if I were to speak, say, of 'Theorem Light Blue', 'Theorem Dark Blue', and so on. Up to '301', as a matter of fact, there would be no difficulty whatever in introducing the so-called positive integers. The first difficulty—overcome in Chapter 1—lies in the totality of the positive integers 1, ... , with the mysterious series of dots after the comma, in defining the arithmetical operations on these numbers, and in the proofs of the pertinent theorems" (Landau, FA, p. v).

[10] The best general account of this alphabetic system I know is Smyly, EAGL; this article is informed with Smyly's experience in transcribing difficult numerical papyri, but note that all of his comments on common fractions, $p/q$, are based entirely on examples drawn from minuscule manuscripts. Also Coulton, TUGTD, 74–89, Neugebauer, *ESA*, Chapter 1, and de Ste. Croix, GRA, are highly recommended. The Phoenician antecedents of the Greek alphabet are described in Coldstream, *GG*, 295–302 and 382.

[11] For further details of these extra letters, see Liddell, Scott, & Jones, *GEL*, s.vv. $F$ (between $\epsilon$ and $\zeta$), $M$, $\varsigma$ (both between $\pi$ and $\rho$), $\Sigma$, $\delta i\gamma\alpha\mu\mu\alpha$, and $\kappa \acute{o}\pi\pi\alpha$.

|       | $\alpha$ | $\beta$ | $\gamma$ | $\delta$ | $\epsilon$ | $\varsigma$ | $\zeta$ | $\eta$ | $\theta$ |
|-------|----|----|----|----|----|----|----|----|----|
| for   | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9,  |

the tens:

|       | $\iota$ | $\kappa$ | $\lambda$ | $\mu$ | $\nu$ | $\xi$ | $o$ | $\pi$ | ϙ or ϙ |
|-------|----|----|----|----|----|----|----|----|----|
| for   | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90,  |

and the hundreds:

|       | $\rho$ | ϲ | $\tau$ | $\upsilon$ | $\phi$ | $\chi$ | $\psi$ | $\omega$ | Ͳ or ϡ |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| for   | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900.  |

An extra place of thousands was often originally indicated by a hooked or looped line extending above the units, $\alpha$ to $\theta$, but this notation gave way, in the fourth century AD, to an oblique stroke below and to the left, right, or under the letter.[12] The notation of a lower left stroke was later adopted in minuscule manuscripts, and thousands are now usually transcribed as either a lower case letter with a lower left accent, or as an upper case letter: $,\alpha$ or $A$ to $,\theta$ or $\Theta$ for 1000 to 9000. The digits may be written in any order in the course of a calculation, but the answer to the calculation or a number in isolation is almost always written in decreasing order of digits, left to right. So, for example 5324 may appear in an edited texts as $,\epsilon\tau\kappa\delta$ or $\overline{E\tau\kappa\delta}$, where the bar over the numeral is a sign, also generally adopted around the fourth century AD (though found not infrequently before then, especially to indicate dates; see Section 7.1(c), below, for an example), and then used in minuscule manuscripts, that is used to distinguish a number from a word or from the letters used to label a figure.[13] The evolution of the letter san into sampi caused some confusion; we shall see a typical example where we find $,\epsilon\tau\kappa\delta$ where we would expect $,\epsilon$ ϡλδ, 5924, and confusion between ϡ and $\lambda$ or $\psi$ is also common.

---

[12] In the Achmîm Mathematical Papyrus (no. 12 in the catalogue in the Appendix) the scribe employs all three usages to denote thousands; see, for example, the Table Paléographique, Planche 1; also see Milne & Skeat, *SCCS*, 62–4. There are valuable discussions of the papyrological evidence concerning thousand, myriad, and fraction indicators in Brashear, *CTAL*, 215 n. 2, and MS.

[13] In one manuscript of Euclid, the scribe has confused the use of the alphabet to label the figure with its use as numerals, and so has translated the letter labels on the figure as Arabic numerals; see Heiberg–Euclid, *Opera*, supplement (= vol. x), ed. Curtze, pp. xvi–xxvii. Numbers can also be confused with words; for example the scribe of Heron's *Metrica* at one place copied the text $\mu\gamma$ $\acute{\gamma}$ $\lambda\hat{\eta}\mu\mu\alpha$ (. . . 43 3′. Lemma . . .) as 43 3′ 38′ 40′ 41′, in the notation to be described in the next section; see Heron–Bruins, *CC* i, folio 77$^r$ line 16 or Heron–Heiberg, Schöne, & others, *Opera* iii, 50. There is no guarantee that an edited and printed text follows the practice of the scribe of the manuscript in its handling of these matters; see how the bars over numerals representing dates are omitted in Plate 7, lines 56, 58, 62, 64 and 66. Also, the numerals that occur in *Elements* XIII 11 (see the opening paragraph, above) are written in the manuscripts B and P, at least, in exactly the same way as the labels on the diagrams: in majuscules in B, minuscules in P, both distinguished by a superior bar; but one cannot tell this from Heiberg's text.

Larger numbers, tens of thousands up to tens of millions, were originally indicated by writing these letters over a mu (for μυριάδες, myriads or ten thousands); then, from the second century AD onwards, other abbreviations and arrangements appear. This system ran out at a myriad myriad, 'ten thousand times ten thousand', the proverbial multitude that no man could number (see Daniel 7:10 and Revelation 5:11).[14] It is now easy for us to see how, in principle, to continue this procedure automatically up to a myriad blocks each containing a myriad (a myriad raised to the power of a myriad) and this appears to be the basis of the system described by Apollonius in a book that is now lost, but which is referred to in Pappus, *Collection* II (excerpted in Thomas, *SIHGM* ii, 352–7). Archimedes' *Sandreckoner* describes an exponential system for arbitrarily large numbers (excerpted in Thomas, *SIHGM* ii, 198–201) in which the process just described—which gives numbers of what he calls the first period—is then repeated to give numbers of the second period.[15] Since Archimedes takes as the limit of his basic system of numbers not the myriad, but the myriad myriad, this yields a myriad myriad raised to the power of a myriad myriad, then raised again to the power of a myriad myriad. At this point he stops his description and gives a theoretical analysis of his procedure.

This interest in very large numbers is clearly theoretical and restricted to mathematics. The only place outside the works of Archimedes, the now lost book of Apollonius, and Diophantus[16] that I know where there is any attempt or need to express a number bigger than a myriad myriad is a very corrupt passage in Theon of Smyrna–Hiller, *ERMLPU*, 126–7, where Theon estimates the volume of the earth. Our only known manuscript, Venice Marciana gr. 303, of this half of the treatise (pp. 120–205) was made in the fourteenth or fifteenth century by a scribe who could not read Greek numerals, so almost all of them

---

[14] So this would be a suitable number to use in my modified $S'_{21}$ in Section 2.3(b), just as I did in the last paragraph of $E_{67}$, in Section 4.4(b). A similar device can be used in manipulating anthyphairetic ratios in a calculating machine, which has the same finite capacity for integers. I have imitated the feature of counting by myriads by splitting transcriptions of numbers that arise in Greek contexts into blocks of four, rather than the now conventional blocks of three or five.

[15] Since Archimedes does not describe how the precise value of a number is got by adding together the various orders of the different periods, it might be more accurate to say that Archimedes' system describes the size of, or an approximation to, a number and this is how he uses it, to obtain an estimate of the number of grains of sand that would fill the universe. Contrast the way we say 'the solution of Archimedes' *Cattle Problem* is $7 \cdot 6 \times \cdots \times 10^{206\,544}$, in which the relation between the approximate and exact values is more straightforward. (The number itself completely fills $46\frac{1}{4}$ sheets of computer print-out, as reproduced in Nelson, *SACP*, 165–76!)

[16] See Diophantus–Heath, *DA*, 47 (but Heath's example can only be found in the apparatus to Diophantus' text, Diophantus–Tannery, *Opera* ii, 323.8, and there are further corruptions too involved to be pursued here). For further comments on Diophantus' notation, see n. 28, below. There is an isolated exception to my statement in a table of squares which goes up to 1 0000 × 1 0000; see no. 68 in the Appendix to this chapter.

are corrupt and have had to be restored by the editor.[17] On the other hand, there is little sign of any difficulty in commercial documents with manipulations which involve numbers in the tens or hundreds of millions, and such large numbers are not uncommon.

### 7.1(b) *Simple and compound parts*

Now let us turn to fractions. First, the word: I shall always use it with its general meaning of "a numerical quantity that is not an integer" (*OED*, s.v. fraction). All other meanings will be qualified: a *common*, or *vulgar* fraction for $p/q$, which is *proper* or *improper* according as $p$ is less than $q$ or not; when $p = 1$, the common fraction is a *unit* fraction; a *decimal* fraction for the decimal expression of a fraction, which may then be *rounded* or not, *periodic* or not, and *terminating* or not; a *sexagesimal* fraction; a *complex* fraction, for a common fraction that has a fraction for its numerator or denominator; a *continued* fraction; and so on. Note that fractions, being 'numerical quantities', belong to what I have called the arithmetised style of mathematics, and the non-arithmetised approach may often be signalled by the use of the alternative terminology of ratios.[18]

I shall argue in this chapter that we have no evidence for any conception of common fractions $p/q$ and their manipulations such as, for example, $p/q \times r/s = pr/qs$ and $p/q + r/s = (ps + qr)/qs$, in Greek mathematical, scientific, financial, or pedagogical texts before the time of Heron and Diophantus; and even the fractional notations and manipulations found in the Byzantine manuscripts of these late authors may have been revised and introduced during the medieval modernisation of their minuscule script. Among the thousands, possibly the tens of thousands, of examples of fractions to be found in *contemporary* Egyptian (hieroglyphic, hieratic, and demotic), Greek, and Coptic texts, all but a few isolated examples in five texts (P. Lond. ii 265 (p. 257); M.P.E.R., N.S. i 1; and three demotic papyri published in Parker, *DMP*), all to be described in detail in Section 7.3, below, use throughout the following 'Egyptian system for expressing fractions:

We take the basic sequence of the *arithmoi*:

$$\text{two, three, four, five}, \ldots,$$

represented in Greek by the letters $\beta, \gamma, \delta, \epsilon, \ldots$, and convert it into the sequence

$$\text{half, third, quarter, fifth}, \ldots,$$

where, after the exceptional cases of the first few terms, for which special symbols (to be described below) are assigned, the derived symbol is got by

---

[17] See Smyly, NTS, for a discussion.
[18] See the example of Aristarchus, *On the Sizes and Distances of the Sun and Moon*, in Section 7.3(b), below.

appending to the numeral a long line, straggling upwards, now generally called a 'fraction indicator' and usually transcribed as an accent, $\gamma$, $\delta$, $\epsilon$, . . . , or a prime, $\gamma'$, $\delta'$, $\epsilon'$, . . . ; clear examples can be seen in Plates 7 and 8. Later a variety of strokes, bars, and dots are found as fraction indicators.

This sequence: half, third, quarter, fifth, . . . is often confused with the nearly homonymous sequence of ordinal numbers: first, second, third, fourth, fifth, . . . , and its terms are often called unit fractions. The universal Greek way of referring to them is as μέροc or μόριον, 'part' (plural μέρη or μόρια, 'parts'). The two Greek words are used interchangeably, sometimes both in the same passage; and they are often translated as 'fractions' or even 'denominators',[19] but I shall always refer to the $\angle$, $\gamma$, $\delta$, $\epsilon$, . . . more accurately as 'parts', and distinguish, where necessary, between simple parts[20] (i.e. isolated members of the sequence) and compound parts (to be described below). One reason why I wish to avoid the name of unit fractions is because of the overtones of the notation it suggests, $1/q$, which immediately introduces associations with the common fraction $p/q$; these associations are completely absent from the Greek description and notation. I shall instead write the translated sequence of parts as $\acute{2}$, $\acute{3}$, $\acute{4}$, . . . , $\acute{10}$, $\acute{11}$, . . . , in imitation of early Greek practice, or as the typographically more convenient and understandable $2'$, $3'$, $4'$, . . . , $10'$, $11'$, . . . , or as 3rd, 4th, . . . , 10th, 11th, . . . ,[21] and I recommend strongly that they be read 'the half, the third, the quarter or the fourth, . . . '[22] and not 'one-half, one-third, one-quarter or one-fourth, . . . '.

The sequence of parts starts with $\acute{\beta}$, 'the two parts', τὰ δύο μέρη, an expression for 'two-thirds'.[23] This, to begin with, was the standard letter beta with the usual fraction indicator (see Plates 7 and 8, to be described below, and the detailed analysis of this point in Fowler & Turner, HP), and it is found later in a range of variants and forms; each text may show its own version. However

---

[19] See, for example, *Timaeus* 36a–b, or Heron–Heiberg, Schöne, & others, *Opera* v, 96.4 and 12. (Liddell, Scott, & Jones, *GEL*, s.vv. μέροc and μόριον, cite the apparently incorrect *Stereometrica* 2.14 and 16 and translate, following Heiberg's *index verborum*, as 'the denominators of fractions'.) Euclid refers to an *arithmos* and its corresponding *meros* as being ὁμώνυμοc, 'called by the same name', in *Elements* VII 37–9.

[20] A simple part is also called an aliquot part; see the *OED*, s.v. aliquot, which cites Billingsley on *Elements* V Definition 1: "This . . . is called . . . a measuring part . . . and of the barbarous it is called . . . an aliquot part".

[21] The notation $\bar{2}, \bar{3}, \bar{4}, \ldots$, based on Egyptian practice, is also used, but this conflicts with the later scribal notation for distinguishing a number or figure label from the text, or for differentiating one kind of number from another. See n. 13, above, and the text to that note,

[22] Curiously, just as we can never say 'second' for 'half' so the Greeks could never write $\beta$ (see below); and just as we have the alternatives 'quarter' and 'fourth', so they had an optional special form for $\delta$. However, Greek has a special terminology for epimoria (described in Section 4.5(b) (see $A_{60}$ and $B_{61}$)) and their reciprocals (see the next note) which has no idiomatic translation.

[23] This principle is sometimes said to be general, with τὰ τρία μέρη, the three parts (sc. of four) for three-quarters, and so on, but I cannot cite any Greek examples, and only two-thirds has an associated symbol. Also see n. 83, below.

extraordinary it may seem to us, it is an incontrovertible fact that the sequence of parts starts with two-thirds, and this term has exactly the same status as the other parts in the symbolic expression of fractions. (This is an additional reason for avoiding the name 'unit fractions' for the sequence $\beta$, $\angle$, $\gamma$, $\delta$, . . . , since $\beta$ is not a unit fraction.) I shall transcribe all variants of this symbol as $\beta$, and translate it as $\overset{=}{3}$ or $3''$. The symbol $\delta$ is often found in a special form in which one side of the $\varDelta$ is extended to form the fraction indicator, and this is sometimes transcribed as d or d'; see for example, the transcription in Plate 8. The various signs for the half may or may not incorporate a fraction indicator which then may or may not be shown on the transcribed text; here they will all be transcribed as $\angle$.

More complicated fractions than simple parts are expressed as sums of an integer and *different* simple parts;[24] I shall refer to them as *compound* parts. They are laid out as follows: first come the digits of the integer part of the number, written out in decreasing order as was explained above; this is followed by the successive parts, in order (i.e. order of increasing 'denominator'), the digits of each part (i.e. 'denominator') being written, as usual, in decreasing order.[25] Some pairs of symbols may be ligatured into a composite symbol, for example $\angle\delta$ for $\frac{3}{4}$ (not unlike an upside-down $\beta$, and sometimes printed as such) and $\beta\varsigma$ for $\frac{5}{6}$.

Let us now look at two texts that will illustrate many of these features.

### 7.1(c)  *P. Hib. i 27, a parapegma*

Many of the features so far described are illustrated in the early papyrus text, P. Hib. i 27; see Plate 7. In addition to its very full original publication in Grenfell & Hunt, *HP* i, 138–57, there are the easily accessible descriptions of its astronomical content in Neugebauer, *HAMA* ii, 599 f. and 686 ff., and of its numerical features in Fowler & Turner, HP, so my account here can be very brief. This text is perhaps our earliest physically surviving Greek scientific or semi-scientific text; it comes from the cartonnage of a mummy and is dated, from internal evidence of the text, to about 300 BC, and this date seems to be corroborated by palaeographical analysis. It opens with a fragmentary and elementary introduction, believed to be based on Eudoxus' astronomical system, and this is followed by a parapegma (see Section 4.4(c)) of which column 4, lines 55–68, is reproduced and transcribed here in Plate 7. The translation of this part of the text is:

[24] The only example I know of an expression that involves repeated parts is in a schoolboy's exercise, P. Louvre [inv.] MND552K side A, described at no. 17 in my catalogue in the Appendix.

[25] The only Greek example I know where the parts are not arranged in decreasing order is P. Oxy. xii 1446, a survey of land and rents. The fractional quantity $\gamma\varsigma\rho\nu\iota\beta$ (3 6' 150' 12') occurs twelve times and $\delta\delta\nu\tau\iota\beta$ (interpreted as 4 4' 50' 300' 12') occurs once. The editors suggest that the anomaly may have been caused by the later addition of an extra 12' artaba of wheat to the rent of the land. Also see n. 88, below.

55 The night is 13 12′ 45′ hours, the day 10 3″ 5′ 30′ 90′.
56 16th, Arcturus rises in the evening,
57 the night is 12 3″ 15′ 45′ hours, the day 11 9′ 10′ 30′.
58 26th, Corona rises in the evening,
59 and the north winds blow which brings the birds, the night
60 is 12 2′ 30′ hours, the day 11 3′ 10′ 30′. Osiris
61 circumnavigates, and the golden boat is brought
62 out. Tybi 5th, the sun enters Aries. 20th, spring equinox,
63 the night is 12 hours, and the day 12 hours
64 and the feast of Phitorois. 27th, Pleiades
65 set in the evening, the night is 11 3″ 6′ 90′ hours,
66 the day 12 10′ 30′ 45′. Mechir 6th, the sun enters
67 Taurus. Hyades set in the evening,
68 the night is 11 2′ 10′ 35′ hours, . . .

As is typical with numerical texts, there are several errors which here can be corrected with some confidence. They are noted in the editors' apparatus:

55 for 13 12′ 45′ read 13 15′ 45′
68 for 11 2′ 10′ 35′ read 11 2′ 10′ 30′ 90′

and the editors have restored some parts, including two smudged numbers on lines 57 and 65 where the scribe has illegibly altered the text.

Note the following features:

(i) The distinctive long straggling fraction indicators.

(ii) The letter $\beta$ in $\overset{\prime}{\beta}$ for $\frac{2}{3}$ (e.g. end of line 55); it is the same as the $\beta$ in $\iota\overset{\prime\prime}{\beta}$, 12′ (e.g. the beginning of this line) or the letter $\beta$ (e.g. line 59).

(iii) The distinction between stigma (fragmentary examples in lines 56 and 58; complete example in line 66) and sigma (no examples occur as numbers; the letter occurs in lines 56 and 58, with no distinction between upper and lower cases, and initial, medial, or final forms). The written letters are completely different, and the transcribed printed letters, in this well-designed font, are slightly different.

(iv) The letter koppa (e.g. at the end of line 55).

(v) The way the dates, the original numbers, are distinguished throughout by a superior bar (not indicated in the transcription), but how the *arithmoi* are not otherwise distinguished from the text.

It is now standard editorial practice to modernise the notation of fractions. For example $\iota\overset{\prime\prime\prime}{\beta}\epsilon\lambda$ at the end of line 55 would now almost always be silently translated as $10\frac{41}{45}$, as was done in Grenfell and Hunt's original edition. There is no harm whatsoever in this as long as it is done consciously; this kind of translation makes texts accessible and comprehensible. But we must remain aware that translations like this can alter the significance of the text, and that similar translations may have occurred in some texts during their long period of transmission and revision, down to our own time.

### 7.1(d) *O. Bodl. ii* 1847, *a land survey ostracon*

There is evidence of Egyptian land surveys from at least 3000 BC, and many detailed Greek examples from the Ptolemaic period, dating from the second century BC onwards, have been discovered and published. They vary from short calculations on potsherd or stone ostraca to long papyrus rolls setting out the survey of a whole village. There are, broadly speaking, two kinds of record: one giving the owner, location, dimensions, and area of each field, while the other, the annual return, gives the area (but not the dimensions), the crop under cultivation, the annual tax paid, and sometimes details of roads, drains, and canals which pass through the land.[26] A typical short example of the first kind of record, O. Bodl. ii 1847[27] from Thebes in Egypt, dated to between 30 BC and AD 14, is given with its transcription in Plate 8. These rapidly written, cursive, highly abbreviated texts present enormous problems to the uninitiated reader, rather as happens today with doctors' prescriptions. However, most abbreviations in this text have been successfully resolved, the added letters being put in parentheses (though, on the bottom line of the apparatus, the second editor notes that one abbreviation in line 2 has defeated her while another is uncertain).

The text reads, in translation:

1   5 2′ 4′ 8 2′ 4′

2   The land measurement of the (land) of Kalliedon (farmed) by Petechonsis, beginning(?) on the west.

3   4′ 8′ 16′ 32′ $\dfrac{2'\ 4'\ 8'}{1}$ 8′ 16′ / 4′ 16′ arouras. Adjacent to the north 32′.

4   Another (field) adjoining the bottom, 2′ 4′ 8′ $\dfrac{1\ 2'\ 8'}{16'}$ 1 32′ / 2′ 4′ 32′.

5   Another (field) adjoining 2′ 8′ $\dfrac{2'\ 4'\ 8'}{2'\ 4'}$ 2′ / 4′ 8′ 16′.

6   Another (field) 4 $\dfrac{2'\ 4'\ 8'\ 16'\ 32'}{1\ 8'\ 32'}$ o / 4 4′.

7   Another (field) adjoining 4 2′ 4′ $\dfrac{2'\ 4'\ 16'\ldots}{2'\ 4'\ 8'\ 16'\ 32'}$ / 3 2′ 4′ 8′ 32′.

8   / 9 2′ 4′ 8′ [16′] arouras.

9   Adjacent on the north, (farmed) by Lysimachos

10   5 $\dfrac{2'\ 4'\ 8'\ 32'}{1\ 4'\ 8'\ 32'}$ o / 5 2′ 4′ 32′.

11   Another (field) adjoining 4 $\dfrac{2'\ 4'\ 8'}{2'\ 4'\ 8'\ 32'}$ o / 3 4′ 8′ 32′.

12   / 9 2′ 8′ 16′.

The basic calculation, which occurs in lines 3, 4, 5, 6, 7, 10, and 11, is as follows. The lengths *a*, *b*, *c*, and *d* of the four sides of a field (with *a* opposite *c*,

---

[26] For a discussion of land surveys in general, and the papyrological evidence in particular, see Crawford, *KEVPP*.

[27] Also known as O. Tait; it is published in Tait, Préaux, & others, *GOBLO* ii, 308.

$b$ opposite $d$), measured in schoinia and half, fourth, eighth, sixteenth, and thirty-second parts of a schoinion (the word means 'measuring cord', and 1 schoinion $= 100$ cubits, 1 cubit $=$ elbow to knuckle, approximately) are written out $a \frac{b}{d} c$. If two sides of a field are equal, then the second will be replaced by an ○, perhaps for *homoion*, or a dot, or even omitted (see lines 6, 7, 10, and 11, and the editor's note to line 7). In the case of a triangular field, *ou* (for *outhēn*, nothing) would replace one side; there are no examples here. The oblique line / is one of the standard symbols in documentary papyri for equality; it is usually described as an abbreviation for the gamma of *ginetai* or *ginontai*, 'becomes', but is more probably derived from an Egyptian sign, which is later often Hellenised into a gamma; then, in the Byzantine period, this symbol disappears completely, and is replaced by much more explicit biliteral or triliteral abbreviations which continue into the minuscule manuscripts.[28] The area of the field follows, given in arouras (1 aroura $= 1$ square schoinion, equal to approximately 0.676 acres, or 0.25 hectares), and it always appears to be calculated by multiplying together the averages of the two pairs of opposite sides,

$$\text{nominal area} = \tfrac{1}{2}(a + c) \times \tfrac{1}{2}(b + d) = \tfrac{1}{4}(ab + ad + bc + cd),$$

---

[28] The text M.P.E.R., N.S. i 1, to be described below in Sections 7.3(d) and 4(a), uses variously this sign, the biliteral abbrevation ﬁ, or the word *ginetai*. On these abbreviations, see Blanchard, *SAPDG*, 29 ff, especially notes 9–11 and the text to these notes, and McNamee, *AGLPO*. Differences between the abbreviations in the papyrological traditions and those found later in Byzantine minuscule texts are very common. For example, consider the abbreviations for subtraction: throughout the Byzantine minuscule manuscripts of Diophantus and once in Heron (at *Metrica* III, 7; see Heron–Bruins, *CC* i, folio 103$^{\mathrm{r}}$, line 5 for a photograph), we find a sign similar to a rounded vertical arrow, ⬆. I am grateful to Alain Blanchard for the following sketch of the general development of the earlier Greek symbol for 'from which': "My impression is that there must again be a break between the papyrological and medieval periods. The papyrological forms, derived from the demotic ∠, *wp.t* (for this word see Mattha, *DO*, 156, no. 190, line 2; Erichsen, *DG*, 85; and Pestman, *GDZA*, 76–8) are relatively faithful to their prototype, with different kinds of deformations. There are various kinds of examples: ⟩, U.P.Z. ii 157.9, 242 BC and M.P.E.R., N.S. i 1 [see Section 7.3(d), below]; ∠, P. Petrie ii 39c.9, iii BC, to P. Lond. i 99 (p. 158).28, iv AD; and ⟩, P. Berl. Leihg. 13.15, beginning of ii AD. The demotic word appears to be used for itemising, with subtracting as a secondary meaning; the Greek expression seems to be ἀφ' ὧν, 'from which'. Despite some reinforcement of this sign at the end of its evolution (for example, ℓ°, O. Bodl. i 1957.3, ii AD? and ℓ·, P. Lond. ii 755 verso (p. 221).38, iv AD), the sign did not survive beyond the Roman period. In the Byzantine period it is replaced by a more explicit abbreviation $\alpha\phi$ (e.g. P. Sorb 61.16, v AD? & P. Lond iv 1419.1396, after AD 716); or $\alpha\phi^{\omega}$ (e.g. P. Lond iv 1413.221, AD 716–21); or $\alpha\phi^{\circ}$ (e.g. P. Lond iv 1419.43, after AD 716)."

The possibility that the sign ⬆ might be an introduction during the Byzantine minusculisation was dismissed by Tannery, the editor of Diophantus' works, in SSCG, 208–9: "A la vérité, comme on ne le rencontre que dans ce passage [Heron, *Metrika* III 7] . . . il pourrait n'y avoir là qu'une abréviation byzantine; cependent cette dernière hypothèse n'est guère vraisemblable." Tannery's argument is that the manuscript of Heron was copied more than two centuries earlier than those of Diophantus, at a time when Diophantus was not being studied; and Diophantus' procedures had been too little vulgarised for them to have been the source of this sign. But, one can object, these are far from the only possibilities. Also see Diophantus–Heath, *DA*, 41–4.

and approximating the result to within the thirty-second or less of an aroura. Look, for example, at line 3: We have $a = 4'\ 8'\ 16'\ 32'$, $c = 8'\ 16'$, so

$$\tfrac{1}{2}(a + c) = 8'\ 8'\ 16'\ 64' = 4'\ 16'\ 64',$$

while $b = 2'\ 4'\ 8'$, $d = 1$, and so

$$\tfrac{1}{2}(b + d) = 2'\ 4'\ 8'\ 16'.$$

Now do a long multiplication sum in what is, in effect, binary representation to get

| 2' | 4' | 8' | 16' | | | | |
|---|---|---|---|---|---|---|---|
| | ×4' | | 16' | | 64' | | |
| | 8' | 16' | 32' | 64' | | | |
| | | | 32' | 64' | 128' | 256' | |
| | | | | | 128' | 256' | 512' | 1024' |
| | 8' | 16' | 16' | 32' | 64' | 128' | 512' | 1024' |

and simplify the carries to give $4'\ 32'\ 64'\ 128'\ 512'\ 1024'$. We do not know just how the taxman has performed this calculation, which here has been rounded up to $4\ 16'$, an approximation in his favour of $256'\ 1024'$ of an aroura.

It is easy to show that the area of the field is actually $\tfrac{1}{4}(ab \sin ab + ad \sin ad + bc \sin bc + cd \sin cd)$, so the taxman's formula will *always* overestimate the area unless the field is rectangular. Sometimes it seems that a field may be split up into pieces, to reduce the error; and sometimes a correction, called a 'difference of measurement' (*diaphoron schoinismou*) may be noted, though this correction is itself likely to be incorrect, and the tax is still calculated on the uncorrected area.[29] There are also frequent arithmetical mistakes; see, for example, line 7, where the answer to the multiplication

$$4\ 2'\ 4' \times 2'\ 4'\ 8'\ 64' = 4\ 8'\ 16'\ 32'\ 128'\ 256'$$

is given by the scribe as $3\ 2'\ 4'\ 8'\ 32'$, and no readjustment of the doubtful numbers seems to lead to any approximation to the scribe's result, although the editor points out that the absence here of either a number, or an o, or an *ou* in the right-hand place may indicate a scribal omission of a dimension. The entry on line 11 also seems to contain an error. But since some annual tax returns seem purely notional, with the same minutely detailed account being returned in consecutive years, perhaps it is irrelevant to pick on such small details as these.[30]

[29] Sir Eric Turner commented on a preliminary version of this, in a letter written in April 1983, a week before his death: "I wonder whether *diaphoron schoinismou* means 'tax-payment for measurement' rather than 'difference of measurement'. There are many taxes under the name of *diaphoron*. Even more in the pocket of a king or inspector!"

[30] See de Ste. Croix, GRA, 59: "Mistakes in operating with fractions are particularly common in papyri. But it is sometimes difficult to tell how far deficiencies observable in

This text illustrates a feature of commercial documents, that there are conventions restricting the range of permitted parts. In recording and manipulating distances and areas, as here, fractions are always expressed in halves, fourths, eighths, sixteenths, . . . of schoinia and arouras; while taxes and rents are expressed in equivalents of artabas of wheat (see Section 7.3(c), below), where the permitted parts are $2', 4', 8', 5', 10', 3'', 3', 6', 12', 24', \ldots$ which simplifies division by various numbers.[31] This means that approximations are necessary—for example, it is not possible to express the third of an aroura of land, or the seventh of an artaba of wheat exactly, within these restrictions— and these approximations are sometimes crude and usually in the taxman's or landlord's favour.

The earliest surviving land registers such as I have described here date from the second century BC, but some evidence for a land survey in Egypt can be traced back to at least 3000 BC. The approximate formula for the area occurs in Problems 51–3 of the Rhind Mathematical Papyrus, where it is applied to isosceles triangles and trapezia, though many commentators conceal the fact that it is an approximation by adapting the translation, and rendering 'side' as 'height'. After the conquest of Egypt by Alexander in 331 BC, the Greeks took over the existing Egyptian administration, including the land survey, merely imposing Greek as the official language for virtually all administrative and financial documents that were of concern to them. On the other hand, the Greeks never seemed to have operated a land survey in Greece itself, though there are suggestions in Proclus' *Commentary on the First Book of Euclid's Elements* that they did use the same procedure for calculating area. I shall return to this topic, in a discussion of Egyptian land surveys and the stories about the development of Greek mathematics, in Section 8.1.

## 7.2 TABLES AND READY RECKONERS

### 7.2(a) *Division tables*

The practice of expressing fractions as sums of parts raises two distinct questions: how were these expressions manipulated, and how were they calculated? I shall deal mainly with the first question,[32] and will here describe

ancient arithmetical calculations are due to the nature of the scripts used and how far to the comparative indifference of the Greeks and Romans to extreme precision in such matters as the calculation of interest on loans." Indeed, extreme precision is unattainable within the conventions of commercial practice; see the next paragraph.

[31] For more details and references, see Fowler, NFA.

[32] My attitude will be analogous to that taken by most users, even most sophisticated mathematical users, of mathematical tables and calculating machines. Few people pause to think how the numbers they use were actually calculated. And this information is usually far from simple; I suspect that any trick or intuition that shows hope of being useful will be exploited, then as now.

some of the extensive published primary evidence most, though perhaps not all, of which derives from school and commercial texts.

A preliminary grammatical observation is that the definite article plays a more important role in Greek than in many modern languages, and the sequences of the *arithmoi* and *merē* should really be thought of as

the duet, the trio, the quartet, the quintet, ...

and

the half, the third, the quarter, the fifth, ....

Since Greek is an inflected language, it is often through the case that the relationship of one word to another is expressed; and since the higher numbers, after four, are indeclinable, it is the article which will carry this information. So a phrase like 'the quintet' (nominative plural) must be 'οἱ ε', 'αἱ ε', or 'τὰ ε' according as the implied noun is masculine, feminine, or neuter; and this implied presence enhances the concreteness of the number system. Also, while the *arithmoi* take plural articles, as in this illustration, the *merē* take neuter singular articles, for example τὸ ε [μέρος], the 5th part.

Now consider, for example, our common fraction $\frac{12}{17}$. This would be expressed in a division table (such as will be described below[33]) as

$$\begin{array}{ccc} τῶν & ιβ & [τὸ ιζ] \quad\quad ∠\,ἰβιζλδναξη \\ \text{of the} & 12 & [\text{the 17th is}] \quad 2\,12\,17\,34\,51\,68 \end{array}$$

for what we would write as

$$\frac{12}{17} = \frac{1}{2} + \frac{1}{12} + \frac{1}{17} + \frac{1}{34} + \frac{1}{51} + \frac{1}{68}.$$

These expressions are far from unique (we could here have, among infinitely many others, 2′ 6′ 34′ 102′, or 3′ 4′ 12′ 51′ 68′ 204′) and they may be written out using different arrangements of diacritical marks or spaces in the original text, or they may be presented differently by a modern editor, with the more convenient and understandable format ∠ ιβ′ ιζ′ λδ′ να′ ξη′ now being the most common. Our ancient evidence clearly indicates that most such expressions would be found in tables that would have been either memorised or consulted.

I shall call these tables 'division tables', rather than their more usual but misleading names 'tables of fractions' or 'multiplication tables'. Many of them have been found and published. An excerpt from a typical table reads:

$$\begin{array}{ll} ἔνατα & \text{ninths} \\ τῆς\,α\,τὸ\,θ\,θ & \text{of the 1 the 9 [is] 9} \\ τὸ\,θ\,χξϛβ & \text{[of the 6000] the 9 [is] 666 3} \end{array}$$

---

[33] This example comes from P. Mich. iii 146, no. 20 in my catalogue of division tables, in the Appendix. Also see the following n. 34.

| | |
|---|---|
| τῶν β ϛίη | of the 2 [the $\acute{9}$ is] $\acute{6}$ ĩ8 |
| τῶν γ γ, | of the 3 [the $\acute{9}$ is] 3⸱ |
| τῶν δ γθ | of the 4 [the $\acute{9}$ is] 3̈ 9̈ |
| τῶν ε ∠ίη | of the 5 [the $\acute{9}$ is] 2̈ ĩ8 |
| τῶν ϛ β, | of the 6 [the $\acute{9}$ is] 3̈ |
| τῶν ζ βθ | of the 7 [the $\acute{9}$ is] 3̈ 9̈ |
| τῶν η ∠γίη | of the 8 [the $\acute{9}$ is] 2̈ 3̈ ĩ8 |
| τῶν θ α, | of the 9 [the $\acute{9}$ is] 1 |
| τῶν ι αθ | of the 10 [the $\acute{9}$ is] 1 9̈ |
| τῶν κ βϛίη | of the 20 [the $\acute{9}$ is] 2 6̈ ĩ8 |
| and so on up to | |
| τῶν ϙ ι, | of the 90 [the $\acute{9}$ is] 10 |
| τῶν ρ ιαθ | of the 100 [the $\acute{9}$ is] 11 9̈ |
| τῶν ϲ κβϛίη | of the 200 [the $\acute{9}$ is] 22 6̈ ĩ8 |
| and so on up to | |
| τῶν Τ ρ, | of the 900 [the $\acute{9}$ is] 100 |
| τῶν Α ριαθ | of the 1000 [the $\acute{9}$ is] 111 9̈ |
| τῶν Β ϲκβϛίη | of the 2000 [the $\acute{9}$ is] 222 6̈ ĩ8 |
| and so on up to | |
| τῶν Θ Α | of the 9000 [the $\acute{9}$ is] 1000 |
| τῶν $\overset{α}{μ}$ Α ριαθ | of the 1 0000 [the $\acute{9}$ is] 1111 9̈ |

A catalogue of published tables is given in the Appendix to this chapter, and hereafter I shall refer to the tables by their serial numbers in this catalogue; this example comes from no. 20, except that I have standardised and corrected the use of the fraction indicators.[34]

This table is in what I shall call 'full' format, which lists expressions up to 10 000 *n*th parts; some tables stop at the entry 'of the *n* [the *ń* is] 1', in what I shall call 'abbreviated' format. This example starts with the 'heading' *enata*, ninths, and it contains an 'initial entry' "[of the 6000] the $\acute{9}$ is 666 3̈" of a kind that is believed to have been used in financial calculations and to date from the time (up to the third century AD at the latest, and possibly earlier) when 6000 drachmae, equalling 1 talent, was the monetary standard. Although inflation and change from a 'gold', then a 'silver', standard rendered this entry obsolete, it continued to be copied into some of the tables for another five hundred years at least. There is also some variation in the wording of the tables. For example, the first line may be written in words:

$$\tau\hat{\eta}\varsigma\ \mu\iota\hat{\alpha}\varsigma\ \tau\grave{o}\ \acute{\theta}\ \acute{\theta},$$

---

[34] There is a photograph of this part of this table in Knorr, TFAEG, 145. The scribe has put fraction indicators on all letters occurring in the answers, irrespective of whether they are *arithmoi* or *merē*.

and the initial entry may be found as either the first or second line, and it may appear as:

$$\tau\grave{o} \ \acute{\theta} \ \mathring{\alpha}\rho\iota\theta\mu\hat{\omega}\nu \ \chi\xi\varsigma\acute{\beta},$$

as:

$$\tau\grave{o} \ \acute{\theta} \ \psi\acute{\eta}\phi\omega\nu \ \chi\xi\varsigma\acute{\beta}$$

or as:

$$\tau\grave{o} \ \acute{\theta} \ \mathring{\epsilon}\xi\alpha\kappa\iota\varsigma\chi\iota\lambda\acute{\iota}\omega\nu \ \chi\xi\varsigma\acute{\beta}.^{35}$$

The actual expressions themselves vary·from table to table; for example the Egyptian Rhind Mathematical Papyrus, no. 2, copied around 1650 BC from an older archetype, contains a table of tenths, in an extremely concise format, laid out exactly as in the following translation:[36]

| | |
|---|---|
| 10′ | 3″ 30′ |
| 5′ | 3″ 10′ 30′ |
| 5′ 10′ | 3″ 5′ 30′ |
| 3′ 15′ | |
| 2′ | |
| 2′ 10′ | |

while the other two published tables of tenths are as follows[37] (I have again listed the example from the Rhind Mathematical Papyrus for comparison):

| | Rhind papyrus (no. 2) | P. Mich. 146 (no. 20) | Achmîm papyrus (no. 12) |
|---|---|---|---|
| 1 | 10′ | 10′ | 10′ |
| 2 | 5′ | 5′ | 5′ |
| 3 | 5′ 10′ | 5′ 10′ | 4′ 20′ |
| 4 | 3′ 15′ | 3′ 15′ | 3′ 15′ |
| 5 | 2′ | 2′ | 2′ |
| 6 | 2′ 10′ | 2′ 10′ | 2′ 10′ |
| 7 | 3″ 30′ | 2′ 5′ | 2′ 5′ |
| 8 | 3″ 10′ 30′ | 3″ 10′ 30′ | 2′ 4′ 20′ |
| 9 | 3″ 5′ 30′ | 2′ 3′ 15′ | 2′ 3′ 15′ |

The disagreements between these tables seem to indicate that we cannot easily argue for a continuous textual tradition: the tables must have been frequently recomputed, when occasion demanded.

[35] There are discussions of these features by Brashear, Crawford, and Sijpesteijn in the references to nos. 11, 24, 25, and 39, of the catalogue.

[36] See the facsimile and transcript in Plate 33 of Chace, *RMP*. Unfortunately the photograph of this part of the text was not reproduced in the abridged reprint of this book.

[37] The tables of tenths in nos. 6, 7, and 14 of the catalogue are incomplete or not transcribed.

These tables are very common. Some examples, like nos. 7, 12, and 20, are systematic and extensive ready reckoners; others, like no. 38, are less apparently coherent collections; many are isolated single examples. Some seem crude, in the sense that their entries give more cumbersome expressions, by any criterion, than are necessary (see, for example, no. 9; several entries of this table need correction and can be simplified, for instance 'of 6 the 31′ is 12′ 20′ 31′ ⟨62′⟩ 155′ 186″' can also be expressed 'of 6 the 31′ is 6′ 62′ 93″'). Many of these texts are school exercises. In all, the Appendix contains details of 97 publications containing 172 individual division tables, distributed as follows:

| 3″ | 2′ | 3′ | 4′ | 5′ | 6′ | 7′ | 8′ | 9′ | 10′ | 11′ | 12′ | 13′ | 14′ | 15′ |
|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|
| 14 | 14 | 12 | 12 | 5  | 4  | 12 | 8  | 9  | 6   | 7   | 7   | 5   | 7   | 7   |

| 16′ | 17′ | 18′ | 19′ | 20′ | 23′ | 24′ | 25′ | 29′ | 30′ | 31′ | 48′ | 49′ | 90′ | 150′ |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| 6   | 9   | 4   | 5   | 2   | 1   | 5   | 2   | 1   | 2   | 1   | 2   | 1   | 1   | 1    |

Any discussion of fractional calculations in Greek antiquity should take place against a background of awareness of this published material.

### 7.2(b) *Multiplication and addition tables*

One feature of the Greek alphabetic system is that the tables for multiplication by 2, 3, 4, . . . , 9, by 20, 30, 40, . . . , 90, and by 200, 300, 400, . . . , 900 are completely different so this makes a complete multiplication table to be very long.[38] Broadly speaking, three different formats are found for these tables; the first, which resembles the format of a division table, is to set out a table of 1-times, 2-times, . . . , 9-times, then sometimes continuing with 10-times, 20-times, . . . , 90-times, 100-times, 200-times, . . . , 900-times, 1000-times, 2000-times, . . . , 9000-times, 1 0000-times. A typical table, for example the 80-times, is then given as:

$$\pi\alpha \; \pi \qquad 80 \; 1 \; 80$$
$$\pi\beta \; \rho\xi \qquad 80 \; 2 \; 160$$
$$\pi\gamma \; \varsigma\mu \qquad 80 \; 3 \; 240$$
$$\text{etc.}$$

---

[38] Many discussions of the consequences of this lack of place-value in Greek arithmetic have no basis in our knowledge, now, of Greek numerical practice. I recommend Smyly, EAGL (see n. 10, above) for an informed opinion: "I am convinced that the [Greek] notation, though it may be inferior in some respects to that now in use, by no means deserves the unlimited contempt which has been heaped upon it" (p. 515). Also de Ste. Croix, GRA, 55–6: "The habitual arrangement of figures in columns in our notation is not an intrinsic virtue of that notation but, on the contrary, an incidental defect, due, somewhat paradoxically, to the combination of its two greatest virtues, namely place-value and the small number of symbols it employs; but . . . this very defect (the necessity for the arrangement of figures in columns) has, paradoxically again, provided a very useful stimulus towards the evolution of the advanced concepts of debit and credit (positive and negative entries) . . . which was an essential preliminary to a coordinated system of book-keeping by double or even single entry"—and so eventually, one might add, perhaps to negative numbers.

In the second format, each entry is commuted and repeated. The following illustration of the 8-times table suggests how such a table might have been read:

| ηα η | *oktakis mia estin oktō* | eight-times one is eight |
| αη η | *hapax oktō estin oktō* | once eight is eight |
| ηβ ιϛ | *oktakis duo estin hekkaideka* | eight-times two is sixteen |
| βη ιϛ | *dis oktō estin hekkaideka* | twice eight is sixteen |
| etc. | | |

(We have some instances of such multiplications described in Socrates' encounter with the slaveboy, *Meno*, 81e–85d; for example, at 87d, we find both *dis duoin* and *duoin dis* in consecutive speeches.) The need to commute each entry surely arises not from ignorance of the commutativity of multiplication, but from the need to recognise automatically the different sounds of the cardinal and repetition numbers.

The third format (which occurs in conjunction with the second format in one of our tables, no. 51) has the following form:

| ηα | η | 8 | 1 | 8 |
| ηι | π | 8 | 10 | 80 |
| ηρ | ω | 8 | 100 | 800 |
| ηα | ͵ω | 8 | 1000 | 8000 |
| ηβ | ιϛ | 8 | 2 | 16 |
| ηκ | ρξ | 8 | 20 | 160 |

etc., up to multiplication by 9, 90, 900, and 9000.

Aristotle makes a reference to multiplication tables at *Topics* 163b24–29:

> For just as in geometry it is useful to have been trained in the elements, and in arithmetic to have a ready knowledge of the multiplication table up to ten times helps much to the recognition of other numbers which are the result of multiplication, so too in arguments it is important to be prompt about first principles and to know your premisses by heart.

From this we can infer that we cannot immediately dismiss the evidence I have been describing as being relevant only to the training of scribes and accountants; Aristotle's audience clearly is aware of the basics of this kind of material, though he implies that they may not be fluent in its practice.

Tables of additions, in the same kinds of formats as multiplication tables, are also found and are listed in the Appendix to this chapter. Most of these examples of multiplication and addition tables seem to be more in the nature of systematic school exercises rather than ready reckoners intended for consultation.

### 7.2(c) *Tables of squares*

In contrast to our evidence about Babylonian mathematics, relatively few examples of other kinds of mathematical tables have been found. This may be a

consequence of the random way in which the papyrus texts have been preserved; on this we can only speculate, but such speculation should be moderated by the fact that a substantial amount of other material, for example astronomical[39] and astrological data, has been preserved on papyrus. Setting aside very fragmentary texts and marginal annotations whose interpretation is often very doubtful, the only examples known to me of systematic tables other than addition, multiplication, division, and astronomical tables are the five tables of squares listed in the Appendix to this chapter, of which no less than three are early, dating from the third to first centuries BC. This is insufficient evidence to support anything more than the most tentative conclusions, so I shall do no more than remark on the importance of squares within the mathematical reconstructions in Part One of this book.

## 7.3 A SELECTION OF TEXTS

### 7.3(a) *Archimedes' Measurement of a Circle*

Consider, now, what is apparently the most blatant piece of evidence against my thesis that common fractions are not found in early Greek mathematics: Archimedes' *Measurement of a Circle*.[40] A typical description of part of this short treatise (Heath, *HGM* ii, 50 ff.) is as follows:

[Proposition 3 proves] that the ratio of the circumference of any circle to its diameter (i.e. $\pi$) is less than $3\frac{1}{7}$ but greater than $3\frac{10}{71}$. ... The circulation starts from a greater and a lesser limit to the value of $\sqrt{3}$ which Archimedes assumes without remark to be known, namely

$$\tfrac{265}{153} < \sqrt{3} < \tfrac{1351}{780}.$$

Let us contrast this description with our received text. Take first the enunciation of the proposition; here, with a word-for-word literal translation, we have:

| παντὸς | κύκλου | ἡ | περίμετρος | τῆς | διαμέτρου | τριπλασίων |
|--------|--------|-----|-----------|-----|-----------|------------|
| Of every | circle | the | perimeter | the | diameter | three-times |

| ἐcτὶ | καὶ | ἔτι | ὑπερέχει | ἐλάccονι μὲν[41] ἢ | ἑβδόμῳ | μέρει |
|------|-----|-----|----------|-------------------|---------|-------|
| is | and | further | it exceeds | by less than | a seventh | part |

| τῆc | διαμέτρου, | μείζονι δὲ ἢ | δέκα | ἑβδομηκοcτομόνοιc |
|-----|-----------|-------------|------|-------------------|
| of the | diameter, | but by more than | ten | seventy-first [parts] |

---

[39] For astronomical material, see Neugebauer, *APO*, the inspiration for my catalogue in the Appendix.

[40] The critical edition of the text is in Archimedes–Heiberg, *Opera* i, 232–43, with the commentary from the early sixth century AD by Eutocius in *Opera* iii, 228–61. The Greek text, without critical apparatus, is reproduced in Thomas, *SIHGM* i, 316–33. The medieval Latin translations are in Clagett, *AMA*; see especially vol. ii. The transmission of the text has been studied in minute detail by Knorr, *TSAMG*, 421–816. Also see n.44 below. On the chronological problems associated with Eutocius, see the *DSB*, s.v. Eutocius.

[41] The Greek expression μέν ... δέ ... merely requires the translation 'but' later in the sentence.

and, in the English word-order

The perimeter of every circle is three-times the diameter and further it exceeds by [a line] less than a seventh part of the diameter but greater than ten seventy-first [parts of the diameter].

Archimedes clearly does *not* here consider the ratio of circumference to diameter as a numerical quantity. Also the phrase 'ten seventy-first parts' seems to refer to the line obtained by concatenating ten copies of the seventy-first part of the diameter, rather than the common fraction $\frac{10}{71}$; it is a *line*, not a number. Moreover, the ratio of 10 to 71 is capable of a far wider range of interpretations than as a common fraction, as has already been illustrated many times in this book. In particular, see Sections 2.2(b) and 4(d), where some mathematical features of Archimedes' calculation are analysed; here I shall consider purely textual issues.

Next, Archimedes' treatment of the lower bound to $\sqrt{3} : 1$ reads, in translation:

Let there be a circle with diameter $A\Gamma$ and centre $E$, and let $\Gamma AZ$ be a tangent and the angle $ZE\Gamma$ one third of a right angle. Then $EZ$ has to $Z\Gamma$ the ratio that 306 has to 153, and $E\Gamma$ has to $\Gamma Z$ the ratio [greater than] that which 265 has to 153.

Here we have ratios, not fractions.

Now turn to the proof of the proposition. One feature can be seen immediately by consulting the admirable translation in Thomas, *SIHGM* i, 316–33.[42] Look at the parallel Greek and English texts on pp. 324 and 325: six lines of Greek text about ratios give rise to more than twenty-four lines of arithmetical manipulations. Most of this translation is set in square brackets; it does not come from Archimedes' text (Archimedes–Heiberg, *Opera* i, 232–43) but is freely adapted from Eutocius' long commentary (ibid. iii, 228–61), and it is these additions that make the translation intelligible to us. Now look in detail at a typical passage in this translation, *SIHGM* i, 329.19–28 = *Opera* i, 242.2–5. With the interleaved material from Eutocius and elsewhere deleted, we read:

$$A\Theta : \Theta\Gamma < 5924\tfrac{3}{4} : 780$$

$$< \tfrac{4}{13} \cdot 5924\tfrac{3}{4} : \tfrac{4}{13} \cdot 780$$

$$< 1823 : 240.$$

Therefore $A\Gamma : \Gamma\Theta < 1838\tfrac{9}{11} : 240$.

---

[42] The following corrections and minor amendments to the translation should be noted:
p. 329  line 11. For '$=1351$' read '$<1351$'.
line 13. For 'Hence' read '[Hence'.
line 23. On '$\tfrac{4}{13} \cdot 5924\tfrac{3}{4} : \tfrac{4}{13} \cdot 780$': this is not really a translation, but a very helpful editorial resolution of a complicated piece of text, to be quoted below.
p. 331  line 3. For '24' read '240'.
line 6. This line '$\tfrac{11}{40} \cdot 3661\tfrac{9}{11} : \tfrac{11}{40} \cdot 240$' is another resolution of even more complicated text.

The corresponding Greek text, with a literal translation, is as follows:

1    ἡ        ΑΘ ἄρα διὰ τὰ αὐτὰ                πρὸς
     The [line] ΑΘ then for the same [reason] in relation to

2    τὴν      ΘΓ ἐλάσσονα λόγον ἔχει ἢ          ὃν      ͵ε ͵ΣκδΔ ∠ δ'
     the [line] ΘΓ a less    ratio  has  than that [ratio] which 5924 2' 4'

3         πρὸς      ͞ψ͞π ἢ               ὃν        ͵αωκγ
     [has] in relation to 780 or [less than] that [ratio] which 1823

4         πρὸς     ͞ς͞μ·      ἑκατέρα     γὰρ  ἑκατέρας
     [has] in relation to 240, each one of the two for[43] of the other [line]

5    ͞δ ͞ιγ' ὥστε   ἡ       ΑΓ    πρὸς   τὴν ΓΘ
     [is] 4 13', so that the [line] ΑΓ in relation to the ΓΘ [is less]

6    ἢ           ὃν        ͵αωλη ͞θ ͞ια'       πρὸς      ͞ς͞μ
     than that [ratio] which 1838 9 11' [has] in relation to 240.

This text can be pursued back to the manuscripts using the apparatus of the critical editions by Heiberg and Clagett.[44] All of the surviving Greek texts and Latin translations of Archimedes, except for those found in the palimpsest rediscovered by Heiberg in 1906, to be described briefly below, derive from an archetype compiled in the ninth or tenth century which may have belonged to Leo the Philosopher[45] and which seems to have contained, in addition to Archimedes' works, the commentary on Archimedes by Eutocius and a book by Heron. This codex disappeared in the sixteenth century, but its contents have been reconstructed by Heiberg from four surviving Greek copies made from it between 1450 and 1564; this reconstructed manuscript is Heiberg's and Clagett's manuscript A.

The autograph copy survives of a careful literal Latin translation of parts of A made in 1269 by the Flemish Dominican monk, William of Moerbeke; it contains the translation of all of the Archimedean books in A except *The Sandreckoner*, and all of Eutocius' commentaries except for that on *Measurement of a Circle*;

---

[43] This particle always comes second in a Greek phrase and first in English: 'for each one of the two [lines] is 4̄ 13′ of the [corresponding] other line'.

[44] The critical apparatus of Archimedes–Heiberg, *Opera*, must be collated with that in Clagett, *AMA* ii₃, 397–9. The textual details in Archimedes–Heath, *WA*, should be ignored since they are based on Heiberg's first edition of 1880–1; there is a summary based on the second edition in Heath, *HGM* ii, 25–7 and a further account in Clagett, *AMA* ii₁, 54–78. For details of Codex A up to its disappearance some time after 1584, see Rosie, *FHCAA*.

[45] So Archimedes–Heiberg, *Opera* iii, pp. x ff. and xxii f., Heath, *HGM* ii, 25 f., and Wilson, *SB*, 83; but Clagett, *AMA* ii₁, 55 is non-committal. Clagett gives details of its probable contents: Archimedes, *Sphere and Cylinder* I and II, *Measurement of a Circle, Conoids and Spheroids, Spiral Lines, Plane Equilibria* I and II, *Sandreckoner, Quadrature of a Parabola*; Eutocius, *Commentaries on Sphere and Cylinder, Measurement of Circle, Plane Equilibria*; and Heron, *On Measures.*

this is manuscript B.[46] The translation was corrected by Moerbeke himself in passages, distinguished by a blacker ink, that I shall refer to as $B^1$. Then, in the fifteenth or sixteenth century, an unidentified hand added further annotations, though none of these are found in *Measurement of a Circle*. Finally the manuscript was acquired by one Andreas Coner in 1508, who added many more, often perceptive, annotations[47] which will here be denoted by $B^2$. It appears that Moerbeke found difficulty in translating the numbers in *Measurement of a Circle*: sometimes he wrote the corresponding Greek text in the margin, leaving a gap in the translation and then, as $B^1$, he subsequently inserted the missing text.[48] The marginal Greek annotation was then erased, but these and other erased portions can now be read under ultraviolet light and they give more information about the lost archetype, manuscript A. However, since some of the numbers in A were already wrong or meaningless, this procedure did not ensure an arithmetically correct result, and Coner has clearly worked through the manuscript, correcting and amplifying it further. Clagett (*AMA* ii$_1$, 55 f.) proposes that Moerbeke may not have translated Eutocius' commentary on *Measurement of a Circle* and *The Sandreckoner* because of his difficulties with Greek numerical practice.

The third Archimedes manuscript, a palimpsest in rather poor condition rediscovered by Heiberg in Constantinople in 1906, is a tenth-century copy that has subsequently been erased to make way for a theological text of the twelfth or thirteenth century.[49] Fortunately the original text is still legible, and Heiberg was able to decipher most of it with the aid of a magnifying glass. He recovered, in decreasing order of completeness, our second Greek version of the extant *Sphere and Cylinder* I and II, *Spiral Lines*, *Measurement of a Circle*, and *Plane Equilibria*; the first Greek texts of *Floating Bodies*, hitherto known only in a Latin translation made by Moerbeke from another Greek manuscript,[50] now lost; the preface and two propositions of the *Stomachion*, which deals with a kind of puzzle; and, most valuable of all, the new and unexpected *Method*. This

---

[46] Clagett calls this manuscript O.

[47] Heiberg dated these annotations to the fifteenth century (*Opera* i, pp. v–vi), although he described Coner's acquisition of the manuscript in 1508 and his subsequent annotations (*Opera* iii, pp. xliii f.). Also Heiberg apparently frequently confused the corrections in the second and third hands; see Clagett, *AMA* ii$_1$, 63 n. 7, 64 n. 8, 68, 73; and ii$_3$, *passim*.

[48] Since the main difference between B and $B^1$ lies in the colour of the ink, I presume that it must sometimes be difficult to distinguish between them. In any case, the distinction between them is not very great since they are both assigned to Moerbeke, and it is significant to us here only as an indication of his difficulties with Archimedes' text.

[49] This manuscript is described and its text of the *Method* is edited in Heiberg, ENA; also see Archimedes–Heiberg, *Opera* i, pp. v–vi and iii, pp. lxxxv–xc, and the supplement to Archimedes–Heath, *WA*, after p. 326. It was formerly in the Library of the Metochion of the Holy Sepulchre, Istanbul but now is in a private collection, except for one leaf, identified by N. G. Wilson, which has strayed to Cambridge University Library, Add. 1879.23; see Wilson, *SB*, 139.

[50] Heiberg calls this lost manuscript ß, Clagett calls it B. It is described in Archimedes–Heiberg, *Opera* iii, pp. liv–lvii and Clagett, *AMA* ii$_1$, 58–60. It was last mentioned in a catalogue of 1311.

palimpsest, which is not directly related to A, is called manuscript C; and those parts of it where Heiberg could detect the presence of text that he was unable to read are identified as (C).

The following variations in the passage above are found in the different manuscripts.

line 2:        $\overline{{}_{,\epsilon\lambda\kappa\delta}}$ $\angle$ $\delta'$

    A had     $_{,\epsilon\tau\kappa\delta}$ $\epsilon'$ $\delta'$

    B[1] has     5324 $\bar{3}$ $\bar{4}$, in an unsuccessful attempt to correct the irregular[51] $\epsilon'$

    B[2] has the correct[52] 5924 $\frac{1}{2}$ $\frac{1}{4}$

    C has     $_{,\epsilon\rho\kappa\delta}$

    Eutocius' commentary has the correct $\overline{{}_{,\epsilon\lambda\kappa\delta}}$ $\angle$ $\delta'$

line 4:        $\overline{\varsigma\mu}$

    A had     $\overline{\varsigma\nu}$

    B has     250

    B[2] has     240

    C has     $\overline{\varsigma\mu}$

    Eutocius has $\overline{\varsigma\mu}$

line 5:        $\bar{\delta}$ $\iota\gamma'$

    A had     $\bar{\delta}$ $\iota\gamma'$ $\alpha'$ according to Heiberg's reconstruction but Moerbeke transcribed the Greek $\bar{\varDelta}$ $\iota\gamma'$ $\alpha'$, now erased, in the margin of B

    B is erased, and is presumably illegible

    B[2] has $\frac{4}{13}$ in the erasure

    (C) is illegible, but the space for $\bar{\delta}$ [or $\bar{\varDelta}$?] $\iota\gamma'$ $\alpha'$[53]

    Eutocius has $\bar{\delta}$ $\iota\gamma'$

line 6:        $\bar{\theta}$ $\iota\alpha'$

    A omitted the $\iota\alpha$

    B has $\theta'$, erased, in the margin

    B[1] has $\bar{9}$, followed by a gap

    B[2] has the correct $\frac{9}{11}$

    (C) has no space for the $\iota\alpha$, and

    Eutocius has $\bar{\theta}$ $\iota\alpha'$, but he interprets it[54] as 9' 11'

These variations illustrate how those parts of the passage that seemed to suggest some kind of manipulation of fractions have all but disappeared as we trace the text back through the surviving manuscripts. For example, the fraction $\frac{4}{13}$ on line 5 exists in the meaningless form $\bar{\delta}$ or $\bar{\varDelta}$ (?) $\iota\gamma'$ $\alpha'$ in A; it is illegible in C; and it is only found explicitly and correctly in a sixteenth-century addition to the Latin translation B; also the operation of multiplying by $\frac{4}{13}$ is

---

[51] Recall that the conventional way of writing this would be $\overset{\prime\prime}{\delta}\overset{\prime}{\epsilon}$ (4' 5'), in decreasing order of size of the parts.

[52] The word 'correct', here and elsewhere, refers of course to the generally accepted interpretation of the passage.

[53] The details of C are understandably sometimes left ambiguous by Heiberg.

[54] Eutocius' commentary gives a long, long multiplication for (1838 9' 11')², evaluated using a mixture of fractional notations, whose answer, and with it the inequality he is purportedly verifying, is fudged. See Archimedes–Heiberg, *Opera* iii, 252–3.

described in Greek by a phrase '*hekatera gar hekateras* . . .' which is capable of a range of alternative interpretations. However, some evidence does survive from this examination: if Eutocius interpreted $\overline{,a\omega\lambda\eta}\ \overline{\theta}\ \iota a'$ as 1839 9′ 11′, then this implies a misinterpretation of some such notation that he found in his manuscript; and there is also the corrupted evidence of a similar kind of notation on line 5.[55]

An even greater complication, which I shall not describe here, pervades the next passage in the text, whose translation appears in Thomas, *SIHGM* i, 331.2–7, as

$$AK : K\Gamma < \tfrac{11}{40} \cdot 3661 \tfrac{9}{11} : \tfrac{11}{40} \cdot 240 < 1007 : 66,$$

and I have suggested, in note 42, that the expression in the middle of this inequality should really be placed in square brackets as an editorial interpretation.

These variant readings and later additions and corrections indicate a range of editorial and scribal interferences that have occurred since the now lost manuscript A was compiled. We have no way of assessing the extent of corruption already introduced in the transmission from Archimedes' own version up to manuscript A, but other features indicate that this could be considerable. For example, the order of the propositions is now clearly incorrect: the first two propositions were unnumbered in A, but the second of these uses the result proved in the third proposition, which is numbered $\gamma' = 3$; and the second half of the proof of this third proposition is erroneously numbered $\delta' = 4$.[56] Then Heron refers, at *Metrica* I 37, to a further proposition on the relationship between the area of a sector of a circle and the arc it subtends, and he says that it is in his version of the text; such a result is now only to be found in a translation from Arabic into Latin made by Gerard of Cremona.[57] Also Archimedes wrote in a Doric dialect, and some of his works still preserve this feature, but all traces of dialect have been purged from the *Measurement of a Circle*. Further, the treatise has no preface or introductory postulates, unlike almost all of Archimedes' other works.

Heiberg's careful assessment of the text is: "In general, in the whole of this short work, the type of language suffers from so careless a brevity that one recognises the hand of an excerptor rather than that of Archimedes" (Archimedes–Heiberg, *Opera* i, 233 n. 3), an opinion that does not presuppose that the original text was meant as an elementary exercise, intended for popular consumption. The bolder common opinion, that Archimedes' *original* text was meant for popular consumption, I find hard to accept. Note the subtleties of Proposition 1, which launches immediately into a very summary description of an exhaustion argument of the type found in *Elements* XII 2, based on the

[55] The reader is also encouraged to explore the textual details of the three references to $\tfrac{10}{71}$ at the conclusion of the proposition.

[56] See Archimedes–Heiberg, *Opera* i, variant readings to 232.1, 234.18, 236.7, and 140.12.

[57] See Claggett, *AMA* ii$_1$, 5, 32, 47, 57.

bisection arguments of X 1; and note also the difficulties that Moerbeke, a translator of enormous experience with Greek philosóphical and scientific works, had with his translation of Proposition 3. Our easy understanding, today, depends on the clarifications, commentaries, explanations, and corrections made by Eutocius, Coner, and other editors and inserted into the received text.[58] Now observe how all of these corrections have been made from within an arithmetised tradition of mathematics. This is very strikingly illustrated by Eutocius' long commentary, a detailed verification of many of the steps, executed throughout in compound parts, though occasionally expressed in the manuscripts in an abbreviated notation that resembles common fractions. I suggest that Archimedes' original intention may have been different; even that the work may originally have been the first steps in the anthyphairetic exploration of the ratio of circumference to diameter of a circle. And I tried to show, in Sections 2.2(b) and 2.4(d), how some of the mathematical features of the calculation which cause problems when interpreted arithmetically fit very naturally within an anthyphairetic context.

### 7.3(b) *Aristarchus' On the Sizes and Distances of the Sun and Moon*

Aristarchus' calculation in his *On the Sizes and Distances of the Sun and Moon* has survived in a less disturbed state than Archimedes' calculation. As usual, all of our sources are minuscule manuscripts and, of these, the oldest and best, and perhaps the ultimate source of all the other surviving manuscripts, is the beautiful ninth-century Vatican gr. 204, described briefly in Section 6.3. A Latin translation by Valla was published in 1488, another by Commandinus in 1572, and the Greek text was first published by John Wallis in 1688; it is not quite clear which manuscript they used. There is a modern critical edition in Aristarchus–Heath, *AS*, with a very useful introduction and commentary, to which I refer the reader. The numbers in the different manuscripts seem to be free of corruption, though there are minor variations in the way they are expressed.

　　We now find it natural and obvious to express Aristarchus' calculation in terms of manipulations of common fractions, but Aristarchus actually employs two distinct procedures:

　　(i) To begin with, in Proposition 4 and part of 7, simple parts are expressed as such. Sometimes they are written out in words, as in Hypothesis 6:

That the moon subtends one fifteenth part ($\pi\epsilon\nu\tau\epsilon\kappa\alpha\iota\delta\epsilon\kappa\alpha\tau\sigma\nu$ $\mu\epsilon\rho\sigma\varsigma$) of the zodiac,

sometimes they are expressed as numerals, as in Proposition 4:

But a 15′ [part] ($\iota\epsilon'$ [$\mu\epsilon\rho\sigma\varsigma$]) of a sign is a 180′ ($\rho\pi'$) [part] of the whole circle of the zodiac.

---

[58] I recommend any reader who remains unconvinced to delete all editorial additions, to replace the numbers in the text by those in A, and then to try to understand the text without using any modern notations and techniques. Also Clagett's comment (*AMA* ii₁, 51) that "Coner made significant corrections of these numbers, no doubt by consulting the *Commentary* of Eutocius" should be expanded slightly since, as we have seen, Coner also made corrections that are not found, or are incorrect, in Eutocius' commentary.

(ii) Compound parts are never used, with two very minor exceptions to be described below. Instead anything that is not a simple part is, right from the start, expressed as a ratio. For example:

The diameter of the sun has to the diameter of the earth a ratio greater than that which 19 has to 3, but less than that which 43 has to 6 [Introduction and Proposition 15].

When this language of ratios is in use, it is not mixed with the language of multiples or parts, even in the most obvious cases. See, for example, Proposition 7:

But it was proved that GE has to EH a ratio greater than that which 15 has to 2; therefore, *ex aequali*, FE has to EH a ratio greater than that which 36 has to 2, that is, than that which 18 has to 1.

Also see the two passages quoted in Section 2.4(c).

We appear to have two distinct approaches: the notation of simple parts, which is treated as an extension of the *arithmoi*, and a notion of ratio, which is never confused with any kind of numerical quantity. While there is no difficulty in translating those passages that are expressed in simple parts, style i, into ratios, style ii, the translation in the opposite direction is less evident. So the simple parts seem here to be playing the role of an informal shorthand, not a developed language that can be used freely or manipulated.

The first exception to this description is found, isolated, in Proposition 7:

... the first angle GBE is a fourth part (τέταρτον μέρος) of a right angle. But the angle DBE is a 30th part (λ′ μέρος) of a right angle; therefore the ratio of the angle GBE to the angle DBE is that which 15 has to two (τὰ ιε πρὸς τὰ δύο): for, if a right angle be regarded as divided into 60 (ξ) equal [parts], the angle GBE contains 15 (ιε) of such [parts], and the angle DBE contains two (δύο) [parts].

The language is curiously tentative ("be regarded as divided"), and seems to indicate that though the operation of taking the ratio and simplifying it mildly is not a standard manipulation, yet it is not difficult to fit within a formally correct procedure.

The second exception is found throughout Proposition 11:

The diameter of the moon is less than the two 45′ (δύο με′ in Vat. gr. 204; δύο τεσσαρακοστόπεμπτα in Wallis), but greater than the 30′ (λ′ in Vat.; τριακοστόν in Wallis), of the distance of the centre of the moon from our eye,

and this quantity, 'two 45″', runs through the first half of the proposition. Again, there is no difficulty in adapting the argument to the alternative language of ratio, at some cost in length.

Despite these exceptions, I do not think that we have the evidence here to justify seeing anything which corresponds to the manipulation of general fractions. And again, as I illustrated in Section 2.4(c), one of the most problematic features of the calculation, the derivation of the inequalities in Propositions 13 and 15, fits very naturally within the context of an anthy-phairetic ratio theory, but it seems to need an anachronistic kind of

explanation in arithmetised mathematics, especially if it is interpreted in terms of inequalities between fractional quantities.

### 7.3(c) *P. Lond. ii* 265 (*p. 257*)

The papyrus, published in Kenyon, Bell, & others, *GPBM* ii, 257 ff. is not typical of the thousands of financial documents that have been found, but it provides a good example for illustrating the style and procedures of a straightforward and repetitive calculation; Plate 9 is a photograph of columns 2 and 3, lines 22–67.[59] The main reason for including it here is that it is sometimes cited as evidence for the Greek use of a notation for common fractions.[60] The papyrus is tentatively assigned, on palaeographical grounds, to the first century AD, and it concerns the conversion between different standards of artabas, a measure of wheat which was also used as a unit for expressing all manner of commercial transactions. This text deals with six different sizes of artabas,[61] called *dromōi, chalkōi, anēlōtikōi, Philippou, Gallou,* and *Hermou,*[62] and it contains twenty-nine examples of conversions of 625 artabas of one standard into another. There are potentially thirty possible different such conversions; three of the conversions in the text are performed in two ways, to illustrate slightly different methods, and part of the text is missing. Other parts of the text are fragmentary but can generally be restored with confidence because of the repetitive nature of the calculations and the stylised verbal formulae describing the operation. In the examples that follow, I shall omit most of the editorial details of restorations and difficult readings, except where they involve the arithmetical calculations or the 'fractional' notation, our main concern here. The scribe has made several errors and approximations, some of which I shall describe. The text contains two standard abbreviations: ō, for artabas; and /, one of the common signs for equality.[63] The names of some of the different artabas, and occasional other words, are also abbreviated.[64]

---

[59] The text starts with a column of demotic, which does not appear to have been published, and which is not mentioned in the edition of the Greek text. Then Greek column 1 contains lines 1–21; 2 22–45; 3 46–67; 4 68–91; 5 92–109; 6 110–30; 7 131–50; and 8 151–62. The verso contains a census return in Greek, also unpublished.

[60] For an influential example, see Heath, *HGM* i, 44, quoted in n. 92, below.

[61] Much as today there are Imperial pints (= 20 fluid ounces: 'A pint of water weighs a pound and a quarter'), the US pint (= 16 fluid ounces; 'A pint's a pound and the world around'), the Scotch pint (= 3 Imperial pints), and other local variants. A suggestive modern translation of artaba might be 'bushel'. Similar conversions can be found in P. Mich iii 145.

[62] The words *dromōi, chalkōi,* and *anēlōtikōi,* all in the dative, are adjectives agreeing with a noun to be understood, probably *metron. Philippou, Gallou,* and *Hermou* are genitives referring either to persons or, less probably, to places. But these grammatical points are unimportant for my argument here.

[63] See n. 28, above.

[64] I am very grateful to T. S. Pattie for examining the original of this papyrus, and for confirming the following corrections to the published text:
Line 40: There is a fraction indicator above the eta of ρκή but no traces over the rho and kappa. But there is a hole over these letters and very little space between this line and the line above. There seems no doubt that the scribe conceived this as ρκή, with fraction indicators.

A typical sum,[65] lines 25–9, with a word-for-word translation, reads:

25  δρομωι      ō    χκε   αι   Γαλλ⟨    ποϲαι
    Dromōi   artabas  625   the   Gallou,   how many

           προϲ     κγ   β́    ⟨ταιϲ  ρ    ō⟩[66]
      at the rate of   23   3    ⟨to the 100 artabas⟩?

26  ωϲ       δει      ποιηϲαι   ποιει   ταϲ      ō      χκε    επι    ταϲ
    As    you should    do.      Do     the    artabas   625    to     the

27   κγ   β́    του    διαφορου   /    μ͡δ͡ψϞα    β̈
     23   3   of the   difference   /    14791     3.

28  τουτων   το    ρ́    /    ρμζ   ∠   γ́   ιβ́    ταυταϲ
    Of these   the   100th   /    147   2   3   12    These

29  προϲθεϲ   ταιϲ     ō      χκε    /    ψοβ   ∠   γ́   μβ̈    ⟨δεδεικται⟩[67]
    add     to the   artabas   625    /    722   2   3   42    ⟨QED⟩

or, in a more liberal translation,

Given 625 dromōi artabas, how many in Gallou will there be, at a rate of 23 3″ per 100 artabas? (I.e. 123 3″ Gallou equals 100 dromōi.) How you should go about it. Multiply 625 by the 23 3″ which represents the difference / 14791 3″. Divide by 100 / 147 2′ 3′ 12′. Add the 625 / 771 2′ 3′ 12′. QED

At six places only in the surviving text (on lines 40, 48, 49, 65, 66, and 139) there is something resembling our notation for common fractions.[68] Let us examine each of these examples, starting with lines 38–40:

38  χαλκωι     ō      χκε    αι    δρομωι    ποϲαι
    Chalkōi   artabas   625   the   dromōi,   how many?

39  ωϲ       δει      ποιηϲαι   επει   εϲτιν   το    διαφορον
    As    you should    do.    Since   it is   the    difference

Line 48: Parts of an alpha with fraction indicator and a theta can clearly be read, so '⟨]θ́ά must be restored.
Line 139: There are no fraction indicators over the first iota alpha, but the second iota has them.
Line 147: The papyrus is broken off after the zeta, so ∠ could easily be restored.
Line 160: Restore [νθ ια
Line 162: Restore [φ α_ι^ιά
All of these corrections have been made to the quotations, below.

[65] This particular example is chosen because it contains more fractional manipulations than others and its text is reproduced on Plate 9.
[66] This phrase is frequently omitted. The conversion rate means that 123⅔ Gallou are equal to 100 dromōi artabas.
[67] The scribe wrote μβ̈ in an evident error for ιβ́, which is given correctly on the previous line. On the previous line he has omitted an accent from the gamma and many such accents are omitted elsewhere; I shall restore them all in all subsequent passages. All but two of the sums finish with dedeiktai, 'it has been shown'.
[68] The 'fractional notation' is also implausibly reconstructed by the editors on line 160, and it should have been reconstructed on line 162. See n. 64, above.

40    προς    κη    ταις    ρ    ō    $\overset{\rho\kappa\acute\eta}{\acute\epsilon\sigma\tau\iota\nu}$  „ ρ    μορια    ο
                                              $\overset{}{\dot{1}\dot{2}\dot{8}}$
      at the rate of  28  to the  100  arbatas,  it is  100  parts  which

41    γινεται  /  ∠ d λβ    λαβε    ουν    το    ∠ d $\overset{\prime\prime}{\lambda\beta}$    των
      become  /  2 4 32.  Take  therefore  the  2 4 32  of the

41/2    ō    χκε  /  υπη  d  $\langle\overset{\prime\prime}{\lambda\beta}\rangle$    τοσαυται    χαλκωι[69]    δεδεικται
      artabas  625  /  488  4  32;    as many  in *dromōi*  QED

Here 128 *chalkōi* artabas are equal to 100 *dromōi*; so the conversion proceeds by multiplying by $\frac{100}{128} = \frac{25}{32}$. But note particularly how, before this operation can be carried out, the fraction has to be converted into a sum of parts, $\frac{100}{128} = 2' \ 4' \ 32'$. This complex operation of taking the $2' \ 4' \ 32'$ part of some quantity then recurs in three subsequent calculations, on lines 71, 102, and 130, where it is treated just like a normal division. The example on line 102 contains an error

101/2  5    ō    χπζ  ∠    τουτων    το    ∠ d $\overset{\prime\prime}{\lambda\beta}$  /  δρομωι    ō    φλς
      artabas  687  2  of these  the  2 4 32  /  *dromōi*  artabas  536

      ∠ d $\overset{\prime\prime}{\iota\varsigma}$ $\overset{\prime\prime}{\lambda\beta}$
      2 4 16 32

where the answer should be 537 16' 32' 64'. These calculations with $2' \ 4' \ 32'$ are straightforward and this kind of manipulation, using only the parts $2', 4', 8', 16', \ldots$, would have been familiar from land measurements; see Section 7.1(d), above.

Now look at the fourteenth sum, lines 63–7. Here the scribe abbreviates the procedure and leaves the answer in a non-standard form:

63  ανηλωτικωι    ō    χκε    αι    χαλκωι    ποσαι
    *Anēlōtikōi*    artabas  625  the    *chalkōi*  how many?

64  ως    δει    ποιησαι    λαβε    το    $\overset{\prime}{\kappa\alpha}$    των    ō    χκε
    As  you should  do.  Take  the  21st  of the  artabas  625

65  / κθ  $\overset{\overset{\prime}{\kappa\alpha}}{\iota\varsigma}$  ταυτα    αφελε    απο    των    ō    χκε
                    $\overset{}{\dot2\dot1}$
    / 29  16.  These  take away  from  the  artabas  625.

66    λοιπαι    φϟε  $\overset{\overset{\prime}{\kappa\alpha}}{[\epsilon]}$    τοσαυται    χαλκωι
                          $\overset{}{\dot2\dot1}$
    There are left  595  5,  as many  in *chalkōi*.

67  δεδεικται
    QED

---

[69] On this line, the scribe has omitted the fraction $\overset{\prime\prime}{\lambda\beta}$ and written *chalkōi* for *dromōi*.

Here there are 100 *chalkōi* artabas to 105 *anēlōtikōi*, so 20 *chalkōi* to 21 *anēlōtikōi*; hence the conversion proceeds by subtracting a 21st part of the *anēlōtikōi*. When the conversion factor is $n$ to $(n + 1)$ or $(n + 1)$ to $n$, the calculation will be simplified, since it can proceed by subtracting an $(n + 1)$th part, as here, or by adding an $n$th part, as in the next example. The ratios $(n + 1)$ to $n$ were well known from financial calculations of interest, and they have a special name, *epimorios logos*, with the particular examples of *epitriton* (4 to 3), *epipempton* (6 to 5), *ephekton* (7 to 6), *epogodoon* (9 to 8), and *epidekaton* (11 to 10) being of special importance.[70]

A similar procedure is followed in lines 46–9:

46   χαλκωι      ō      χκε    αι    Φιλιππ⸍     ποϲαι
     *Chalkōi*   artabas   625    the    *Philippou,*   how many?

47   ωϲ     δει      ποιηϲαι   λαβε   το    ίά    των      ō      χκε
     As   you should    do.     Take   the   11th   of the   artabas   625

48   /   [νϛ ]θ̄   ταυτας    αφελε    απο   των     ō      χκε
     /   56   9.   These    take away   from    the    artabas   625.

49   /   φξη   β̄   τοϲαυται    Φιλιππ⸍    δεδεικται
     /   568   2   as many    in *Philippou.*     QED

In these last two calculations, the scribe has performed two arithmetical operations with the 'fractional' notation: $1 - \frac{16}{21} = \frac{5}{21}$ and $1 - \frac{9}{11} = \frac{2}{11}$. He performs a similar but more elaborate manipulation in the next example, and then gives an approximation to the result. So here, finally, are lines 136–41; I have left in the editorial details to indicate just how much of this part of the text is editorial restoration:

136   [Γαλλ⸍     ō     χκε    αι    Φι]λιππ⸍    [ποϲαι]
     *Gallōu*   artabas   625    the    *Philippou*   how many?

137   [ωϲ     δει     ποιηϲαι]   ταυτα[ϲ     χαλκιϲον]
     As   you should    do.    These   convert to *chalkōi*

138   [ωϲ   προκειται]   /   [χ]αλκωι     ō     [χμβ  ∠ γ̇ κδ̇]
     as   set out     /   *chalkōi*   artabas   642   2   3   24.

139   τουτων   [το]    ια   /   νη   έ   [ταυτας αφελε]
     Of these    the    11th   /   58   5.    These take away

---

[70] See Sections 4.5(b) ($A_{60}$ and $B_{61}$) and 4.5(d), and Burkert, *LS*, 439; but Burkert's argument that these epimoric ratios appear in the musical fragments of Archytas and Philolaus as borrowings from the financial calculations of everyday speech seems very doubtful.

140   απο   των   [χ]αλκωι   [ō]   χμβ   [∠´ γ´ κδ´´ /]
      from   the   *chalkōi*   artabas   642   2̇ 3̇ 2̇4̇ /                    ⸱

141   Φιλιππ´   ō   φπδ γ´ ιβ´´   [δεδεικται]
      *Philippou*   artabas   584 3̇ 1̇2̇   QED

More than half of the calculations proceed by passing first to *chalkōi* artabas, and then into the required standard; the process is described by the verb χάλκιcον ('chalkise'). In this case, the conversion from *Gallou* to *chalkōi* was performed in the 23rd sum, lines 122–6, and it contains a mistake and a skilful approximation: $64687 \div 100$ is given[71] and subsequently used as 642 2′ 3′ 24′, where it should be 646 2′ 3′ $\frac{11}{300}$, closer to 646 2′ 3′ 27′. The mistake, in a division by 100, is less obvious in Greek than in Arabic notation:

$$των\ μ\ ͵δχπζ\ τουτων\ το\ ρ\ /\ χμβ\ ∠´ γ´κδ´,\ \text{instead of}\ χμς\ ∠´ γ´κδ´.$$

In line 139, $\frac{5}{11}$ is a reasonable and straightforward approximation for 4 2′ 3′ 24′ ÷ 11, since 4 2′ 3′ 24′ is already close to 5, and this difference will then be divided by 11. In line 141, 3′ 12′ is a reasonable and straightforward approximation for 2′ 3′ 24′ − $\frac{5}{11}$ since $\frac{5}{11}$ is less than 2′ by 22′, close to 24′. The approximations are not always so adept; for example, in lines 120–1, the result of $68570 \div 104$ is given as 661 12′, where the correct fractional result is $\frac{3}{52}$, which is between 18′ and 17′.

The arithmetical operation in the papyrus that generates the fractions is division, and these divisions are always expressed exactly as we find them in division tables, for example as on lines 28, 102, 64, 47, and 139 in the examples quoted above. The following divisions are performed in the papyrus; the examples cited above are marked with an asterisk:

| Division by | On lines |
|---|---|
| 4 | 35 |
| 6 | 22 |
| 10 | 96 |
| 11 | 47*, 76, 139*, 160 |
| 20 | 12, 44, 134, 154 |
| 21 | 64* |
| 100 | 4, 6, 15, 28*, 33, 126 |
| 104 | 61, 93, 120, 149 |
| 128 | 40* |
| 207 | 55, 84, 111 |
| 2′ 4′ 32′ | 41*, 71, 102*, 130 |

The 'fractional' notation is associated *only* with the divisions by 11, 21, and 128; I have quoted in detail all of these divisions except for that on line 76

---

[71] In fact this part of the text has to be reconstructed, but the scribe subsequently clearly uses the wrong answer.

(where 595 $4'$ ÷ 11 is approximated as 54 $,11'$ with an error of $44'$ which arises from ignoring the fractional part $4'$ in the dividend) and in line 160 (where, in text that is lost and has been completely restored, $650 ÷ 11 = 59$ $11'$ is then subtracted from 650 to laeve 591 $\frac{11}{10}$). Of the other divisions, those by the more commonplace numbers, 4, 6, 10, 20, and 100, are *all* evaluated in a straightforward way,[72] while the seven divisions by the more unusual numbers 104 and 207 are *all* expressed in conventional simple and compound parts.

Section 7.2 described the extensive evidence we have concerning division tables, the ready reckoners and exercises that deal with the expressions of divisions as sums of parts. These division tables are sometimes quite extensive, though few would contain tables for division by 11 or 21 (see how the sets of tables in nos. 7 and 14 omit these) and none would contain tables for division by 104 and 207. I propose that the 'fractional' notation may plausibly have been used here by the scribe in the following circumstances: by convention, the final answer to any sum should always be presented in a standard form, as simple or compound parts.[73] Whenever possible, this expression should be precise, but approximations were permitted; and the exact value for all commonly occurring divisors could be got from memory or tables. The scribe did not have the tables for division by 11 and 21 to hand, or he could not remember the answers though he was aware that these tables existed; so where the sums required divisions by 11 or 21, the scribe approximated the answer or left a shorthand abbreviation in the text, to be expanded later. Where the scribe knew that the tables did not exist, for example for division by 104 and 207, he always approximated the result. And I think that a close examination of the papyrus leads to one firm overall conclusion: that we have no good evidence here for the notation or manipulation of anything similar in conception to common fractions, beyond simple subtractions like '1 minus the 21st of 16 leaves the 21st of 5'', which are conceived throughout as simple or compound parts. I shall return to this kind of explanation at several places below.

### 7.3(d) *M.P.E.R., N.S. i* 1

This text[74] is a collection of thirty-eight easy problems in three-dimensional metrical geometry very much in the style of the Heronian *Stereometrica*; it has been edited with full mathematical commentary in Gerstinger & Vogel, *MNWPER* i, 11–76. It was dated by its editors, on palaeographic grounds, to the second half of the first century BC, but this date has now been revised by W. E. H. Cockle and H. Harrauer to the first century AD, at the earliest.[75] Plate

---

[72] With one error on division by 100 (see the comment on lines 137–8, above), and one approximation in the division by 20 on line 134, in a passage that has been completely restored by the editors (642 $2'$ $3'$ $24'$ ÷ 20 / 32 $6'$).

[73] Analogous conventions apply today to the presentation of answers to school mathematical problems.

[74] Also known as P. Vindob. [inv.] 19996 and P. Rainer (N.S.) i 1.

[75] I am grateful to Walter Cockle and Hermann Harrauer for answering my many queries about this text and for examining, respectively, a photograph and the original. Several of the

11 is a photograph of column 6,[76] which contains seven of the nine examples of the general 'fractional notation' that occur in this papyrus, and the remaining two examples (in Problem 10) arise in exactly the same context as examples in Problem 12, as will be described below. Hence a discussion of column 6 will cover all occurrences of the notation in this papyrus. The column starts in the middle of Problem 12 (whose beginning is lost), and Problem 13 begins in the middle of line 3; this is indicated by a forked paragraphus or *diplē obelismenē*[77] under the beginning of line 3. There is another such paragraphus, not indicated by the editors, under the beginning of line 11, marking the end of Problem 13.

The seven instances of the 'fractional notation' in this column are as follows:

Problem 11: $\tau\upsilon\varsigma$ $\overset{\bar{\epsilon}}{\alpha}$ is used as a label to its figure, not legible on Plate 9. This is an unusual way of writing the part 5′, which in any case here needs emendation to yield anything meaningful.

Problem 12: $\overset{\acute\epsilon}{\beta}$ and $\overset{\acute\epsilon}{\delta}$ (twice) occur in the text, in lines 2 and 3, and $\kappa\alpha\delta^{\gamma\epsilon}$ (in a doubtful reading) occurs as a label to its figure, but this also needs emendation to make mathematical sense. The epsilons in line 2 are written cursively while that in line 3 is written in two strokes, and all of them have fraction indicators.

Problem 13: $\overset{\acute\epsilon}{\beta}$ and $\overset{\acute\kappa}{\gamma}$ occur in the text, in lines 9 and 10. The epsilon is written in two strokes, and both the epsilon and the kappa have fraction indicators.

Here now is a transcription and literal translation of column 6, together with its text figures and their labels, in which all abbreviations have been expanded. Note, in particular, that $\text{Ꝺ}$ is $\lambda o\iota\pi\acute o\nu$, '(there is) left'; $\text{ᐱ}$ is $\hat\omega\nu$, 'of which';[78] and $\gamma\acute\iota\nu\epsilon\tau\alpha\iota$ is written in full or abbreviated as / or ⫫. Further remarks on the abbreviations will be found in Section 7.4(a), below.

[End of problem 12:]

1      ($\lambda o\iota\pi\grave o\nu$)   $\alpha$   $\hat\omega\nu$   [$\pi\lambda\epsilon(\upsilon\rho\grave\alpha)$⫫]  $\alpha$  [$\tau\eta\lambda\iota\kappa\alpha(\acute\upsilon\tau\eta)$]  $\mathring\eta$  $\kappa\acute\alpha\theta\epsilon\tau o\varsigma$
   . . . there is left 1 of which   a side   /  1.  So much [is]  the  height.
       $\epsilon\mathring{\iota}\tau[\epsilon\nu$ $\grave\alpha]\nu\alpha\mu\acute\epsilon\tau\rho\epsilon\iota$ $\tau\rho\acute\iota\gamma\omega(\nu o\nu)$ $\iota\sigma\acute o\pi\lambda\epsilon\upsilon\rho(o\nu)$
       Then   measure   a triangle,   equilateral,

details described below, particularly of the use of fraction indicators in the 'fractional notation', are not included in the *editio princeps* and are not discernible on the photograph, reproduced here as Plate 11; but they can be securely read on the original. Also there is a revised reading in column 6, line 10: for $\tau.\zeta$ now read $\pi\zeta$. These corrections have been incorporated in the transcription, below. Further discussion can be found in the edition of the verso, M.P.E.R., N.S. xv 151, which contains an extensive addition table in a later, school-boy's hand; see no. 56 in the Appendix.

[76] There is a photograph of column 10, Problems 24 and 25 and their figures, in the *editio princeps* and another photograph in Weitzmann, *ABI* of columns 12 (part), 13, and 14 (part), Problems 28–32 and the figures 12b and c, and 13.

[77] See Turner, *GMAW*, 14–5.

[78] For details of the second abbreviation, see n. 27, above. On the first, A. Blanchard notes: "Here again there is borrowing from the demotic ᑐ, *sp*, rationalised into a lambda by the Greeks and then, since the Greek abbreviations contain at least two letters, into $\text{Ꝺ}$, $\text{Ꝺ}$, and even later, into $\lambda_o^\pi\iota$."

2       οὗ αἱ πλε[υ]ρα[ὶ] ἀνὰ     ιβ γίνεται ξδ $\overset{\acute{\epsilon}}{\beta}$
                                                  5
of which the    sides     each [are] 12. It makes 64 2,

      (ὧν)    τὸ γ́ ⟨⟩ κα $\overset{\acute{\epsilon}}{\delta}$ ταῦτ' ἐπὶ τὴν κάθε-
                                 5
of which the 3rd / 21 4. This to the height

3 τον α / κα $\overset{\acute{\epsilon}}{\delta}$ τηλικα(ύτη) [⟨ἡ⟩ στ]ήλη
                               5
   1 / 21 4. So much is the block.

                 ἐὰν    δ(οθ)ῆι     ἄλλη{ι}     τρ[ίγ]ωνο(ς) ἡμιτελὴς ἡ κορυ
[Problem 13:] If there is given another     triangle, half-complete, the top

4 φῆ ἀνὰ    γ ἡ δὲ βάς[ις][ἀν]ὰ    ιη τὰ δὲ κλίματα ἀνὰ      ια
each [is] 3, the base each [is] 18, the    slopes each [are] 11.
    ἄφελε    τὴν κορυφ(ὴν)
    Take away the     top,

5 τὰ γ ἀπὸ τῶν ιη    [(λοιπὸν)] ιε τὰ ιε] ἐφ'ἑα(υτὰ) ⟨⟩ σκε
the 3 from the 18, there is left 15. The 15 to itself / 225,
     ὧν    τὸ γ́ ⟨⟩ οε καὶ τὰ ι ἐφ' ἑα(υτὰ) ⟨⟩ ρ[
of which the 3rd / 75, and the 10 to itself / 100.

6 ἀπὸ τούτων ἄφελε    τὰ [οε] . . . (λοιπὸν) κε
From    these take away the 75      there is left 25,
     ὧν    πλε(υρὰ) ⟨⟩ ε τηλικα(ύτη) ἡ κάθετος εἶτεν
of which    a side / 5. So much is the height. Then

7    σύνθες    τὰ γ καὶ τὰ [ιη] ⟨⟩ κα[ ] (ὧν) τὸ ∠ ⟨⟩ ι ∠
put together the 3 and the 18 / 21 of which the 2 / 10 2.
    ἀναμέτρησον τρίγω(νον)    οὗ    αἱ
    Measure     a triangle, of which the

8 πλευραὶ ἀνὰ     [ι ∠]⟨⟩ [μ]ζ ∠ [δ] εἶτεν ἄφελε τὰ γ ἀπὸ
sides each [are] 10 2 / 4 7 2 4. Then take the 3 from
    τῶν ιη [(λοιπὸν)] ιε    ὧν τὸ ∠ ⟨⟩ ζ ∠
    the 18, there is left 15 of which the 2 / 7 2

9 ἐγμέτρει     ἄλλο ⌐ [τ]ρίγ(ωνον) ἰσόπ[λ]ευρ(ον)     οὗ     αἱ πλευ(ραὶ)
Measure out another    triangle,    equilateral, of which the    sides

$$\overset{\acute{\epsilon}}{\mathstrut}$$

ἀνὰ      ζ ∠ Ι κε έ    ὧν    τὸ γ́ / η β ταῦτα       ₀

$$\overset{\acute{5}}{\mathstrut}$$

each [are] 7 2̇ / 25 5̇, of which the 3rd / 8 2. These

$$\overset{\acute{\kappa}}{\mathstrut}$$

10 πρόϲθεϲ τῆι πρώτηι ἄ[λ]λο ἐχθήϲει τοῖϲ πζ ∠ δ/ νϲ γ ταῦτ' ἐπὶ τὴν

$$\overset{2\overset{.}{0}}{\mathstrut}$$

add to the first, elsewhere set-out, to the 87 2̇ 4/56 3. This to the

11   κάθετον ε/ ϲπ ∠ d τρϲων [τ]ῶν ποδῶν ἐϲτὶν ἡ πυραμίϲ
&gt; ‾‾‾‾‾‾

height 5/ 280 2̇4̇. Of so many feet   is   the pyramid.

Line 2: For ξδ, 64, read ξβ, 62.
                    έ    5̇        ίέ   1̇5̇      έ    5̇
Lines 2 and 3: For κα δ̇, 21 4, read κα ζ̇, 21 7, or κ δ̇, 20 4.
Line 4: For ια, 11, read ι, 10.
Line 10: For πζ, 87, read μζ, 47.
For notes on the labels on the figures see the *editio princeps*.

The scribe is applying the same procedures for calculating pyramids that we find set out in the Heronian *Stereometrica*; see Heron–Heiberg, Schöne, & others, *Opera* v, 137–61 or Heron–Bruins, *CC*, folios 55ʳ–61ʳ. The following notes will explain what is happening, and the reader is referred to these texts and commentaries for further details:

(i) The area of an equilateral triangle is (approximately) the 3rd plus the 10th of the square on the side; the scribe performs this operation, without explanation and with errors and approximations, on lines 2, 7–8, and 9, and each such operation will be analysed below. Heron's explanation of this procedure, for the case of an equilateral triangle of side 30 (Heron–Heiberg, Schöne, & others, *Opera* V, 148.25–6 or Heron–Bruins, *CC*, folio 58ʳ.19–20), is:

εὑρεθήϲεται τὸ ἐμβαδὸν οὕτωϲ· τὰ λ̅ ἐφ' ἑαυτά· / ϡ. τούτων τὸ γ́ καὶ τὸ ί· / τϟ̅.
The base will be found thus: the 30 to itself / 900. Of these the 3rd and the 10th / 390.
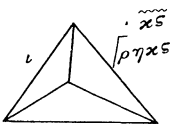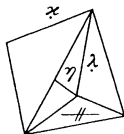


[Problem 10]                 [Problem 12]
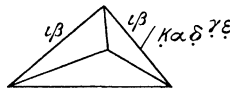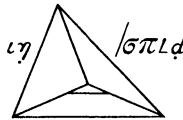


[Problem 11]                 [Problem 13]

(ii) The square on the circumradius of an equilateral triangle is the 3rd of the square on the side (*Elements* XIII 12); see line 5.

(iii) The height of these regular pyramids is calculated using the right-angled triangle of inclined edge, height, and circumradius of the base; see lines 5–6.

(iv) The volume of a pyramid is the third of the base, multiplied by the height; see lines 2–3.

(v) The procedure for evaluating the volume of a truncated pyramid is clearly described in lines 7–11.

Every occurrence of the 'fractional notation' in the text is associated with an evaluation of the area of an equilateral triangle, as was described above, and none of these operations is explained by the scribe:

Line 2: When the side is 12, its square is 144, of which the 3rd and 10th is 48 and 14 and the 10th of 4, viz. 62 and the 5th of 2. The identical operation was performed correctly in Problem 10, column 5.3–4, but here, in Problem 12, the scribe writes 64 $\overset{5}{2}$ instead of 62 $\overset{5}{2}$.

Line 8: When the side is 10 2′, its square is 110 4′, of which the 3rd and 10th is 36 2′ 4′ and 11 and the 10th of 4′, which the scribe approximates as 47 2′ 4′ by neglecting the 40′.

Line 9: When the side is 7 2′, its square is 56 4′, of which the 3rd and 10th is 18 2′ 4′ and 5 and the 10th of the 4th of 25, which is 23 2′ 4′ and the 8th of 5, which is 24 4′ 8′. The scribe makes an error and gets 25 $\overset{5}{5}$′, which he then divides by 3 to give 8 2. He then adds this to 47 2′ 4′ to give 56 $\overset{20}{3}$; the fractional part of this can be evaluated by adding the 20th of 8 to the 20th of 10 and of 5.

*All* other calculations in the text are carried out in the usual simple and compound parts. The text contains divisions by 3, 5 and its multiples, and various powers of 2. The examples of the 'fractional notation' *only* occur in the divisions into 5th parts or its multiples, and *every* such division by 5 or its multiples introduces the notation. Also *every* such 'denominator' has a fraction indicator. Thus an explanation such as was offered for the examples in P. Lond. iii 265 again seems to apply: that the scribe used the notation only when he did not know, could not find, or did not want to break off to compute, a table of division by 5ths or 10ths.

Fractions, expressed in parts, also arise in this papyrus when square roots are quoted: without any explanation, in Problem 14 the square root of 66 3″ is taken to be 8 6′, and in Problem 29 that of 200 as 14 7′. Also the rule for evaluating the area of an equilateral triangle is based on an approximation of $\frac{1}{4}\sqrt{3}$ as 3′ 10′. All of these are good approximations.[79]

---

[79] The anthyphaireses are as follows: $\sqrt{200}:\sqrt{3} = [8, \overline{6, 16}]$, $\sqrt{200}:1 = [14, \overline{7, 28}]$, and $\sqrt{3}:4 = [0, 2, \overline{3, 4}]$ where $[0, 2, 3, 4] = 13:30$, or $\sqrt{3}:1 = [1, \overline{1, 2}]$ where $[1, 1, 2, 1, 2, 1] = 26:15$. See Chapters 2 and 3; here again we have another example of unexplained fluency with square root operations.

### 7.3(e)  *Demotic mathematical papyri*

The small corpus of known mathematical works written in demotic comprises eight texts: five of them, containing in all seventy-two short problems, published in Parker, *DMP*; two in Parker, DMPF and ME; and the division tables, no. 6 in the catalogue in the Appendix below.[80] All of my examples here will be taken from *DMP*, and will be referred to by the problem numbers therein. Readers are directed to this book for photographs, transcriptions, and full commentaries.

Here I shall consider only the demotic treatment of fractions. As in Greek, there is no specific word for 'fraction', and the demotic word *r3* means 'part' or 'fraction', or even 'number' or 'amount', according to context.[81] The treatment of fractions follows the familiar pattern: there are special signs for 3″, 2′, 3′, and 4′, and thereafter the parts are indicated by a fraction indicator, a small stroke, which may often be omitted, attached to the normal numeral.[82] A ligatured 3″6′ is used to represent $\frac{5}{6}$, but no surviving text seems to contain a similar ligatured 2′4′ for $\frac{3}{4}$, common in Greek documentary papyri.[83]

While the operation of division, 'the *n*th of *m*', is described in Greek by very minor variations, for example in the word order, of the basic phrase 'τῶν *m* τὸ ń', there is a much wider range of expressions in demotic. Here are the translations of some examples:[84]

---

[80] The inventory numbers, locations, and dates of these texts are as follows. In Parker, *DMP*: P. (dem.) Cairo J.E. [inv.] 89127–30, 89137–43 verso; iii BC, perhaps from Hermopolis: Problems 1–40. P. (dem.) B.M. [inv.] 10399, Ptolemaic, later than P. Cairo, provenance unknown: Problems 41–52. P. (dem.) B.M. [inv.] 10520, early(?) Roman, provenance unsure, possibly Memphis: Problms 53–65. P. (dem.) B.M. [inv.] 10794, uncertain date, unknown provenance: Problems 66–7, the division tables no. 4 in my catalogue. P. (dem.) Carlsberg [inv.] 30, probably ii AD, from Tebtunis: Problems 68–72. In Parker, DMPF: P. (dem.) Griffiths Institute [inv.] I E.7. In Parker, ME.: P. (dem.) Heidelberg [inv.] 663, Ptolemaic, perhaps ii or i BC, provenance unknown: three fragments of problems on a trapezoidal field.

Also see Neugebauer, *ESA*, 91: "One large demotic text was found in Tûna el-Gebel, according to *The Illustrated London News* 104 (1939) and *Chronique d'Égypte* 14 No. 28 (1939) 278. No information about this text could be obtained". (Hermopolis West is modern Tûna el-Gebel; this is surely P. Cairo, above.) I would like to thank R. A. Parker and John Tait for further explanations of some of this material.

[81] I was tempted to change the word 'fraction' into 'part' in several of the passages cited below, but have resisted this blind interference with the translations. See, for Egyptian numerals and fractions, Erichsen, *DG*, 694–706, and Gardiner, *EG*, 191–203.

[82] In the passages cited below, I have replaced Parker's translations $\frac{2}{3}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ by 3″, 2′, 3′, 4′, . . . .

[83] This ligatured 3″6′ is very common, and the expression 2′ 4′ seems to be avoided; for example, in Problem 23, the sum of 1, 2′, 4′, and 8′ is then expressed as $1\frac{5}{6}$ 30′ 120′, and 2′ 4′ only seems to occur in the fragmentary and uncertain readings in Problems 33 and 70, while Problem 13 uses the alternative expression 3″ 12′. Also see n. 88, below.

[84] See the discussion in Parker, *DMP*, 8 for further details and examples. Even within the text of the single scribe of Problems 1–40, we find a wide variety of different equivalent expressions.

Problems 66 and 67, the division tables, are expressed thus:

> The way of taking 150′ to 10. Viz.[85]
> 1 result 150′
> 2 result 90′ 450′ (*sic*)
> 3 result 60′ 300′ (*sic*)

Problem 13: Carry 15 3″ into 100,
Problem 3: You shall take the number 47 to 100,
Problems 9, 10, and 46–52: You shall say: '4 is what of 5?'. Result, its 3″ 10′ 30′.

Six problems in *DMP*, numbers 2, 3, 10, 13, 51, and 72, coming from three different manuscripts, show the special feature that the results of divisions are not expressed in parts, but employ some more general 'fractional notation'. I shall start with a complete citation of the translation of Problem 3, and with brief observations on these other examples. I have adjusted the quotations to conform to the style of editorial indications used here (described in the index to the plates, p. xxv), and the notation for parts (see n. 82, above); otherwise the quotations are word-for-word from *DMP*:

Problem 3:
1  If it is said to you: '[Carry 15 3″ into 100]'.
2  you shall subtract 6 fr[om 100: remainder 94]
3  You shall say: '[6] customarily go[es] into 94 (with the) result 1]5 3‴'.
4  Bring it (to the) number 3′: result 47.
5  You shall take the number 47 to 100.
6  47 to 1
7  94 to 2
8  Remainder 6: result $\frac{6}{47}$
9  Result $2\frac{6}{47}$.
10  You shall reckon it 3 times: result $6\frac{18}{47}$.
11  You shall say: '$6\frac{18}{47}$ is its number.'
12  You shall reckon $6\frac{18}{47}$, 15 3″ times.
13  $6\frac{18}{47}$ to 1
14  $63\frac{39}{47}$ to 10
15  31 2′ $\dfrac{19\ 2'}{47}$ to 5.
16  Total 95 2′ $\dfrac{11\ 2'}{47}$ to 15.
17  $4\frac{12}{47}$ to 3″
18  Total 99 2′ $\dfrac{23\ 2'}{47}$ to 15⟨3″⟩.
19  Its remainder is 2′, which is equivalent to $\dfrac{23\ 2'}{47}$
20  Total 100.

[85] This word 'viz.' is the translation of ' / ', the transliteration of 𒑱. See n. 28 and its text, above.

These fractions, all with 'denominators' 47, are indicated here by underlining the 'numerator'. For example, in line 15,

$$31 \ 2' \ \frac{19 \ 2'}{47} \ \text{is written} \ 47 \ \underline{2'} \ 19 \ 2' \ 31$$

(read from right to left!). This same system applies to Problems 2, 3, 10, and 15, all from the same papyrus. In Problems 51 and 72, below, the 'denominators' are underlined; for example, in Problem 51, $\frac{5}{11}$ is written $\underline{11} \ 5$.

Problem 2: A fragmentary part of the text, very similar to Problem 3, based on dividing 100 by 17 $3''$ and involving divisions by 53.

Problem 10: A problem on division of cloth that involves division by 11, in which the complex expression $1 \ 2' \ \dfrac{1 \ 2'}{11}$ appears. Problems 8 and 9 are very similar; they involve division by 6 (giving an answer $\frac{5}{6}$, written as the ligatured $3''6'$) and by 5 (see the discussion below).

Problem 13: A fragmentary and obscure problem on the division and valuation of pieces of cloth, in which, without any explanation, a further valuation at a rate described by the complex division $3' \ 15' \ \dfrac{7 \ 2' \ 10'}{131}$ is introduced. This yields a set of complex expressions involving divisions by 131.

Problem 51: The last of a sequence of six problems. The previous five problems all ask: given the fraction $n'$ which is added to 1, determine what fraction of $1 \, n'$ must be subtracted from it to give 1 again. The answer $(n + 1)'$ is verified in each case, for $n = 5$, 6, 7, 8, and 9. Then Problem 51 asks: given the fraction $\frac{5}{6}$ which is added to 1, determine what fraction of 1 must be subtracted from it to give 1 again. Two of these problems will be quoted in full below.

Problem 72: A fragmentary and obscure part of a problem involving proportions of gold, silver, copper, and lead. Without explanation in the surviving fragment, division by 76 is introduced and then not expressed in parts.

Problems 13 and 72 seem to be so fragmentary and obscure as to defy elucidation, and I shall not consider them further here. Take now the following parallel extracts, from Problems 9 and 10:[86]

| Problem 9 | Problem 10 |
|---|---|
| You shall say: '4 is what | You shall say: '3 is what |
| (fraction) of 5?' | (fraction) of 11?' |
| Result, its[87] $3'' \ 10' \ 30'$. Add it to 4: | Result $\frac{3}{11}$. Add it to 3: |
| result 4 $3'' \ 10' \ 30'$. . . . | result $3\frac{3}{11}$. |
| You shall reckon 4 $3'' \ 10' \ 30'$, | You shall reckon $3\frac{3}{11}$ |
| 5 times: result 24. . . . | $2'$ times: result 1 $2' \ \frac{1 \ 2'}{11}$. . . . |

[86] The editorial apparatus, apart from indications of editorial additions, has been omitted from these and most of the following examples.

[87] I do not know whether the inclusion of the possessive article in this phrase in Problems 8 and 9 (demotic "$r \ p3y \cdot f \ \ldots$", translated "Result, its . . .") is noteworthy, but its exclusion from 10 (demotic "$r \ \ldots$", translated "Result . . .") is significant.

In the left-hand column, the calculation involving division and multiplication by 5 is carried out in compound parts; in the right-hand column, division by 11 is left in its raw state, but it is clearly conceived in simple and compound parts. This is further corroborated by the corresponding extract from the similar Problem 8, which starts:

You shall say: '5 is what fraction of 6?' Result, its $\frac{5}{6}$ . . . ,

in which the $\frac{5}{6}$ is written as the ligatured $3''6'$, again a compound part.

Now consider Problem 51. This follows exactly the same stereotyped verbal formula as the previous five problems 46–50, as can be seen in the following parallel quotations.

|  | Problem 48 | Problem 51 |
|---|---|---|
| 1 | You are told: '$7'$ is (the) addition: what is (the) subtraction?' | You are told: '$\frac{5}{6}$ is (the) addition: what is (the) subtraction?' |
| 2 | The way of doing it. Viz. | The way of doing it. Viz. |
| 3 | You shall add $7'$ to 1: result 1 $7'$. | You shall add $\frac{5}{6}$ to 1; result $1\frac{5}{6}$. |
| 4 | You shall say: '$7'$ is what (fraction) of 1 $7'$?' Result, its $8'$. | You shall say: '$\frac{5}{6}$ is what (fraction) of $1\frac{5}{6}$?' Result, its $\frac{5}{11}$. |
| 5 | You shall say: '$7'$ is (the) addition; $8'$ is (the) subtraction.' | You shall say: '$\frac{5}{6}$ is (the) addition; $\frac{5}{11}$ is (the) subtraction.' |
| 6 | To cause that you know it. Viz. | To cause that you know it. Viz. |
| 7 | You shall subtract $8'$ from 1: remainder $3''$ $12'$ $8'$. | You shall subtract $\frac{5}{11}$ from 1: remainder $\frac{6}{11}$. |
| 8 | The fraction $3''$ $12'$ $8'$. Viz. | The fraction $\frac{6}{11}$. Viz. |
| 9 | Its $7'$ (is) $8'$. You shall add it ($8'$) to it ($3''$ $12'$ $8'$): result 1 again. | Its $\frac{5}{6}$ (is) $\frac{5}{11}$. You shall add it $\left(\frac{5}{11}\right)$ to it $\left(\frac{6}{11}\right)$: result 1 again. |

In each of Problems 46–50, the subtraction in line 7 is expressed as a compound part, with the feature that, in all of the cases other than the one quoted, these parts involve the ligatured $3''6'$, translated as $\frac{5}{6}$. These subtractions are as follows:

Problem 46:    $1 - 6' = 3''$ $6'$

Problem 47:    $1 - 7' = 3''$ $6'$ $42'$

Problem 48:    $1 - 8' = 3''$ $12$ $8'$       (see above[88])

Problem 49:    $1 - 9' = 3''$ $6'$ $30'$ $45'$

Problem 50:    $1 - 10' = 3''$ $6'$ $15'$

Problem 51:    $1 - \frac{5}{11} = \frac{6}{11}$         (see above)

It seems clear from this block of examples that again, while the scribe prefers, whenever he can, to express the divisions as compound parts, here also, as in

[88] Parker observes (*DMP*, 61, note to line 16): "For no discernible reason the scribe uses the clumsy fraction $3''$ $12'$ $8'$, with $8'$ in incorrect order after $12'$, instead of the more convenient $\frac{5}{6}$ $24'$." I suggest that $3''$ $12'$ may be the standard expression for the 4th of 3. So to subtract $8'$ from 1, 1 is mentally decomposed into $3''$ $12'$ $4'$, then $8'$ is subtracted from $4'$ to give $8'$; hence the form of the scribe's answer. Also see n. 83, above.

Problem 10, division by 11 defeats him. Similarly, in Problems 2 and 3, division by 53 and 47 leads to expressions which he also cannot express in parts. In all of these cases, he resorts to abbreviations for expressing the incomplete divisions; and it is these abbreviations that we interpret as a notation for common fractions.

This explanation is again consistent with the proposed explanation of the previous texts, that, while an expression in parts is the preferred form of answer, in cases for which tables are not available, the divisions may very occasionally be left in an incomplete state. I shall now draw these conclusions together and indicate some of their wider implications.

## 7.4 CONCLUSIONS AND SOME CONSEQUENCES

### 7.4(a) *Synthesis*

The evidence for the proposal that early Greek mathematicians conceived and used manipulations of common fractions is said to come from two sources: from the medieval manuscripts of scientific treatises, and from contemporary texts on papyrus which are mainly of a commercial and pedagogic nature; and the proposal is often reinforced by a more or less unconscious retrojection of our own ways of thinking, which usually manifests itself in wonder that anyone of any mathematical insight could ever be so obtuse as to reckon in unit fractions.[89]

However, the evidence of calculations in the papyri displays almost unanimous evidence against this proposal. The very few instances there that can be cited as illustrating notions for common fractions appear, on closer scrutiny, more probably to be abbreviations of unresolved descriptions of divisions that are still conceived as sums of unit fractions, and all can be more naturally explained as relaxations of stylistic conventions about how these divisions should be evaluated and expressed. Possibly the abbreviations did then evolve into our conceptions of common fractions, and certainly the practice and popularity of common fractions developed, particularly among Italian mathematicians, from the ninth or tenth centuries onwards. These new fractional notations and conceptions may then have been adopted by the scribes and readers of the medieval manuscripts, and so infiltrated and corrupted the evidence to be found there; on this more work needs to be done.[90]

[89] In this section, only, I shall sometimes use the common but, I believe, misleading description 'unit fractions', to ensure that my proposals are clear to everybody. But I encourage any sympathetic reader to conceive these unit fractions, in Greek style, as the half, the third, the quarter, . . . and their combinations, that is as *meros ē merē*, part or parts.

[90] For example, it would be interesting to know if there is any systematic difference between contrasting early manuscripts, like Paris gr. 2389 and Vatican gr. 1594, which contain the same numerical data (here the *Almagest*) but expressed differently, in the old capital and the new minuscule scripts. Also see the notes in Diophantus–Sesiano, *BFSDA*, on the treatment of numbers and fractions in the newly discovered Arabic translation of part of the *Arithmetica*, especially on pp. 37–42 and 437 s.v. *juz'*. This translation was probably made from an earlier exemplar than the surviving minuscule Greek manuscripts (see p. 67).

Of course, the manipulations of fractions expressed as unit fractions are (arithmetically) equivalent to the same manipulations when expressed as common fractions; but they will be conceived differently in the two systems. Some operations will come very close to each other; for example, the manipulation of 'the $n$th of $m$' into the equivalent 'the $kn$th of $km$' is parallel to the identity $m/n = km/kn$, but very far from it in conception; and a frequent step in our evidence is the manipulation 'the $n$th of $m$ plus or minus the $n$th of $p$ is the $n$th of $(m \pm p)$', equivalent to $m/n \pm p/n = (m \pm p)/n$. But reciprocation gives an example of an operation which manifests itself very differently in the two systems; for an illustration, the relation between $\frac{7}{17}$ and $2\frac{3}{7}$ is much more evident than that between 3′ 17′ 51′ and 2 3′ 15′ 35′. Just one example of some operation such as the addition, subtraction, multiplication, or division of two fractional quantities, expressed directly as something like, 'the $n$th of $m$ multiplied by the $q$th of $p$ gives the $nq$th of $mp$' and *clearly unrelated, by context, to any conception in terms of simple and compound parts*, could be fatal to my thesis that we have no good evidence for the Greek use or conception of common fractions. I know of no such example.[91]

We can take my explanation further. First consider the form of the abbreviation, in which the stereotyped phrase '$\tau\tilde{\omega}\nu$ $m$ $\tau\grave{o}$ $\acute{n}$' is written $\overset{\acute{n}}{m}$ with, as is often expressed, 'the numerator and denominator reversed'.[92] The practice of writing one letter above another is very common in the abbreviations found

---

[91] For a recent, useful, detailed study of Greek and Egyptian fractional techniques, see Knorr, TFAEG. But what are presented there as manipulations of common fractions $\frac{m}{n}$ are clearly manipulations of the descriptions 'the $n$th of $m$' which are conceived throughout the texts under discussion in terms of unit fractions. See pp. 148–51 on the Archmîm Mathematical Papyrus, for example on p. 150: "Although the problems are invariably expressed in terms of unit-fractions and the final solutions are given in this same mode, the actual execution of the arithmetic operations first introduces their conversion to terms of the form $\frac{a}{b}$" (but here $\frac{a}{b}$ invariably abbreviates the expression $\tau\tilde{\omega}\nu$ $a$ $\tau\grave{o}$ $b$); and "Thus, the scribes in the arithmetic tradition of late antiquity present themselves as *virtuosi* in the art of manipulating unit-fractions; yet their very methods reveal this to be a superfluous art. How was it possible that they failed to perceive this, embellishing these techniques and so distracting from the teaching and development of the more general techniques of [common] fractions employed within their computations?" (The answer here is surely that the scribes were actually conceiving, teaching, and developing unit fraction techniques, and had little or no conception of common fractions.)

[92] See, for an influential example, Heath, HGM i, 43 f.: "The most convenient notation of all [for common fractions] is that which is regularly employed by Diophantus, and occasionally in the *Metrica* of Heron. In this system the numerator of any fraction is written in the line, with the denominator *above* it, without accents or other marks (except where the numerator or denominator itself contains an accented fraction); the method is therefore simply the reverse of ours, but equally convenient. In Tannery's edition of Diophantus a line is put between the numerator below and the denominator above ... but it is better to omit the horizontal line (cf. $\overset{\rho\kappa\eta}{\rho} = \frac{100}{128}$ in Kenyon's Papyri ii, No. cclxv.40, and the fractions in Schöne's edition of Heron's *Metrica*)." In fact, these letters $\rho\kappa\eta$ and all of the other examples in P. Lond. ii 265 and M.P.E.R., N.S. i 1 incorporate fraction indicators; see nn. 64 and 75, above. *Every* known instance of this abbreviation in contemporary papyri has been discussed in Sections 7.3(c), (d) and (e), above.

in Greek papyri. See, in Plate 8, how practically every word in O. Bodl. ii 1847 is abbreviated in this way. For example, at the beginning of lines 4 and 5, $\overset{\prime\prime}{A}\lambda(\lambda\eta)$ $\dot{\epsilon}\chi o(\mu\epsilon\nu\eta)$ is written $\overset{\lambda}{a}\,\epsilon\,\overset{o}{\chi}$, while, in line 7, the same expression is written $\overset{\lambda}{a}\,\overset{\chi}{\epsilon}$; and, in Plate 11 in the first line of column 6 of M.P.E.R. i 1, $\tau\rho\acute{\iota}\gamma\omega(\nu o\nu)$ is written $\tau\rho\iota\overset{\omega}{\gamma}$ and $\dot{\iota}\sigma\acute{o}\pi\lambda\epsilon\nu\rho(o\nu)$ is written $\iota\sigma o\pi\lambda\epsilon\nu\overset{\diagup}{\rho}$.[93] It may be that it was this common practice of the scribes of documentary papyri which led to the writing of numerical abbreviations by superposition. For example:

(i) In the alphabetic system of numerals, $\overset{a}{\mu}$, $\overset{\beta}{\mu}$, etc. denote 1 myriad (1 0000), 2 myriads, etc.

(ii) In describing amounts of money measured in the $\chi\alpha\lambda\kappa o\hat{\upsilon}\varsigma$, one-eighth of an obol, $\overset{a}{\chi}$, $\overset{\beta}{\chi}$, etc. is very often used for 1 *chalkous*, 2 *chalkoi*, etc.

(iii) In describing amounts of wheat, or its equivalent, measured in the $\chi o\hat{\iota}\nu\iota\xi$, one-thirtieth of an artaba, the same notation $\overset{a}{\chi}$, $\overset{\beta}{\chi}$, etc. is very often used for 1 *choinix*, 2 *choinikes*, etc.

Moreover, in the 'fractional notation' the 'denominator' would tend to be placed above the 'numerator' in such an abbreviation since this 'denominator' is written with its long fraction indicator which proclaims its different grammatical and conceptual status from that of the 'numerator'. As this fraction indicator would get in the way of anything written too close and above it, it would be much more conveniently situated on top. Further, in these examples of abbreviation by superposition, the later part of the abbreviation is usually written on top, and the same goes for the 'fractional' abbreviation of '$\tau\hat{\omega}\nu\ m\ \tau\grave{o}$ $\acute{n}$' as $\overset{\acute{n}}{m}$.

A further argument in support of my thesis can be drawn from the full-format division tables, described in Section 7.2(a), above. Consider how such a table might be used to evaluate a division like the 9th of 842. From the table, we can read off that the 9th of 800 is 88 2′ 3′ 18′, of 40 is 4 3′ 9′, and of 2 is 6′ 18′. When these are added, we get 92 2′ 3′ 3′ 6′ 9′ 18′ 18′; but the repeated parts now have to be removed, by several more steps of working. However, it does not take much insight to see that we can greatly simplify the evaluation as follows: the 9th of 800 is 88 2′ 3′ 28′, and, also from the table, we recognise the fractional expression here as being the 9th of 8; similarly the 9th of 40 is 4 and the 9th of 4. Hence the 9th of 842 is 88 + 4 and the 9th of 8 + 4 + 2, so 93 and the 9th of 5, so 93 2′ 18′. I cannot believe that some such simplification was not used in practice; but that it seems never to have been expressed in the tables is yet another indication of the absence of any notation

---

[93] There are about thirty different such abbreviations in M.P.E.R., N.S. i 1; they are listed in Gerstinger & Vogel, *MNWPER*, 47. For abundant other early examples of these abbreviations, see McNamee, *AGLPO* and the index of symbols and abbreviations in P. Lugd. Bat. xxi (= Pestman, *GZA*). The whole question of abbreviations in papyri is very fully discussed in Blanchard, *SAPDG*; see n. 28, above.

for it, and so the absence of any conception of division in terms other than as sums of parts.

Now consider the contrast between the rather rigid forms of the Greek expressions for division, all very minor variations of '$\tau\hat{\omega}\nu$ $m$ $\tau\grave{o}$ $\acute{n}$', that are found in the wealth of Greek papyrus texts,[94] and the wealth of different expressions that have been found in the small handful of demotic texts; and the associated contrast of the minute proportion of Greek uses of a 'fractional notation' in surviving calculations with the very much greater proportion of demotic occurrences. The abundance of our school exercises in division indicates how the very rigid Greek style for expressing division seems to have been drummed into schoolchildren and thereafter to have determined their conception of fractions while, in demotic, our evidence seems to indicate a much looser sense of style. My explanation of the occurrences of the 'fractional notation' concerns this imposed sense of style: that the results of properly executed calculations should always be presented in standard form, as expressions in unit fractions, either to be recalled from memory, or looked up in tables, or evaluated, or approximated. If this sense of style was relaxed, so would the necessity of this final, often inconsequential, step be removed, and so unresolved phrases for 'of $m$ the $n$th' could begin to appear, expressed or abbreviated in a variety of ways. The adoption of the demotic script rather than hieratic for Egyptian mathematical texts may well have left scope for the development of new summary notations for stereotyped common operations like division. In this way, it is perhaps not too surprising that our small clutch of surviving demotic texts should contain so many different ways of describing division; but I do not think this, in itself, gives evidence for an early familiarity with common fractions, since the expressions themselves always seem still to be conceived there in terms of unit fractions.

Another way of expressing this is to view division as an implied question which anticipates some conventional reply, thus 'the $n$th of $m$ is . . .', or 'the quotient (literally *quotiens*, how many times?) of $m$ by $n$ is . . .'. Answers can be given in many different ways—division with remainder, unit fractions, sexagesimal numbers,[95] decimal numbers, anthyphairetic ratios, astronomical ratios, etc.—and only the context can determine what will be acceptable. Common fractions correspond to the sophisticated step of ignoring the interrogative aspect, and realising that the question itself can be treated as its own answer. In order to see that this step is meaningful and useful, we need to be convinced that the questions themselves can be freely manipulated; for

---

[94] My catalogue in the Appendix gives only a small part of this evidence, that dealing with systematic tables.

[95] We have the example of the completely different reaction of another ancient culture in the Babylonian need for and use of reciprocal and multiplication tables to resolve divisions in their arithmetised sexagesimal mathematics; see Neugebauer, *ESA*, 31–4 and 50. Observe how there is simply no need for a Greek reciprocal table to correspond to these Babylonian examples!

example that 'the $n$th of $m$ is . . .' and 'the $q$th of $p$ is . . .' can be added to give 'the $nq$th of $(mq + np)$ is . . .'. Modern mathematics has shown, time and time again, how difficult such a step of treating a question as its own answer can be, however natural it might subsequently become, and how productive it may be of new concepts and developments. I do not believe that this step was taken by Greek mathematicians in their calculations with fractions.

The contempt that is often expressed for the system of unit fraction calculations is, I believe, not entirely justified. Unit fraction expansions can be evaluated and manipulated by a wealth of algorithms[96] and they can convey some information, for example of magnitude and approximation, as efficiently as any other representation.[97] The problems that do arise with unit fraction arithmetic are often exaggerated and distorted: on the one hand, unit fraction arithmetic is quite feasible, as our evidence of ancient commercial practice demonstrates abundantly, while on the other hand the problems inherent in any formally correct and complete description of arithmetic with real numbers are far greater than many mathematicians seem to believe.[98] I do not deny that the introduction first of common fractions, then of decimal fractions, introduced a dramatic new fluency and confidence into mathematics, with profound consequences. I only wish to insist that the old unit fraction representation does not suffer from quite as many disadvantages as is often supposed.

Much more could be said on the *merē* in Greek mathematics. For example, I have just alluded to the problem of the computation of the entries in the division tables, and there is also the issue of Euclid's use of the word *merē*, especially at the beginning of *Elements* VII. These issues must be deferred to another place. Here I shall finish with my final dialogue, a little fantasy to top off this heavy chapter.

### 7.4(b)  *The slaveboy meets an accountant*

ACCOUNTANT[97]: I was out at the garden of Academe on business, and I heard Plato and his friends discussing some of your ideas about mathematics. Then one of my scribes

---

[96] To begin to discuss this issue properly would double the length of this chapter. See n. 32, above.

[97] For example, if sums and differences of unit fractions are permitted, the expression $\pi = 3 + 7' - 791' - 3\,748\,549' + \cdots$ conveys the information about the size and best approximations to $\pi$ as succinctly and directly as any other method; more directly, for example, than the anthyphairesis $\pi = [3, 7, 15, 1, 272, 1, 1, 1, \ldots]$. But the anthyphairesis can contain a lot more additional information.

[98] Of course a formal description of arithmetic with rational numbers is straightforward and so, in this sense, the unit fraction description of rational numbers does introduce unnecessary problems. But the theoretical description of arithmetic with decimal fractions is less straightforward than is generally thought; see the remarks in Sections 1.2(e), 4.2, and 9.3(a), below, and the discussion in Fowler, FHYDF. On the practical difficulties with sexagesimal arithmetic, see Neugebauer, *HAMA* ii, 590 ff. ("The further one moves away from the productive period of Greek astronomy towards the didactic phase the more one finds long and clumsy explanations of sexagesimal operations. Such is the case already in the Commentaries to the Almagest by Pappus and Theon. Byzantine treatises are full of such trivia.") and pp. 968 ff.

had to copy out your dialogue with Socrates, so I got to read that, and I have a copy here. Could we go through it together, with you taking the part of Socrates?

BOY$_{98}$: All right, though it seems a curious thing to do. Here goes. (Reading:) Tell me, sir, what is the relationship of size between this heap of sixty stones, and that heap of twenty-six stones?

ACCOUNTANT$_{99}$: Do you mean the number of times the smaller goes into the larger?

BOY$_{100}$: Try it.

ACCOUNTANT$_{101}$: It goes more than twice, but less than three-times.

BOY$_{102}$: Can you be more precise?

ACCOUNTANT$_{103}$: It goes twice with eight stones left over.

BOY$_{104}$: Those 'eight stones' aren't related to anything now.

ACCOUNTANT$_{105}$: I just omitted to say that they're still in relation to the other heap of twenty-six stones.

BOY$_{106}$: Go on.

ACCOUNTANT$_{107}$: Give me time; I can go on to describe that relationship. So let's say: twice at the first step, and now I'll tell you about the relationship between eight stones and twenty-six stones.

BOY$_{108}$: Go on.

ACCOUNTANT$_{109}$: (Putting the script to one side:) Those eight stones are less than the third and more than the fourth of the twenty-six stones; so the relationship is: first step, twice; second step, the fourth; and then a bit more.

BOY$_{110}$: (Aside:) I see what he's up to; this could be interesting. But what happens to *arithmētikē* if part of a stone is still called a stone? Does it still make sense? (To the accountant:) Can you be more precise?

ACCOUNTANT$_{111}$: I can work out how much more it is as follows: I need to take the 4th of 1 away from the 26th of 8, which is the 13th of 4; so I need to take the $4 \times$ 13th of 13 away from the $13 \times$ 4th of $4 \times 4$, which leaves the 52nd of 3, which is less than the 17th and more than the 18th. So the ratio of sixty to twenty-six is expressed by: first step, twice; second step, the fourth; third step, the eighteenth, and a bit more which I can easily work out in the same way to be the four hundred and sixty-eighth exactly. Or, as we learned at school to say and write it:

$$\text{of 60 the 26th is 2 } 4' \text{ } 18' \text{ } 468'.$$

There are lots more equivalent ways of describing the same ratio, some much more convenient than this: for example, a preliminary division with remainder and then a look at my ready reckoner gives:

$$\text{of 60 the 26th is 2 } 4' \text{ } 26' \text{ } 52'.$$

What's wrong with all that? It seems just as logical and convenient as your anthyphairetic ratios, so why did Socrates dismiss it so contemptuously at the end of S$_{35}$?

BOY$_{112}$: There's the awkward operation of cutting up the unit that the philosophers don't like, though if the mathematicians find something that interests them, they might nevertheless still go ahead. You also have a problem in deciding when two ratios are equal; so you can define ratio, but its not entirely clear how you can go on to describe proportionality! But let's see what might be involved in a theory of these 'accountant's ratios'. For a fully developed theory we'll have to investigate:

    (i)    Evaluating these kinds of ratios (cf. S$_1$–B$_{14}$),

    (ii)   Describing equality and order (cf. S$_{17}$–B$_{22}$),

(iii) Translating backwards (cf. $B_{26}$–$S_{29}$), and

(iv) Arithmetic (cf. $B_{32}$–$S_{35}$),

and, on the way, we hope to pick up insights about yet undiscovered phenomena or appealing new ways of handling mathematical ideas and proofs. Then, we must relate this new definition to the other definitions of ratio.

ACCOUNTANT$_{113}$: But before you start your abstruse investigations, and explain why some of these things are so difficult and so unilluminating, I'd just like to say that I've got an office full of scribes who are calculating with these kinds of expressions all the time, and they don't seem to have many problems in getting by.

BOY$_{114}$: Wait a minute. Instead of doing a reciprocal subtraction process, we could always subtract from the same larger term. Take that ratio of eight to twenty-six that we have to describe [see $A_{107-11}$]: we could say eight goes into *twenty-six* three-times with two left over; two goes into *twenty-six* thirteen times. Work that out and I think you'll find

<p align="center">of 60 the 26th is 2  3′ less (3.13)′.</p>

Think about that, and you'll see that if you don't want subtractions like that "less 39′" to occur you will have to overshoot at each step; four-times eight overshoots twenty-six by six; five-times six overshoots twenty-six by four; seven-times four overshoots twenty-six by two; and thirteen-times two is twenty-six. So now

<p align="center">of 60 the 26th is 2  4′ (4.5)′ (4.5.7)′ (4.5.7.13)′.</p>

Do you see what's happening?

ACCOUNTANT$_{115}$: I give up! Will you people never stick to the point?

BOY$_{116}$: I wonder if we can relate anthyphairetic ratios to these kinds of accountant's ratios. . . .

<p align="center">7.5 APPENDIX:<br>A CATALOGUE OF PUBLISHED TABLES</p>

See Section 7.2, above, for a general description of the tables in this appendix.

These lists contain all the published examples I have been able to locate in Egyptian (hieratic and demotic), Coptic, and Greek. They are identified as being on papyrus (P.), ostraca (O.), writing tablets (WT.), waxen wooden tablets (WWT.), and as graffiti (G.). The lists are arranged alphabetically by language and script; museum inventory numbers are distinguished from catalogue numbers; where the text appears in a standard edition (like *The Oxyrhynchus Papyri, Papyri in the University of Michigan Collection*, etc.) no further publication details are given, but otherwise they are given in detail (and if a reference that occurs here is not cited elsewhere, it will be given here in full and not listed in the bibliography); the usual name of each text is set in bold type; if the text appears in Pack, *GLLTGRE* (second edition), its reference number there is given; where a date or provenance is assigned or known, it is given; a note about the publication may be added; and each entry finished with a brief description.

I have restricted this list to systematic examples of tables; there are many other published examples of arithmetical exercises involving calculations

with, or verifications of, these kinds of expressions which have not been included.

The following abbreviations are used:

ff = full format, up to the $n$th part of 1 0000.
af = abbreviated format, up to the $n$th part of $n$.
ie = initial entry, the $n$th part of 6000.
*BASC* = *Bulletin de la Société d'Archéologie Copte.*
*ZPE* = *Zeitschrift für Papyrologie und Epigraphik.*
*RA* = *Revue Archéologique.*

### 7.5(a) *Division tables*[99]

Egyptian, hieratic:

1.     **P. Kahun 8**, in Griffith, *HPKG*. A fragmentary collection of mathematical problems and a '2/$n$-table', with verifications, for $n = 3, 5, 7, \ldots, 21$, identical with the expressions in the Rhind Mathematical Papyrus; photograph.

2.     **Rhind Mathematical Papyrus** (= P. B.M. [inv.] 10057 and 10058, and a fragment in the Brooklyn Museum), in Chace, *RMP*; copied *c.* 1575 BC by Ahmose from an archetype some 300 years older. An extensive collection of 86 mathematical problems, a table of 10′, in af no ie, quoted above, and verifications of $n$th parts of 2 for $n = 3, 5, 7, \ldots, 99$, and 101; photographs and facsimiles. This text, especially the '2/$n$-table', has provoked an enormous literature (bibliography to 1929 in Chace, *RMP*, abridged and extended randomly in the 1979 reprint), but with only occasional references to other division tables.

3.     **O. Sen-Mūt 153**, in W. C. Hayes, *Ostraka and Name Stones from the Tomb of Sen-Mūt (No. 71) at Thebes*, New York: Metropolitan Museum, 1942. Fragments of a table of 7′ with auxiliary red numbers; photograph.

Egyptian, demotic:

4.     P. B.M. [inv.] 10794, Problems 66 and 67 in **Parker, *DMP*,** 72–3. Tables of 90′ and 150′, with entries from 1 to 10; photograph.

---

[99] A preliminary version of this table was published as Fowler, TP, and an addendum, FAT, consisting of a further 38 sets of tables, has now appeared; I have here revised my original proposal to refer to these as 'tables of parts', and now suggest the more conventional and, I hope, acceptable 'division tables'. I would like to thank the many scholars, especially R. S. Bagnall, B. Boyaval, W. Brashear, A. Bülow-Jacobsen, P. Cauderlier, W. E. H. Cockle, B. R. Goldstein, L. Koenen, P. Mertens, P. J. Parsons, T. S. Pattie, R. Pintaudi, P. J. Sijpesteijn, and Sir Eric Turner, who have provided information and references. Yet more examples are quite possible from the unpublished residues of other papyrological collections around the world, or from excavations.

5.    O.   Private collection, in **G. Belli & B. Costa, Una tabellina arith-metica per uso elementare scritta in demotico**, *Egitto e Vicino Oriente* 4 (1981), 195–200. Table of 2′; photograph.

6.    P.   Unidentified example, in **E. Revillout, *Mélanges sur la métrologie, l'économie politique, et l'historie de l'ancienne Égypte***, Paris: Maisonneuve, 1895, pp. lxix–lxxiii. (Revillout describes it as "un papyrus mathématique qui m'a été communiqué par l'Exploration Fund".) Tables of 7′, 8′, . . . , 14′, and 15′ in af.

**Coptic:**

7.    P.   B.M. MS 528, in **W. E. Crum, *Catalogue of Coptic Manuscripts in the British Museum*** London, 1905, and J. Drescher, A Coptic Calculation Manual, *BASC* 13 (1948–9), 137–60; *c.* AD 900. A palimpsest parchment codex; the later Coptic text contains multiplication tables (see no. 47, below), mathematical exercises, and division tables for 2′, 3″, 3′, 7′ [*sic*], 4′, 5′, 6′, 8′, 9′, 10′, 12′, 15′, 16′, 20′, 24′, and 48′, in af with ie; only the first few entries of the tables for 2′, 15′, and 48′ are transcribed. Described as "identical with those in Greek mathematical papyrus of Achmîm", no. 12, below.

8.    P.   Cambridge University Library [inv.] T.-S. Ar 39 : 380, in **B. R. Goldstein & D. Pingree, More Horoscopes from the Cairo Geniza**, *Proceedings of the American Philosophical Society* 12 (1981), on 186–9. Two folia containing multiplication tables (see no. 48, below) and tables for 12′, 24′, and 48′, in ff; photograph.

9.    **O.**   **Crum 480**, in W. E. Crum, *Coptic Ostraca from the Collections of the Egypt Exploration Fund, the Cairo Museum, and Others*, London, 1902. A fragmentary potsherd table of 31′ in af with ie. Crum was unable to identify its purpose in 1902 (but see no. 7, above). There is a description in K. H. Sethe, *Von Zahlen und Zahlworten bei den Alten Ägyptern*, Strassbourg: Trübner, 1916, pp. 71–2, but some of his restorations, including the ie, are clearly wrong; see W. Brashear, Quisquiliae, *BASC* 26 (1984), 19–22.

10.    P.   Strasbourg Bibliothèque Universitaire [inv.] 4110, same publication as no. 8 above, on 176–7. Table of 24′ in ff, a marginal annotation to an astrological text; photograph.

11.    **O.**   **Wadi Sarga 24–28** (= *Coptica* iii 24–28). Five fragmentary pot-sherds with brief traces of division tables, four with ie (on this, see W. Brashear, Quisquiliae, *BSAC* 26 (1984), 19–22) for 7′ (3 examples), 11′, 25′, 49′, and one unidentified table.

**Greek:**

12.    **Achmîm Mathematical Papyrus** (= P. Cair. [inv.] 10758), Pack 2306, in J. Baillet, *Mémoires de la Mission Archéologique Française* ix 1, Paris: Leroux, 1892; vii AD. Leather-covered papyrus codex containing division tables for 3″, 3′, 4′, . . . , 8′, 9′,

and 10′ in ff and for 11′, 12′, ... , 19′, and 20′ in af, all with ie, and mathematical exercises; complete facsimile.

13.  P.  Berol. [inv.] 21296, in **W. Brashear, Greek Papyri: Fractions and Tachygraphy**, *Anagennesis* 3 (1983), 167–77; ii BC. Fragments of tables of 3′ and 4′ in ff; photograph.

14.  WT.  Bodleian Gr. Inscription [inv.] 3019 in **P. J. Parsons, A School-Book from the Sayce Collection**, *ZPE* 6 (1970), on 142–3; late iii AD. Seven wooden school tablets containing a wide variety of exercises in Greek and Coptic; three sides contain division tables for 2′, 3′, 4′, 5′, 6′, 8′, 9′, 10′, and 12′, in ff. Only the table for 2′ is transcribed.

15.  P.  **Freib. i 1**; also see W. Brashear, Greek Papyri: Fractions and Tachygraphy, *Anagennesis* 3 (1983), 167–77; ii or i BC. Fragments of a table of 3″.

16.  WT.  Louvre [inv.] AF 1196¹, AF1196², AF 1196³, AF 1197¹, in **B. Boyaval, Tablettes mathématiques du musée du Louvre**, *RA* (1973), 243–60 and Le Cahier de Papnouthion et les autres cahires scolaires grecs, *RA* (1977), 215–30; P. Cauderlier, Cinq tablettes en bois au musée du Louvre, *RA* (1983), 259–80; and W. Brashear, Corrections à des tablettes arithmétiques du Louvre, *Revue des Études Grecques* 97 (1984), 214–17 (all publications must be consulted; photographs in *RA*); vi AD. Parts of a wooden schoolbook containing multiplication (see no. 54, below) and division tables:
AF 1196¹ side B: table of 9′ in ff with ie;
AF 1196² side B: table of 3″, with ie and verifications (or possibly derivations; see nos. 26 & 44, below) of each successive entry, up to 900. Also see AF 1196³ side B, below;
AF 1196³ side A: table of 11′ in af with ie, and multiplication tables;
AF 1196³ side B: on the basis of three legible letters, Cauderlier proposes a restoration of the continuation of AF 1196² side B, a table of 3″ with verifications from 1000 up to 10000;
AF 1197¹ side A: table of 14′ in af with ie;
AF 1197¹ side B: table of 17′ in af with ie.

17.  WWT.  Louvre [inv.] MND 552, MND 552h, i, k, and l; same publications by B. Boyaval as the previous entry; v or vi AD. Wax tablets containing a collection of school exercises, some arithmetical:
MND 552k: Headed and dated, "Papnouthion son of Iboïs, Mechir 21st" (incorporating a corrected reading by P. Cauderlier; personal communication), a table of 24′, in af with ie, with six alternative entries for 2 (of 2 the 24′ is 12′, or 15′60′, or 20′30′, or 30′30′60′ [sic; this is the only example I know of a repeated part], or 45′48′60′80′90′, and one lost entry).

MND 552i: Again headed and dated Mechir 21st, a fragmentary table of 25′ with ie, again with duplicated but largely illegible expressions for the 25′ of 2.

18.    WT.   Louvre [inv.] MND 551dl, in **B. Boyaval, *CRIPEL* 2** (1974), 270–1 and W. Brashear, Trifles, *ZPE* 56 (1984), on 64–5; iv–vi AD. Wooden school tablet with fragmentary table of 3″ in ff.

19.    P.   **Mich. iii 145**, Pack 2309; ii AD. Fragments of tables of 23′ and 29′, and mathematical problems.

20.    P.   **Mich. iii 146**, Pack 2310; iv AD. Long narrow papyrus roll (about 106·75 × 9·2 cm) broken at the beginning, incomplete at the end, containing tables for 7′ (fragmentary), 8′, 9′ (quoted above in Section 7.2(a)), and 10′ in ff, then 11′, 12′, 13′, . . . , 17′, 18′, and 19′ (heading εννεακαιδεκατα to a blank column) in af, all with ie and headings.

21.    P.   **Mich. iii 147**, Pack 2311; early ii AD. Fragments of a table of 4′.

22.    P.   **Mich. xv 686**; ii–iii AD. Multiplication table (see no. 55, below) and fragmentary division tables including a table of 30′ increasing by halves.

23.    O.   Mich. [inv.] 9733, in **H. C. Youtie, A Table of Fractions**, *ZPE* 18 (1975), 17–19; first half of iii AD. Potsherd table of 3″ in ff with ie.

24.    WT.   **(P.) Michael. 62**, Pack 2308; vi AD? Wooden tablet containing tables of 2′, 3″, 3′, and probably 4′ (now illegible) in ff with ie and, on the other side, mathematical problems. More complete publication in D. S. Crawford, A mathematical tablet, *Aegyptus* 33 (1953), 222–40.

25.    WT.   **M.P.E.R., N.S. xv 154**; vii AD. Wooden tablet containing multiplication tables (no. 59, below) and a table of 2′, in ff with ie τὸ ∠ ἑξακιςχιλίων ͵Γ; photograph.

26.    **M.P.E.R., N.S. xv 156**, first published in P. J. Sijpesteijn, Wiener Mélange, *ZPE* 40 (1980), on 97–8, and re-edited in W. Brashear, *Enchoria* 12 (1984), 1–6; v–vi AD. Fragment containing only multiplications, but very possibly coming from a table of 9′ with verifications (or calculations) similar to those in Louvre AF 1196[2] side B and P. Würzburg K 1024 (nos. 16 & 44, here); photograph.

27.    **M.P.E.R., N.S. xv 158**; iii AD? Minute fragment of a table of 3″; photograph.

28.    **M.P.E.R., N.S. xv 159**, first published in P. J. Sijpesteijn, Wiener Mélange, *ZPE* 40 (1980), on 98–9; iii or iv AD. Table of 3″ in ff followed by a calculation with money; photograph.

29.    **M.P.E.R., N.S. xv 160**; ii AD. Tables of 2′ and 3′ in ff; photograph.

30.    **M.P.E.R., N.S. xv 161**; vii–viii AD. Tables of 2′ and 4′ in ff with

ie, but not written out in tabular form—the only such example known to me; photograph.

31.    **M.P.E.R., N.S. xv 162**; first half of ix AD. Fragment of a table of 3″ in ff with ie; photograph.

32.    **M.P.E.R., N.S. xv 163**; vi AD. Fragment of tables of 3′ and 4′; photograph.

33.    **M.P.E.R., N.S. xv 164**; vi–vii AD. Fragment of tables of 3′ and 4′ in ff; photograph.

34.    **M.P.E.R., N.S. xv 165**; vi–vii AD. End of table of 5′ and table of 3″, in ff with ie; photograph.

35.    **M.P.E.R., N.S. xv 166**; ii BC. End of a table of 4′ in ff; photograph.

36.    **M.P.E.R., N.S. xv 167**; ii AD. Fragments of tables of 7′, 8′, and 9′ in ff with headings; photograph.

37.    **M.P.E.R., N.S. xv 168** verso; vii AD. Fragment of a table of 14′; photograph.

38.    **M.P.E.R., N.S. xv 169** recto and verso; viii AD. Fragment of tables of 13′, 14′, 11′, 17′, and 19′; photograph.

39.    WT.   Moen [inv.] 602, in **P. J. Sijpesteijn, A Wooden Tablet in the Moen Collection**, *Chronique d'Égypte* 56 (1981), 97–101; vi or vii AD. Half of a wooden tablet containing part of a table of 7′ in ff with ie.

40.    **P.**   **Oxy. xxxiii 2656**, first published in E. G. Turner, *New Fragments of the Misoumenos of Menander*, Bulletin of the Institute of Classical Studies Supplement 17 (1965), on 18–19; iv or v AD. Tables of 13′, 14′, 15′, 16′, and perhaps 17′ and 18′, in af with ie, headed $\tau\rho\epsilon\iota\sigma\kappa\alpha\iota\delta\epsilon\kappa\alpha\tau\alpha$, etc. These tables are found in a papyrus codex that otherwise carried a lost play of Menander; there would have been space for tables of 10′ to 19′. Only the table of 13′ is transcribed. The republication as P. Oxy. xxxiii 2656 omits the tables.

41.    **P.**   **Oxy. xlix 3456**; iii or iv AD. Tables of 7′ and 8′ in ff, followed by metrological definitions. Only part of each table is transcribed.

42.    WT.   University College [inv.] 36114, Pack 2312, in **H. Thompson, A Byzantine Table of Fractions**, *Ancient Egypt* 1 (1914–15), 52–4. Wooden school tablet containing tables of 15′ and 16′ in af with ie; facsimile. The editor misunderstood the initial entry, but it can clearly be guessed from his drawing and I have checked the original.

43.    WT.   Würzburg [inv.] K 1014, in **W. Brashear, Holz- und Wachstafeln der Sammlung Kiseleff**, *Enchoria* 13 (1985), 13–23 and Tafel 1–13; vi–vii AD. Four writing tablets containing school exercises in multiplication and division. Tables of 6′, 3″, 5′, 3′ (in that order) in ff with ie; photographs.

44.    WT.   Würzbûrg [inv.] K 1024, in **W. Brashear, Neue Griechische Bruchzahlentabellen**, *Enchoria* 12 (1984), 1–6 and Tafel 1–2;

viii AD. Wooden tablet containing tables of 3″ in ff and 11′ in af, both with ie, with verifications (or calculafions) similar to those in Louvre AF 1196² side B and M.P.E.R. 156 (nos. 16 & 26, above); photograph.

45.      WT.    **SB iii 6219**, first published in G. Plaumann, Antike Schultafeln aus Ägypten, *Amtliche Berichte aus den Königlichen Kunstamm-lungen* (Berlin) 34 no. 11 (1913), 210–23, on 222 f; vii AD. Wooden school tablet with tables for 2′ and 3″, in ff with ie, and problems in addition and multiplication; photograph.

### 7.5(b) *Multiplication and addition tables*

Egyptian, demotic:

46.      P.    B.M. [inv.] 10520, Problem 54, in **Parker, *DMP***. List of multiples of 16 from 1 to 16; photograph.

Coptic:

47.      P.    B.M. MS 528 (= no. 7, above). Table, presumably with first page missing, of 7×, 8×, 9×, 10×, 20×, . . . , 900×, in third format.

48.      P.    Cambridge University Library [inv.] T.-S. Ar. 39 : 380 (= no. 8, above). Table of 1×, 2×, . . . , 9×, 10×, 20×, . . . , 90×, 100×, 200×, . . . , 1000× in third format; photograph.

49.      G.    Monastère de Phoebammon 5 (mural no. 10), 153 (mural no. 186), and possibly 6 (mural no. 12), published by R. Rémondon in **C. Bachatly, *Le monastère de Phoebammon dans la Thébaïde***, Tome ii: *Graffiti, inscriptions et ostraca*, Cairo: Publications de la Société d'Archéologie, Copte, 1965, with review by J. Schwartz in *Chronique d'Égypte* 42 no. 83 (1967), 251–4. Tables of 3× and 4× in first format, and an unidentified table; photograph.

50.      O.    **Wadi Sarga 22 and 23** (see no. 11, above). No. 22: tables for 6× and 7× in first format; no. 23: tables for 7× in second format, a palimpsest, with this the younger text.

Greek:

51.      **P.**    **Baden iv 644**, complete publication in F. Bilabel, *Berichtigung-sliste* ii part 2, 177–81. Table of 2×, 3×, . . . , 9×, 10×, 20×, . . . , 90×, 100×, 200×, 300×, 400×, 500× in second format; then 2×, 3×, . . . , 9×, 10×, 20×, 30× in third format.

52.      WWT.    Leiden Papyrological Institute, in **E. Boswinkel, Schulübungen auf 5 Leidener Wachstäfelchen**, *Proceedings of the XIV International Congress of Papyrologists, Oxford 1974*, London: British Academy, 1975, on 25–28; after 212 AD. Five school tablets, one containing a table of 1 × 40 to 10 × 40 in second format.

53.      P.    Lond. iii 737 (p. xxx), published as **M.P.E.R., N.S. xv 150**; iii AD. Table of $n + m$ for $1 \leqslant n \leqslant m \leqslant 9$, then similarly for 10 + 10, . . . , 10 + 90, 20 + 20, . . . , 20 + 90; photograph.

54.    WT.    Louvre [inv.] AF $1196^1$, $1196^2$, and $1196^3$ (= no. 16, above).
AF $1196^1$A: tables for $2\times$, $3\times$, . . . , $8\times$, in second format.
AF $1196^2$A: traces of tables, illegible at beginning, for [$8\times$], $9\times$, $10\times$, $20\times$, . . . , $50\times$, in second format.
AF $1196^3$A: tables of $5000\times$, $6000\times$, . . . , $9000\times$ in first format.

55.    **P.**    **Mich. xv 686** (= no. 22, above). Fragments of a multiplication table up to $1\,0000 \times 4$ in first format.

56.    **M.P.E.R., N.S. xv 151**, the verso of M.P.E.R. i 1; i AD. Addition table in same format as no. 53, above, continuing (with gap for 30+ and 40+) up to $1000 + 1000$; photograph.

57.    **M.P.E.R., N.S. xv 152**; ii–iii AD. Fragments of tables of multiples and squares. The multiplication table gives $7 \times 8$, $7 \times 9$, $7 \times 10$, $8 \times 1, \ldots, 8 \times 10$, $9 \times 1, \ldots, 9 \times 10$, $10 \times 1, \ldots, 10 \times 10$ in first format, followed immediately by the squares (see no. 68, below); photograph.

58.    **M.P.E.R., N.S. xv 153**; ii AD. Fragment of a table of $20 \times 500$, $20 \times 5000$, $20 \times 6$, $20 \times 60$, . . . , $20 \times 9000$ in third format; photograph.

59.    WT.    **M.P.E.R., N.S. xv 154**; vii AD. Wooden tablet with tables of $7\times$, $8\times$, and $9\times$, for $1, 2, \ldots, 9$ in second format, and a table of $2'$ (no. 25, above); photograph.

60.    **M.P.E.R., N.S. xv 157**; ix–x AD. Fragment of tables for $2\times$ and $3\times$ in third format; photograph.

61.    **P.S.I. viii 958**, Pack 2307; iv AD. Table headed πολυπλαcιαcμοc, 'multiplication', starting with some unsystematic entries for smaller numbers, then systematically for $30\times$, $40\times$, . . . , $90\times$, $100\times$, $200\times$, $300\times$, . . . , $1000\times$, $2000\times$, $3000\times$, and $4000\times$ in second format.

62.    WWT.    **SB iii 6215**, first published in G. Plaumann, Antike Schultafeln aus Ägypten, *Amlichte Berichte aus den Königlichen Kunstsammlungen* (Berlin) 34 no. 11 (1913), 210–23, on 216. Wax tablet with addition table, $8 + 1$ to $8 + 9$; photograph.

63.    WWT.    Würzburg [inv.] K1013, in **W. Brashear, Holz- und Wachstafeln der Sammlung Kiseleff**, *Enchoria* 13 (1985), 13–23 and Tafel 1–13; iv–v AD. Parts of five wax tablets containing school exercises including two simple addition tables, second format; photographs.

64.    WT.    Würzburg [inv.] K1014, see previous entry and no. 43 for publication and details. The multiplications are: side 1B, $60 \times 1$, $60 \times 2$, . . . , $60 \times 10$, in first format, then similarly $70\times$, $80\times$, $90\times$, $100\times$, $200\times$, and $300\times$; side 2A, $400\times$, $500\times$, $600\times$, then reverting to $20\times$, $30\times$, $40\times$, $50\times$; Side 3B (fragment), $700 \times 3, \ldots,$ $700 \times 8$, $800 \times 8$, . . . , $900 \times 4$, $1000 \times 4$, . . . , $1000 \times 10$, $3000 \times 1, \ldots, 3000 \times 7$; photographs.

7.5(c)  *Tables of squares*

Coptic:

65.        G.   Monastère de Phoebammon 16 (mural no. 97) (see no. 49, above). Tables of squares for 10, 11, . . . , 29, and 30; photograph.

Greek:

66.        P.   *Un livre d'écolier du III$^e$ siècle avant J.C.*, ed. O. Guéraud & P. Jouguet, Cairo: Institut Français d'Archéologie Orientale, 1938 (= P. Cairo J. E. [inv.] 65445), lines 216–34. Tables of squares of 1, 2, . . . , 10, 20, . . . , 100, 200, . . . , 800; complete facsimile, and photograph in Neugebauer, *ESA*.

67.        P.   **Haun. iii 49**; ii BC? Table of squares for 1, 2, . . . 9, 10, 11, 12, . . . , 19, 20, 30, and 40; photograph.

68.             **M.P.E.R., N.S. xv 152**; ii–iii AD. Multiplication table (see no. 57, above) and two tables of squares, first in the form 1 1, 4 2, 9 3, . . . , 100 10, 400 20, . . . , 6400 80, 1 0000 100, 4 0000 200, . . . , 25 0000 500, then in the form $1 \times 1$ 1, $2 \times 2$ 4,          .          .          .          ,          1 0000 × 1 0000  1 0000 0000

$$\left(\text{written } \begin{matrix} \alpha & \alpha & \alpha \\ \mu & \mu & \mu \\ & \mu & \end{matrix}\right); \text{ photograph.}$$

69.             **P.S.I. vii 763**, Pack 2315; i BC. Fragment of a mathematical lesson, partly in question and answer form. Lines 5 to 11 deal with squares, but not in tabular form: *hapax hen α dis duo δ tris tria ennea θ tetrakis tessara ις pentakis pente κε . . . ,* up to *dekakis deka ρ.*

This catalogue has now been extended to Item no. 107 in Fowler, 'Further arithmetical tables'.

# PART THREE

---

# LATER DEVELOPMENTS

Doubtless a vigorous error vigorously pursued has kept the
embryos of truth a-burning: the quest for gold being at the
same time a questioning of all substances.
George Eliot, *Middlemarch*, Chapter 48.

The topic of this third part may be paraphrased as ancient and modern
reactions to anthyphairesis. In Part One, I proposed that early Greek
mathematicians made remarkable investigations and discoveries about ratio,
especially anthyphairetic ratios; but when we look to the later Greek com-
mentators to find their reactions to this work, we draw a blank. So, in Chapter
8, as an illustration of what we do find, I take three themes that are discussed in
later antiquity, and set them against the early evidence and my interpretation
here. I would have liked to have included discussions of several other topics
such as analysis and synthesis, the nature of the early *Elements* (*stoicheia*), the
role of problems, the deductive method, and so on, and I do not even begin
the evaluation of the effects of my proposals on the chronology of the period;
but restrictions of space impose, once again, their limits. The purpose of
these analyses is to illustrate the quite serious discrepancies between the
preoccupations of later commentators and the earlier evidence, such as we
now possess.

With the exception of Pappus and Eutocius, the later Greek commentators
were not mathematicians. By contrast, many Arab commentators of the
eleventh and twelfth centuries were mathematicians in their own right, and I
would also have liked to have included a discussion of four known Arab
commentators who do refer to anthyphairesis. Unfortunately the time is not
yet ripe: more work needs to be done on the manuscripts before one can do
more than give a précis of published summaries of the topic. Meanwhile I
commend the interested to the discussion and translation in Plooij, *ECR*, and
to the translation of Omar Khayyam–Amir-Móez, DDE.

Chapter 9 gives detailed proofs of the main mathematical results underlying
the reconstruction, and sketches the history of their development since the
seventeenth century. Although it is highly mathematical and is presented in the
modern idiom, and its historical content relates only to the period from the
seventeenth century up to today, it is far from irrelevant to my historical
argument about Greek mathematics. For we can see how anthyphairesis, in this

new manifestation of continued fractions, intrigued and inspired many of the most powerful and influential mathematicians of the seventeenth, eighteenth, and nineteenth centuries. yet if we take a modern textbook of mathematics or a general history of mathematics (our latter-day equivalents of Pappus' *Collection* and Eudemus' *History*, respectively) and search for any hint of the richness of this theory or its historical importance, we are very very unlikely to find anything. A satisfactory explanation would perhaps lead even further afield, into a study of the psychology and sociology of mathematicians, but some awareness of this modern void should take the edge off our surprise about this lack of comment in late antiquity. Chapter 10 consists of new material for this second edition, and the Epilogue looks back on my personal involvement in this story.

# LATER INTERPRETATIONS

## 8.1 EGYPTIAN LAND MEASUREMENT AS THE ORIGIN OF GREEK GEOMETRY?

It has been an oft-repeated story since antiquity that the origins of Greek geometry are to be found in Egyptian land measurement; for a discussion of some of the texts, see Heath, *HGM* i, 121–8. The earliest such possible reference comes from Herodotus, *Histories* ii, 109, written in the fifth century BC:

This king [the semi-mythical Sesostris] moreover (so they said) divided the country among all the Egyptians by giving each an equal square parcel of land, and made this his source of revenue, appointing the payment of a yearly tax. And any man who was robbed by the river of a part of his land would come to Sesostris and declare what had befallen him; then the king would send men to look into it and measure the space by which the land was diminished, so that therefore it should pay in proportion to (*kata logon*) the tax originally imposed. From this, to my thinking, the Greeks learnt the art of *geōmetrica*; the sunclock and the sundial (*polis kai gnōmōn*), and the twelve divisions of the day, came to Hellas not from Egypt but from Babylonia [Herodotus–Godley, *H* i, 397–9].

(The "equal square parcel of land" here should not be taken too literally, and the "sunclock and sundial" is more probably the hemispherical bowl and pointer that make up an early sundial; see Goldstein & Bowen, NVEGA, 332 n. 9.) Another version of this story is given in Proclus' *Commentary on the First Book of Euclid's Elements*, written and compiled from various source materials a thousand years later:

We say, as have most writers of history, that *geōmetria* was first discovered among the Egyptians and originated in the remeasuring of their lands. This was necessary for them because the Nile overflows and obliterates the boundary lines between their properties. It is not surprising that the discovery of this and the other sciences had its origin in necessity, since everything in the world of generation proceeds from imperfection to perfection. Thus they would naturally pass from sense-perception to calculation (*logismos*) and from calculation to reason. Just as among the Phoenicians the necessities of trade and exchange gave the impetus to the accurate study of number, so also among the Egyptians the invention of *geōmetria* came from the cause mentioned. Thales, who had travelled to Egypt, . . . [Proclus–Morrow, *CFBEE*, 64–5],

and there now follows the celebrated summary of early Greek mathematics that is believed to derive from Eudemus. However, this introduction shows clear neo-Platonic characteristics and we have no reason to doubt that it was written by Proclus himself.

Let us confront this story with what we know about land measurement in Ptolemaic Egypt, at a period when records become available. In view of what

we shall find, it is worth noting immediately that Aristotle proposes a different explanation at *Metaphysics* 981b23, in the context of his critical history of the development of philosophy:

Thus the mathematical sciences originated in the neighbourhood of Egypt, because there the priestly class was allowed leisure,

while Plato, at *Phaedrus* 274c–275b (quoted above in Section 1.2(e)), has Socrates relate the myth that geometry was one of the arts invented by the Egyptian god Theuth. To this Phaedrus replies:

It's easy for you, Socrates, to make up tales from Egypt or anywhere else you fancy.

The style of early Greek mathematics was described in Section 1.2(b), where *Elements* I 35:

Parallelograms which are on the same base and in the same parallels are equal to one another,

was taken as a characteristic example. Proclus says of this proposition:

It may seem a great puzzle to those inexperienced in this science that the parallelograms constructed on the same base [and between the same parallels] should be equal to one another. For when the sides of the areas constructed on the same base can be increased indefinitely . . . we may well ask how areas can remain equal when this happens. For if the breadth is the same (since the base is identical) while the side becomes greater, how could the area fail to become greater? This theorem, then, and the following one about triangles belong among what are called the 'paradoxical' theorems in mathematics [Proclus–Morrow, *CFBEE*, 396].

It may seem a great puzzle to us that Proclus should express such surprise over this result and go on to discuss it and the succeeding propositions at such great length, when he is able to regard, for example, the fact that the angle sum of a triangle is equal to two right angles (see *Elements* I 17 and his commentary on it) as reasonably straightforward.

But we can understand something of this surprise when we consider the evidence of Egyptian land surveying, as was described and illustrated in Section 7.1(d). If the nominal area of a field is taken to be the product of the average of the lengths of opposite sides, then there is something to reconcile between the propositions of the mathematicians and the practice of the accountants and tax collectors. It might be objected that this evidence relates to Egyptian, not Greek, practice, and that we have no evidence of land surveys in Greece. But Proclus, born in Lycia, resident in Athens, goes on to tell a story that indicates that it was not unknown to the Greeks:

The participants in a division of land have sometimes misled their partners in the distribution by misusing the longer boundary line; having acquired a lot with a longer periphery, they later exchanged it for lands with a shorter boundary and so, while getting more than their fellow colonists, have gained a reputation for superior honesty [Proclus–Morrow, *CFBEE*, 403].

Herodotus gives another example in which size was measured by perimeter: three times he called Sardinia "the biggest island in the world" (*Histories* i, 170, v, 106, and vi, 2) when, even in Herodotus' world, Sicily had a greater area though a smaller perimeter. (Sardinia: 9187 square miles and 830 miles circumference *v.* Sicily: 9860 square miles and 680 miles circumference; see Rowland, BIW for further literary and geographical details.) Thucydides indicates the obvious way in which such perimeters are estimated by a maritime nation; at the beginning of Book 6 of his *History of the Peloponnesian War* we read:

The voyage round Sicily takes rather under 8 days in a merchant ship, yet, in spite of the size of the island. . . .

Of course, the mathematicians whose work was later compiled in the *Elements* knew better. But, within a tradition of non-arithmetised geometry and given Plato's attitude to practical mathematics (see Section 2.2 and the end of my S35, above), the issue of land measurement would have held little interest for them. Heron, who worked in the later tradition of arithmetised, metrical geometry that arose as a combination of Greek geometrical and Babylonian arithmetical mathematics, describes correctly how to calculate the precise numerical area of irregular quadrilaterals at *Metrica* I, 14–16.

Did, then, the Greeks discover geometry from Egyptian land measurement? We have so little reliable information about this question that we shall probably never be able to give a sure answer, but I very much doubt the story. Even Plato and Aristotle do not seem to know enough to be able to go beyond myths and generalities. And go back and look again at the texts quoted above: Herodotus only says it if the Greek word *geōmetria* is translated as 'geometry' (as in Heath's discussion, at *HGM* i, 121–8) rather than 'land measurement' (as, for example, in the translation of Godley, in the Loeb Classical Library edition, which I have used here); the same ambiguity is found in the French *géomètre*. In Proclus, the word is surely correctly translated as 'geometry', but in line 2 of O. Bodl. ii 1847 (see Plate 8 and Section 7.1(d)) it clearly must refer to land measurement. So, for example, Herodotus could be contrasting the way that the Greeks learned the measurement of space from Egypt and the measurement of time from Babylon. As to any proposal that the Greeks realised that the basic formula of Egyptian land measurement was wrong, and discovered geometry by trying to correct it, we have no evidence to justify this interpretation, and it runs counter to what evidence we do have.

It may be of interest to describe briefly the medieval European tradition of land surveying. Elaborate surveys of large estates were being made in France in the ninth century, and a survey of unparalleled scope and completeness covering practically the whole of England was made by order of William the Conqueror in 1086: the Domesday Survey. These are usually described as land surveys (see, for example, the *OED*, s.v. Domesday), but the precise measurement of land plays little or no part in them and they are more accurately

described as surveys of actual or potential sources of revenue. Take, for example, the Domesday Book, which embodies the result of the Domesday Survey. There are regional variations in the entries; this having been said, a typical entry reads:

The Bishop of Worcester holds Hantone [Hampton Lucy]. There are xii hides. There is land for xxii ploughs. ii are in the demesne, and iiii serfs. And there are xxii villeins and ix bordars with a priest who have xxiiii ploughs. There is a mill worth vi shillings and viii pence, and xv furlongs of meadow in length and one furlong in breadth. In Warwick iii houses worth xvi pence. Wood i league long and another broad. In King Edward's time it was worth iiii pounds, and afterwards the same, now it is worth xx pounds. The same bishop holds and held Stratforde. There are xiv and a half hides . . . [translation adapted from *Victoria County History*, Warwickshire i, p. 302].

The land is described as being for so many ploughs; often woods are described as for so many pigs, while in this extract the 'league' used to describe the wood probably had no precise measured length and the wood itself is very unlikely to have been rectangular, or even four-sided. The 'hide' is a unit of assessment, divided into four 'virgates', each of which contained thirty 'acres'; but these were fiscal units, not measures of area, and the acre was the smallest unit used to describe the tax on a holding. There are no maps.

During the thirteenth and early fourteenth centuries, local standards of linear measure were gradually introduced and applied in England, but they were not generally converted into measurements of area; nor could they be, in the absence of any precise way of measuring angles; nor did they need to be in the context of medieval farming, in which valuable arable land was cultivated in narrow strips and other small parcels, mostly rectangular. No further new techniques were applied to surveying during the fourteenth and fifteenth centuries, and there is no trace of the growth of a class of professional surveyors. The earliest English printed treaties on surveying (by Sir Anthony Fitzherbert in 1523 and Richard Benese in 1537) show little advance on thirteenth-century practice, but new methods, particularly of measuring angles, were developed in Germany and the Netherlands, and reached England through books printed in the 1550s. However, two points about these works should be emphasised: first, as some of the writers themselves remark, there is an enormous gap between the theory and practice of surveying; and, second, the increasing accuracy was probably directed towards the new techniques of map-making. Up to the mid-sixteenth century, the result of any kind of survey would be written text, much in the medieval tradition of the Domesday Book; even during the reign of Elizabeth I, most estate surveys were of this form, and very few estate maps survive from before 1550. However, from the 1580s onwards, the results of detailed surveys began to be embodied in maps, without any accompanying written text. There was then an explosion in the use of maps in general, and, with it, a new phase of development began.

The development of the linear measurement of land is described in Jones, LME, and a history of map-making and description of all surviving medieval

English local maps, in Harvey, *HTM* and Harvey & Skelton, *LMPME*. I would like to thank Paul Harvey for help; this summary is almost entirely based on information provided by him.

## 8.2 *NEUSIS*-CONSTRUCTIONS IN GREEK GEOMETRY

It is widely believed that Greek geometry was restricted to (unmarked) ruler-and-compass constructions. In fact, our evidence is much more diverse than this, and the repertoire of basic constructions should certainly also include the *neusis*-construction, in which a line is drawn 'verging' or 'inclining' (*neuein*) towards—that is, passing through—a given point and intercepting two given curves or straight lines in a given line segment. In vulgar terms, one might imagine putting a nail at the given point, marking off the given segment on a ruler, and then rotating the ruler against the nail until the intercept between the two given curves is this marked segment. An example will illustrate clearly what is involved.

Our earliest detailed description of a piece of Greek geometry comes at third hand. It is found in Simplicius' commentary on Aristotle's *Physics*, where Simplicius quotes, with his interpolations, a passage from Eudemus' *History of Mathematics* which describes Hippocrates' *Quadrature of Lunes*. (A lune is a plane region bounded by two arcs of circles; Hippocrates described three such regions that can be made equal to rectilinear plane figures, and hence can be squared.) For a discussion of the historical, textual, and mathematical aspects of this passage, see Heath, *HGM* i, 183–200. A purged version of the text is given in Thomas, *SIHGM* i, 234–53, from which the following extract is taken.
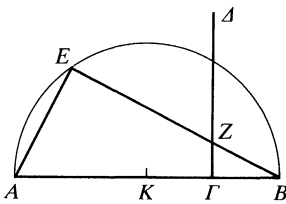


If [the outer circumference of the lune] were less than a semicircle. Hippocrates solved (*kateskeuasen*, constructed) this also, using the following preliminary construction. Let there be a circle with diameter $AB$ and centre $K$. Let $\Gamma\Delta$ bisect $BK$ at right angles; and let the straight line $EZ$ be placed between this and the circumference verging (*neuousa*) towards $B$ so that the square on it is one-and-a-half (*hēmiolos*, half-whole) times the square on one of the radii. . . .

In other words, draw the line $EZB$, through $B$ and such that $EZ$ is equal to the specified line.

This particular construction can in fact be accomplished by ruler-and-compass methods, as follows. For any line $EB$, verging towards $B$, the angle $E$ in the semicircle will be a right angle, so the triangle $AEB$ and $Z\Gamma B$ will be similar. Hence $(BZ + EZ).BZ = AB.B\Gamma = r^2$, where $r$ is the radius of the semicircle. Thus, if so required, we can perform the following

ruler-and-compass construction. First construct, separately, a line $E'Z'$ such that $E'Z'^2 = \frac{3}{2}r^2$ (using the techniques of *Elements* II 14), and a line $B'Z'$ such that $B'Z'^2 + B'Z' \cdot E'Z' = r^2$. (This is the elliptic application of areas discussed in Section 5.2(d), and performed in Fig. 5.4(c). Alternatively a construction can be manufactured around the configuration of *Elements* II 6; see Heath, *TBEE* i, 386 ff. for details.) If we then locate $Z$ on $\Delta\Gamma$ so that $BZ = B'Z'$, we shall have $EZ^2 = \frac{3}{2}r^2$, as required. However, neither Hippocrates nor Simplicius makes any attempt to provide such a ruler-and-compass justification of this *neusis*-construction, even though the text elsewhere gives elaborate descriptions of matters that are far less obvious than this. The *neusis*-construction seems to be invoked as a permitted construction in its own right.

Each *neusis*-construction needs some kind of verification that a solution is possible. For example, here in Hippocrates' *Quadrature*, we need to observe that the line $AB$ will intercept a segment $A\Gamma$ equal to $\frac{3}{2}r$, greater than $\sqrt{(\frac{3}{2})}r$, while a line through the intersection of $\Gamma\Delta$ and the circle will have a zero intercept. So, by some principle of continuity, an intermediate line with the required intercept will exist. Such an implied condition is analogous to the *diorismos* in the formal division of a Euclidean proposition (see the description in Proclus–Morrow, *CFBEE*, 203 f. quoted in Heath, *TBEE* i, 129 f. and the examples in *Elements* I 22 and VI 28), but no such condition is, to my knowledge, considered in any surviving *neusis*-construction. Nor do I know of any ancient discussion of the underlying principle of continuity, in any context; see, for example, how *Elements* I 1 assumes without comment that the two circles will intersect—but the circles will intersect only if their radii satisfy some implied *diorismos*; and then, here again, something like the principle of continuity is also needed.

Archimedes uses *neusis*-constructions freely in his *Spiral Lines*: the construction appears in Propositions 5, 6, 7, 8, and 9, and some of these propositions play crucial roles later in the book. This time the *neusis*-construction in Propositions 5, 6, and 7 cannot be carried out by ruler-and-compass methods. This has provoked comment and criticism from Pappus, in late antiquity, that Archimedes was using methods inappropriate to the problem (see below), and ancient and modern commentators have shown how Archimedes' material can be reworked by reformulating the later propositions of the treatise which depend on these *neusis*-constructions so that the properties of the spiral can be established differently, by ruler-and-compass methods; see Archimedes–Heath, *WA*, Chapter 5; Heath, *HGM* ii, 556–61; and Knorr, *ANCSL* for examples, discussions, and references.

Proposition 8 of another work loosely attributed to Archimedes, the *Book of Lemmas*, which survives only in an Arabic translation, describes a geometrical configuration that can immediately be used to give a *neusis*-construction for trisecting an angle. The statement of this proposition is:
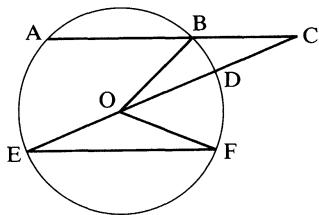
If AB is any chord of a circle whose centre is O, and if AB be produced to C so that BC is equal to the radius; if further CO meet the circle in D and be produced to meet the circle a second time in E, the arc AE will be equal to three-times the arc BD.

The proof is straightforward: if EF is parallel to AB then, since OE = OF = OB = BC, we easily see that the angle DOF is twice BOD, from which the result follows. We can use this configuration to trisect an angle as follows: set up the figure so that AOE is the angle to be trisected and EOD is the diameter produced beyond D. Then use a *neusis*-construction to insert the line CB, verging towards A, between the circle FDB and the extension of line BC; BOD is then one-third of AOE. This construction cannot now be performed by ruler-and-compass methods, though it can be related to construction involving conic sections. For discussion of the *Book of Lemmas* and this kind of construction, see Archimedes–Heath, *WA*, pp. xxii f. and 301–18, and Hogendijk, HTATGIG.

Apollonius, a younger contemporary of Archimedes, wrote a work *On Neuses*, in two books. This is lost, but we have a description of Pappus which seems to indicate that it was an investigation of the techniques that would be needed to carry out different kinds of *neusis*-constructions. See Heath, *HGM* ii, 189–92 for a description.

Nicomedes, a younger contemporary of Apollonius, defined the conchoid, a curve based on a *neusis*-construction, which he used to provide a construction for duplicating the cube; the procedure is described both in Pappus' *Collection* and Eutocius' *Commentary on Archimedes' Sphere and Cylinder* and is excerpted in Thomas, *SIGHM* i, 296–309. The conchoid is the locus of points which, along with a given straight line, cut off a given segment on all lines which verge towards a given point, and it can then be used to determine a particular *neusis*-construction. Nicomedes, as reported by Eutocius, describes in some detail an instrument for drawing the conchoid on the basis of which it is possible to reconstruct, with some confidence, his arrangement of sliding slotted rulers; see Archimedes–Heiberg, *Opera* iii, 99, or Heath, *HGM* i, 239.

In broad terms, we see three phases of development. The early geometers, from Hippocrates to Archimedes, seem to have used the *neusis*-construction freely, as a basic construction of mathematics. This was followed by a period, represented by Apollonius and Nicomedes, in which the scope and implications of the construction were investigated, and its relation with other constructs was explored. Finally, in late antiquity, we find Pappus laying down formal criteria that the use of constructions must satisfy. In this last phase, Pappus does not mince his words in his explicit criticism of Archimedes and Apollonius:

It seems to be a grave error into which geometers fall whenever anyone discovers the solution of a plane problem by means of conics or linear curves, or generally solves it by means of a foreign kind, as is the case, for example, (i) with the problem in the fifth Book of Conics of Apollonius relating to the parabola, and (ii) when Archimedes assumes in his work on the spiral a *neusis* of a solid character with reference to a circle; for it is possible without calling in the aid of anything solid to find the [proof of] the theorem given by the latter [sc. Archimedes], that is, to prove that . . . [Archimedes–Heath, *WA*, pp. ciii f. or Heath, *HGM* ii, 68 or Thomas, *SIHGM* i, 246–53].

I shall not digress here to describe and explore Pappus' description of plane, solid, and linear constructions; nor shall I discuss the status of other kinds of constructions such as Archytas' duplication of the cube; nor shall I describe how tenuous is the evidence for the criticism that Plato is said to have directed against the use of mechanical constructions in geometry; nor shall I describe the other evidence we have for more exotic mechanical devices and constructions in geometry, such as the device attributed by Eutocius to Plato for duplicating the cube, or the mysterious 'horn ruler' invoked in Proposition 4 of Diocles' *On Burning Mirrors*. (On these matters see the references already cited; Niebel, *UBGKA*; Steele, *URZLGM*; and the further references contained therein.) I shall take the well-founded example of the *neusis*-construction as an illustration of the proposal, accepted by most modern commentators, that more general kinds of constructions than ruler-and-compass constructions were admitted freely into early Greek geometry (see Knorr, *ANCSL*, 86–9 for a discussion), and will now turn to a consideration of Euclid's *Elements*.

Although the *Elements* is set in the middle of the period when *neusis*-constructions were freely used, yet we find there no reference to or use of them. Let me dispose first of one common explanation of this, that the *Elements* provides only an introduction to the basic techniques of mathematics, preparing the way for further treatises that will deal with the more advanced problems where *neusis*-constructions will find their proper place. First, to describe the *Elements* as restricted to basic mathematics seems to me to be a short-sighted judgement based only on the contents of the first few books. The material in Books X, XII, and XIII can scarcely be called straightforward, basic, or easy; see, for example, the discussion in Chapter 5, above. Next, if the aim of the *Elements* were pedagogical, with Euclid the supreme teacher (but see Section 5.2(a) for damning comments on Book X, about a quarter of the bulk of the *Elements*), then surely his treatment would anticipate the later enlargement of the geometrical repertoire, and would prepare the reader for the introduction to more general constructions. But the *Elements* does precisely the opposite: in Book I, Propositions 1–3, we find the most restricted possible interpretation of the Postulates, in which the notional, unmentioned 'compass' (see Postulate 3) cannot be used as dividers, to move around freely the endpoints of a line. Far from anticipating a relaxation of the scope of geometrical constructions, this sets the scene for the most severely constrained interpretation of Euclidean compasses. (A possible explanation of these propositions is, I feel,

to set up a complete theory of transformation and application of lines founded on the most basic interpretation of the first three postulates, as a model for the later theory of application of rectilinear areas and for the unachieved theory of application of volumes.) Further, the argument that *neusis*-construction belongs only to the more advanced parts of geometry does not stand close scrutiny. See, for example, Aristotle, *Posterior Analytics*, I 10, 76a30 f.:

> By first principles in each genus I mean those the truth of which it is not possible to prove. What is *denoted* by the first [terms] and those derived from them is assumed; but, as regards their *existence*, this must be assumed for the principles but proved for the rest. Thus what a unit is, what the straight [line] is, or what a triangle is [must be assumed]; and the existence of the unit and of magnitude must also be assumed, but the rest must be proved. Now of the premisses used in demonstrative sciences some are peculiar to each science and others common [to all], the latter being common by analogy, for of course they are actually useful in so far as they are applied to the subject-matter included under the particular science. Instances of first principles peculiar to a science are the assumptions that a line is of such-and-such a character, and similarly for the straight [line]; whereas it is a common principle, for instance, that, if equals be subtracted from equals, the remainders are equal. But it is enough that each of the common principles is true so far as regards the particular genus [subject-matter]; for [in geometry] the effect will be the same even if the common principle be assumed to be true, not of everything, but only of magnitudes, and, in arithmetic, of numbers.
>
> Now the things peculiar to the science, the existence of which must be assumed, are the things with reference to which the science investigates the essential attributes, e.g. arithmetic with reference to units, and geometry with reference to points and lines. With these things it is assumed that they exist and that they are of such-and-such a nature. But, with regard to their essential properties, what is assumed is only the meaning of each term employed: thus arithmetic assumes the answer to the question what is [meant by] 'odd' or 'even', 'a square' or 'a cube', and geometry to the question what is [meant by] 'the *alogon*' or 'deflection' (*keklasthai*) or [the so-called] 'verging' (*neuein*) [to a point]; but that there are such things is proved by means of the common principles and of what has already been demonstrated. Similarly with astronomy. For every demonstrative science has to do with three things, (1) the things which are assumed to exist, namely the genus [subject-matter] in each case, the essential properties of which the science investigates, (2) the common axioms so-called, which are the primary source of demonstration, and (3) the properties with regard to which all that is assumed is the meaning of the respective terms used. There is, however, no reason why some sciences should not omit to speak of one or other of these things [Aristotle–Heath, *WA*, 50 f.; there is a longer extract in Thomas, *SIHGM* i, 418 ff].

Aristotle's reference to *neusis*-construction as an essential attribute for study is all the more relevant when we consider that he probably learned his mathematics during his nearly twenty-year association with the Academy, up to Plato's death, exposed to the ideas of Theodorus, Archytas, Theaetetus, and Eudoxus. Thus this passage gives yet further direct evidence of the basic role of *neusis*-constructions in geometry during the pre-Euclidean period.

Consider the constructions invoked by Euclid. First, the scope of the *Elements* is not restricted to ruler-and-compass constructions. While the mean, third, and fourth proportionals for lines are constructed in detail by

(implied) ruler-and-compass methods in Book VI 11–13, there are not, and there cannot be, similar constructions to back up the use of the fourth proportional to two circles and a square in XII 2, and to various three-dimensional figures in XII 5, 11, 12, and 18. While these latter fourth proportionals do not have ruler-and-compass constructions, their existence is assumed, without any constraints. (A separate issue, which I have set to one side above and shall continue to ignore here, is how one can talk of three-dimensional figures being 'constructed' in any sense of the word; here the same kinds of difficulties that apply to, for example, Archytas' duplication of a cube are also found throughout the construction of the regular polyhedra in *Elements* XIII.)

We can perhaps discern a later historical stratum: those parts of the *Elements* that invoke these more general kinds of constructions might be associated with the innovations due to Eudoxus. Even if this is indeed the case, our question still applies in a modified form: why do the earlier, pre-Eudoxan parts of the *Elements* make no reference to *neusis*-constructions?

An answer to this can only be very speculative. With this qualification, and viewing the issue from the vantage point of the reconstruction proposed here, we see how the basic geometry that is developed in Books I, III, VI, and XI is applied, in Books II, IV, X, and XIII, in contexts which have very strong anthyphairetic overtones. I have illustrated throughout Chapters 2 and 5 and will go on to describe in Chapter 9 in more detail how quadratic problems and only quadratic problems are amenable to anthyphairetic analysis. The geometrical translation of what I have here loosely called 'quadratic problems' is 'ruler-and-compass constructions'; so, I propose, the bias of the *Elements* towards ruler-and-compass constructions and its exclusion of *neusis*-constructions may be a consequence of its anthyphairetic origins.

As the anthyphairetic preoccupations faded away, as new techniques were introduced, particularly by Eudoxus, as the geometrical language that had been developed was applied to wider and wider contexts, particularly by Archimedes and Apollonius, so would these underlying constructions move into prominence and be of interest in their own right. There would then be a tendency towards an exploration and classification of the underlying geometrical procedures, as in what I called above the second and third phases of development. It is a commonplace in mathematics for the initial developments to be prompted by some quite specific problem which then fades away; that which is left, if successful, then often acquires an autonomous existence as a 'theory' in its own right. The eighteenth- and nineteenth-century investigations of a vibrating string, which provoked the creations of analysis and eventually Cantor's set theory, give one good such example; many more could be cited without difficulty. Thus the apparent concern for ruler-and-compass constructions in the *Elements* may well be the influence of the language developed around the successful treatment of anthyphairetic phenomena, and the absence of higher degree constructions may reflect the paucity, even today,

of any anthyphairetic results about anything other than linear or quadratic phenomena.

## 8.3   THE DISCOVERY AND ROLE OF THE PHENOMENON OF INCOMMENSURABILITY

### 8.3(a)   *The story*

Part of every literate person's intellectual baggage, along with the second law of thermodynamics and the principles of relativity and indeterminacy, is some version of the story of the discovery of incommensurability by Pythagoras or the Pythagoreans and of the role this discovery played in Greek mathematics. Here, to end this last chapter on Greek mathematics, I shall describe and discuss our evidence on this topic. I shall continue to take pains to maintain the perspective of this book, and so will distinguish between that evidence which comes, or purports to come, from the early period, up to the time of Euclid and Archimedes, and that found only in later sources.

I shall only give a summary treatment, and direct the reader to the account in Knorr *EEE*, Chapter 2, for further details and references. I will not attempt to deal further with the philological issues concerning the words (*a*)*summetros*, (*ar*)*rhētos*, (*a*)*logos*, *diametros*, and *diagōnios*, etc., beyond the comments given earlier and listed in the Index. I start by reviewing the treatment of the subject in authors up to Proclus, in late antiquity. Every reference to anything that may be connected with the topic will, I hope, be included somewhere in this list, and I shall also include occasional examples of the many important witnesses whose surviving writings do not mention the topic.

### 8.3(b)   *The evidence*

(i)   Democritus (fl. *c.* 460 BC) wrote a treatise in two books called *Peri alogōn grammōn kai nastōn*, 'On unreasonable/unutterable/irrational/ . . . lines and solids'. Our source is Diogenes Laertius, *Lives of the Eminent Philosophers*, IX 47 (Diels & Kranz, *FV* ii, Fragment A33, p. 91.20 = Fragment B11p, p. 141.24; Diogenes–Hicks, *LEP* ii, 458–9). Diogenes is a gossipy and uncritical compiler of anecdotes, perhaps of the third century AD, though we know nothing with confidence of his dates and origins. Here he is quoting, perhaps accurately, a catalogue of titles of Democritus' works by Thrasyllus of Alexandria (died AD 36), who also catalogued and arranged Plato's dialogues. We know of Democritus' ideas only through secondary reports and brief quotations, and none of the ninety treatises in this catalogue, twelve of them described as mathematical, have survived. In particular, we know nothing about the contents of this treatise, and it is not even certain that it had anything to do with the topic of incommensurability. A conspectus of scholarly attitudes to Democritus' mathematics can be found in Guthrie, *HGP* ii, 484–8; for a succinct assessment see the *OCD*, s.v. Democritus: "Little is known (though much is written) about the mathematics of Democritus."

(ii) Plato: Our first unequivocal and explicit reference to incommensurability is in Plato, at *Theaetetus* 147d–148b, where the topic is handled confidently in the context of Theodorus' geometry lesson to the young Theaetetus. This passage gives no indication that the phenomenon itself posed any fundamental conceptual difficulty for mathematicians; rather, it is presented as a source of interesting and fruitful problems. There is a playful metaphorical allusion at *Republic* VII, 534d; it enters the construction of the nuptial number, *Republic* VIII, 546b–d, in a passage which prompts the later descriptions of side and diagonal numbers in the commentaries of Theon of Smyrna and Proclus; it is discussed at *Parmenides* 140b–d; and there is a reference in the possibly spurious *Hippias Major*, 303b–c, to sums of incommensurable lines that is reminiscent of the classification of *Elements* X. Finally, in Plato's last dialogue, at *Laws* VII 817e–820e, there is a celebrated passage that talks of "Ignorance [which] seemed to me worthy not of human beings but of pigs, and I felt ashamed not for myself alone but for all the Greeks". The context is a curriculum similar to that of *Republic* VII, but at a lower level, for the basic education of the citizens of the state. The Athenian Stranger talks not only of the measurability of line with line, surface with surface, and volume with volume, but also:

Again, what of the relations of surface and line to volume, or of surface and line to one another; do not all we Greeks imagine that they are measurable (*metrēta*) in some way or another. . . . Then if this is absolutely impossible, though all we Greeks, as I was saying, imagine it possible, are we not bound to blush for them all as we say to them, "Worthy Greeks, this is one of the things of which we said that ignorance is a disgrace and that to know such necessary matters is no great achievement" [820a–b; translation from Thomas, *SIHGM* i, 20–7].

The matter at issue here is not simply the topic of commensurability and incommensurability of homogeneous magnitudes, and it may instead be related to the everyday techniques of measuring and comparing areas that I described in Sections 7.1(d) and 8.1, above.

(iii) One of Aristotle's favourite mathematical illustrations is the incommensurability of the diagonal, which he cites about thirty times, at *Prior Analytics* 41a23–30, 46b26–37, 50a37–8, 65b16–21; *Posterior Analytics* 71b25–6, 89a29–32; *Topics* 106a38–b1, 163a11–13; *Sophistical Refutations* 170a25–6, 176b20; *Physics* 221b23–25, 222a3–7; *De Caelo* 281a6–7, b5–6, b12–14; *De Anima* 430a30–1; *Generation of Animals* 742b27–28; *Metaphysics* 983a12–21, 1012a31–3, 1017a34–5, 1019b24–7, 1024b19–21, 1047b6–12, 1051b20–1, 1053a17–8; *Nicomachean Ethics* 1112a21–3; *Eudeman Ethics* 1226a2–4; and *Rhetoric* 1392a15–8, while in a few other places such as *Metaphysics* 1004b11–12, 1021a3–6 (see Section 2.1 above), and 1061b1, he introduces (in)commensurability in other contexts. A typical example occurs in *Metaphysics* 983a12–20:

Yet the acquisition of it [sc. knowledge, *epistēmē*] must in a sense end in something which is the opposite of our original inquiries. For all men begin, as we said, by

wondering that things are as they are, as they do about self-moving marionettes, or about the solstices or the incommensurability of the diagonal; for it seems wonderful to all who have not yet seen the reason, that there is a thing which cannot be measured even by the smallest unit. But we must end in the contrary and, according to the proverb, the better state, as is the case in these instances too when men learn the cause; for there is nothing which would surprise a geometer so much as if the diagonal turned out to be measurable (*metrēta*).

(This and the other translations from Aristotle have been taken from the edition of W. D. Ross, and they have all been adapted slightly as will be explained below. I have here also adjusted the translation of the last word, which is the same as that found at Plato's *Laws* VII, 819e–820c, quoted above.) Often Aristotle's reference is very short and stereotyped: *estin he diametros asummetros*, 'the diagonal is incommensurable'. Sometimes he adds bits of extra perplexing information, as at *Metaphysics* 1053a14–20:

But the measure is not always one in number—sometimes there are several; e.g. the quarter-tones (not to the ear, but as determined by the ratios) are two, and the articulate sounds by which we measure are more than one, and the diagonal and its side are measured by two quantities, and all spatial magnitudes reveal similar varieties of unit. . . .

Here there is variation between the manuscripts that makes the original text unsure; the critical editions should be consulted for details. But one point seems clear from the grammar of the undisputed text of the Greek, and it seems to be supported by the example of the quarter-tones: that the meaning is not that there is one measure for the side, and another for the diagonal. The translation should better read "the diagonal *is* measured by two quantities *and [so is] the side . . . * . . . ", where the asterisks indicate the uncertain text.

In the large majority of these examples where Aristotle writes of 'the diagonal', he does not mention 'the side,' *hē pleura*. The only places where these words occur are *Topics* 106a38, 163a13; *Physics* 221b25; *Generation of Animals* 742b28; *Nicomachean Ethics* 1112a23; and the doubtful passage of *Metaphysics* 1053a18, just cited. However, there seems no doubt that 'the side' is always understood, as passages such as *Topics* 163a11–13 make clear.

For example if he had to show that the diagonal is incommensurable with the side, and were to beg that the side is incommensurable with the diagonal. . . .

More curious is that Aristotle nowhere says explicitly that he is talking of a square. I have excised these words 'the side' and 'of the square' from the quoted translations whenever they are not found in Greek text.

Three examples, all in the *Prior Analytics*, give more information about Aristotle's method of proof. One of these, at 65b16–21, is of doubtful use for further reconstruction since it is an example of an incorrect proof:

The most obvious case of the irrelevance of an assumption to a conclusion which is false is when a syllogism drawn from middle terms to an impossible conclusion is independent of the hypothesis, as we have explained in the *Topics*. For to put that which is not the cause as the cause, is just this: e.g. if a man, wishing to prove that the diagonal is

incommensurate, should try to prove Zeno's theorem that motion is impossible, and so establish a *reductio ad impossibile*: for Zeno's false theorem has no connexion at all with the original assumption.

The other two passages give more details of the argument. At *Prior Analytics* 41a23–30 he writes:

For all who effect an argument *per impossibile* infer syllogistically what is false, and prove the original conclusion hypothetically when something impossible results from the assumption of its contradictory; e.g. that the diagonal is incommensurate, because odd numbers are equal to evens if it is supposed to be commensurate. One infers syllogistically that odd numbers come out equal to evens, and one proves hypothetically the incommensurability of the diagonal, since a falsehood results through contradicting this:

and at 50a35–8:

In the latter [sc. arguments which are brought to a conclusion *per impossibile*], even if no preliminary agreement has been made, men still accept the reasoning, because the falsity is patent, e.g. the falsity of what follows from the assumption that the diagonal is commensurate, viz. that then odd numbers are equal to evens.

So Aristotle has in mind an indirect argument *per impossibile* (i.e. *reductio ad absurdum*) involving odd and even numbers; but, beyond that, he gives no details.

(iv) Eudemus: The lost *History of Mathematics* by Eudemus is cited several times, either directly or indirectly, by Proclus in his *Commentary on the First Book of Euclid's Elements*. One passage, the so-called Catalogue of Geometers, speaks of Pythagoras and the discovery of incommensurability:

Following upon these men, Pythagoras transformed mathematical philosophy into a scheme of liberal education, surveying its principles from the highest downwards and investigating its theorems in an immaterial and intellectual manner. He it was who discovered the *alogōn pragmateia* [see below] and the structure of the cosmic figures [65.15–21]

(This translation is adapted from Proclus–Morrow *CFBEE*, 52–3.) The word *alogōn* is frequently emended to *analogōn*, as Proclus–Friedlein, *PEEC*, 65.19 and Morrow's translation (there "the doctrine of proportionals"), but *alogōn* is in fact the reading of all manuscripts; see Wasserstein, THTN, p. 165 n. 3 (on p. 166) and the following reference. The first sentence was shown, by Vogt and Sachs, not to have been in Eudemus' original text, but to have been an addition copied from Iamblichus (see (xiv), below), and the second sentence is almost certainly a neo-Platonic gloss, perhaps by Proclus himself; see Burkert, *LSAP*, 409–12 for a discussion and references. So Eudemus' summary of the beginnings of Greek mathematics may have contained no reference to the discovery of incommensurability.

On the evidence of Proclus' writings, Eudemus' *History* does not seem to have contained any direct reference to incommensurability. For Proclus cites Eudemus by name in his *Commentary* on pp. 125.7, 299.3, 333.6, 352.14, and

419.15, and discusses topics related to incommensurability on pp. 6, 60, 65 (see above), and 74; but Proclus never cites anything *by* Eudemus *on* incommensurability. I know of no other fragment of Eudemus that introduces the topic, but unfortunately Eudemus–Wehrli, *ER*, does not contain any indexes so I cannot be certain.

(v) Euclid: The topic of incommensurability appears in the *Elements* only in Books X and XIII, which were described in Chapter 5, above. The first place that exhibits an incommensurable magnitude is X 10, a proposition that may have been interpolated; see the note in Heath, *TBEE* iii, 32–3. At the end of Book X, we find a proposition with two alternative demonstrations that the diagonal and side of a square are incommensurable; this proposition is unnumbered in the manuscripts but is now generally known as X 117 (Euclid–Stamatis, *EE* iii, 231–4). It is very rarely translated, though a complete French translation can be found in Euclid–Peyrard, *OE*, and a translation of the first demonstration and paraphase of the second are given in Knorr, *EEE*, 23 and 230–1. This proposition is almost certainly an interpolation made after the time of Alexander of Aphrodisias; see (xiii) below.

(vi) Archimedes only refers to (in)commensurability once, when describing the principle of the balance, at *Plane Equilibria* I, Propositions 6 and 7; these texts are excerpted in Thomas, *SIHGM* ii, 208–17. Proposition 7 is corrupt and garbled, and has provoked an enormous literature; see Knorr, APEPT for a full discussion and references.

(vii) Apollonius wrote an extension of the theory of *Elements* X. We have reports of this from Proclus, who refers to:

Material that Apollonius has elaborated at considerable length about unordered irrationals (Proclus–Morrow, *CFBEE*, 74),

and from Pappus (see (xvi), below). Nothing of this work survives, and I know of no reference to incommensurability anywhere in Apollonius' extant writings.

(viii) Heron of Alexandria (first century AD): The only places in the large amount of heterogeneous material that has been transmitted to us under the name of Heron where the topic of incommensurability is mentioned are passages appended, perhaps during the eleventh century AD, to the *Definitions* (Heron–Heiberg, *Opera* iv, 84–6 and 136–40). This appended material also contains an abbreviated version of Eudemus' Catalogue of Geometers (*Opera* iv, 108; see (iv), above) which does not mention the discovery of incommensurability.

(ix) Plutarch (AD *c.* 50–*c.* 120) often introduces mathematical topics of popular interest in his writings. For example, *Moralia* viii 2 is a dialogue: "What Plato meant by saying that God is always doing geometry", (to which Plutarch adds "if indeed this statement is to be attributed to Plato"!), where we read:

Therefore Plato himself censured Eudoxus and Archytas and Menaechmus for endeavouring to solve the doubling of the cube by instruments and mechanical

constructions, thus trying *dicha logou* [see below] to find two mean proportionals, so far as that is allowable [translation from Thomas, *SIHGM* i, 386 ff.].

This passage, whose syntax is fragmented, has led editors to propose the emendations: δι' ἀλόγου, 'by irrational means', as in this translation, or διαλόγου, 'by means of reason', or ἀναλογον 'in proportion', for the manuscript's δίχα λόγου, 'in separation from reason'; see the critical edition Plutarch–Huberg, *M* iv, 262, for details. The passage has the appearance of a gloss, which disturbed the syntax of the sentence when someone later tried to fit it into the text.

His *Numa* 22 contains the first but the least precise mention of the secrecy of the Pythagoreans, their reluctance to write things down, and the betrayal of this secrecy: "And when their treatment of the abstruse and mysterious processes of geometry had been divulged to a certain unworthy person, they said the gods threatened to punish such lawlessness and impiety with some signal and widespread calamity"; see Burkert, *LSAP*, 457 and 461.

I know of no other passage in Plutarch that could be making any reference to incommensurability.

(x) Nicomachus of Gerasa (AD c. 50–150), an early neo-Pythagorean writer of introductory books on arithmetic, geometry, and music theory, which are then often quoted by later neo-Pythagoreans, does not mention incommensurability in any of his surviving writings.

(xi) Theon of Smyrna (fl. AD c. 125) does not mention incommensurability in his surviving complete work, *Expositio Rerum Mathematicarum ad Legendum Platonem Utilium*, or, to my knowledge, in any fragments.

(xii) Athenagoras of Athens (second half of the second century AD) reports an aphorism attributed to Lysis, a Pythagorean of c. 425 BC: *arrhētos arithmos theos*, 'God is an irrational/ineffable/inexpressible/ . . . number'. Athenagoras' own direct or indirect source is almost certainly a neo-Pythagorean forgery which may have been influenced by a Jewish intermediary. (See Diels & Krantz, *FV* i, 421.5; and Thesleff, *PTHP*, 114 and *IPWHP*.) A translation of the complete report is:

Lysis and Opsimus define as God, the former irrational number, the latter the excess of the greatest of the numbers over the [number] nearest to it; and if the greatest number, according to the Pythagoreans, is the ten, this being the *tetraktys* and containing all the harmonic ratios, and the nine lies close beside it, then God is [a] unit, that is [a] one; for the greatest [number] exceeds the nearest [number] by [a] one, that being the smallest number.

This fragment has no value for matters connected with the discovery and early role of incommensurability, and it may have had nothing to do with mathematics, in the sense of my use of the word in this book.

(xiii) Alexander of Aphrodisias (early third century AD): The first explicit proof of the incommensurability of the diagonal and side of a square is found in Alexander's commentary on *Prior Analytics* 41a26 (Alexander–Wallies, *IAP*

260.7–261.28). There is no generally available translation of this long proof, but there is a paraphrase and full discussion in Knorr, *EEE*, 228–31, which should be consulted for further details of the following summary. Alexander's proof is not the same as those at *Elements* X 117, though it includes three citations from the *Elements* that agree, word for word, with our manuscripts; from this we infer that his proof must post-date the composition of those parts of the *Elements*. Also, since Alexander's other dozen citations of the *Elements* also correspond to our text, we infer that his version of the *Elements* did not include X 117; so this proposition was probably interpolated sometime between Alexander and Theon of Alexandria (late fourth century AD), whose recension of the *Elements* has been transmitted to us in all but one of our Euclidean manuscripts. Knorr proposes, very plausibly, that X 117 was interpolated into the *Elements* as a result of the continuing activity of later Aristotelian commentators. For example Philoponus (sixth century AD) also discusses incommensurability in his commentary on the *Posterior Analytics*, on 71a17 (Philoponus–Wallies, *IAP*, 12.4–17.9; see especially 16.22–25) and on 71b25 (*IAP* 26.17–27.11; see especially 26.25–27).

(xiv) Iamblichus (AD *c.* 250–*c.* 325): The first surviving stories of the circumstance and impact of the discovery of incommensurability are found in the neo-Pythagorean commentator Iamblichus, who gives five mutually inconsistent accounts which permute stories of the discovery, the construction of a dodecahedron, the career of Hippasus, and the death at sea of a Pythagorean; these passages are quoted in Section 10.1(c) below. Also see Burkert, *LSAP*, Chapter 6; Knorr, *EEE*, Chapter 3; O'Meara, *PR*; and Mueller, MPPC.

(xv) Pappus of Alexandria (fl. AD *c.* 320) opens his *Commentary on Book X of Euclid*'s *Elements*, which survives only in an Arabic translation, with a version of the story. Since Pappus is rare among these later commentators in being an able mathematician, and since this reference occurs in a commentary on Euclid's treatment of incommensurable magnitudes, it is worth quoting the passage in full (Pappus–Thomson & Junge, *CPBXEE*, 63–4, with the editor's emphases, parentheses, and expansions):

Book I of the treatise of Pappus on the rational and irrational continuous quantities, which are discussed in the tenth book of Euclid's treatise on the Elements: translated by Abū 'Uthmān Al-Dimishqī.

§1. The aim of Book X of Euclid's treatise on the Elements is to investigate the commensurable and incommensurable, the rational and irrational continuous quantities. This science (or knowledge) had its origin in the sect (or school) of Pythagoras, but underwent an important development at the hands of the Athenian, Theaetetus, who had a natural aptitude for this as for other branches of mathematics most worthy of admiration. One of the most happily endowed of men, he patiently pursued the investigation of the truth contained in these [branches of] science (or knowledge), as Plato bears witness for him in the book which he called after him, and was in my opinion the chief means of establishing exact distinctions and irrefragable proofs with respect to the above-mentioned quantities. For although later the great

Apollonius whose genius for mathematics was of the highest possible order, added some remarkable species of these after much laborious application, it was nevertheless Theaetetus who distinguished the *powers* (i.e. the squares) which are commensurable in length, from those which are incommensurable (i.e. in length), and who divided the more generally known irrational lines according to the different means, assigning the medial line to geometry, the binomial to arithmetic, and the apotome to harmony, as is stated by Eudemus, the Peripatetic. Euclid's object, on the other hand, was the attainment of irrefragable principles, which he established for commensurability and incommensurability in general. For rationals and irrationals he formulated definitions and (specified) differences; determined also many orders of the irrationals; and brought to light, finally, whatever of finitude (or definiteness) is to be found in them. Apollonius explained the species of the ordered irrationals and discovered the science of the so-called unordered, of which he produced an exceedingly large number by exact methods.

§2. Since this treatise (i.e. Book X of Euclid.) has the aforesaid aim and object, it will not be unprofitable for us to consolidate the good which it contains. Indeed the sect (or school) of Pythagoras was so affected by its reverence for these things that a saying became current in it, namely, that he who first disclosed the knowledge of surds or irrationals and spread it abroad among the common herd, perished by drowning: which is most probably a parable by which they sought to express their conviction that firstly, it is better to conceal (or veil) every surd, or irrational, or inconceivable in the universe, and, secondly, that the soul which by error or heedlessness discovers or reveals anything of this nature which is in it or in this world, wanders [thereafter] hither and thither on the sea of non-identity (i.e. lacking all similarity of quality or accident), immersed in the stream of the coming-to-be and the passing-away, where there is no standard of measurement. This was the consideration which Pythagoreans and the Athenian Stranger held to be an incentive to particular care and concern for these things and to imply of necessity the grossest foolishness in him who imagined these things to be of no account.

§3. Such being the case, he of us who has resolved to banish from his soul such a disgrace as this, will assuredly seek to learn from Plato, the distinguisher of accidents, those things that merit shame, and to grasp those propositions which we have endeavoured to explain, and to examine carefully the wonderful clarity with which Euclid has investigated each of the ideas (or definitions) of this treatise (i.e. Book X). . . .

Pappus here tells a version of the story of the discovery and then, in §2, he reports another story that "became current" among the Pythagoreans about the disclosure of the discovery to which he assigns a symbolic meaning, since he says that it is unlikely to be literally true. So Pappus, who also had access to Eudemus' *History* in some form or another, is uncertain of the details and circumstances of the discovery of incommensurability, and of its effect on the wider mathematical community; this also confirms further that Eudemus' *History* did not seem to contain details of the story (see (iv), above). Pappus' story also occurs as an anonymous scholium to *Elements* X (see Euclid–Heiberg, *Opera* v, 417).

Unfortunately Pappus' commentary is of little help in understanding *Elements* X.

(xvi) Proclus (AD *c.* 410–85); see (iv), above, and the passage from his commentary on Platos' *Republic*, quoted in Section 3.6(b), above.

### 8.3(c)  *Discussion of the evidence*

I shall focus my brief discussion of the evidence on two issues that play a prominent role in the received interpretation: did the discovery of incommensurability reveal a fatal flaw in the Pythagorean programme of describing everything by number, and what might have been the details of Aristotle's proof *per impossibile* of the incommensurability of the diagonal? These questions are usually taken as basic to any reconstruction of pre-Euclidean mathematics, and some interpretations of the received answers are often taken as the fundamental facts of the subject.

First, the Pythagorean programme. Our most important early source of information is Aristotle, especially in the *Metaphysics*. For example, we find there several versions of what has become the concise and celebrated epitome of Pythagorean philosophy:

They [sc. the Pythagoreans] reduce all things to numbers [1036b13],

The Pythagoreans, also, believe in one kind of number—the mathematical; only they say it is not separate but sensible substances are formed out of it. For they construct the whole universe out of numbers—only not numbers consisting of abstract units; they suppose the units to have spatial magnitude. But how the first one was constructed so as to have magnitude, they seem unable to say [1080b16–21],

and

These thinkers [sc. the Pythagoreans] identify number with real things; at any rate they apply their propositions to bodies as if they consisted of those numbers [1083b17–19].

(For further comments, see Huffman, *PCPP*, and Zhmud, *AN*.)

Aristotle's polemical method, particularly in the *Metaphysics*, is to give a prejudiced summary of his predecessors' philosophies, through which he can point out the flaws in their point of view. See, for a further example, one of his longer descriptions of Pythagorean mathematics:

Contemporaneously with these philosophers [sc. Leucippus and Democritus] and before them, the so-called Pythagoreans, who were the first to take up mathematics, not only advanced this study, but also having been brought up in it they thought its principles were the principles of all things. Since of these principles numbers are by nature the first, and in numbers they seemed to see many resemblances to the things that exist and come into being—more than in fire and earth and water (such and such a modification of numbers being justice, another being soul and reason, another being opportunity—and similarly almost all other things being numerically expressible); since, again, they saw that the modifications and the ratios of the musical scales were expressible in numbers;—since, then, all other things seemed in their whole nature to be modelled on numbers, and numbers seemed to be the first things in the whole of nature, they supposed the elements of numbers to be the elements of all things, and the whole heaven to be a musical scale and a number. And all the properties of numbers and scales which they could show to agree with the attributes and parts and the whole arrangement of the

heavens, they collected and fitted into their scheme; and if there was a gap anywhere, they readily made additions so as to make their whole theory coherent. E.g. as the number ten is thought to be perfect and to comprise the whole nature of numbers, they say that the bodies which move through the heavens are ten, but as the visible bodies are only nine, to meet this they invent a tenth—the 'counter-earth'. We have discussed these matters more exactly elsewhere [985b23–986a12].

This and other passages, and his further objections to Platonic philosophy (see, for example, 987b23–988a16, 1080b16–35, and 1090a16–1091a23) make it quite clear that Aristotle is vehemently opposed to such a numerical explanation of nature. For example:

Mathematics has come to be identical with philosophy for modern thinkers [sc. Plato and the Academy], though they say that it should be studied for the sake of other things [992a32–b1],

followed, a few lines later, by the complaint that in such a programme:

The whole study of nature has been annihilated [992b8–9].

I have described in the previous section how Aristotle's favourite mathematical illustration is the incommensurability of the diagonal, which he cites eight times in the *Metaphysics* alone. It is therefore curious that Aristotle nowhere points to the phenomenon of incommensurability as a fatal flaw in the Pythagorean programme, of which he is so unequivocally critical.

Not only Aristotle, but every witness up to the time of Iamblichus, is silent on the issue of the impact of the discovery; the only possible exception is Plato, at *Laws* VII, 817e–820e, in a passage which deals with topics more basic and more general than incommensurability. Finally, we have no explicit evidence about the details of the Pythagorean technique for handling proportions or ratios which is, today, often said to have been shown to be inadequate by the discovery of incommensurability; and to talk of such a theory is to appeal to a speculative reconstruction. Throughout this book, I have tried to respect the overall texture of this early evidence; for example, since our best and earliest sources do not make the topic of incommensurability a subject of special notice, no more have I. It is a fact of mathematical life, as Aristotle in effect points out repeatedly, that the ratio of two lines may not be expressible as a *ratio* of two *arithmoi*; but there is no difficulty in expressing any ratio of lines *using only the arithmoi* (and of completely expressing the ratio of any two expressible lines in a finite number of steps). Only when mathematics becomes fully arithmetised and deductive does the dichotomy between commensurable and incommensurable ratios pose a serious and subtle problem for the foundations of this arithmetised mathematics. I have given here, in this book, an extended illustration of a kind of non-arithmetised mathematics where the split between the tractable and intractable occurs at a different place, just beyond the ratios of expressible lines, and have argued that we do have abundant evidence, in Plato and Euclid, of explorations of the nature of this dichotomy.

Now consider Aristotle's comment on the proof of incommensurability *per impossibile* via odd and even numbers. To widen the context of the discussion I shall start by describing two further proofs which, I shall propose, might also fit Aristotle's remarks. (Another relevant proof is in Waskies, ENH.) The first begins in the conventional way, so I will model it on the first demonstration of *Elements* X 117 (in the translation of Knorr, *EEE*, 23):

Let it be proposed to us to prove that in square figures the diagonal is incommensurable in length with side.

Let ABCD be a square, of which AC is the diagonal. I say that AC is incommensurable in length with AB. For if possible, let it be commensurable. I say that it will follow that one and the same number is odd and even. Now it is manifest that the square on AC is the double of that on AB. Since AC is commensurable with AB, then AC will have the ratio to AB of one number to another. Let these numbers be [m and n]. . . . And since AC is to AB as [m] is to [n], so also the square on AC is to that on AB as the square of [m] is to that of [n]. The square on AC is double that on AB, so the square of [m] is double that of [n]. [So the odd number of factors of $2n^2$ is equal to the even number of factors of $m^2$.]

The prolixity of an argument often hides the point of the proof, so here, again, is this first proof. The square number $m^2$ will have an even number of factors, since each factor is repeated twice; but $2n^2$ will have an odd number of factors, since it has the additional factor of 2. Hence, if $m^2 = 2n^2$, an odd number will be equal to an even number.

The second proof is based on the method of diagonals and sides explained in Section 2.2(a). Consider the example of the pentagon. From Fig. 8.1(a), we see that, starting from a smaller pentagon with side $s$ and diagonal $d$, we construct a larger pentagon with side $S = s + d$ and diagonal $D = s + 2d$. Now suppose that the ratio of the diagonal to side of the smaller pentagon is commensurable, so we can express $s = mb$ and $d = nb$ for some common measure $b$. Then $d : s = n : m$ will have one of the following forms: either even : even, or even : odd, or odd : even, or odd : odd. Suppose, for example, that $d : s$ is even : odd. Then, for any other pentagon whatsoever with side $s'$ and diagonal $d'$, $s'$ and $d'$ will be commensurable and if $s' = m'b'$ and $d' = n'b'$,
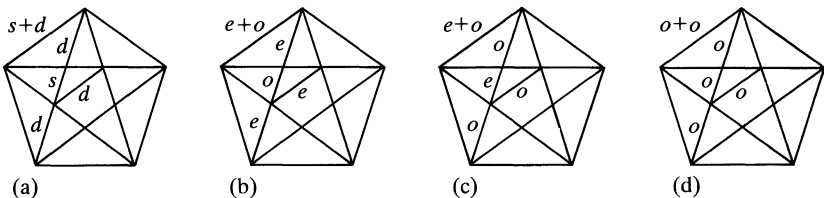


FIG. 8.1

| (a) | (b) | (c) | (d) |
|---|---|---|---|
| $S = s + d$ | *If    $d : s = e : o$ | If    $d : s = o : e$ | If    $d : s = o : o$ |
| $D = s + 2d$ | then $D : S = o : o$ | then $D : S = e : o$ | then $D : S = o : e$ |

(a)                                              (b)
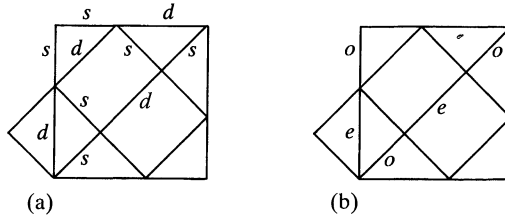
FIG. 8.2

| (a) | (b) |
|-----|-----|
| $S = s + d$ | If    $d : s = e : o$ |
| $D = 2s + d$ | then  $D : S = e : o$ |

then $d' : s' = n' : m'$ will be either also even : odd or even : even, because, if it were odd : even or odd : odd, then odd numbers would be equal to even numbers. (For example, if $e_1 : o_1 = o_2 : e_2$, then $e_1 . e_2 = o_1 . o_2$.) But, from Fig. 8.1(b), we see that our assumption that $d : s$ is even : odd implies that $D : S = (s + 2d) : (s + d)$ is odd : odd; and so that this case cannot occur. Of the four possible cases for $d : s$, we can eliminate the case of even : even by removing factors of two, in the usual way; but then all of the other three cases lead to contradictions, as is illustrated in Figs 8.1(b), (c), and (d). So, *per impossibile*, the ratio $d : s$ cannot have been commensurable in the first case.

Let us try the same procedure on the diagonal and side of a square. This time we appeal to the construction of Fig. 8.2(a), described by the prescription $S = s + d, D = 2s + d$. But, now, when $d : s$ is even : odd, we again have $D : S$ is even : odd, so no contradiction can be deduced by this method for the diagonal of a square.

If we analyse this kind of argument further, we find that the only kind of relations $S = as + bd$, $D = cs + fd$ that will yield contradictions are of the forms:

$$S = os + od \qquad \text{and} \qquad S = es + od$$
$$D = os + ed \qquad\qquad\qquad D = os + od.$$

(The argument is summarised in Table 8.1; for example, the case we have just considered, that of the square, is inconclusive because of the exception listed in row 2, column 4, of Table 8.1(b).) However, the generalised side and diagonal relationships that were explored in Section 3.6 are all of the form $S = as + bd$, $D = nbs + ad$ and so none of these can give rise to a contradiction since, for them, $S = as + \ldots, D = \ldots + ad$.

With these two different proofs *per impossibile* in mind, I shall draw my discussion to a close. Our first detailed evidence about Aristotle's proof comes from Alexander of Aphrodisias, and it seems clear from the way he constructs his argument that neither he nor his source had any account of Aristotle's proof before them. It is quite possible that Aristotle's proof did indeed deal with the

TABLE 8.1. The inconclusive cases in the proof of incommensurability of $d : s$ by the generalised method of diagonal and sides: Proofs *per impossibile* can only be constructed for the cases marked *

| $d$ | $s$ | $es + ed$ | $es + od$ | $os + ed$ | $os + od$ |
|---|---|---|---|---|---|
| $e$ | $o$ | $e$ | $e$ | $o$ | $o$ |
| $o$ | $e$ | $e$ | $o$ | $e$ | $o$ |
| $o$ | $o$ | $e$ | $o$ | $o$ | $e$ |

(a)  Odd and even sides and diameters:
the parity of $as + bd$ as $a$ & $b$ and $s$ & $d$ are even or odd

| $D=$ \\ $S=$ | $es + ed$ | $es + od$ | $os + ed$ | $os + od$ |
|---|---|---|---|---|
| $es + ed$ | $e:o \to e:e$<br>$o:e \to e:e$<br>$o:o \to e:e$ | $e:o \to e:e$ | $e:o \to e:o$<br>$o:e \to e:e$ | $e:o \to e:o$<br><br>$o:o \to e:e$ |
| $es + od$ | $e:o \to e:e$<br>$o:e \to o:e$ | $e:o \to e:e$<br>$o:o \to o:o$ | $e:o \to e:o$<br>$o:e \to o:e$<br>$o:o \to o:o$ | $e:o \to e:o$ |
| $os + ed$ | $o:e \to e:e$<br>$o:o \to o:o$ | $o:o \to o:o$ | $o:e \to e:e$<br>$o:o \to o:o$ | * |
| $os + od$ | $o:e \to o:e$<br>$o:o \to e:e$ | * | $o:e \to o:e$ | $o:o \to e:e$ |

(b)  The inconclusive cases for $d : s \to D : S$

square, and that a brief description of it had been passed down as something like 'if the diagonal of a square is commensurable with the side then, since the square on the diagonal is twice the square on the side, odd numbers will be equal to even numbers'. But, as my first alternative proof shows, such a summary description is ambiguous, since it could equally well refer either to the numbers themselves, or to the number of their factors; and it is not inconceivable that one proof could be confused with the other during transmission. There is a historical observation to be made about the alternative proof by factorisation, that Euclid does not give a theory of factorisation in the *Elements* and, indeed, the topic of unique factorisation does not become fundamental in number theory before the time of Gauss (see Knorr, PIGNT, for a brief discussion); but this does not, I think, completely rule the alternative proof out of serious consideration as a plausible reconstruction.

Also, it could be significant that Aristotle never says that he is referring to the diagonal of a square since there is at least one simple proof that does not work for the square, as my second alternative proof illustrates. (Indeed I do not know of any figure other than a pentagon for which this method does work as a proof of incommensurability.) One could then go on to invent explanations for Aristotle's omission, for example that if he had specified the pentagon, then this might have complicated his illustration by introducing Pythagorean overtones which, for some further reason about which we can then speculate further, Aristotle wished to avoid. But our early evidence is so scanty, our late evidence so doubtful, that it is difficult to pin these speculations down to anything more substantial.

Given the tenuous nature of this evidence and the range of its possible interpretations, it seems unwise to base a reconstruction of early Greek mathematics on speculative assumptions about Aristotle's proof of the incommensurability of the diagonal, or the circumstances and effects of the first realisation of the basic fact of incommensurability. And since our earliest and more direct sources do not seem to give any special attention to the discovery and its impact, perhaps it also should play a subordinate role in our reconstruction. There lies behind my alternative reconstruction here the suggestion of excitement over another discovery, that of the power, surprise, and promise of anthyphairetic mathematics, especially the potential in exploring and generalising the periodicity of the anthyphairetic ratios $\sqrt{n}:\sqrt{m}$. But this promise has still to be realised, even today. In the next chapter I shall describe some of the very few further steps that have been carved out in the uphill struggle to anthyphairetic enlightenment.

# CONTINUED FRACTIONS

> I am sure that, after close examination, you will like the article unless you allow yourself to be put off by the words 'continued fractions', which I too used to recoil from.
>
> (Minkowski to Hilbert, 30 December 1899, translated from Minkowski, *Briefe an David Hilbert*, 118.)

The mathematical techniques developed and deployed so far in the historical reconstructions of this book have not been arithmetised; that is, they have not been expressed in terms of some model, explicit or implicit, of the real numbers and its arithmetic. This chapter will be different. My aim is to prove the main results underlying the reconstruction and to describe some of the more recent investigations into the subject, and I shall give this modern theory in the arithmetised style in which it is now conceived, expressed in the modern algebraic idiom. So the formulae in this chapter, and this chapter only, should be regarded as abstractions of the underlying arithmetic of real numbers.[1]

The mathematico-historical essay that follows makes no pretensions to being more than an introductory sketch to a selection of accessible topics relating to continued fractions and their role in my reconstructions of Greek mathematics. Some parts of the picture have been explored and described in detail; in particular, all aspects up to 1731 (the date of Euler's first reference to continued fractions), and the role of continued fractions in number theory up to 1801 (the date of Gauss's *Disquisitiones Arithmeticae*). Beyond that, the subject cries out for adequate historical investigation.[2]

I shall start the historical account at the beginning of the seventeenth century, and so will omit any discussion of the ascending continued fractions

---

[1] The difficulties in describing the real numbers lie not so much in labelling the actual numbers themselves—for some of the techniques described here, like anthyphairetic and astronomical ratios, will provide immediate solutions to this—but, first, in describing how to perform arithmetical operations on these numbers, and, second and much later, in establishing the completeness of the reals. It does not generally seem to be appreciated that Dedekind was the first person to give a satisfactory solution to *both* of these problems; see Section 1.2(e), above, and, for an informal illustration of the problems with arithmetic, my FHYDF.

[2] Brezinski, *HCF* is a first step in this direction, but the book is mainly a catalogue of names and references, with some extracts, and it contains little analysis of this material. My articles ATUWT, DT, NCRR2, and CF deal with developments up to the time of Euler.

of Fibonacci (see n. 11, below), or the work of Bombelli and Cataldi. My aim is to describe some of the contributions of Euler, Lagrange, and Gauss in the context of the work which more closely preceded and succeeded them.

## 9.1 THE BASIC THEORY[3]

### 9.1(a) *Continued fractions, convergents, and approximation*

Let $x_0$ be a real number, and set

$$x_0 = n_0 + \phi_0,$$

where $n_0$ is the largest integer less than or equal to $x_0$, written $n_0 = \lfloor x_0 \rfloor$; so $\phi_0$ is the fractional part of $x_0$, $0 \leqslant \phi_0 < 1$. If $\phi_i \neq 0$ then, similarly, write

$$1/\phi_0 = x_1 = n_1 + \phi_1,$$

$$1/\phi_1 = x_2 = n_2 + \phi_2,$$

$$\text{etc.;}$$

but if some $\phi_K = 0$, the process will terminate with

$$1/\phi_{K-1} = x_K = n_K.$$

Now eliminate the $\phi_i$s:

$$x_0 = n_0 + \frac{1}{x_1} = n_0 + \cfrac{1}{n_1 + \cfrac{1}{x_2}} = n_0 + \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{x_3}}} = \text{etc.,}$$

to get the continued fraction expansion for $x$. For typographical convenience, this will hereafter be set out as follows:

$$x_0 = n_0 + \frac{1}{x_1} = n_0 + \frac{1}{n_1+} \, \frac{1}{x_2} = n_0 + \frac{1}{n_1+} \, \frac{1}{n_2+} \, \frac{1}{x_3} = \text{etc.}$$

Note that, as thus described, $n_0$ is an integer, $n_i$ $(i \geqslant 1)$ is a positive integer, and, if the continued fraction terminates with $n_K$, then $n_K \geqslant 2$; these conditions describe a simple continued fraction in its standard form. We shall consider these simple continued fractions, together with the similar expressions in which

---

[3] For a lucid introductory account of the mathematical theory of continued fractions, see Davenport, *HA*; more details can then be sought in the standard references such as Chrystal, *TA*; Hardy & Wright, *ITN*; and Perron, *LK*. However great may be the general neglect of the history of continued fractions, most of the developments of this section have received the attention they deserve. My brief historical notes can be supplemented by, for instance, Dickson, *HTN* ii, Chapters 2 and 7; Diophantus–Heath, *DA*, 277–92; Whiteside, *PMTLSC* (see especially pp. 195–6, 207–13, and 241); and for a magisterial treatment that sets the material on number theory in the widest mathematical and historical context, Weil, *NT* (see the General Index, s.vv. 'continued fractions', ' "Pell's equation" ', and 'Ricatti').

the terms $n_i$ are regarded as indeterminates ('variables') that might take real (i.e. non-integral, as with the $x_k$ above, and even negative) values, and will only make occasional reference to general continued fractions:

$$n_0 + \cfrac{m_1}{n_1 + \cfrac{m_2}{n_2 + \cdot_{\cdot_\cdot}}},$$

which will be set out as $n_0 + \frac{m_1}{n_1+} \frac{m_2}{n_2+} \dots$.

We immediately relate the continued fraction algorithm to the Euclidean algorithm[4] applied to $a$ and $b$. Suppose

$$a = n_0 b + c \quad \text{where} \quad c < b,$$

$$b = n_1 c + d \quad \text{where} \quad d < c,$$

$$c = n_2 d + e \quad \text{where} \quad e < d,$$

$$\text{etc.,}$$

where the process terminates if any remainder is zero. Then

$$a/b = n_0 + c/b, \quad \text{i.e.} \quad a/b = n_0 + \phi_0 \text{ where } \phi_0 = c/b,$$

$$b/c = n_1 + d/c, \quad \text{i.e.} \quad 1/\phi_0 = n_1 + \phi_1 \text{ where } \phi_1 = d/c,$$

$$c/d = n_2 + e/d, \quad \text{i.e.} \quad 1/\phi_1 = n_2 + \phi_2 \text{ where } \phi_2 = e/d,$$

$$\text{etc.,}$$

and the same integers $n_0$, $n_1$, $n_2$, ... now describe either the pattern of the Euclidean algorithm applied to $a$ and $b$ or the continued fraction expansion of the real number $a/b$.[5] So the expressions $n_0 + \frac{1}{n_1+} \frac{1}{n_2+} \dots$ and $[n_0, n_1, n_2, \dots]$ now refer to the same thing. I shall, very shortly, say precisely what the three dots mean.

PROPOSITION (see *Elements* X2 & 3)   The continued fraction of a real number $x_0$ terminates if and only if $x_0$ is rational.
PROOF   Since $x_k = n_k + 1/x_{k+1}$, $x_{k+1}$ will be rational if and only if $x_k$ is rational so, in particular, the algorithm can terminate only if $x_0$ is rational. Now suppose $x_0$ is rational, $x_0 = a/b$, where $a$ and $b$ are integers, $b$ positive. Since the remainders of the Euclidean algorithm, applied to $a$ and $b$, decrease in size (see B24 in Section 1.3, above), the algorithm must terminate.                                                                          QED

---

[4] It is ironic that the name 'Euclidean algorithm' is now used for the arithmetised version of Euclid's geometrical anthyphairesis!.

[5] In arithmetised geometry these letters $a$, $b$, $c$, ... might, for example, denote ambiguously either lines or the lengths of these lines, and $a/b$ always denotes the quotient of their lengths. There has been a tendency since the seventeenth century to regard $a : b$ and $a/b$ as synonymous; but $a/b$ has no conventional meaning in non-arithmetised geometry, such as has been described in the historical reconstructions in this book.

Apart from the ambiguity in the last term of a terminating continued fraction, that

$$[n_0, n_1, \ldots, n_K] = [n_0, n_1, \ldots, n_K - 1, 1],$$

the expression as a simple continued fraction is unique.[6] This follows because, if

$$n_0 + \frac{1}{n_1+} \frac{1}{n_2+} \ldots = m_0 + \frac{1}{m_1+} \frac{1}{m_2+} \ldots,$$

then, since the expression $(1/n_1+)(1/n_2+) \ldots$ will always be less than 1 and both $n_0$ and $m_0$ are integers, we must have $n_0 = m_0$. Subtracting this term and reciprocating, we get

$$n_1 + \frac{1}{n_2+} \frac{1}{n_3+} \ldots = m_1 + \frac{1}{m_2+} \frac{1}{m_3+} \ldots,$$

and the argument repeats itself. To complete this proof we must show, in the standard style of arithmetised mathematics, that any such non-terminating expression $[0, n_1, n_2, \ldots]$ converges to some real number less than one; this result will be established below.

The convergents (see Sections 2.3(b) and (c)) of the continued fraction $[n_0, n_1, n_2, \ldots]$ are got by truncating and then evaluating the resulting expression as a common fraction in its lowest terms:

$$[n_0, n_1, \ldots, n_k] = p_k/q_k.$$

In fact, any straightforward evaluation of a terminating simple continued fraction will give a common fraction in its lowest terms, or, when the $n_i$ are regarded as indeterminates, a quotient of two polynomials in the $n_i$ with no common factor. This more general point of view, that the $n_i$ are regarded as indeterminates, will often be taken in the following propositions.

PROPOSITION (see Section 2.3(c) and Table 2.4)   For $k \geqslant 0$

$$p_k = n_k p_{k-1} + p_{k-2}$$

$$q_k = n_k q_{k-1} + q_{k-2},$$

where $p_{-1} = q_{-2} = 1$, $p_{-2} = q_{-1} = 0$.

PROOF   I shall adopt the approach of Euler (see, for example, UNAPPS, §19 ff.) and evaluate the $p_k$ and $q_k$ as polynomials in the indeterminates $n_0, n_1, \ldots, n_k$. To emphasise this point of view, write

$$[n_0, n_1, \ldots, n_k] = \frac{P_k(n_0, n_1, \ldots, n_k)}{Q_k(n_0, n_1, \ldots, n_k)},$$

---

[6] It is easily seen that this result is false if the $n_i$ can take zero or negative values. For example $[\ldots, x, 0, y \ldots] = [\ldots, x + y, \ldots]$; $[\ldots, x, -y] = [\ldots, (x-1), 1, (y-1)]$ (Galois refers to this, at the end of his DTFCP, as "la formule connue"); $[\ldots, x, -y, -z] = [\ldots, x - 1, 1, y - 2, 1, z - 1]$; etc. In fact we can take any sequence of non-zero numbers $n_0, n_1, \ldots, n_k$, write $x = [n_0, n_1, \ldots, n_k, y]$, and solve this for $y$.

where $P_k(n_0, \ldots, n_k)$ and $Q_k(n_0, \ldots, n_k)$ are the polynomials. For example, we immediately evaluate that

$$\frac{P_0(n_0)}{Q_0(n_0)} = \frac{n_0}{1}, \qquad \frac{P_1(n_0, n_1)}{Q_1(n_0, n_1)} = \frac{n_0 n_1 + 1}{n_1},$$

$$\text{and} \qquad \frac{P_2(n_0, n_1, n_2)}{Q_2(n_0, n_1, n_2)} = \frac{n_0 n_1 n_2 + n_0 + n_2}{n_1 n_2 + 1}.$$

The result that

$$Q_k(n_0, n_1, \ldots, n_k) = P_{k-1}(n_1, n_2, \ldots, n_k)$$

now suggests itself; it can be verified by observing that

$$\frac{P_k(n_0, \ldots, n_k)}{Q_k(n_0, \ldots, n_k)} = n_0 + \left(\frac{1}{n_1 +} \cdots \frac{1}{n_k}\right) = n_0 + \frac{Q_{k-1}(n_1, \ldots, n_k)}{P_{k-1}(n_1, \ldots, n_k)}$$

$$= \frac{n_0 P_{k-1}(n_1, \ldots, n_k) + Q_{k-1}(n_1, \ldots, n_k)}{P_{k-1}(n_1, \ldots, n_k)}.$$

(We suppose here and will shortly prove that the $P_k$ and $Q_k$, as defined by the proposition, do not have any common factor.) We also deduce that

$$P_k(n_0, \ldots, n_k) = n_0 P_{k-1}(n_1, \ldots, n_k) + P_{k-2}(n_2, \ldots, n_k),$$

with $P_{-1} = 1$, $P_{-2} = 0$; and a similar result for the $Q_k$s with $Q_{-1} = 0$, $Q_{-2} = 1$. This is very close to what we want to prove.

We now deduce that the polynomial $P_k(n_0, \ldots, n_k)$ is got by adding together the following terms: the product $n_0 n_1 \ldots n_k$; every product that arises by omitting any pair of consecutive terms; every product that arises by omitting any two separate pairs of consecutive terms $n_i$ and $n_{i+1}$, $n_j$ and $n_{j+1}$; products omitting three separate pairs, etc., where, if $k + 1$ is even, the empty product that arises by omitting all the terms has the conventional value 1. Since $P_0(n_0)$ and $P_1(n_0, n_1)$ are of this required form, and $P_k$ satisfies the recurrence relation just established, we get an immediate verification of this description of $P_k$.

From this it follows that

$$P_k(n_0, n_1, \ldots, n_k) = P_k(n_k, n_{k-1}, \ldots, n_0)$$

and hence we can manipulate the recurrence relation into the form

$$P_k(n_0, n_1, \ldots, n_k) = n_k P_{k-1}(n_0, \ldots, n_{k-1}) + P_{k-2}(n_0, \ldots, n_{k-2}),$$

as required.                                                                 QED

REMARK   The $P_k$ are sometimes called Euler or Gauss brackets. They were introduced by Euler, who used the notation $(n_0, n_1, \ldots, n_k)$ for them; Gauss wrote them as $[n_0, n_1, \ldots, n_k]$ (see Section 2(d), below). They satisfy many identities, first explored by Euler. For example the basic generating relation:

$$P_k(n_0, \ldots, n_k) = n_k P_{k-1}(n_0, \ldots, n_{k-1}) + P_{k-2}(n_0, \ldots, n_{k-2})$$

can be generalised to

$$P_k(n_0, \ldots, n_k) = P_l(n_0, \ldots, n_l) P_{k-l}(n_{l+1}, \ldots, n_k)$$

$$+ P_{l-1}(n_0, \ldots, n_{l-1}) P_{k-l-1}(n_{l+2}, \ldots, n_k).$$

Later the corresponding polynomials for a general continued fraction $n_0 + \frac{m_1}{n_1+} \frac{m_2}{n_2+} \ldots$ were seen by Sylvester (1853), Muir (1874), and others to be expressible as special kinds of determinants, called continuants, from which many further identities can be generated. For a historical account see Muir, *TD*.

COROLLARY    For $k \geqslant 2$, $p_{k+1} > p_k$, $q_{k+1} > q_k$.

COROLLARY    For $k \geqslant 0$,

$$x_0 = \frac{x_{k+1} p_k + p_{k-1}}{x_{k+1} q_k + q_{k-1}}.$$

PROOF    Apply the proposition to $[n_0, n_1, \ldots, n_k, x_{k+1}] = x_0$.              QED

COROLLARY (see Section 2.4(f))    For $k \geqslant 0$, $p_{k+1} q_k - p_k q_{k+1} = (-1)^k$. Hence $p_k$ and $q_k$, as calculated by the recurrence relation of the proposition, are coprime, either as numbers or as polynomials.
PROOF

$$p_{k+1} q_k - p_k q_{k+1} = (n_{k+1} p_k + p_{k-1}) q_k - p_k (n_{k+1} q_k + q_{k-1})$$

$$= -(p_k q_{k-1} - p_{k-1} q_k)$$

$$= (-1)^k (p_1 q_0 - p_0 q_1) = (-1)^k.$$

Hence any common factor of $p_k$ and $q_k$ must also divide 1; so $p_k$ and $q_k$ are coprime.                                                                        QED

COROLLARY (see $S_{21}-B_{24}$ in Section 1.3)

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \frac{p_4}{q_4} < \cdots < x_0 < \cdots < \frac{p_5}{q_5} < \frac{p_3}{q_3} < \frac{p_1}{q_1}$$

where if $x_0$ is rational and so its continued fraction terminates, then $x_0$ is equal to the last convergent, $x_0 = [n_0, n_1, \ldots, n_K] = p_K/q_K$; while if $x_0$ is irrational, then $p_k/q_k$ converges to $x_0$ as $k$ increases.
PROOF

$$\frac{p_{k+1}}{q_{k+1}} - \frac{p_k}{q_k} = \frac{p_{k+1} q_k - p_k q_{k+1}}{q_k q_{k+1}} = \frac{(-1)^k}{q_k q_{k+1}},$$

and so this difference is positive if $k$ is even, negative if $k$ is odd. Moreover, the numbers $q_k$ increase as $k$ increases, so the difference between consecutive terms will decrease as $k$ increases. Hence we get the chain of inequalities.

If $x_0$ is rational, its continued fraction will terminate and, by the definition of the convergents, it will be equal to its last convergent. If $x_0$ is irrational, the $q_k$

will increase without limit, so the difference between consecutive terms, which bracket $x_0$, will tend to zero. Hence the convergents $p_k/q_k$ will, indeed, converge to $x_0$ as $k$ increases.                                                                         QED

REMARK   This corollary describes the usual meaning in arithmetised mathematics of the three dots in expressions like

$$x_0 = [n_0, n_1, n_2, \ldots],$$

that either the sequence terminates, in which case it is an arithmetical (or algebraic) identity, or it does not terminate and the numbers $[n_0], [n_0, n_1]$, $[n_0, n_1, n_2], \ldots$ tend arithmetically to $x_0$. The chain of inequalities implies that $n_0 \leqslant x \leqslant n_0 + 1$, where the left-hand equality can only hold if $x = [n_0]$, the right-hand if $x = [n_0, 1] = [n_0 + 1]$ in standard form.

COROLLARY (see Section 2.2(b)(vii))   If $n_0$, $n_1$, $n_2$, ... is any non-terminating sequence of integers with $n_i \geqslant 1$ for $i \geqslant 1$, then the continued fraction $[n_0, n_1, n_2, \ldots]$ converges to some irrational number $x_0$.

PROOF   The convergents satisfy the inequalities of the previous corollaries; so $p_0/q_0$, $p_2/q_2$, $p_4/q_4$, ..., being a bounded increasing sequence, converges to a limit; also $p_1/q_1$, $p_3/q_3$, $p_5/q_5$, ..., being a bounded decreasing sequence, also converges to a limit; and these two limits must be equal since we can make $p_k/q_k - p_{k+1}/q_{k+1}$ as small as we please. Finally, since the continued fraction expansion is unique, we must have $x_0 = [n_0, n_1, n_2, \ldots]$ and so, since this does not terminate, $x_0$ is irrational.                                                         QED

In fact, consecutive convergents also improve in accuracy, and we can give close estimates of their error and deduce that, in some very precise sense, they give the best approximations to $x_0$ by rational numbers; see Sections 2.2(b)(iii) and (iv) above. For, since

$$x_0 = \frac{x_{k+1} p_k + p_{k-1}}{x_{k+1} q_k + q_{k-1}},$$

either we can rewrite this to give

$$x_0 - \frac{p_k}{q_k} = \frac{q_{k-1}}{x_{k+1} q_k} \left( \frac{p_{k-1}}{q_{k-1}} - x_0 \right)$$

whence, since $0 < q_{k-1} < x_{k+1} q_k$, we get:

COROLLARY   When $k \geqslant 1$,

$$\left| x_0 - \frac{p_k}{q_k} \right| < \left| x_0 - \frac{p_{k-1}}{q_{k-1}} \right|.$$

Or, alternatively, we can evaluate

$$x_0 - \frac{p_k}{q_k} = \frac{p_{k-1} q_k - p_k q_{k-1}}{q_k(x_{k+1} q_k + q_{k-1})} = \pm \frac{1}{q_k(x_{k+1} q_k + q_{k-1})},$$

and hence, since $n_{k+1} \leqslant x_{k+1} < n_{k+1} + 1$ and $q_{k-1} < q_k$ for $k \geqslant 2$, we have the estimates of Sections 2.2(b)(iii) and (iv):

COROLLARY    If $k \geqslant 2$,

$$\frac{1}{(n_{k+1} + 2)q_k^2} < \left| x_0 - \frac{p_k}{q_k} \right| < \frac{1}{q_k q_{k+1}} < \frac{1}{n_{k+1} q_k^2}.$$

Or we can manipulate the formulae to prove that the convergents provide the best approximations:

COROLLARY    If $\left| x_0 - \frac{a}{b} \right| < \left| x_0 - \frac{p_k}{q_k} \right|$, where $\frac{p_k}{q_k}$ is a convergent of $x_0$, then $b > q_k$.

PROOF    Suppose, first, that $a/b$ and $p_k/q_k$ lie on the same side of $x_0$, for example that $p_k/q_k < a/b < x_0 < p_{k+1}/q_{k+1}$. Then

$$0 < \frac{a}{b} - \frac{p_k}{q_k} < \frac{p_{k+1}}{q_{k+1}} - \frac{p_k}{q_k} = \frac{1}{q_k q_{k+1}}.$$

Hence

$$0 < aq_k - bp_k < \frac{b}{q_{k+1}},$$

and since $aq_k - bp_k$ is an integer, $b > q_{k+1}$.

Now suppose that they lie on opposite sides, for example that $p_k/q_k < x_0 < a/b < p_{k-1}/q_{k-1}$, where the final inequality follows from the previous corollary but one. Then, by what we have just proved, $b > q_k$.                          QED

HISTORICAL NOTES    The convergents to a simple continued fraction, as such, seem to have first been defined in the West since the Renaissance by Daniel Schwenter in his *Geometriae Practicae* of 1618 and *Deliciae Physico-Mathematicae* of 1636, and for a general continued fraction by Wallis, in his *Arithmetica Infinitorum* (1658)—where the expression 'continued fraction' first appeared—though no general algebraic formulae are given before Euler. Approximation methods which turn out to be closely related to the continued fraction process are rather older; see the next section for an early example.

Wallis devoted Chapters 10 and 11 of his *Treatise of Algebra* of 1685 to the problem of the "Reduction of Fractions or Proportions to smaller Terms, as near as may be to the just Value", that is to approximating a given fraction or decimal number by a fraction whose denominator should not be greater than some given number, say 999. The problem, he said, was sent to him in 1663 or 1664; it is due to a certain Dr Davenant and it arose in connection with Metius's approximation of $\pi$ by $\frac{335}{113}$. Wallis' procedure is tantamount to the computation of the terms, convergents, and intermediate convergents (see the next section) of the continued fraction expansion, and it is usually described in these terms, but I have argued in my ATUWT that this is not the case; indeed

Wallis himself says as much, in a letter to John Collins (see my article, pp. 211–12). His procedure is based on manipulations of the decimal expansion of the given common fraction. Then, in Chapter 11, he demonstrates his procedure on van Ceulen's thirty-five decimal-place value of $\pi$, augmented by a final approximate place:

$$3{\cdot}14159\,26535\,89793\,23846\,26433\,83279\,50288\tfrac{1}{2}$$

from which, at first sight, it seems that he deduced that

$$\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, 2, 1, 1, 2, 2, 2, 1, 84,$$

$$2, 1, 1, 15, 3, 13, 1, 4, 2, 6, 6, 1 \text{ (should be 99)}, \ldots]$$

with a tabulation of 494(!) convergents and intermediate convergents. However, if one follows through his calculation in detail (he explains that he has suppressed the numerical evaluations, for "it is not necessary to trouble the Reader with so oft repeating those long Numbers. But it is easy[!] for any who will give himself that trouble to restore them when he pleaseth"), one can see that he knows precisely when he has reached the limit of the accuracy of van Ceulen's approximation, and he disowns the accuracy of the final place of his expansion[7] which should, in any case, read $[\ldots, 6, 6, 3$ or more$, \ldots]$.

Approximation using the continued fraction process is handled fluently, with descriptions of proofs, by Huygens, in his *Descriptio Automati Planetarii* (written between 1680 and 1687 but only published posthumously in 1703), in connection with the problem of selecting gear ratios for a planetarium.[8] The procedure is then formalised and developed by Euler and Lagrange; the latter called it "une des principales découvertes de ce grand géomètre [sc. Huygens]" in §7 of his *Additions* to Euler's *Algebra*; see Weil, *NT*, 120, for further references and a brief comment.

As illustrations of the understanding and use of continued fractions just before the introduction of Euler's formalism, here are two extracts from Roger Cotes. First, his discovery of the surprising behaviour of *e*, published first in his article 'Logometrica' and repeated in his posthumous *Harmonia Mensurarum* (1722):

The major term 2·71828 &c. [= *e*] should be divided by the minor 1 . . . and once more the minor by the number which is left, and this again by the last remainder, and so continue forward; and the quotients 2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, 1, 1, 12, 1, 1, 14, 1, 1, 16, 1, 1 &c. will be produced. Having made these calculations, we must set out two columns of ratios, of which one contains the terms which have a ratio greater than the true one (*rationes vera majores*), and the other contains the terms which have a ratio less than the true one (*rationes vera minores*), beginning the computation with the ratios 1 to 0 and 0 to 1, which are most remote from the true one, and having started there,

---

[7] The error in this final term seems to have been noticed first in Lehmer, *EALN*.

[8] The technique can now be found in any good advanced manual of engineering workshop practice. Some such procedure is also needed for the Antikythera mechanism; see Price, *GG*.

continuing to deduce the remaining ratios, which approach ever closer to the true one. . . .
[There follows here a long verbal description of the following table:]

| Rationes Vera Majores | | | Rationes Vera Minores | |
|---|---|---|---|---|
| | | | 0 | 1 |
| 1 | $0 \times 2$ | | 2 | 0 |
| 2 | 1 | | 2 | $1 \times 1$ |
| 3 | $1 \times 2$ | | 6 | 2 |
| 8 | 3 | | 8 | $3 \times 1$ |
| 11 | $4 \times 1$ | | 11 | 4 |
| 76 | 28 | | 19 | $7 \times 4$ |
| 87 | $32 \times 1$ | | 87 | 32 |
| 106 | 39 | | 106 | $39 \times 1$ |
| 193 | $71 \times 6$ | | | |
| . . . | | | | |

The application of these approximations is widespread, wherefore I have given a somewhat prolix exposition of their invention, by the method which seems to me simplest and easiest. The celebrated men Wallis and Huygens have dealt with the same arguments, slightly differently [*Harmonica Mensurarum*, 7–9].[9]

Next, there is an extract from Cotes's correspondence with Newton over the second edition of the *Principia*:[10]

I find that 25 & 21 express the proportion of $\sqrt{\sqrt{2}}$ to 1 as nearly as it is possible for so small numbers to do it, whence it is probably y$^t$ the exact proportion of the diameter of the Hole to y$^e$ diameter of y$^e$ Stream is that of $\sqrt{\sqrt{2}}$ to 1, & then y$^e$ proportion of 44 to 37 will be much nearer the truth than y$^t$ of 25 to 21.

In fact $\sqrt{\sqrt{2}} = [1, 5, 3, 1, 1, 40, \ldots]$, with convergents $\frac{6}{5}$, $\frac{19}{16}$, $\frac{25}{21}$, and $\frac{44}{37}$ and, because of the next large term of 40, $\frac{44}{37}$ will be a very good estimate: $\frac{1}{42 \times 37^2} < (\sqrt{\sqrt{2}} - \frac{44}{37}) < \frac{1}{40 \times 37^2} < 1 \cdot 8 \times 10^{-5}$.

Finally, as a postscript to this sketch of the early history of continued fractions and approximations, I note that, rather later, Lagrange wrote a modest little article, EANTF (1799), in which he described different approaches to a similar problem, that of approximating a given common fraction by sums and differences of other particular kinds of fractions:

$$\frac{p}{q} = \frac{m_0}{n_0} \pm \frac{m_1}{n_1} \pm \frac{m_2}{n_2} \pm \ldots,$$

[9] I am grateful to Ronald Gowing for sending me material on Cotes; the translation is by him. I have slightly rearranged the order and presentation of Cotes's text: the table precedes its description in the original, and I have introduced a one-line displacement between the two columns of the table to clarify his procedure. Cotes also gives a similar description, with table, in which the intermediate convergents are calculated. Incidentally, I do not believe that Cotes continued his numerical calculation up to the terms ". . . 16, 1, 1, &c"; this is surely an 'induction', seventeenth-century style, because the sequence given exactly fills one line of text in the original.

[10] Quoted from Westfall, *NR*, 711, where an explanation of the context can be found. The letter is dated 31 March 1711. For more details, see my NCRR2.

under various different conditions on the $m$s and $n$s. Within this general framework, he described expressions like

$$\frac{p}{q} = m_0 \pm \frac{m_1}{n} \pm \frac{m_2}{n^2} \pm \frac{m_3}{n^3} \pm \ldots ,$$

which is a generalisation of the sexagesimal or decimal expansion to base $n$ with positive or negative terms; or its more general form

$$\frac{p}{q} = m_0 \pm \frac{m_1}{n_1} \pm \frac{m_2}{n_1 n_2} \pm \frac{m_3}{n_1 n_2 n_3} \pm \ldots ,$$

an expansion to changing bases, as with Imperial measures;[11] then

$$\frac{p}{q} = m_0 \pm \frac{1}{n_1} \pm \frac{1}{n_1 n_2} \pm \frac{1}{n_1 n_2 n_3} \pm \ldots ,$$

an expression in sums and differences of unit fractions (precisely the procedure described at $B_{114}$ in Section 7.4(d), above), one example of which he gave being

$$\pi = 3 + \frac{1}{7} - \frac{1}{7 \times 113} - \frac{1}{7 \times 113 \times 4739} + \frac{1}{7 \times 113 \times 4739 \times 47\,051} + \cdots$$

(contrast note 97 to Chapter 7, above). He then gave a treatment of continued fractions approached via the problem of constructing approximations, similar in spirit but different in its details from the algorithms of Section 2.3. The article is not of great importance mathematically and seems generally to have been forgotten, but parts of it echo and develop themes that have been raised here, particularly in some of my slaveboy's remarks and questions.

### 9.1(b)  The Parmenides proposition and algorithm[12]

The historical reconstructions in this book have not been based on the continued fraction algorithms that I have so far been describing,[13] but on the

---

[11] These kinds of expansions are, in fact, expressions as ascending continued fractions,

$$m_0 \pm \frac{m_1 \pm}{n_1} \frac{m_2 \pm}{n_2} \ldots .$$

They were considered by Fibonacci and other Italian mathematicians up to the sixteenth century; see my DT for a brief discussion.

[12] This section is inspired by Fletcher, AV, and owes a lot to the meticulous account in Stark, INT, Chapter 7. Readers should consult these for more detailed explanations and illustrations.

[13] This, it seems to me, is a fatal flaw in most of the other detailed anthyphairetic reconstructions of early Greek mathematics known to me. I repeat, yet again, a theme that runs throughout this book: we have no evidence to indicate that early Greek mathematics was arithmetised, either in conception or in underlying motivation, and we know much which suggests that it was not. Therefore the description so far in this chapter is simply irrelevant to early Greek mathematics. Of course, the later Greek mathematics, for example of Heron, Ptolemy, and Diophantuṣ, is clearly deeply influenced by Babylonian arithmetised techniques, but anthyphairesis does not seem to play any practical or theoretical part during this later period. See Section 10.5, below.
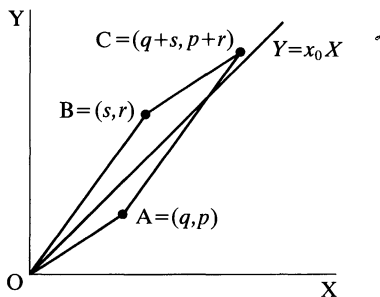
FIG. 9.1. The *Parmenides* algorithm

generation of approximations by the *Parmenides* proposition; see Section 2.2, above, where the use of the *Parmenides* proposition to compute approximations and continued fraction expansions is explained. Let us now investigate these propositions and algorithms in more detail. We suppose, for simplicity, that $x_0$ is positive.[14]

THE *PARMENIDES* PROPOSITION    If $\frac{p}{q} < \frac{r}{s}$ then $\frac{p}{q} < \frac{p+r}{q+s} < \frac{r}{s}$.

PROOF    $\frac{p+r}{q+s} - \frac{p}{q} = \frac{qr-ps}{q(q+s)} > 0$ since $\frac{p}{r} < \frac{q}{s}$.          QED

THE *PARMENIDES* ALGORITHM    If $p/q$ is an underestimate and $r/s$ is an overestimate for $x_0$, then $(p+r)/(q+s)$ will be a better approximation than that original estimate which lies on the same side of $x_0$. Start the algorithm with the formal universal estimates $0/1 < x_0 < 1/0$.

I shall develop the properties of the *Parmenides* algorithm using coordinate geometry and properties of the integer lattice, the set of all points having integer coordinates, a description first set out in Smith, NCF. Take the line $Y = x_0 X$ in the positive quadrant of the $(X, Y)$-plane (see Fig. 9.1); then our aim is to find lines joining the origin to points of the integer lattice whose slopes will closely approximate the slope $x_0$ of this given line.

If OA, where $A = (q, p)$, is the underestimate in the algorithm, and OB, where $B = (s, r)$, is the overestimate, then the *Parmenides* proposition states that the slope of the diagonal OC, the vector sum of OA and OB, lies between the slopes of OA and OB. In the case illustrated in Fig. 9.1, OC will then be a better estimate than OB. So, at the next step of the algorithm, we repeat the operation with the points A and C. Figure 9.2 illustrates the process applied to $x_0 = \sqrt{3}$, and the annexed table lists the successive approximations; it is identical in content to Table 2.1, save for the mild inconvenience that the slope $p/q$ corresponds to the point $(q, p)$ but the ratio $p : q$. Note also that we

---

[14] When $x_0$ is negative, a similar description can be given for the fourth quadrant $X \leqslant 0$, $Y \geqslant 0$. Or we can work with $-x_0$.

Y

Y=√3X

(4,7)

(3,5)

(2,3)

(1,2)

(0,1)

(1,1)

(1,0)

X

| under | over |
| --- | --- |
| (1, 0) | (0, 1) |
| (1, 1) | |
| | (1, 2) |
| (2, 3) | |
| (3, 5) | |
| | (4, 7) |
| (7, 12) | |
| (11, 19) | |
| | (15, 26) |
| etc. | |
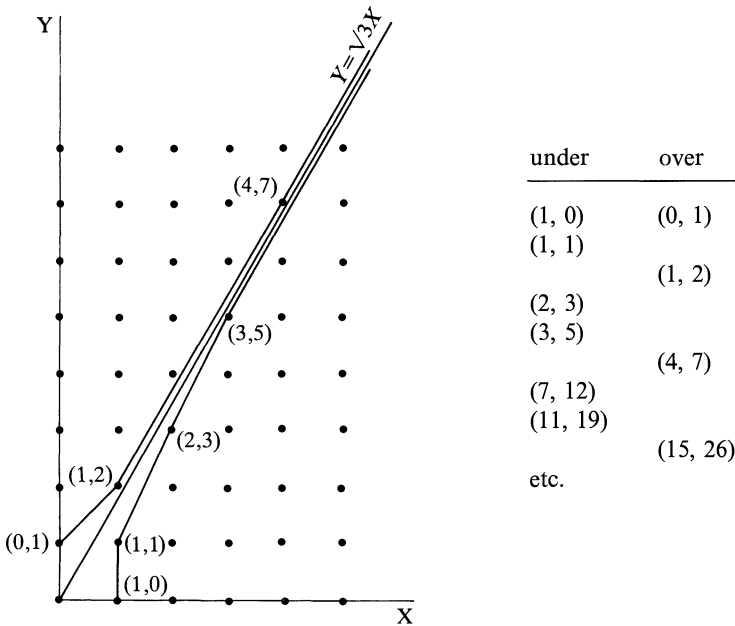
FIG. 9.2. Approximations to $\sqrt{3}$

distinguish $p/q$, $2p/2q$, $3p/3q$, etc., since they correspond to different approximating points in the integer lattice; for this reason, I shall generally refer to $(q, p)$ rather than the fraction $p/q$. Also observe that if $(q, p)$ and $(s, r)$ are the current estimates, then $q \leqslant s$ and $p \leqslant r$, or $q \geqslant s$ and $p \geqslant r$, where both inequalities are strict after the initial estimates have been superseded, so thereafter any estimate will lie above and to the right of any earlier estimate. I always suppose that this configuration holds.

If A and B are the current approximations, and C the new approximation (as in Fig. 9.1), then call OAB the approximating triangle and OABC the approximating parallelogram. I shall base the analysis of the *Parmenides* algorithm on:

PICK'S THEOREM   The area of a simple polygon whose vertices are lattice points is equal to half of the number of points on the perimeter, plus the number of points inside, minus one.

A satisfactory discussion and proof of this result would lead us too far from our main interest; for further details see, for example, Coxeter, *IG*, 209–10, or Varberg, PTR.

PROPOSITION   The approximating triangle and parallelogram contain no points of the integer lattice.
PROOF   Each step of the algorithm replaces the approximating triangle OAB by OAC or OBC, and each of these triangles is half of the approximating

parallelogram OABC; so the area of the approximating triangle is unchanged. Since it starts with $A = (1, 0)$ and $B = (0, 1)$, its area will always be $\frac{1}{2}$. Since its three vertices are lattice points, Pick's theorem tells us that it can have no lattice points inside.                                                                QED
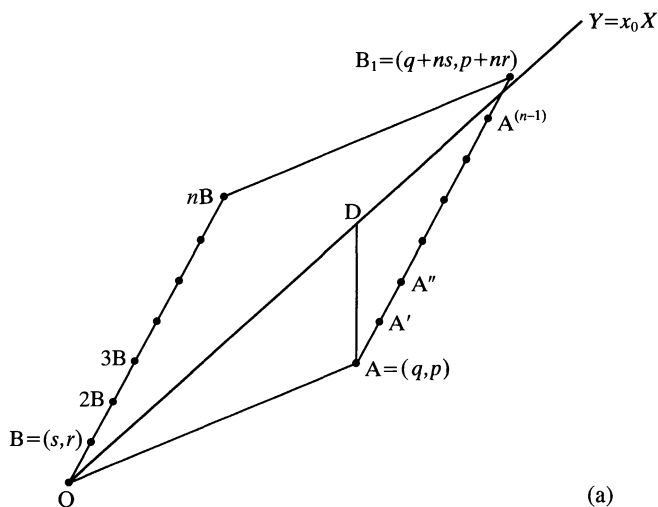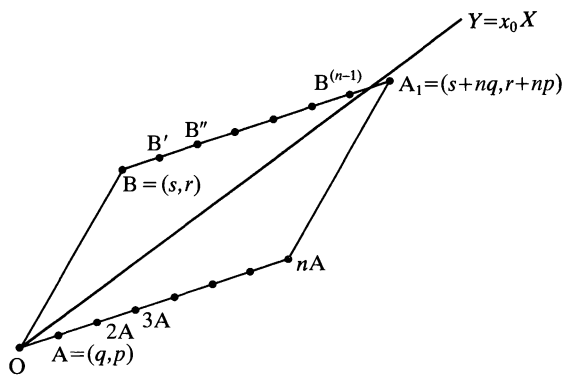


FIG. 9.3. Typical runs of estimates in the *Parmenides* algorithm. (The thickness and areas of the parallelograms in all figures except Fig. 9.2 have been greatly exaggerated and distorted to make the drawings possible. The axes have been omitted since they will also have been distorted in direction and scale)

(a) A run of $n$ underestimates $A$, $A'$, . . . , $A^{(n-1)}$, followed by a new overestimate
$$B_1 = (q + ns, p + nr)$$
(b) A run of $n$ overestimates $B$, $B'$, . . . , $B^{(n-1)}$, followed by a new underestimate
$$A_1 = (nq + s, np + r)$$

Characteristic sequences of steps are illustrated in Fig. 9.3. Consider Fig. 9.3(a): the estimates $A = (q, p)$ and $B = (s, r)$ give rise to a run of $n$ underestimates $A$, $A'$, $A''$, ..., $A^{(n-1)}$, followed by a new overestimate $B_1 = (q + ns, p + nr)$. Each of the triangles $OAA', OA'A'', \ldots, OA^{(n-1)}B_1$, being half of its corresponding approximating parallelogram, will contain or meet no lattice point other than its three vertices. Hence neither the triangle $OAB_1$ nor $OB_1 nB$—which has the same area and the same number of perimeter lattice points—will contain or meet any other lattice points than $O$, $A$, $A'$, $A''$, ..., $A^{(n-1)}$, $B_1$, $nB$, ..., $3B$, $2B$. $B$. Figure 9.3(b) illustrates the similar case of a run of $n$ overestimates followed by an underestimate.

Hence the procedure will generate the configuration of Tables 2.2 and 2.3 from which, if we extract only the run-end estimates, we get the more schematic configuration of Table 2.4; but this describes precisely the calculation of the convergents of our basic continued fraction algorithm. Hence the two procedures are equivalent. Thus the estimates that occur at the end of the runs of the *Parmenides* algorithm are the convergents of the continued fraction expansion of $x_0 = [n_0, n_1, n_2, \ldots]$, while the other estimates are the so-called intermediate convergents, defined by

$$\frac{p_k^{(m)}}{q_k^{(m)}} = [n_0, n_1, \ldots, n_{k+1}, m] \quad \text{where} \quad 1 \leqslant m < n_k.$$

The way the current and run-end estimates alternate around the line $Y = x_0 X$ illustrates clearly the way the intermediate convergents and convergents alternate around $x_0$.

The approximating parallelograms in a correctly drawn figure quickly become very, very long and thin; see, for example, Fig. 9.2. To make drawings possible, gross distortions have had to be introduced into almost all of the figures here. However the proofs ultimately depend on the careful use of words like 'above', 'below', 'inside', and 'outside' in contexts that are not affected by these distortions. This point is examined carefully in Stark, *INT*.

So far we have been considering angular approximations to the line $Y = x_0 X$, i.e. approximations of the form $|x_0 - p/q|$. We now show that the continued fraction process actually locates those points of the integer lattice which lie closest to the line $Y = x_0 X$, although it is in practice more convenient to work with the vertical distance from the point $(q, p)$ to $Y = x_0 X$, given by $|x_0 q - p|$ (see, for example, the line AD in Fig. 9.3(a)); this is the shortest distance multiplied by a constant, $\sec x_0$. Clearly points which are close to the line $Y = x_0 X$ will generate good angular approximations, though not necessarily conversely: consider $(nq, np)$ for large $n$.

PROPOSITION If $A = (q, p)$ and $B = (s, r)$ are the current under- and overestimates to $x_0$, with $q < s$ (so A is a convergent) then

$$|x_0 q - p| < |x_0 u - t|$$

for every $F = (u, t)$ different from A and B with $t \leqslant s$. If, further, B is a convergent, then

$$|x_0 s - r| < |x_0 q - p|.$$

PROOF We suppose that A lies below $Y = x_0 X$ and B lies above. Since $q < s$, A must be a run-end estimate (i.e. a convergent), while B may or may not be a run-end estimate (i.e. it is either an intermediate convergent or a convergent). We therefore have a situation analogous to that shown in Fig. 9.3(b) in which any one of the B, B', . . . , $B^{(n-1)}$ can now play the role of B. This figure is incorporated in Fig. 9.4, in which $B^{(iv)}$ is chosen to typify the point B. We have shown, in Fig. 9.3(b), that the only lattice points in or on the parallelogram $OnAA_1B$ are the marked points O, A, 2A, . . . , nA, $A_1$, $B^{(n-1)}$, $B^{(n-2)}$, . . . , B. If, as in Fig. 9.4, we now tile the plane with translates of the parallelogram, the points of the integer lattice will lie, similarly, on the sides of the translated parallelogram. (We see clearly here an effect of the distortion in this figure. The axes have been omitted from these figures since they, too, will be distorted in direction and scale.)

Suppose first that the point $F = (u, t)$ lies below $Y = x_0 X$. Then, since $t \leqslant s$, it must lie on or below the line OA and before the line $X = s$; and, by hypothesis, it is not equal to A. Since the slope of $Y = x_0 X$ is greater than the slope of OA, F must therefore be further away than A from $Y = x_0 X$.

Next suppose that F lies above $Y = x_0 X$, in which case it must lie on or above the line containing B, and it is not the point B itself. Now the lattice
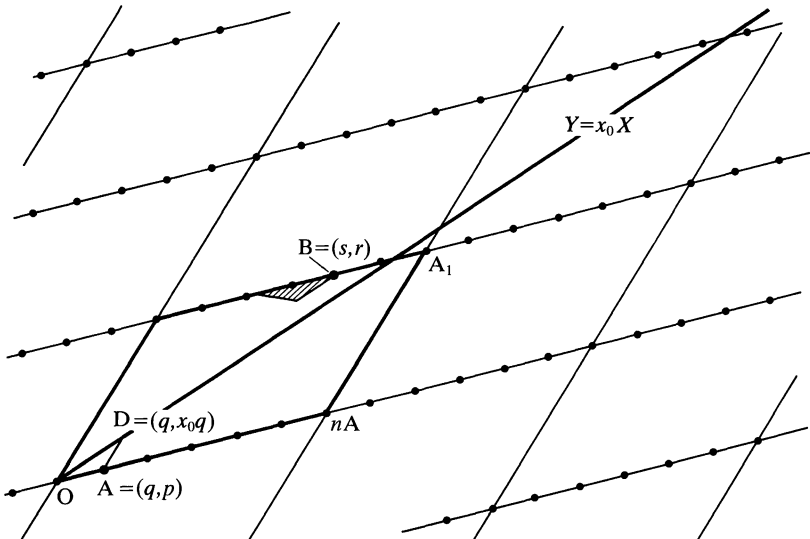


FIG. 9.4

points before B are further from $Y = x_0 X$ than is A, because $Y = x_0 X$ cannot pass through the small shaded triangle in Fig. 9.4, congruent to OAD, where $D = (q, x_0 q)$, the point vertically above A. Hence F must again lie further from $Y = x_0 X$ than does A, and again we have

$$|x_0 q - p| < |x_0 u - t|.$$

Finally, if B is a run-end estimate, then line $Y = x_0 X$ will pass through the corresponding small triangle congruent to OAD with vertex at $A_1$. Hence

$$|x_0 s - r| < |x_0 q - p|. \qquad \text{QED}$$

PROPOSITION (see Section 2.2(b)(ii)) If $A = (q, p)$ and $B = (s, r)$ are successive convergents to $x_0$ then

$$\left| x_0 - \frac{p}{q} \right| < \frac{1}{q^2} \quad \text{and} \quad \left| x_0 - \frac{r}{s} \right| < \frac{1}{s^2},$$

and at least one of them satisfies

$$\left| x_0 - \frac{p}{q} \right| < \frac{1}{2q^2} \quad \text{or} \quad \left| x_0 - \frac{r}{s} \right| < \frac{1}{2s^2}.$$

PROOF See Fig. 9.5, in which we suppose that A is an underestimate, B is an overestimate, and $q < s$. Then the error in each estimate is less than $\left| \frac{r}{s} - \frac{p}{q} \right| = \frac{|qr - ps|}{qs} = \frac{2}{qs} \times$ area of the approximating triangle OAB $= \frac{1}{qs}$. So, since $q < s$,

$$\left| x_0 - \frac{p}{q} \right| < \frac{1}{q^2}.$$

Now suppose that both

$$\left| x_0 - \frac{p}{q} \right| \geqslant \frac{1}{2q^2} \quad \text{and} \quad \left| x_0 - \frac{r}{s} \right| \geqslant \frac{1}{2s^2};$$

we shall derive a contradiction by analysing Fig. 9.5 more closely, and showing that these inequalities imply that the triangle OAB is too big.
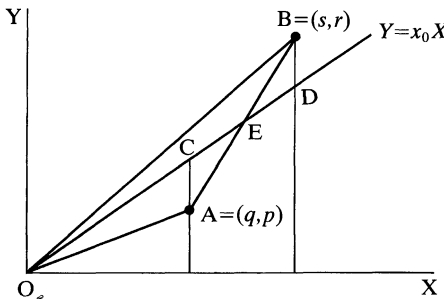


FIG. 9.5

Let C be the point $(q, x_0 q)$, vertically above A, and D be $(s, x_0 s)$, vertically below B; and let AB intersect $Y = x_0 X$ in E. Then, by our assumed inequalities

$$\text{AC} = |x_0 q - p| \geqslant 1/2q \quad \text{and} \quad \text{BD} = |x_0 s - r| \geqslant 1/2s.$$

Since $q < s$, and A and B are both convergents, AC > BD; so, since the triangles ACE and BDE are similar, ACE is larger than BDE. Next

$$\text{area OAC} = \tfrac{1}{2} q \,.\, \text{AC} \geqslant \tfrac{1}{4} \quad \text{and} \quad \text{area ODB} = \tfrac{1}{2} s \,.\, \text{BD} \geqslant \tfrac{1}{4}.$$

Hence

$$\text{area OAB} = \text{area OAC} + \text{area OBD}$$

$$+ (\text{area ACE} - \text{area DBE}) > \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{1}{2},$$

which gives the required contradiction.          QED

There is an important converse to this result:

PROPOSITION    If $r/s$ satisfies

$$\left| x_0 - \frac{r}{s} \right| < \frac{1}{2s^2}$$

then $r/s$ is a convergent of $x_0$.

PROOF    Suppose that $r/s$ is in its lowest terms, and that $A = (q, p)$ is the last convergent (i.e. that with largest $q$) such that $q \leqslant s$. We will then prove, by contradiction, that $r/s = p/q$. For suppose that $r/s \neq p/q$, i.e. $ps - qr \neq 0$; then, since $ps - qr$ is an integer, $|ps - qr| \geqslant 1$, so triangle OAB (where $B = (s, r)$) will have area $\geqslant \tfrac{1}{2}$. But we shall show that the assumption on B does not allow this. As in the previous proposition, this assumption implies that $|sx_0 - r| < 1/2s$, i.e. BD $< 1/2s$ in Figs 9.6(a) and (b).

We consider the case where $p/q < x_0$, i.e. A lies below $Y = x_0 X$; the case of $p/q > x_0$ is similar. We investigate separately the three situations where $r/s \leqslant p/q < x_0$; $p/q < r/s \leqslant x_0$; and $p/q < x_0 < r/s$.
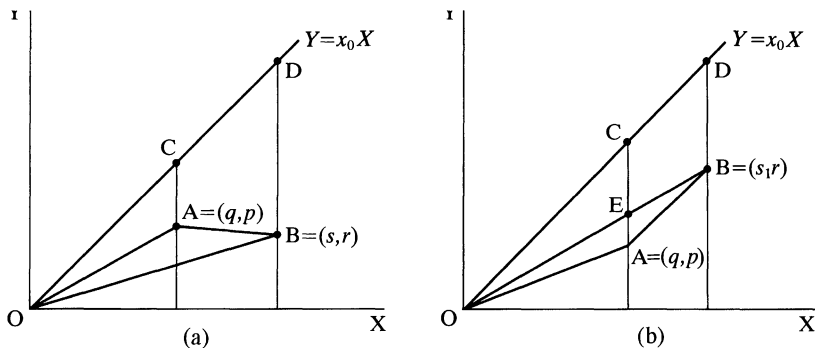


FIG. 9.6

First suppose $r/s \leqslant p/q < x_0$, Fig. 9.6(a). Then A lies inside or on the edge OB of triangle OBD, so

$$\text{area OAB} < \text{area OBD} = \tfrac{1}{2}s \,.\, \text{BD} < \tfrac{1}{4},$$

which shows this case cannot occur.

Next suppose $p/q < r/s \leqslant x_0$, Fig. 9.6(b). Then

$$\text{area OAB} = \text{area OAE} + \text{area AEB}$$
$$= \tfrac{1}{2}q \,.\, \text{AE} + \tfrac{1}{2}(s - q) \,.\, \text{AE}$$
$$= \tfrac{1}{2}s \,.\, \text{AE} < \tfrac{1}{2}s \,.\, \text{AC} < \tfrac{1}{2}s \,.\, \text{BD} < \tfrac{1}{4},$$

which again cannot occur. (Note that both of these contradictions would follow from the weaker hypothesis that $|x_0 - r/s| < 1/s^2$.)

Finally, suppose that $p/q < r/s < x_0$; here see Fig. 9.5, but in a different interpretation from that in the previous proposition. Let $(p', q')$ be the convergent following $(p, q)$, so that $q \leqslant s < q'$. Then, by the last proposition but one, AC < BD. So, since triangles ACE and BDE are similar, ACE is smaller. Then

$$\text{area OAB} = \text{area OAC} + \text{area ODB} - (\text{area BDE} - \text{area ACE})$$
$$< \tfrac{1}{2}q \,.\, \text{AC} + \tfrac{1}{2}s \,.\, \text{BD} = \tfrac{1}{2}$$

and again we have a contradiction.                                          QED

HISTORICAL NOTES   Chuquet described the *Parmenides* algorithm under the name *La règle des nombres moyens* in his *Triparty* of 1484:

This rule serves to find as many numbers intermediate between two neighbouring numbers as one desires. By its means it is possible to find many more numbers and to do more calculations than are found by the rule of three or by one position or by two positions. . . . Numerator is added to numerator, and denominator to denominator [Chuquet–Flegg, Hay, & Moss, *NC*, 90].

There was some sixteenth-century exploration of this technique by Juan de Ortega and Buteo (see Dickson, *HTN* ii, 350–1 for references), but modern commentators often seem not to have realised its potential. For example, the *DSB*, s.v. Chuquet, says of it: ". . . this rule, together with a lot of patience, makes it possible to solve any problem allowing of a rational solution. . . . Since he was little concerned about rapid methods of approximation, Chuquet throughout his work used nothing but this one rule. . . ."

The geometrical description of continued fractions in terms of the integer lattice is sometimes attributed to Klein, but the idea is to be found twenty years earlier in Smith, NCF, and then, shortly afterwards, in Poincaré, GFC. Klein made the following additional observation:

Imagine pegs or needles affixed at all the integral points, and wrap a tightly drawn string about the sets of pegs to the right and then to the left of the ray $[Y = x_0 X]$, then the

vertices of the two convex string-polygons which bound our two points sets will be precisely the points $(q_k, p_k)$ whose coordinates are the numerators and denominators of the successive convergents to $x_0$, the left polygon having the even convergents, the right one the odd. This gives a new and, one may well say, an extremely graphic definition of a continued fraction [Klein, *EMAS* i, 44; he illustrates it by Fig. 9.2, above].

The approach of the *Parmenides* algorithm is closely related to the properties of Farey series, on which see, for example, Dickson, *HNT* i, 155–8, or Hardy & Wright, *ITN*, Chapter 3 (and do not miss the note to §3.1).

Lagrange opened his *Additions* to Euler's *Algebra* with a discussion of approximations of the form $|x_0 q - p|$, and he showed that the best such approximations correspond to the convergents of the continued fraction expansion of $x_0$. This was extended in Smith, *NCF*, to the more complicated situation of approximations of the form $|x_0 - p/q|$, whose solution involves the intermediate convergents. This, incidentally, gives a complete solution to Dr Davenant's problem, which was described in the Historical Notes to the previous section.

### 9.1(c)  *The quadratic theory*

Let us start with an example such as delighted eighteenth-century (and, in a different guise, seventeenth-century) mathematicians. This is taken from Euler, UNAPPS (1767) §14, but expressed in the notation that has been established here.

Take $x_0 = \sqrt{54}$; then

$$n_0 = 7 \quad \text{and} \quad x_1 = \frac{1}{\sqrt{54} - 7} = \frac{\sqrt{54} + 7}{5};$$

$$\text{so } n_1 = 2 \quad \text{and} \quad x_2 = \frac{5}{\sqrt{54} - 3} = \frac{5(\sqrt{54} + 3)}{45} = \frac{\sqrt{54} + 3}{9};$$

$$\text{so } n_2 = 1 \quad \text{and} \quad x_3 = \frac{9}{\sqrt{54} - 6} = \frac{9(\sqrt{54} + 6)}{18} = \frac{\sqrt{54} + 6}{2};$$

$$\text{so } n_3 = 6 \quad \text{and} \quad x_4 = \frac{2}{\sqrt{54} - 6} = \frac{2(\sqrt{54} + 6)}{18} = \frac{\sqrt{54} + 6}{9};$$

$$\text{so } n_4 = 1 \quad \text{and} \quad x_5 = \frac{9}{\sqrt{54} - 3} = \frac{9(\sqrt{54} + 3)}{45} = \frac{\sqrt{54} + 3}{5};$$

$$\text{so } n_5 = 2 \quad \text{and} \quad x_6 = \frac{5}{\sqrt{54} - 7} = \frac{5(\sqrt{54} + 7)}{5} = \sqrt{54} + 7;$$

$$\text{so } n_6 = 14 \quad \text{and} \quad x_7 = \frac{1}{\sqrt{54} - 7} = x_1.$$

Hence $\sqrt{54} = [7, \overline{2, 1, 6, 1, 2, 14}]$, where the bar indicates that the term will repeat indefinitely.

Euler frequently gives such examples to illustrate the algorithm. Here, in UNAPPS, he starts with his favourite but less illuminating example $\sqrt{13}$, and then gives $\sqrt{61}$ and $\sqrt{67}$, all in complete detail; then he goes on to $\sqrt{31}$, $\sqrt{46}$, and $\sqrt{54}$ in slightly abbreviated format; then he gives a table of the expansions of $\sqrt{n}$ for $2 \leqslant n \leqslant 120$, $n$ not a square (see Table 3.1, above) and describes their form when $n = m^2 + 1$, $m^2 + 2$, $m^2 + m - 1$, $m^2 + 2m$, $4m^2 + 4$, $9m^2 + 3$, and $9m^2 + 6$; these correspond, of course, to the hypotheses of Section 3.3, above.[15]

We can describe the process in general as follows:

Let $x_0 = (a_0 + b_0\sqrt{n})/c_0$, where $n$ is not a square (or, equivalently, $x_0$ is the irrational solution of some quadratic equation with integer coefficients).[16] At each step of the algorithm we write

$$x_k = \frac{a_k + b_k\sqrt{n}}{c_k} = n_k + \frac{(a_k - n_k c_k) + b_k\sqrt{n}}{c_k},$$

where $n_k = \lfloor x_k \rfloor$. Then

$$x_{k+1} = \frac{c_k}{(a_k - n_k c_k) + b_k\sqrt{n}} = \frac{c_k((a_k - n_k c_k) - b_k\sqrt{n})}{(a_k - n_k c_k)^2 - b_k^2 n},$$

and so we get a quadratic surd of the same kind, in which the new coefficients $a_{k+1}$, $b_{k+1}$, and $c_{k+1}$ are simple functions of the $a_k$, $b_k$, and $c_k$. However, it will now be found that these coefficients grow rapidly in size as they pick up new common factors at each step. In the example of $\sqrt{54}$, above, these factors were divided out at each step; in the general case we must be a bit more subtle to ensure that this will happen.

Manipulate the surd into the form $x_0 = (a_0 + \sqrt{n})/c_0$ where $n$ is not a square and $c_0$ divides $n - a_0^2$; this can be done by writing

$$\frac{a_0 + b_0\sqrt{n}}{c_0} = \frac{\pm a_0 c_0 + \sqrt{(b_0^2 c_0^2 n)}}{\pm c_0^2},$$

and note that, by fortunate chance, $\sqrt{n} = (0 + \sqrt{n})/1$ is already in this required form. Now look at the step from $x_k$ to $x_{k+1}$:

$$x_{k+1} = \frac{c_k}{(a_k - n_k c_k) + \sqrt{n}} = \frac{c_k((n_k c_k - a_k) + \sqrt{n})}{n - (n_k c_k - a_k)^2};$$

---

[15] An understanding of the material that follows will be greatly enhanced by exploring some numerical examples of this, and other such, algorithms. A programmable pocket calculating machine gives an advantage undreamt of by previous generations of mathematicians.

[16] All letters except $x$, $y$, and $z$ will hereafter denote integers; $n$ and $m$ will denote positive integers, except that $n_0$, the initial term of the expansion, may sometimes be negative; and $\sqrt{n}$ is always the positive square root.

if $c_k$ divides $n - (n_k c_k - a_k)^2$, we can then write this as

$$x_{k+1} = \frac{a_{k+1} + \sqrt{n}}{c_{k+1}} \quad \text{where} \quad a_{k+1} = n_k c_k - a_k \quad \text{and} \quad c_{k+1} = \frac{n - a_{k+1}^2}{c_k}.$$

But $c_{k+1} = (n - a_{k+1}^2)/c_k$ divides $n - a_{k+1}^2$ if and only if $c_k$ divides $n - a_{k+1}^2 = n - a_k^2 - c_k(n_k - c_k)$, i.e. if and only if $c_k$ divides $n - a_k^2$, which we have arranged to be so. Hence $x_{k+1}$ will indeed again be a surd of the required form. Let us state this formally:

ALGORITHM FOR QUADRATIC SURDS    Express the quadratic surd in the form $x_0 = (a_0 + \sqrt{n})/c_0$ where $n$ is not a square and $c_0$ divides $n - a_0^2$. For $k \geqslant 0$ let

$$n_k = \lfloor x_k \rfloor, \quad \text{the integer part of } x_k,$$

$$a_{k+1} = n_k c_k - a_k, \quad \text{and}$$

$$c_{k+1} = (n - a_{k+1}^2)/c_k.$$

Then $x_{k+1}$ will again be of the required form, and hence $x_0 = [n_0, n_1, n_2, \ldots]$.

There is a remarkable theory of continued fractions of quadratic surds. We start with an easy result:

PROPOSITION    If the continued fraction of $x_0$ is eventually periodic,

$$x_0 = [n_0, n_1, \ldots, n_k, \overline{m_0, m_1, \ldots, m_l}],$$

then $x_0$ is a quadratic surd.

PROOF    Observe first that

$$y_0 = [\overline{m_0, m_0, \ldots, m_l}] = [m_0, m_1, \ldots, m_l, y_0] = \frac{y_0 r_l + r_{l-1}}{y_0 s_l + s_{l-1}}$$

where $r_{l-1}/s_{l-1}$ and $r_l/s_l$ are the last two convergents of the first period of $y_0$. Hence $y_0$ satisfies the quadratic equation

$$s_l y^2 + (s_{l-1} - r_l)y - r_{l-1} = 0.$$

Since the coefficient of $y^2$ and the absolute term have opposite signs, there will be one positive and one negative root; so $y_0$ must be the positive root. Finally

$$x_0 = [n_0, n_1, \ldots, n_k, y_0] = \frac{y_0 p_k + p_{k-1}}{y_0 q_k + q_{k-1}}$$

is again a quadratic surd.                                          QED

The converse of this proposition is also true, but the proof is much more difficult.

PROPOSITION (Lagrange, AMREN (1770)) The continued fraction of a quadratic surd is eventually periodic,

PROOF We shall prove that, in the notation above for the algorithm for quadratic surds, there is an index $K$ such that, for all $k \geqslant K$, $1 \leqslant a_k < \sqrt{n}$ and $1 \leqslant c_k < 2\sqrt{n}$. Hence some $x_k$ must eventually recur as $x_l$ with $l > k$, and we shall then have $x_0 = [n_0, n_1, \ldots, \overline{n_k, \ldots, n_{l-1}}]$.

The surd $x_0 = (a_0 + \sqrt{n})/c_0$ is one root of some quadratic equation $Ax^2 + Bx + C = 0$; the other root, called the (algebraic) conjugate of $x_0$, will then be $x_0' = (a_0 - \sqrt{n})/c_0$. Write, similarly, $x_k'$ for the conjugate of $x_k$. Now call $x_k$ a reduced quadratic surd if $x_k > 1$ and $-1 < x_k' < 0$. Then, if $x_k$ is reduced:

$$x_k - x_k' > 0, \quad \text{so} \quad \sqrt{n}/c_k > 0, \quad \text{so} \quad c_k > 0,$$
$$x_k + x_k' > 0, \quad \text{so} \quad a_k/c_k > 0, \quad \text{so} \quad a_k > 0,$$
$$x_k' < 0, \quad \text{so} \quad a_k < \sqrt{n}, \quad \text{and}$$
$$x_k > 1, \quad \text{so} \quad c_k < (a_k + \sqrt{n}) < 2\sqrt{n}.$$

These are precisely the inequalities we wish to establish.

Now suppose $x_k$ is reduced. Then $x_{k+1} = 1/(x_k - n_k) > 1$, and $x_{k+1}' = -1/(n_k - x_k')$, where $-1 < x_k' < 0$ and $n_k \geqslant 1$. Hence $-1 < x_{k+1}' < 0$. Thus if $x_k$ is reduced, $x_{k+1}$ will also be reduced.

It remains to show that some $x_k$ is reduced. Observe first that $x_k \geqslant 1$ for all $k \geqslant 1$. Next, we have shown that

$$x_0 = \frac{x_{k+1} p_k + p_{k-1}}{x_{k+1} q_k + q_{k-1}},$$

and so the similar relation holds between $x_0'$ and $x_{k+1}'$. Hence

$$x_{k+1}' = \frac{x_0' q_{k-1} - p_{k-1}}{x_0' q_k - p_k} = -\frac{q_{k-1}}{q_k} \left( \frac{x_0' - p_{k-1}/q_{k-1}}{x_0' - p_k/q_k} \right).$$

Now both $p_k/q_k$ and $p_{k-1}/q_{k-1}$ tend to $x_0$ as $k$ increases, and they lie on opposite sides of $x_0$. Hence the quantity in brackets will tend to 1, and will eventually alternate on each side of 1. Also $0 < q_{k-1} < q_k$. So, eventually, some $x_K'$ will satisfy $-1 < x_K' < 0$ and the corresponding $x_K$ will be reduced. Hence the proposition is proved. QED

PROPOSITION (Galois, DTFCP (1829)) A quadratic surd $x_0$ has a purely periodic continued fraction expansion if and only if it is reduced, and the purely periodic number got by reversing the period is then $-1/x_0'$.
PROOF Suppose that $x_0 = [\overline{n_0, n_1, \ldots, n_k}]$; observe that for this to be a simple continued fraction, we must have $n_0 \geqslant 1$, so $x_0 > 1$. Now consider the two equations

$$x = [n_0, n_1, \ldots, n_k, x] \quad \text{and} \quad y = [n_k, n_{k-1}, \ldots, n_0, y].$$

We have seen in the previous proposition but one that these are each quadratic equations of which one root is positive, the other roof is negative. We now relate their roots.

Manipulate the first equation as follows:

$$x = n_0 + \cfrac{1}{n_1+} \cdots \cfrac{1}{n_k+} \cfrac{1}{x},$$

so

$$\frac{1}{n_0 - x} = -n_1 - \cfrac{1}{n_2+} \cdots \cfrac{1}{n_k+} \cfrac{1}{x},$$

so

$$\cfrac{1}{n_1+} \cfrac{1}{n_0 - x} = -n_2 - \cfrac{1}{n_3+} \cdots \cfrac{1}{n_k+} \cfrac{1}{x},$$

and so, eventually,

$$n_k + \cfrac{1}{n_{k-1}+} \cdots \cfrac{1}{n_1+} \cfrac{1}{n_0 - x} = \frac{1}{-x},$$

or, writing $y = -1/x$, $y = [n_k, n_{k-1}, \ldots, n_0, y]$. Hence, if the roots of the first equation are $x_0$ and $x_0'$, then the roots of the second are $-1/x_0$ and $-1/x_0'$. Since $-1/x_0$ is negative, its conjugate $-1/x_0'$ must be positive; it is also greater than 1. Hence $-1 < x_0' < 0$, and $x_0$ is reduced.

Conversely, suppose that $x_0$ is reduced; we show that its continued fraction expansion is purely periodic. We know, from Lagrange's theorem, that it is eventually periodic, so that $x_k = x_l$ for some $k$ and $l$ with $k < l$. Now $x_k = n_k + 1/x_{k+1}$, so $x_k' = n_k + 1/x_{k+1}'$. Write $y_k$ for $-1/x_k'$; then, since $x_k'$ and $x_{k+1}'$ lie between $-1$ and 0, $y_k$ is greater than 1, and we have the relation

$$-\frac{1}{y_k} = n_k - y_{k+1}, \quad \text{i.e.} \quad y_{k+1} = n_k + \frac{1}{y_k}.$$

Hence $n_k$ is the integer part of $y_{k+1}$. Now if $x_k$ and $x_l$ are equal, then $x_k'$ and $x_l'$ will be equal, hence $y_k$ and $y_l$ will be equal, and hence $n_{k-1}$ and $n_{l-1}$ will be equal. But

$$x_{k-1} = n_{k-1} + 1/x_k \quad \text{and} \quad x_{l-1} = n_{l-1} + 1/x_l,$$

so $x_{k-1}$ and $x_{l-1}$ will be equal. Repeating this argument, we will get eventually that $x_0$ and $x_{l-k}$ are equal, so the continued fraction is purely periodic. QED

PROPOSITION If $n/m$ is a positive rational number greater than 1 whose square root is irrational, then

$$\sqrt{(n/m)} = [n_0, \overline{n_1, n_2, \ldots, n_2, n_1, 2n_0}],$$

and conversely, if $x_0$ has an expansion of this form, then it is the irrational

square root of a rational number greater than 1. (For numbers between 0 and 1, the expansion starts with a zero term.)

PROOF   Consider $y_0 = \sqrt{(n/m)} + n_0$, where $n_0 = \lfloor \sqrt{(n/m)} \rfloor$; its conjugate $n_0 - \sqrt{(n/m)}$ lies between $-1$ and $0$, and so $y_0$ is a reduced quadratic surd. Hence, by Galois' theorem, its continued fraction will be purely periodic, of the form

$$\sqrt{(n/m)} + n_0 = [\overline{2n_0, n_1, n_2, \ldots, n_k}],$$

while

$$\frac{1}{\sqrt{(n/m)} - n_0} = [\overline{n_k, n_{k-1}, \ldots, 2n_0}].$$

From this second expansion we get, by reciprocating and adding $2n_0$,

$$\sqrt{(n/m)} + n_0 = [2n_0, \overline{n_k, n_{k-1}, \ldots, 2n_0}],$$

and since the expansion is unique, we see that $n_1 = n_k$, $n_2 = n_{k-1}$, etc.

Conversely, suppose that

$$x_0 = [n_0, \overline{n_1, n_2, \ldots, n_2, n_1, 2n_0}].$$

Then $y_0 = x_0 + n_0$ is purely periodic, so its conjugate root $y_0'$ is given by

$$-\frac{1}{y_0'} = [\overline{n_1, n_2, \ldots, n_2, n_1, 2n_0}],$$

so $y_0' = 2n_0 - y_0$. Hence $y_0$ is the positive root of $(y - y_0)(y - 2n_0 + y_0) = 0$, i.e. of $y^2 - 2n_0 y + y_0 y_0' = 0$. Hence $y = n_0 + \sqrt{z_0}$ for some rational number $z_0$, and $x_0 = \sqrt{z_0}$.                          QED

PROPOSITION (see Section 2.4(e))   If

$$\sqrt{n} = [n_0, \overline{n_1, n_2, \ldots, n_2, n_1, 2n_0}]$$

and $p_k/q_k = [n_0, n_1, \ldots, n_2, n_1]$ is the penultimate convergent of the first period of the expansion (so the period contains $k + 1$ terms), then

$$p_k^2 - nq_k^2 = (-1)^{(k+1)}.$$

Similarly, for the penultimate convergent of the second period,

$$p_{2k+1}^2 - nq_{2k+1}^2 = (-1)^{2k+2} = 1,$$

and so on for each succeeding period; so the signs alternate if $k$ is even, and are still all positive if $k$ is odd.

PROOF   We have

$$\sqrt{n} = x_0 = \frac{x_{k+1} p_k + p_{k-1}}{x_{k+1} q_k + q_{k-1}},$$

where

$$x_{k+1} = [2n_0, \overline{n_1, n_2, \ldots, n_2, n_1, 2n_0}] = \sqrt{n} + n_0.$$

Hence
$$\sqrt{n}((\sqrt{n}+n_0)q_k + q_{k-1}) = (\sqrt{n}+n_0)p_k \mp p_{k-1},$$
that is
$$\sqrt{n}(n_0 q_k + q_{k-1} - p_k) + (nq_k - n_0 p_k - p_{k-1}) = 0.$$
Since $\sqrt{n}$ is irrational, this implies that
$$n_0 q_k + q_{k-1} - p_k = nq_k - n_0 p_k - p_{k-1} = 0.$$
Now substitute for $p_{k-1}$ and $q_{k-1}$ in $p_k q_{k-1} - p_{k-1} q_k = (-1)^{k+1}$. We get
$$p_k^2 - nq_k^2 = (-1)^{k+1}. \qquad\qquad \text{QED}$$

PROPOSITION   Every solution of $x^2 - ny^2 = \pm 1$ arises as a convergent of $\sqrt{n}$.
PROOF   First suppose $p^2 - nq^2 = 1$. Then, since $p > q\sqrt{n} > q$,
$$\frac{p}{q} - \sqrt{n} = \frac{1}{q(p+q\sqrt{n})} < \frac{1}{2pq} < \frac{1}{2q^2}.$$

Hence, by an earlier proposition, $p/q$ is a convergent of $\sqrt{n}$.
   Next, if $p^2 - nq^2 = -1$, then $nq^2 - p^2 = 1$. Again, since $q\sqrt{n} > p > q$,
$$\sqrt{n} - \frac{p}{q} = \frac{1}{q(q\sqrt{n}+p)} < \frac{1}{2pq} < \frac{1}{2q^2},$$

and the result again follows.                                    QED

HISTORICAL REMARKS   The material of this section first manifested itself in connection with Fermat's challenge of 1657 concerning the so-called Pell's equation, $x^2 - ny^2 = 1$ (see Section 2.4(e), above, for quotations and further details). Euler was the first to publish a continued fraction description of the 'English solution' (by Brouncker & Wallis) in his UNAPPS, written in 1759 but not published until 1767. The first actual proof that Pell's equation can always be solved was by Lagrange, SPA, written in 1768 but not published until 1773; it was very long and, after an initial step in which continued fractions were used to find a subsidiary equation $x^2 - ny^2 = r$ with an infinite number of solutions, it was based on his rediscovery of the 'Indian solution'. (The introduction to this article is quoted in n. 55, below). By the time this article was published, Lagrange had found another, simpler and clearer, proof, based on the properties of the continued fraction expansion of quadratic surds; the most succinct exposition of the idea of this proof is in §37 of Lagrange's *Additions* to Euler's *Algebra*. The story continues with further contributions by Euler, Lagrange, Legendre, Galois, and others. A masterly account, with its full mathematical and historical context, can be found in Weil, *NT*.
   I shall not attempt to summarise or select further episodes from this rich picture, beyond the following comment. An important development in this

story occurs with the publication of Euler's UNAPPS, in which we find one of the first published systematic explorations of the properties of the continued fraction expansion of a quadratic surd. Euler's motivating problem is the solution of Pell's equation, but most of the article is taken up in describing the algorithm for evaluating the continued fraction expansion of $\sqrt{n}$, in tabulating many results, and in inferring some properties, exactly parallel to those that underlie my exploration of the dimension of squares, in Section 3.3, above. Euler then turns to Pell's equation, and the first cases he considers are again directly analogous to the further developments I proposed in Section 3.4. He explores those cases when the period of the expansion of $\sqrt{n}$ contains 1, then 2, then 3, and so on up to 8 terms, displaying the solution in each case, and then finishes by inferring the general procedure and tabulating the solutions. The first complete proofs of the behaviour he has uncovered, by Lagrange, follow a completely different path just as, in Section 3.6, the general proofs of the dimension of squares are resolved by other very circuitous means. Euler is rare among mathematicians in having the desire and facilities to lay before his readers his infectious enthusiasm for the initial steps of exploration and conjecture that precede much of deductive mathematics; if all we possessed on number theory up to the early nineteenth century was Gauss's *Disquisitiones Arithmeticae* (see Section 9.2(d), below), then it would be very difficult to reconstruct this early heuristic phase.

For Greek mathematics, we know very little of the developments that led up to the *Elements* and any reconstruction must therefore of necessity be very speculative. This being admitted, I have tried to show how one can find hints of a similar substrate of problems, explorations, and conjectures that eventually give rise to some of the most formal and arid parts of the *Elements*, for example Books II and X. And I was again surprised to discover subsequently how the analogous explorations by eighteenth-century mathematicians followed a very similar path; and then to see the extent to which the greater part of these explorations and manipulations has already been set aside and forgotten by later generations of mathematicians.

### 9.1(d) *Analytic properties*[17]

As so far described, most of the developments can be loosely categorised as arithmetical and number theoretical. In addition, alongside, and even in advance of, these discoveries, there was a developing analytic theory of continued fractions, starting with Brouncker's discovery of 1654 that

$$\frac{\pi}{4} = \frac{1}{1+} \frac{1}{2+} \frac{3^2}{2+} \frac{5^2}{2+} \cdots,$$

---

[17] Here I use the words 'analysis' and 'analytic' in their modern sense, as connected with limiting processes, infinite operations, the theory of functions, the use of the calculus, etc. Seventeenth-century use of 'analysis' was different, and started from speculations about the lost methods used by the Greeks in their mathematical discoveries.

an illustration of the promise in passing between expressions as infinite products, continued fractions, and infinite series. ʿAnother such hint, perhaps the first suggestion of interesting analytic behaviour involving simple continued fractions and in itself as surprising as the expansion of $\sqrt{n}$, was the expansion

$$e = [2, 1, 2, 1, 1, 4, 1, 1, 6, 1, \ldots]$$

first uncovered by Cotes in 1714. These results in arithmetic and analysis excited and inspired Euler and Lambert no less than the number-theoretic explorations; in fact Euler's first references to continued fractions are found in his correspondence in connection with the Riccati differential equation and the expansion of $e$.[18] This is not the place to attempt a survey of these and later analytic developments,[19] and I shall restrict myself here to two very brief observations about the analytic theory during the eighteenth century.

First, it is an immediate consequence of a proof of the expansion of $e$ that it is irrational (see Section 2.2(v)(vii), above), and indeed this is how it and associated results were first used by Lambert and Euler. Second, as an illustration of this developing analytic theory, let me describe Chapter 18, 'De Fractionibus Continuis', of Euler's *Introductio in Analysin Infinitorum*, his influential introductory text on the techniques of the newly developing subject of analysis.

The chapter opens with a description of the convergents of both simple and general continued fractions, and then passes immediately to transformations between the convergents of continued fractions and the partial sums of infinite series. These are then applied to several examples:

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots = \frac{1}{1+} \frac{1}{1+} \frac{4}{1+} \frac{9}{1+} \frac{16}{1+} \cdots ;$$

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots = \frac{1}{1+} \frac{1}{2+} \frac{9}{2+} \frac{25}{2+} \frac{49}{2+} \cdots ;$$

$$\frac{\pi}{n} \cot \frac{m\pi}{n} = \frac{1}{m} - \frac{1}{n-m} + \frac{1}{n+m} - \frac{1}{2n-m} + \frac{1}{2n+m} - \cdots$$

$$= \frac{1}{m+} \frac{m^2}{n-2m+} \frac{(n-m)^2}{2m+} \frac{(n+m)^2}{n-m+} \frac{(2n-m)^2}{2m+} \frac{(2n+m)^2}{n-2m+} \cdots ;$$

$$\frac{1}{e} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots, \quad \text{whence} \quad \frac{1}{e-1} = \frac{1}{1+} \frac{2}{2+} \frac{3}{3+} \frac{4}{4+} \cdots ;$$

---

[18] See Euler, FCD, or Euler–Wyman & Wyman, ECF, for further details. The reference to the correspondence with Goldbach can be found in the note in the *Opera Omnia* to §28, or in Weil, *NT*, 183 or 230.

[19] At this point, historical work on the topic starts to become seriously deficient. The broad lines of development can be inferred from Perron, *LK*; and the later part of the sketch in Brezinski, LHCF, expanded in his *HCF*, is biased towards analytic properties. I know of no other general historical surveys of the subject.

and more. Only at this stage does Euler turn to periodic simple continued fractions and quadratic surds, which he introduces with a view to making some remarks about convergence and approximation; for elsewhere in this chapter, as throughout his works, he spends little time on such justifications of his manipulations, though he operated with a sure intuition for what was permitted or not. The chapter finishes with some comments on the conversion of decimal fractions into continued fractions, which he illustrates on the expansion of

$$\frac{e-1}{2} = [0, 1, 6, 10, 14, 18, 22, \ldots],$$

$$\pi = [3, 7, 15, 1, 292, 1, 1, \ldots],$$

and an astronomical illustration, the ratio of the day to the solar year, [365, 4, 7, 1, 6, 1, 2, 2, 4, . . .], and the Julian and Gregorian calendars.

Such was the basic introduction that Euler set out. It is attractive, accessible, and directed more to describing manipulations than offering rigorous proofs. It is really more arithmetic than analysis, for he considers only numbers that arise in analytic contexts, like $e$ and $\pi$, rather than the functions of analysis like

$$\cotan x = [x^{-1}, -3x^{-1}, 5x^{-1}, -7x^{-1}, \ldots],$$

such as appear abundantly in his articles, and which were also used by Lambert. It is an inspiration to further investigation.

I refer those readers who are rightly dissatisfied with this *hors d'oeuvre* to the comprehensive descriptions of the theory that has subsequently developed, for example in Jones & Thron, *CF*; Henrici, *ACCA* ii; Perron, *LK*; or Wall, *ATCF*; and I shall give examples of more ambitious analytic manipulations in Sections 9.3(b) and (c) below.

### 9.1(e) *Lagrange and the solution of equations*

The solution of a general polynomial equation with integer coefficients is a pervasive theme in mathematics. One approach—solution in radicals—reaches a climax in the work of Galois, and Lagrange played no small part in this development with his long survey article RRAE. This successful resolution, albeit negative, of this much studied problem has subsequently overshadowed many of the previous developments and rendered them obsolete or irrelevant. So, for example, the already obsolete theory of the construction of equations, initiated by Descartes, has now disappeared almost completely from the view of even all but the most specialised histories of mathematics (for discussion, see Bos, AMRDMT); and Lagrange's other approach, which I shall now briefly describe, has now also almost vanished, even though it is a very serviceable technique for computing the roots of equations.

Lagrange's work is contained in a series of memoirs that start with his REN (1769), followed a year later by AMREN, almost twice as long as the original memoir, then followed by further memoirs, all of them published by the Berlin

Academy; then later all of this material was brought together into a full-length treatise, *TREN* (1798), with a second edition in 1808. His treatment ranges widely, but its focus is the following observation. Take the equation

$$a_k x^k + a_{k-1} x^{k-1} + \cdots + a_1 x + a_0 = 0$$

and suppose, for simplicity, that it has only one positive root $x_0$, which we can further suppose is greater than 1 (otherwise replace $x$ by $1/x$). Then $n_0 \leqslant x_0 < n_0 + 1$ for some $n_0$. So, writing $x_0 = n_0 + 1/x_1$, we see that $x_1$ will be the unique positive root, in fact greater than 1, of a new equation

$$b_k x^k + b_{k-1} x^{k-1} + \cdots + b_1 x + b_0 = 0,$$

where

$$b_k = a_k n_0^k + a_{k-1} n_0^{k-1} + \cdots + a_0,$$

$$b_{k-1} = k a_k n_0^{k-1} + (k-1) a_{k-1} n_0^{k-2} + \cdots + a_1,$$

etc.

This equation, again, will have only one positive root $x_1$, again greater than 1, so we can again express this as $x_1 = n_1 + 1/x_2$, etc. Thus we evaluate the solution of the original equation as a continued fraction, $x_0 = [n_0, n_1, n_2, \ldots]$. Lagrange dealt with many other aspects of this procedure, and also described procedures for evaluating the approximations $n_0$, for computing coefficients of the new equation, for separating roots that lie close together, for finding complex roots, and so on.

The first illustration that he gives is of the time-honoured cubic equation $x^3 - 2x - 5 = 0$ which, since Newton's time, had been the standard test for methods of solving equations (see Whiteside, PMTLSC, 206). He gets $x = [2, 10, 1, 1, 2, 1, 3, 1, 1, 12, \ldots]$ from which he evaluates the convergents and gives estimates of their errors. But here, in a deliberate reference to what I have earlier called the problem of the dimension of cubes (see Section 4.3, above), I shall illustrate the process on the equation $x^3 - 2 = 0$.

We start by writing out the formulae explicitly for $k = 3$. The equation

$$f(x) = a_3 x^3 + a_2 x^2 + a_1 x + a_0 = 0,$$

in which we may suppose $a_3 > 0$, becomes, under the transformation $x \to n + 1/x$, the equation

$$-(a_3 n^3 + a_2 n^2 + a_1 n + a_0) x^3$$
$$- (3a_3 n^3 + 2a_2 n + a_1) x^2 - (3a_3 n + a_2) x - a_3 = 0.$$

Since we have supposed that $f(x) = 0$ has only one positive root which satisfies $n < x_0 < n + 1$, and since $a_3 > 0$, we must have $f(n) < 0$; thus again $b_3 = -f(n) > 0$. We then iterate the procedure.

Consider now $x^3 - 2 = 0$. Then $1 < x_0 < 2$ so $n_0 = 1$; $b_3 = 1$, $b_2 = -3$, $b_1 = -3$, $b_0 = -1$; then the positive root $x_1$ of $x^3 - 3x^2 - 3x - 1 = 0$ satisfies

TABLE 9.1. Evaluation of $\sqrt[3]{2} = [1, 3, 1, 5, 1, 1, \ldots]$
by Lagrange's algorithm

| $k$ | Equation | $n_k$ |
|---|---|---|
| 0 | $x^3 - 2 = 0$ | 1 |
| 1 | $x^3 - 3x^2 - 3x - 1 = 0$ | 3 |
| 2 | $10x^3 - 6x^2 - 6x - 1 = 0$ | 1 |
| 3 | $3x^3 - 12x^2 - 24x - 10 = 0$ | 5 |
| 4 | $55x^3 - 81x^2 - 33x - 3 = 0$ | 1 |
| 5 | $62x^3 + 30x^2 - 84x - 55 = 0$ | 1 |
| etc. | | |

$3 < x_1 < 4$, so $n_1 = 3$, etc. The first six steps are set out in Table 9.1; we get $\sqrt[3]{2} = [1, 3, 1, 5, 1, 1, \ldots]$. The arithmetical operations can be considerably refined and economised to make this a practical algorithm for hand or machine calculation.[20]

Lagrange's treatise contains a rich study of general properties of polynomial equations, dealing, for example, with positive and negative, and real and complex, roots, but all directed towards describing and analysing his algorithm for expressing the solutions as continued fractions—a procedure that he clearly considered as important and valuable. (Some of this work is taken up and extended by Legendre, Fourier, Vincent, and others; see, for example, Serret, *CAS*, Chapters 6 and 7. It has even been argued, in Grabiner, CB, that Lagrange's statement and use of the intermediate value theorem led Bolzano and Cauchy to their pioneering work on the foundations of analysis.) Lagrange's treatises also contain the first proof of the periodicity of the expansion of a quadratic surd, at AMREN, §§32–43 or *TREN*, §§52–59. Hence, he observes, the method will identify all linear and quadratic rational irreducible factors of the polynomial, since the corresponding roots, being rational numbers or quadratic surds, will be revealed by his procedure.[21] He then concludes:

It is to be hoped that one may also find some characteristic that may serve to reveal the rational divisors of third, fourth, . . . degrees, which the given equations may have. This

---

[20] For a succinct modern description of this and other continued fraction algorithms, see the exercises to Knuth, *ACP* ii, §4.5.3; this is exercise 20. A detailed practical implementation is given in Rosen & Shallit, CFA. Also see n. 32, below.

[21] It must be added that this information will not be effectively computable, since we have no idea how far the computation must be extended for the periodicity to reveal itself. Lagrange was well aware of this problem, and described mathematical techniques for identifying periodicity, based on the idea of generating functions. For comments on this problem of recognising ˘periodic behaviour, see Ferguson, RRQN, and Fowler, BTEE2, 34.

seems to me to be a topic very worthy of the attention of mathematicians [AMREN §49, or *TREN*, §64].[22]

As I shall describe in Section 9.3(b), below, the search for some such criterion, based on some generalisation of continued fractions, is still active, but has not really yielded the results that Lagrange had hoped.

## 9.2 GAUSS AND CONTINUED FRACTIONS

### 9.2(a) *Introduction*

At the end of the eighteenth century, a chorus of mathematicians sang forth the wonders of continued fractions. If we restrict our attention to material published more or less in the previous fifty years, we see, in the front rank: Euler, with his *Introductio in Analysin Infinitorum* (1748) and *Anleitung zur Algebra* (1770), as well as a bundle of articles on the topic; Lambert, who, in several articles, discussed the expansion of functions as continued fractions, including details of convergence, and then proved the irrationality of numbers connected with exponential and trigonometric functions; Lagrange, with his *Additions, De l'analyse indéterminée* (1773) to the French translation of Euler's *Algebra*, his *Traité de la résolution des équations numériques de tous les degrés* (1769–98), and a series of memoirs that build on and establish rigorously results uncovered by Euler; and Legendre, with *Recherches d'Analyse Indéterminée* (1785) and *Essai sur la Théorie des Nombres* (1798) which are, in great part, based on continued fraction techniques. Around them are a host of supporting figures.

Gauss (1777–1855) was heir to this material, and one can be sure that he assimilated it. For example, continued fractions play a significant part in his published work in analytic contexts, in connection with hypergeometric series and numerical integration, and we know from his unpublished notes that he investigated the probability theory of continued fractions over a twelve-year period, at least. But his enormously influential book on number theory, the *Disquisitiones Arithmeticae*, makes almost no reference to continued fractions. I shall now look in turn at Gauss's treatment of each of these topics.

### 9.2(b) *Continued fractions and the hypergeometric series*

Many elementary functions of analysis are particular cases of the hypergeometric series

$$F(\alpha, \beta, \gamma, x) = 1 + \frac{\alpha\beta}{\gamma} x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{2!\gamma(\gamma+1)} x^2 + \cdots,$$

---

[22] "Il serait à souhaiter que l'on pût trouver aussi quelque caractère qui pût servir à faire reconnaitre les diviseurs commensurables du troisième, quatrième, . . . degré, lorsqu'il y en a dans l'équation proposée; c'est du moins une recherche qui me parait très-digne d'occuper les Géomètres."

which, for $\alpha$, $\beta$, and $\gamma$ different from $0, -1, -2, -3, \ldots$, is an infinite series which converges for $|x| < 1$. The proposal to generalise the geometric series goes back to Wallis' *Arithmetica Infinitorum*, while the study of this particular generalisation was initiated by Euler. Gauss's mastery of the topic was complete.[23]

The hypergeometric satisfies many identities. Here we shall start from

$$F(\alpha,\beta,\gamma,x) = F(\alpha,\beta+1,\gamma+1,x) - \frac{\alpha(\gamma-\beta)}{\gamma(\gamma+1)} F(\alpha+1,\beta+1,\gamma+2,x)x,$$

which we write as

$$\frac{F(\alpha,\beta+1,\gamma+1,x)}{F(\alpha,\beta,\gamma,x)} = \left(1 - \frac{\alpha(\gamma-\beta)}{\gamma(\gamma+1)} \frac{F(\alpha+1,\beta+1,\gamma+2,x)}{F(\alpha,\beta+1,\gamma+1,x)}x\right)^{-1},$$

whence, by exchanging $\alpha$ and $\beta$, and then replacing $\beta$ by $\beta+1$ and $\gamma$ by $\gamma+1$, we get

$$\frac{F(\alpha+1,\beta+1,\gamma+2,x)}{F(\alpha,\beta+1,\gamma+1,x)}$$

$$= \left(1 - \frac{(\beta+1)(\gamma-\alpha+1)}{(\gamma+1)(\gamma+2)} \frac{F(\alpha+1,\beta+2,\gamma+3,x)}{F(\alpha+1,\beta+1,\gamma+2,x)}x\right)^{-1}.$$

Now apply these identities turn and turn about to give

$$\frac{F(\alpha,\beta+1,\gamma+1,x)}{F(\alpha,\beta,\gamma,x)} = \cfrac{1}{1-} \cfrac{\dfrac{\alpha(\gamma-\beta)}{\gamma(\gamma+1)}x}{1-} \cfrac{\dfrac{(\beta+1)(\gamma-\alpha+1)}{(\gamma+1)(\gamma+2)}x}{1-}$$

$$\cfrac{\dfrac{(\alpha+1)(\gamma-\beta+1)}{(\gamma+2)(\gamma+3)}}{1-} \cfrac{\dfrac{(\beta+2)(\gamma-\alpha+2)}{(\gamma+3)(\gamma+4)}x}{1-} \ldots.$$

This is the formal derivation of Gauss's continued fraction, to be found in his DGCSI; it is a development of a procedure of Lambert's, who described an algorithm for manipulating two power series analogous to the Euclidean algorithm. While the hypergeometric functions converge only for $|x| < 1$, the continued fraction converges for most complex values of $x$ (specifically, when $x$ is not real with $1 \leqslant x < \infty$, or not equal to an isolated set of values), and so it provides an analytic continuation of $F(\alpha,\beta+1,\gamma+1,x)/F(\alpha,\beta,\gamma,x)$ outside the unit circle, thus possibly providing a meaning for the series in regions where it is divergent. (This convergence theory was not considered by Gauss; it was later investigated by Riemann, in SQSI, a posthumously published paper

---

[23] For details of the early and rich theory of the hypergeometric function, from Wallis to Gauss, see Dutka, EHHF. The manipulations that follow are described and explained at greater length in Wall, *ATCF*, Chapter 18.

completed by Schwarz, and by Heine; then the question was finally settled in Thomae, KGQ.)

Put $\alpha = 1$, $\beta = 0$, $\gamma = 1$, and write $-x$ for $x$. Then, since

$$F(1,1,2,-x) = 1 - \frac{x}{2} + \frac{x^3}{3} - \cdots = \frac{\log(1+x)}{x}$$

and $F(1,0,1,-x) = 1$, we get, after a little adjustment,

$$\log(1+x) = \frac{x}{1+} \frac{1^2 x}{2+} \frac{1^2 x}{3+} \frac{2^2 x}{4+} \frac{2^2 x}{5+} \cdots,$$

provided $x$ is not real and $\leqslant -1$. Or, put $\alpha = \frac{1}{2}$, $\beta = 0$, $\gamma = \frac{1}{2}$, and write $x^2$ for $x$. Then since

$$F\left(\tfrac{1}{2}, 1, \tfrac{3}{2}, x^2\right) = \frac{1}{2x} \log \frac{1+x}{1-x}$$

and $F\left(\tfrac{1}{2}, 0, \tfrac{1}{2}, x^2\right) = 1$, we get, again after adjustment,

$$\log \frac{1+x}{1-x} = \frac{2x}{1-} \frac{1^2 x^2}{3-} \frac{2^2 x^2}{5-} \frac{3^2 x^2}{7-} \cdots.$$

For another illustration, write $x/a$ for $x$,

$$F\left(a, \beta, \gamma, \frac{x}{a}\right) = 1 + \frac{\beta}{\gamma} x + \frac{\beta(\beta+1)}{\gamma(\gamma+1)} \frac{x^2}{2!} \cdot \frac{a+1}{a}$$
$$+ \frac{\beta(\beta+1)(\beta+2)}{\gamma(\gamma+1)(\gamma+2)} \frac{x^3}{3!} \cdot \frac{(a+1)(a+2)}{a^2} + \cdots,$$

and now let $a$ tend to infinity, to give the closely related function

$$\Phi(\beta, \gamma, x) = 1 + \frac{\beta(\beta+1)}{\gamma(\gamma+1)} \frac{x^2}{2!} + \cdots.$$

In particular $\Phi(1,1,x) = e^x$. So do the same to Gauss's continued fractions. We first take the expansion for $F(a, \beta+1, \gamma+1, x/a)/F(a, \beta, \gamma, x/a)$ and let $a$ tend to infinity to get

$$\frac{\Phi(\beta+1, \gamma+1, x)}{\Phi(\beta, \gamma, x)} = \frac{1}{1-} \frac{\dfrac{(\gamma-\beta)}{\gamma(\gamma+1)} x}{1+} \frac{\dfrac{(\beta+1)}{(\gamma+1)(\gamma+2)} x}{1-}$$
$$\frac{\dfrac{(\gamma-\beta+1)}{(\gamma+2)(\gamma+3)} x}{1+} \frac{\dfrac{(\beta+2)}{(\gamma+3)(\gamma+4)} x}{1-} \cdots,$$

then set $\beta = 0$, write $\gamma$ for $\gamma+1$, and adjust to give

$$\Phi(1, \gamma, x) = \frac{1}{1-} \frac{x}{\gamma+} \frac{x}{\gamma+1-} \frac{\gamma x}{\gamma+2+} \frac{2x}{\gamma+3-} \frac{(\gamma+1)x}{\gamma+4+} \frac{3x}{\gamma+5-} \cdots.$$

Now set $\gamma = 1$ to give

$$\Phi(1,1,x) = e^x = \frac{1}{1-}\ \frac{x}{1+}\ \frac{x}{2-}\ \frac{x}{3+}\ \frac{x}{2-}\ \frac{x}{5+}\ \frac{x}{2-}\ \cdots .$$

If, finally, we set $x = 1$ and adjust the expression to eliminate the negative terms and zero terms, exactly as Lagrange tells us in his AMREN §§50–67 or *TREN* §§65–76,[24] we get

$$e = [2, 1, 2, 1, 1, 4, 1, 1, 6, 1, \ldots].$$

These manipulations are offered here as illustrations of Gauss's fascination with the hypergeometric function and mastery of the idiom of continued fractions, as displayed in his memoir DGCSI of 1813.

In the memoir MNIVAI (1816) on numerical integration, Gauss used continued fractions to initiate what has developed into an important and far-reaching technique. He took the expansion of $\log((1 + x)/(1 - x))$ from his earlier memoir DGCSI and rewrote it, by replacing $x$ by $x^{-1}$, as

$$\log \frac{x+1}{x-1} = \frac{1}{x} + \frac{1}{3x^3} + \frac{1}{5x^5} + \cdots = \frac{1}{x-}\ \frac{\frac{1.1}{1.3}}{x-}\ \frac{\frac{2.2}{3.5}}{x-}\ \frac{\frac{3.3}{5.7}}{x-}\ \frac{\frac{4.4}{7.9}}{x-}\ \cdots .$$

The $k$th convergent of this continued fraction can then be expressed as a rational function, a quotient of a polynomial of degree $k - 1$ by a polynomial of degree $k$; and if this rational function is developed as a power series, its terms will agree with the original power series up to degree $2k$, the best that one can hope for in general. (This kind of approximation is now called a Padé approximant, and rapidly growing interest in this topic, together with the stimulus provided by automatic computation, has given rise to a recent revival of interest in the analytic theory of continued fractions.) Gauss then calculated the poles and residues of this convergent, which he used as the nodes and weights for what is now known as the Gaussian method of numerical integration. The procedure was later developed by Jacobi and Christoffel, and eventually led to the Runge–Kutta method of integrating differential equations; and it is connected with the general theory of orthogonal polynomials.

Two entries in Gauss's celebrated *Tagebuch* bear further witness to his interest in the manipulation of infinite series into continued fractions, and in divergent series. The sixth note, dated 23 May 1796 (and so written less than two months after the notebook was started, and just after his nineteenth birthday), recorded: The transformation of the series $1 - 2 + 8 - 64 + \ldots$ into the continued fraction

$$\frac{1}{1+}\ \frac{2}{1+}\ \frac{2}{1+}\ \frac{8}{1+}\ \frac{12}{1+}\ \frac{32}{1+}\ \frac{56}{1+}\ \frac{128}{1+}\ \cdots ,$$

$$1 - 1 + 1.3 - 1.3.7 + 1.3.7.15 - \cdots = \frac{1}{1+}\ \frac{1}{1+}\ \frac{2}{1+}\ \frac{6}{1+}\ \frac{12}{1+}\ \frac{28}{1+}\ \cdots ,$$

and others,

---

[24] These kinds of identities are described in n. 6, above.

while, on 16 February 1797, he returned to the subject:

An amplification of the penultimate proposition on page 1, namely

$$1 - a + a^3 - a^6 + a^{10} \ldots = \cfrac{1}{1+} \ \cfrac{a}{1+} \ \cfrac{a^2 - a}{1+} \ \cfrac{a^3}{1+} \ \cfrac{a^4 - a^2}{1+} \ \cfrac{a^5}{1+} \quad \text{etc.}$$

From this all series where the exponents form a series of the second order are easily transformed.

For discussion of these entries, I refer the reader to the commentary by Schlesinger in Gauss, *Werke* x$_1$, 490–3 and 513–14.

### 9.2(c)  *Continued fractions and probability theory*

Gauss initiated the probability theory of continued fractions, though he published nothing on the topic. An entry in his *Tagebuch* for 25 October 1800 records: *Problema e calculo probabilitatis circa fractiones continuas olim frustra tentatum solvimus* (We are solving problems in the calculus of probabilities about continued fractions that once we attempted in vain). A workbook for 5 February 1799 records an earlier attempt at a solution, in which he calculated and tried to explain some functions $P(n, x)$ which will be described below in his words. His calculation finishes with *Tam complicatae evadunt, ut nulla spes superesse videatur* (They come out so complicated that no hope appears to be left); see *Werke* x, 552–6 for the text and commentary. On 30 January 1812, the day he read his DGCSI to the Göttingen Scientific Society, he described the problem in a letter to Laplace:

. . . I do recall however an intriguing problem which I worked on 12 years ago, but which I did not then succeed in resolving to my satisfaction. Perhaps you might take the trouble of spending some moments with it: in that case I am sure that you will find a more complete solution. Here it is. Let $M$ be an unknown quantity, between 0 and 1, for which all values are either equally probable, or are distributed according to a given law. Suppose it converted into a continued fraction

$$M = \cfrac{1}{a' + \cfrac{1}{a'' + \text{etc.}}}$$

What is the probability, if we terminate the expansion at a finite term $a^{(n)}$, that the remaining fraction

$$\cfrac{1}{a^{(n+1)} + \cfrac{1}{a^{(n+2)} + \text{etc.}}}$$

should lie between 0 and $x$? I denote this by $P(n, x)$. When all the values of $M$ are equally probable, so that

$$P(0, x) = x,$$

then $P(1, x)$ is a transcendental function depending on the function

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{x}$$

which Euler called inexplicable and on which I have just given some researches in a memoir presented to our scientific society which will soon be published.[25] But for larger values of $n$, the exact value of $P(n, x)$ seems intractable. However I have found by very simple arguments that for infinite $n$ we have

$$P(n, x) = \log(1 + x)/\log 2,$$

but my attempts to evaluate

$$P(n, x) - \log(1 + x)/\log 2$$

when $n$ is very large but not infinite have been unfruitful [Gauss, *Werke* $x_1$, 371–4].[26]

He never published his *"raisonnements tres simples"*. The first proof of the limiting distribution $\log(1 + x)/\log 2 = \log_2(1 + x)$ was given by Kuz'min in 1928, while the final question about the asymptotic behaviour was only resolved in 1974, by Wirsing. The classical exposition is in Khinchin, *CF*; and a convenient account of recent developments can be found in Knuth, *ACP* ii, §4.5.3, where it is interesting to compare the treatments in the first edition of 1969 (which concludes "The world's most famous algorithm [sc. Euclid's algorithm] deserves a complete analysis!") and the second edition of 1981 ("Fortunately it is now possible to supply rigorous proofs, based on careful analysis by several mathematicians"). To say more here would take much space, so I shall restrict myself to a few illustrations.

First, an illustration of Gauss's result: if we take a real number $x$ at random[27]—and it must be emphasised that most of the numbers we have considered so far, such as rational numbers, quadratic surds, $e$, tanh 1, etc. do not behave like these randomly chosen numbers—and look at the terms of its continued fraction, then we should expect that:

$$\text{about } \log_2 \frac{4}{3} = 41 \cdot 50\% \text{ of the } n_k \text{ should be 1,}$$

$$\text{about } \log_2 \frac{9}{8} = 16 \cdot 99\% \text{ of the } n_k \text{ should be 2,}$$

$$\text{about } \log_2 \frac{16}{15} = \phantom{0}9 \cdot 31\% \text{ of the } n_k \text{ should be 3,}$$

$$\text{etc.}$$

[25] This is Gauss, DGCSI; Euler describes this 'inexplicable function' in, for example, his *ICD*, Chapter 16. Just as the gamma function is defined by extending $n! = 1.2.3\ldots n$ to non-integral arguments, so this inexplicable function is the extension of $1 + 2^{-1} + 3^{-1} + \cdots + n^{-1}$ to non-integral arguments.

[26] Gauss's original, in French, is also quoted in my REGM, 826–7.

[27] In modern terms, the following results will hold for all $x$ not belonging to some set of Lebesgue measure zero. On the other hand, it must be emphasised that no example is known of any number, arising in any other context, which has been proved to display any of the following kinds of typical behaviour. See Section 9.2(e), below.

So we should expect more than two-thirds of the terms of the expansions of randomly chosen numbers to be 1, 2, or 3. Next, Lévy proved in 1929 that we should expect that the convergents should satisfy

$$\sqrt[k]{(n_1, n_2 \ldots n_k)} \to c$$

where

$$c = \prod_{r=1}^{\infty} \left[ 1 + \frac{1}{r(r+2)} \right]^{\log_2 r} = 2 \cdot 685 \, 55 \ldots,$$

while the arithmetic mean should grow like $\log_2 k$. Finally, the work of Dixon, Heilbron, Knuth, Mendes France, and Wirsing, since around 1970, enables us to give some kind of answer to my slaveboy's reflection at B24 of Section 1.3, above: we should expect the number of terms in the terminating expansion of $m/n$, averaged over all $m$ less than and prime to $n$, for large $n$, to be approximately equal to $(12/\pi^2) \log_2 n + 1 \cdot 467 \, 08 \ldots$, and further closer estimates can be given for the error.[28]

### 9.2(d) *Gauss's number theory*

Gauss had a complete mastery of the techniques of continued fractions and a profound insight into their properties; the descriptions so far illustrate this clearly. And continued fractions played a very significant role in the number theory of Gauss's predecessors; we only need to look at Legendre's *Essai sur la Théorie des Nombres* of 1798 to see this illustrated at length. Yet there is only *one* explicit reference to continued fractions in Gauss's *Disquisitiones Arithmeticae* of 1801, and this found in a historical passage and contains a confusing typographical error! Three other passages introduce material closely related to continued fractions, without explicating the connections. Here are the details.

In §27, Gauss describes how to solve $ax = by \pm 1$ in integers by the Euclidean algorithm, and he defines the Euler brackets.[29] But he says nothing of continued fractions or convergents, and only identifies the process as "the known algorithm for finding the greatest common divisor of two numbers". A footnote gives two further identities for these brackets, without explanation of context or proof. Then, in §28, which I shall now quote in full with my annotations, we find the only explicit mention of continued fractions.

Euler was the first[a] to give the general solution for indeterminate equations of this type (*Comment. Petrop.* T.VII p. 46[b]). The method he used consisted in substituting other unknowns for $x$, $y$ and it is a method that is well known today. Lagrange treated the

---

[28] Simple upper bounds for the number of steps can be given, based on the fact that the longest expansions occur with the quotient of successive Fibonacci numbers $[1, 1, \ldots, 1] = F_k / F_{k-1}$.

[29] See Section 9.1(a), above: $[n_0, \ldots, n_k] = P_k(n_0, \ldots, n_k)/P_{k-1}(n_1, \ldots, n_k)$. Euler uses the notation $(n_0, \ldots, n_k)$ and Gauss uses the notation $[n_0, \ldots, n_k]$ for my $P_k(n_0, \ldots, n_k)$.

problem a little differently. As he noted, it is clear from the theory of continued fractions that if the fraction $b/a$ is converted into the continued fraction

$$\cfrac{1}{\alpha + \cfrac{1}{\beta + \cfrac{1}{\gamma + \text{etc.} \cfrac{}{+ \cfrac{1}{\mu + \cfrac{x}{n}}}}}}$$

and if the last part, $x/n,^c$ is deleted and the result reconverted into a common fraction, $x/y$, then $ax = by \pm 1$, provided $a$ is prime relative to $b$. For the rest, the same algorithm is derived from the two methods. The investigations of Lagrange appear in *Hist. de l'Ac. de Berlin Année* 1767 p. 175,[d] and with others in *Supplementis versioni gallicae Algebrae Eulerianae adiectis*[e] [Gauss–Clarke, *DA*, 28].

*Notes:*
   [a] This is corrected in Gauss's *Additamenta* as follows: "The solution of the indeterminate equation $ax = by \pm 1$ was not first accomplished by the illustrious Euler (as stated in this section) but a geometer of the seventeenth century, Bachet de Meziriac, the celebrated editor and commentator of Diophantus. It was the illustrious Lagrange who restored this honour to him (*Add. à l'Algèbre d'Euler* p. 525 [§45], where at the same time he indicates the nature of the method). Bachet published his discovery in the second edition of the book *Problèmes plaisans et délectables qui se font par les nombres* (1624). In the first edition (Lyon, 1912) which was the only one I saw, it was not included, although it was mentioned" [Gauss–Clarke, *DA*, 461].
   [b] I have given all of the citations in the form used by Gauss, and will identify them in notes; abbreviated titles can be found in the Bibliography. This is 'Solutio problematis arithmetici de inveniendo numero, qui per datos numeros divisus, relinquat data residua', *Opera* I ii, 18–31.
   [c] This final $x/n$ should, of course, be $1/n$; and this error is not corrected in the two pages of errata bound in the end of the volume, nor in the version reprinted in *Werke* i. This is the only explicit continued fraction to be found in the *Disquisitiones Arithmeticae*, though the topic is implicitly contained in his occasional use of Euler brackets, as I shall explain below.
   [d] 'Sur la solution des problèmes indéterminés du second degrée', *Oeuvres* ii, 377–535.
   [e] I.e. Lagrange's *Additions* to Euler's *Algebra*, §§42–5.

Section 5, §§153–265, of the *Disquisitiones* is devoted to solving equations of the second degree in two variables. A hint of continued fractions appears in §199 which starts "In practice, more suitable formulae can be developed". These formulae are expressed in terms of the Euler brackets, which were introduced in §27 and developed further in §189, and which Gauss identifies retrospectively in a footnote to the passage to be quoted next.

In §202, we find another historical review. Again I shall quote in full, with my annotations:

A particular case of the problem of solving the equation $t^2 - Du^2 = 1$ had already been treated in the last century. That extremely shrewd geometer Fermat proposed the problem to the English analysts, and Wallis called Brouncker the discoverer of the solution which he reported in *Alg. Cap. 98, Opp.* T. II p. 418 sqq;[a] Ozanam claims that it was Fermat;[b] and Euler, who treated of it in *Comm. Petr.* VI p. 175, *Comm. nov.* XI p. 28,[c] *Algebra P.* II p. 226, *Opusc. An.* I p. 310[d] claims that Pell was the discoverer, and for

that reason the problem is called *Pellian* by some authors. All these solutions are essentially the same as the one we would have obtained if in §198 we had used the reduced form with $a = 1$; but no one before Lagrange was able to show that the prescribed operation necessarily comes to an *end*, that is that the problem is *really soluble*.[e] Consult *Mélanges de la Soc. de Turin T.* IV p. 19;[f] and for a more elegant presentation, *Hist. de l'Ac. de Berlin*, 1767, p. 237.[g] There is also an investigation of this question in the *supplementis ad Euleri Algebram*[h] which we have frequently commended. But our method (starting from totally different principles and not being restricted to the case $m = 1$) gives many ways of arriving at a solution because in §198 we can begin from any reduced form $(a, b, -a')$ [Gauss–Clarke, *DA*, 185].

*Notes:*
   [a] This is Chapter 98, 'A method of Approaches for Numerical Questions; occasioned by a problem of Mons. Fermat' of the Latin translation of Wallis' *Treatise of Algebra . . .* (1685), which appeared in his *Opera Omnia Mathematica* ii (1693).
   [b] J. Ozanam, *Nouveaux éléments d'algèbre* (1702), described a solution of Pell's equation identical with Brouncker's, illustrated it with the cases $n = 23$ and 19, and attributed it to Fermat. Weil, *NT*, 93, observes: "Did he know more than we do, or had he merely misread the *Commercium Epistolicum*?"
   [c] Gauss inserts a footnote that reads: "In this commentary the algorithm we considered in §27 is used with similar notation. We neglected to note it at that time." The algorithm in question is the basic algorithm for computing the Euler brackets: $P_{k+1}(n_0, \ldots, n_k, n_{k+1}) = n_{k+1} P_k(n_0, \ldots, n_k) + P_{k-1}(n_0, \ldots, n_{k-1})$, $P_{-1} = 1$, $P_0(n_0) = n_0$.
   [d] These articles by Euler are, respectively, 'De solutione problematum Diophanteorum per numeros integros', *Opera* I ii, 6–17; UNAPPS; *VAA* II ii, Chapter 7; and 'Nova subsida pro resolutione formulae $axx + 1 = yy$', *Opera* I iv, 76–90.
   [e] Gauss's italics. He inserts a footnote: "What Wallis, *loc. cit.* pp. 427–8, proposes for this purpose carries no weight. The paralogism consists in the fact that on p. 428, line 4, he presupposes that, given a quantity $p$, integers $a$, $z$, can be found such that $z/a$ is less than $p$ and that the difference is less than an *assigned* number. This is true when the assigned difference is a *given quantity* but not when, as in the present case, it depends on $a$ and $z$ and this is variable." Also see n. 55, below.
   [f] Lagrange, SPA.
   [g] Lagrange, 'Sur la solution des problèmes indéterminées du second degrée', *Oeuvres* ii, 377–535.
   [h] Lagrange, *Additions* to Euler's *Algebra*, §§85–7.


   What has happened?
   The majority of the significant discoveries and developments in the theory of continued fractions in number theory, in the seventh and eighteenth centuries, have centred on the solution of Pell's equation $x^2 - ny^2 = 1$. But Pell's equation is only one example of a very general class of problems of finding for what integers $p$ the equation $ax^2 + bxy + cy^2 = p$ can be solved in integers. This more general approach focuses attention on the so-called binary quadratic forms

$$F(x, y) = ax^2 + bxy + cy^2,$$

and, as number theory developed, so more and more general examples of binary quadratic forms came to be considered: Fermat only dealt with some specific examples $x^2 + ay^2$; Euler extended this systematically to the class $ax^2 + by^2$; and Lagrange, in his *Recherches d'Arithmétique* (1775) with the

supplement of 1777, embraced the set of all binary forms. Lagrange is still inspired by his use of continued fractions in Pell's equation, which he called "la clef de tous les autres problèmes de ce genre",[30] and Legendre, in his *ETN* (Part 1, §XIII), made their role even more prominent. But Gauss's intuition seems to have been that little further significant progress was possible using this technique; and, with his eyes on the most general theory, he seems not to want to distract his reader into any byways, especially since he may have come to believe that however charming these byways may be, further progress along them will be extremely difficult if not impossible. He reformulated the proofs of his number theory so as to omit any use of or intuitions about continued fractions, and only included the briefest mention of them in those passages which sketch the history of the subject and give algorithms for finding solutions.

For tiros, I recommend the basic introduction to binary quadratic forms in Davenport, *HA*, Chapter 6; and for a combined historical and mathematical treatment, I again direct everybody to Weil, *NT*. The final appendix of this latter book ('A proof of Lagrange's on Indefinite Binary Quadratic Forms', pp. 350–9) and its back-references give a detailed description of the evolution of the continued fraction process into the new theory of binary quadratic forms that rendered it obsolete.

### 9.2(e) *Gauss's legacy in number theory*

We see number theory today through Gauss's eyes. For example, unique factorisation into products of primes is, today, *the* fundamental theorem of arithmetic (i.e. *arithmētikē*!), and it is generally referred to as such; but this theorem and attitude is not to be found explicitly before the *Disquisitiones Arithmeticae*, and it may be a historical error to look for it in Greek mathematics, behind what is preserved in the *Elements*.[31] So, as Legendre's *Essai sur la théorie des nombres* went through revisions, emerging as a two-volume treatise *Théorie des Nombres* in its third edition of 1830, its influence declined, and slowly Gauss's point of view came to dominate the subject. We see here, once again, the same evolutionary process that operated earlier with Euclid's *Elements* and Ptolemy's *Syntaxis*, in which new definitive treatments supersede, dominate, and can even annihilate their predecessors. Continued fractions moved from being a central source of insight and problems to being merely a collection of special techniques to be applied in a few restricted contexts; and, for the most part, that is how they are perceived today. Moreover, during the eighteenth and early nineteenth centuries, number theory had provided a central, coherent, and accessible nucleus to a developing theory of continued fractions; when this base was dismantled, the other topics in the theory were marginalised. But many natural problems connected with the process remain unresolved, and are still studied ardently by mathematicians.

---

[30] In the introduction to his SPA (quoted in n. 55, below), his first article on the topic.
[31] For a discussion, see Knorr, PIGNT.

Let me illustrate one such problem, what I called, in Section 4.3, the problem of the 'dimension of cubes': Formulate and prove any hypothesis about the continued fraction expansion of $\sqrt[3]{n}$. A natural first step is to compute some terms of some simple example, and Lagrange's algorithm (see Section 9.1(e) above) provides the means:[32]

$$\sqrt[3]{2} = [1, 3, 1, 5, 1, 1, 4, 1, 1, 8, 1, 14, 1, 10, 2, 1, 4, 12, 2, \ldots].$$

But, faced with such irregular behaviour, we cannot at first formulate anything other than very coarse questions like: Are these terms bounded, or can limits be given their growth? Even then, it is only very recently that any kind of answer to such basic questions has been found, and the bounds that have been found are themselves also very coarse: $n_k \leqslant a^{b^k}$ (see Davenport & Roth, RAAN, and Baker, RACTAN). These proofs depend on Roth's theorem, a deep result on Diophantine approximation, which means that, in general, the values of $a$ and $b$ cannot be effectively computed. Recently an elementary proof of such a result for some cube roots, which does not use much beyond the techniques described in Section 1, has been given in Wolfskill, GBPQCN. But let me emphasise: we still do not know any example of an algebraic irrational of degree higher than two whose terms either are bounded, or are not bounded.

Lagrange's optimistic hope about recognising algebraic numbers, quoted at the end of Section 9.1(e), suggests another approach: Can the classical continued fraction algorithm be extended so as to identify other algebraic numbers than the rationals and quadratic surds? This runs against the spirit of Gauss's work in number theory, and so I shall defer comment on it to Section 9.3(b).

Another approach to the expansion of $\sqrt[3]{2}$ is suggested by the later developments in Gauss's unpublished work on probability theory: Does this expansion conform to the expected behaviour of a typical number? I know no evidence

---

[32] In fact there is no real difficulty in extending the procedure for computing a quadratic surd to the case of a cubic surd. Here is an algorithm which covers three indices and expresses the $(k + 1)$th coefficient in terms of the $k$th and $(k - 1)$th: If $x_k = (a_k \xi^2 + b_k \xi + c_k)/d_k$ where $\xi = \sqrt[3]{n}$ and $n$ is not a cube, and $x_k = n_k + x_{k+1}^{-1}$, then, for $k \geqslant 1$,

$$a_{k+1} = n_k a_k + a_{k-1}$$

$$b_{k+1} = n_k b_k + b_{k-1}$$

$$c_{k+1} = n_k^2 d_k - 2n_k c_k - n_{k-1} d_{k-1} + c_{k-1}$$

$$d_{k+1} = -n_k^3 d_k + 3n_k^2 c_k - 3n_k(c_{k-1} - n_{k-1} d_{k-1}) + d_{k-1}.$$

When the algorithm is started with $a_0 = c_0 = 0$, $b_0 = d_0 = 1$, i.e. $x_0 = \sqrt[3]{n}$, so $a_1 = 1$, $b_1 = n_0$, $c_0 = n_0^2$, $d_0 = n - n_0^3$, then $a_k = q_{k-1}$, $b_k = p_{k-1}$, and $b_k^3 - na_k^3 = (-1)^k d_k$. Such three index algorithms are well known for quadratic surds—see, e.g., Chrystal, $TA$ ii, 454–5—and their derivation would have been immediately accessible to Euler, but I know of no published description of them for cubic or higher order irrationalities. The higher order versions are most conveniently expressed in terms of determinantal identities, as Bernard Teissier has explained to me.

that Gauss computed any terms of the expansion of $\sqrt[3]{2}$, but I would be very surprised if he, and Lagrange and Euler before him, did not do so;[33] but it would be very difficult, using only hand calculations, to get enough statistical evidence to be suggestive. However, since the development of automatic computers, the numerical exploration of continued fractions has become a minor industry. The first such investigation, a weekend's calculation on one of the first electronic computers, is reported in von Neumann & Tuckerman, CFE (1955); their one-page article is inconclusive ("We do not know whether this deviation [from the expected distribution] is significant"). These explorations were extended in Richtmyer, Devaney, & Metropolis, CFEAN. A more recent example of the genre, Lang & Trotter, CFSAN, is longer and more comprehensive (they list one thousand places of the expansions of $\sqrt[3]{2}$, $\sqrt[3]{3}$, $\sqrt[3]{4}$, $\sqrt[3]{5}$, and $\sqrt[3]{7}$, and other numbers, together with statistical and theoretical analyses), but is scarcely more conclusive in proportion to its length and electronic effort. A systematic search by Delone & Faddeev, TITD, of all cubic forms $x^3 + ax + b = 0$, where $|a|$ and $|b| \leqslant 9$, yielded nothing of interest (though some examples, like $\sqrt[3]{2}$ and $\sqrt[3]{5}$, seem to contain more large terms than one might expect). Then, in 1964, Brillhart hit on one spectacular example that had just evaded the earlier search, that the expansion of the real root $x^3 - 8x - 10 = 0$ has some extraordinarily large terms: $n_{121} = 16\,467\,250$, $n_{33} = 1\,501\,790$, $n_{161} = 325\,927$; and five other $n_k$ with $k \leqslant 139$ are greater than 20 000. Moreover the discriminant of $x^3 - 8x - 10$ is $-4 \times 163$, and $-163$ is a very significant discriminant for *quadratic* forms. A beautifully lucid explanation of why these results are so surprising and how they are connected is given in Churchhouse & Muir, CFANMI; and a more sophisticated mathematical exploration can be found in Stark, EECFFB.[34] This one extraordinary example enables us to perceive dimly some of the subtle and

---

[33] See how, when there is some systematic behaviour to be found, Euler will often describe it. For example in his FCD, §§21–2, he calculates or infers that $\sqrt{e} = [1, 1, 1, 1, 5, 1, 1, 9, 1, 1, 13, \ldots]$ and $(\sqrt[3]{e} - 1)/2 = [0, 5, 18, 30, 42, 54, \ldots]$.

[34] The discriminant of the binary quadratic form $f(x, y) = ax^2 + bxy + cy^2$ is the polynomial $D = b^2 - 4ac$. If $D = 0$, then $f$ is a perfect square or, equivalently, $f(x, 1) = 0$ has coincident roots. If $D < 0$ and $a > 0$, then $f(x, y) > 0$ for all $(x, y) \neq (0, 0)$; it is called positive definite. The theory of binary quadratic forms divides the forms into equivalence classes under a very natural transformation of coordinates by unimodular transformations and the class number is the number of equivalence classes. It has been long known that, for negative discriminants, the class number is finite and is equal to one for $D = -1, -4, -7, -8, -11, -19, -43, -67$, and $-163$; also that class number one implies that the algebraic numbers in $Q(\sqrt{D})$ have unique factorisation; but it was only in 1967 that Stark gave a generally acceptable proof that $-163$ is the largest negative class one discriminant.

The discriminant of a binary cubic form $f(x, y, z) = ax^3 + bx^2y + cxy^2 + dy^3$ is $D = 18abcd + b^2c^2 - 27a^2d^2 - 4ac^3 - 4b^3d$; hence the discriminant for Brillhart's cubic is $-4 \times 163$; but $-163$ has no special significance for the discriminants of cubic forms. Not unrelated to the eventual explanation is also the fact that $\exp(\pi\sqrt{163})$ is very close to an integer, $\exp(\pi\sqrt{163}) = 262\,537\,412\,640\,768\,743 \cdot 999\,999\,999\,999\,250. \ldots$ I hope that these tantalising titbits encourage the interested reader to consult, at least, Churchhouse & Muir, CFANMI.

complicated features that may eventually be involved if continued fraction expansions of higher degree irrationals are ever to be understood.

## 9.3 TWO RECENT DEVELOPMENTS

### 9.3(a)  *Continued fraction arithmetic*

I have already described some of the difficulties that seem to lie in the way of an algorithm for continued fraction arithmetic; see Sections 4.2 (especially the discussion of $B_{32-34}$) and 4.5(b). Lagrange refers to the problem in his UFCCI: "Since the form of continued fractions is ill-adapted to algebraic manipulations, we shall reduce these fractions to common fractions. . . ." (*Oeuvres* iv, 327–8). The first approach to any such algorithm seems to be in Hurwitz, KEZE (1981) and KTAR (1986). We find there the basic rules:[35]

$$2[n_0, 2n_1, n_2, x_3] = [2n_0, n_1, 2n, \tfrac{1}{2}x_3],$$

and

$$2[n_0, 2n_1 + 1, x_2] = [2n_0, n_1, 1, 1, \tfrac{1}{2}(x_2 - 1)],$$

to which, to make the calculation explicit, we can add

$$\tfrac{1}{2}[2n_0, x_1] = [n_0, 2x_1],$$

and

$$\tfrac{1}{2}[2n_0 + 1, x_1] = [n_0, 1, 1, \tfrac{1}{2}(x_1 - 1)].$$

With these, and also the relations described in n. 6, above, we can now double or halve any simple continued fraction. For example, take one of Euler's favourite examples:

$$\tfrac{1}{2}(e - 1) = [0, 1, 6, 10, 14, \ldots].$$

This gives

$$
\begin{aligned}
e - 1 &= 2[0, 1, 6, 10, \ldots] \\
&= [0, 0, 1, 1, \tfrac{1}{2}[5, 10, 14, \ldots]] \\
&= [1, 1, 2, 1, 1, \tfrac{1}{2}[9, 14, \ldots]] \\
&= [1, 1, 2, 1, 1, 4, 1, \ldots, 1, 2t, 1, \ldots]
\end{aligned}
$$

---

[35] I know of no earlier occurrence of these rules, though they would have been quite accessible to Euler and Lagrange, and Euler several times juxtaposed the expansions of $e$ and $\tfrac{1}{2}(e - 1)$; see, for example, Section 9.1(d), above.

and so the expansion of $e$. Hurwitz also investigated these kinds of numbers $x_0$, which have the property that there is a set of polynomials $p_1, p_2, \ldots, p_m$ such that, after some initial acyclic terms, their expansion becomes

$$x_0 = [\ldots, p_1(1), p_2(1), \ldots, p_m(1), p_1(2), p_2(2), \ldots, p_m(2), p_1(3) \ldots].$$

For example, for $e$, $m = 3$ and $p_1(t) = p_3(t) = 1$, $p_2(t) = 2t$; while for $\frac{1}{2}(e - 1)$, $m = 1$ and $p_1(t) = 4t + 2$; and for quadratic surds, the polynomials are all constant. He showed that if $x_0$ is such a number, and $y_0 = (a + bx_0)/(c + dx_0)$ where the integers $a$, $b$, $c$, and $d$ satisfy $ad - bc \neq 0$, then $y_0$ will also have the same property, though usually for a different value of $m$.[36] This, together with the fact that $e$, $e^2$, and $\sqrt[n]{e}$ are of this kind, explains in part the expansions of many exponential and hyperbolic numbers. Numbers which display this kind of regular behaviour are now known as Hurwitz numbers.

The problem of arithmetic was again raised in Whittaker, TCF (1915):

The great impediment to the use of continued-fractions in Theory of Functions and Differential Equations is the want of algorithms for adding, multiplying, and differentiating them. The object of the present paper is to supply in some measure this deficiency. I think it would be a mistake to propose the problem in the form: *Given two continued-fractions, to find a third continued-fraction which is equal to their sum or product* . . . for I doubt if the problems so formulated possess any simple solutions. But Sylvester showed long ago [in 1853] that any continued-fraction may be regarded as the quotient of two determinants [see the remark in Section 9.1(a), above, on continuants]; and if we regard continued-fractions in this light, advancing boldly from their theory to the theory of determinants, and aiming to express the products or sums of derivates of continued-fractions *in the form of determinants*, the situation becomes much more promising . . .

Then, in the body of the article, Whittaker discusses only differentiation of a certain type of function expressed as a continued fraction, and makes no attempt to consider continued fraction arithmetic.

General algorithms for evaluating sums and products are given in Hall, SPCF (1947), and Raney, CFFA (1973), but the description of Khinchin, *CF*, quoted above in Section 4.2, applies to these: they are "exceedingly complicated and unworkable in computational practice". Yet a simple algorithm exists! It was found by R. W. Gosper in the 1970s and has not yet been properly published, though enough details are given in Knuth, *APC* ii, §4.5.3, Exercise 15, with its answer, to reconstruct the procedure.[37] Here it is:

[36] If $ad - bc = 0$, then the transformation degenerates in $y_0 = $ constant. There is an older result, due to Serret, that the tails of the expansions of two numbers $x_0$ and $y_0$ are eventually equal (i.e. $x_{K+j} = y_{L+j}$ for $j = 0, 1, 2, \ldots$.) if and only if there exist integers $a$, $b$, $c$, and $d$ such that $ad - bc = \pm 1$ and $y_0 = (a + bx_0)/(c + dx_0)$; see the third edition of Serret, *CAS* i, 34–7. (This textbook contains an influential account of continued fractions which then forms the basis of later textbooks, for example Chrystal, *TA*. Serret was also the editor of Lagrange's collected works.)

[37] I would like to thank Bill Gosper warmly for sending me unpublished material on continued fractions and allowing me to describe his algorithm here. Also many thanks to Mike Paterson for explaining much, here and elsewhere, about continued fractions and algorithms.

Consider the problem of evaluating the continued fraction expansion of the real number

$$z = \frac{a + bx + cy + dxy}{a' + b'x + c'y + d'xy},$$

given the continued fraction expansions of $x$ and $y$; it will become clear that the algorithm itself can be described purely in terms of integer arithmetic and manipulations of the expansions $x_0 = [n_0, n_1, n_2, \ldots]$ and $y_0 = [m_1, m_2, m_3, \ldots]$ of $x$ and $y$.[38] By choosing appropriate initial values of $a$, $b, \ldots, d'$, we then get procedures for evaluating the sum $(0 + 1x + 1y + 0xy)/(1 + 0x + 0y + 0xy)$, the difference, the product, etc., of two continued fractions. I shall describe the algorithm in terms of inputting the terms of $x$ and $y$, and outputting the terms of $z$. For convenience, suppose that everything is positive, though there is no real difficulty in adapting the procedure to negative values.

First consider the simpler case of $z = (a + bx)/(a' + b'x)$. Inputting a term of the expansion of $x$ corresponds to supplying the information that $x = n + 1/x'$ where $1 \leqslant x'$, and so

$$z = \frac{a + b\left(n + \dfrac{1}{x'}\right)}{a' + b'\left(n + \dfrac{1}{x'}\right)} = \frac{b + (a + nb)x'}{b' + (a' + nb')x'}.$$

This step has the effect of replacing the matrix of coefficients $\begin{bmatrix} a & b \\ a' & b' \end{bmatrix}$ by $\begin{bmatrix} b & a + nb \\ b' & a' + nb' \end{bmatrix}$, which contains more information about $z$. This step can be repeated indefinitely; it generalises the procedure of calculating the convergent, when $z = x = \frac{0 + 1x}{1 + 0x}$, and it can be conveniently set out as in the following example of the calculation of $z = \frac{9 + 4\sqrt{2}}{1 + 2\sqrt{2}}$. (Recall that $\sqrt{2} = [1, 2, 2, 2, \ldots]$.)

|  | $x$ |  | 1 | 2 | 2 | $\ldots$ |
|---|---|---|---|---|---|---|
| Matrix of | 9 | 4 | 13 | 30 | 73 | |
| coefficients | 1 | 2 | 3 | 8 | 19 | |

(For example, at the last step an input of 2 has changed the coefficients $\begin{bmatrix} 13 & 30 \\ 3 & 8 \end{bmatrix}$ into $\begin{bmatrix} 30 & 73 \\ 8 & 19 \end{bmatrix}$; here $73 = 2 \times 30 + 13$ and $19 = 2 \times 8 + 3$.)

Next consider the output of $z$. We know that $0 \leqslant x \leqslant \infty$; hence $z$ will lie between $a/a'$ and $b/b'$.[39] Hence if $a/a'$ and $b/b'$ have the same integer part $p$,

---

[38] I shall drop the suffixes on $x_n$ and $x_{n+1}$ and refer to the basic step of the continued fraction algorithm as $x = n + 1/x'$.

[39] In fact $1 \leqslant x \leqslant \infty$, so $z$ will actually lie between $(a + b)/(a' + b')$ and $b/b'$, but the

TABLE 9.2

| | | | Input $x$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 2 | 2 | 2 | | ... |
| | 9 | 4 | 13 | 30 | 73 | | | |
| Output $y$ | 1 | 2 | 3 | 8 | 19 | | | |
| 3 | | | | 6 | 16 | 38 | 92 | |
| 1 | | | | 2 | 3 | 8 | 19 | |
| 4 | | | | | | 6 | 16 | |
| 1 | | | | | | 2 | 3 | |
| $\vdots$ | | | | | | | | |

we know that $z = p + 1/z'$ with $z' > 1$, and so the next term of the expansion of $z$ will be $p$. This gives

$$p + \frac{1}{z'} = \frac{a + bx}{a' + b'x}, \quad \text{so} \quad z' = \frac{a' + b'x}{(a - pa') + (b + pb')x},$$

and hence the effect of outputting the term $p$ will be to replace the matrix $\begin{bmatrix} a & b \\ a' & b' \end{bmatrix}$ by $\begin{bmatrix} a' & b' \\ a - pa' & b - pb' \end{bmatrix}$. Look again at the example of $\frac{9 + 4\sqrt{2}}{1 + 2\sqrt{2}}$; sufficient input has been made above to determine that the initial term of the output is 3. We can now incorporate the operation of extracting this output and continue the calculation further, as in Table 9.2 where the input is written horizontally and the output vertically. In this example, with the last displayed output, the matrix of coefficients reverts to an earlier value of $\begin{bmatrix} 6 & 16 \\ 2 & 3 \end{bmatrix}$, so an input of $2, 2$ will then again generate an output of $4, 1$. Hence the output is periodic, with

$$\frac{9 + 4\sqrt{2}}{1 + 2\sqrt{2}} = [3, 1, \overline{4, 1}].$$

Now consider the case of

$$z = \frac{a + bx + cy + dxy}{a' + b'x + c'y + d'xy}.$$

Here we can write the coefficients in $2 \times 2 \times 2$ array, $\begin{bmatrix} a_{a'} & b_{b'} \\ c_{c'} & d_{d'} \end{bmatrix}$ in which we think of the primed coefficients of the denominator as lying below the

weaker condition is more convenient for hand calculation. These statements are slightly more complicated if any of the $a$, $b$, $a'$, and $b'$ can take negative values.

coefficients of the numerator. (It is sometimes convenient, in hand calculations, to use two colours to distinguish alternate layers.) We then can immediately verify the following operations:

$$\text{Input } x = n + \frac{1}{x'}, \text{ and } \begin{bmatrix} a_{a'} & b_{b'} \\ c_{c'} & d_{d'} \end{bmatrix} \text{ becomes } \begin{bmatrix} b_{b'} & a + nb_{a' + nb'} \\ d_{d'} & c + nd_{c' + nd'} \end{bmatrix}.$$

$$\text{Input } y = m + \frac{1}{y'}, \text{ and } \begin{bmatrix} a_{a'} & b_{b'} \\ c_{c'} & d_{d'} \end{bmatrix}$$

$$\text{becomes } \begin{bmatrix} c & c' & d & d' \\ a + mc_{a' + mc'} & b + md_{b' + md'} \end{bmatrix}.$$

$$\text{Output } z = p + \frac{1}{z'}, \text{ and } \begin{bmatrix} a_{a'} & b_{b'} \\ c_{c'} & d_{d'} \end{bmatrix} \text{ becomes } \begin{bmatrix} a'_{a - pa'} & b'_{b - pb'} \\ c'_{c - pc'} & d'_{d - pd'} \end{bmatrix}.$$

Again, if the coefficients are all positive, the value of $z$ will lie within the smallest interval containing $a/a'$, $b/b'$, $c/c'$, and $d/d'$. When sufficient input has been made to determine the value of $p$, it can be outputted. Observe that the previous one-variable calculation is, in effect, a vertical slice through the plane $y = 0$ of the two-variable calculation.

Finally, if $x$ terminates, the most convenient way of incorporating this is, after the final input, to replace the matrix of coefficients $\begin{bmatrix} a_{a'} & b_{b'} \\ c_{c'} & d_{d'} \end{bmatrix}$ by $\begin{bmatrix} b_{b'} & b_{b'} \\ d_{d'} & d_{d'} \end{bmatrix}$; and similarly for terminating $y$.

Table 9.3 sets out the first few steps of the calculation of $\sqrt{2} \times \sqrt{3}$; alternate layers of the calculation are set in roman and italic type. The periodic inputs provoke an output of $[2, 2, 4, 2, 4, \ldots]$, but the apparent periodicity of this output will not show up, this time, in the recurrence of a matrix of coefficients, and I cannot conceive of any way of proving, within the restriction of integer arithmetic, that the output will indeed be periodic.

The algorithm can easily be extended to accept a general continued fraction

$$x_0 = n_0 + \frac{m_1}{n_1 +} \frac{m_2}{n_2 +} \cdots$$

as input, and Gosper has used it in this form to compute and analyse 204 103 terms of $\pi$, by converting a Ramanujan expansion for $\pi/4$, similar to Brouncker's expansion but converging much more quickly, into a simple continued fraction; see Gosper, TSCFP. This calculation has now been modified by him to exploit fast Fourier transform integer arithmetic and extended to 17 million terms, and on this evidence, $\pi$ appears to exhibit the

TABLE 9.3. $[1, 2, 2, 2, 2, 2, 2, 2, 2, \ldots] \times [1, 1, 2, 1, 2, 1, 2, 1, 2, 1, \ldots] = [2, 2, 4, 2, 4, \ldots]$. Alternate layers are shown in roman and italic type; outputs are encircled

Input $x$

| | | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | etc. |
|---|---|---|---|---|---|---|---|---|---|---|---|

$$0_1 \quad 0_0 \quad 0_1 \quad 0_2 \quad 0\ 5$$
$$0_0 \quad 1_0 \quad 1_0 \quad 3_0 \quad 7\ 0$$

Input $y$:

- 1 $\quad 3_2 \quad 7\ 5$
- 1 $\quad 6_2 \quad 14\ 5$
- 2 $\quad 15_{6\ 3} \quad 35_{15\ 5}$
  - ②
- 1 $\quad 21_{8\ 5} \quad 49_{20\ 9_2} \quad 48_{23\ 2} \quad 55\ 6$
  - ②
- 2 $\quad 22_{13} \quad 55_{23_9} \quad 132_{59_{14}} \quad 141\ 37$
- 1 $\quad 82_{16} \quad 196\ 43$
- 2 $\quad 223_{46_{39}} \quad 533_{123_{41}} \quad 292_{121\ 50} \quad 707\ 283_{141}$
  - ④ ②
- 1 $\quad 305_{62_{57}} \quad 729_{166_{65}} \quad 394_{187\ 20} \quad 954\ 439\ 76$
- 2 $\quad 495\ 90 \quad 1161_{293} \quad 2817_{676_{113}} \quad 6795_{1645\ 215}$
  - ④
- 1 $\quad 682_{110} \quad 1600_{369} \quad 3882_{848_{490}} \quad 9364_{2065\ 1104}$

etc.

behaviour of a typical number, in the sense of Section 9.2(c), above; its only anomaly may be that it conforms too closely to this typical behaviour!

### 9.3(b) *Higher dimensional algorithms*

We can describe the classical Euclidean algorithm as follows: given $(x_0, y_0)$— two positive integers, or two positive real numbers, or two lines, or two areas, etc.—then, at each step, if the smaller term is non-zero, subtract it from the larger; if it is zero, terminate. We have seen that this simple algorithm possesses a host of remarkable properties, some of which we can enumerate as follows:

(i) When $x_0$ and $y_0$ are integers, it will terminate with the non-zero term equal to the greatest common factor or measure of $x_0$ and $y_0$; this is *Elements* VII 1.

(ii) It will terminate if and only if $x_0/y_0$ is rational. In this case, the last non-zero term will be the greatest common measure of $x_0$ and $y_0$, and the last convergent will be equal to $x_0/y_0$, expressed as a common fraction in its lowest terms; see *Elements* X 2, 3, and 5. We can generalise this by saying that the algorithm will terminate if and only if there are integers $n_0$ and $m_0$ such that $n_0 x_0 - m_0 y_0 = 0$, a condition now expressed today by saying that $x_0$ and $y_0$ are linearly dependent over the integers; and the algorithm can be extended to evaluate the least such $(n_0, m_0)$, from which all others can be calculated.

(iii) The procedure will generate best possible approximations. There are various ways of describing this precisely. For example, in the geometrical description of Section 9.1(b), above, we can say that the algorithm will locate those lattice points $(p, q)$ that lie closest to the ray through 0 and $(x_0, y_0)$.

(iv) If $x_0$ and $y_0$ are not linearly dependent over the integers, but are quadratically dependent, so there are integers $a_0$, $b_0$, and $c_0$ such that $a_0 x_0^2 + b_0 x_0 y_0 + c_0 y_0^2 = 0$, then the algorithm will become periodic.

How far can we generalise some or all of these properties to apply to higher-dimensional situations $(x_0, y_0, \ldots, z_0)$?[40] For example, there is no difficulty in generalising (i), and Euclid already presents one such generalisation at *Elements* VII 3: given three integers $(x_0, y_0, z_0)$, apply the classical two-dimensional algorithm to $(x_0, y_0)$, which will then terminate with the non-zero term $w_0$. Then apply the algorithm to $(w_0, z_0)$; this will terminate with the highest common factor of $x_0$, $y_0$, and $z_0$. In this manner, it is clear that Euclid now knows that this algorithm will apply in any dimension $(x_0, y_0, \ldots, z_0)$, and he uses the general form in VII 33. Then he applies the algorithm to three mutually commensurable magnitudes in *Elements* X 4 and this time he notes, in a porism, that the procedure is general.

It is easy to see that, in three or more dimensions, (ii), generalised as 'termination if and only if linearly dependent', and (iii) are incompatible.

Look at the problem geometrically, as described in Section 9.1(b), above. In the two-dimensional case, if $n_0 x_0 - m_0 y_0 = 0$, then the line $Y = (x_0/y_0)X$ will pass through the lattice point $(n_0, m_0)$, and a terminating algorithm will reveal this best possible approximation. But in three or more dimensions, we can have examples like $(1, \sqrt{2}, 1 + \sqrt{2})$ which are linearly dependent but which do not lie on any ray through the origin and a lattice point.[41] Hence any algorithm that reveals this linear dependence by terminating will not go on to generate an infinite sequence of improving rational approximations by lattice points.

Before going further, let me describe two more algorithms. First, Brun's algorithm (see Brun, GK and AETQN, and, for a nice elementary exposition and proof, Ferguson, RRQN[42]): Given $(x_0, y_0, \ldots, z_0)$, at each step of the algorithm, if every term is non-zero, subtract the second largest term from the largest; if any term is zero, terminate. The algorithm will satisfy (i) for any set of integers; and it will satisfy (ii) for triples, but may fail to satisfy it for quadruples or more of real numbers. (Ponder that: if $x_0$ is the solution of a quadratic equation, and we apply the algorithm to $(1, x_0, x_0^2)$, then it will terminate; if $x_0$ is a solution of a cubic equation and we apply the algorithm to $(1, x_0, x_0^2, x_0^3)$, it may or may not terminate. For a specific example, the algorithm will not terminate if $x_0$ is the largest root of $x^3 - 3x^2 + 1 = 0$.) The algorithm will only generate best approximations in special circumstances, and I do not know when it will exhibit periodic behaviour.

Contrast Brun's algorithm with the following: given $(x_0, y_0, \ldots, z_0)$, at each step of the algorithm, if every term is non-zero, subtract the smallest non-zero term from the largest; if any term is zero, terminate. Again it is easy to see that this will satisfy (i), but, beyond that, its behaviour is completely different from Brun's algorithm. For example, applied to $(1, \sqrt{2}, 1 + \sqrt{2})$, it is easy to see that it will not terminate. I do not know its further properties.

So far I have been describing Euclidean-type algorithms based on subtraction. There are also continued-fraction-type algorithms which involve division, generalisations of the basic step $x_k = \lfloor x_k \rfloor + 1/x_{k+1}$, and which may also possibly involve renormalisation (such as replacing $(x_0, y_0, \ldots, z_0, w_0)$ by $(x_0/w_0, y_0/w_0, \ldots, z_0/w_0)$, thus seeming to deal with one less dimension). The Jacobi–Perron algorithm is an example of such a division algorithm. In its normalised form it is as follows. Given $(x_0, y_0, \ldots, z_0)$, if $x_0$ is not an integer,

---

[41] We might write $1 : \sqrt{2} : (1 + \sqrt{2}) \neq p : q : r$ for any integers $p$, $q$, and $r$, in a notation that suddenly stretches all the discussions of ratio in this book. Also note that simultaneous approximation has entered briefly into the discussion of Callippus' calendar in Section 2.4(b) and could very easily have been introduced by my Eudoxus in Section 4.4(b); and it is built into the well-tempered diatonic scale which is based on the approximation of $4 : 7 : 12$ for $\log \frac{5}{4} : \log \frac{3}{2} : \log 2$, which could have entered the discussion with my Archytas in Section 4.5(b). Euclid appears to introduce proportionality (or ratio?) between three or more terms in VII 33, without any further comment or explanation.

[42] Ferguson comments, in his RRQN: "It seems to be easier to reprove Brun's algorithm independently than to find and read the original!"

replace each term by its fractional part, shift one place to the left, insert a 1 in the final place, and divide by $x_0 - \lfloor x_0 \rfloor$. Thus

$$(x_k, y_k, \ldots, z_k) \rightarrow \left( \frac{y_k - \lfloor y_k \rfloor}{x_k - \lfloor x_k \rfloor}, \ldots, \frac{z_k - \lfloor z_k \rfloor}{x_k - \lfloor x_k \rfloor}, \frac{1}{x_k - \lfloor x_k \rfloor} \right).$$

If we perform this curious ritual on the singleton $(x_0)$, we get $x_{k+1} = (x_k - [x_k])^{-1}$, the classical continued fraction algorithm.[43] But why should such a complicated generalisation have provoked so much investigation and counterproposal, as indeed the Jacobi–Perron algorithm has?

The periodicity of the Euclidean algorithm for quadratic surds, property (iv), is a most remarkable mathematical phenomenon. It lies at the heart of my proposed reconstructions of Greek mathematics. It was clearly a property that fascinated Euler and Lagrange, as I have tried to illustrate in Section 9.1 above. That Gauss may have thoroughly understood this mathematical fact and its implications, but yet been unable to extend this kind of behaviour to any other algebraic numbers, may have been one of the reasons why he purged continued fractions from his number theory in favour of the more general and promising theory of quadratic forms, and turned instead towards the probability theory of continued fractions. Lagrange posed his problem about higher-degree algebraic numbers, quoted in Section 9.1(e) above, from firmly within the context of periodic continued fractions. Seventy years later, in 1839, Hermite posed exactly the same kind of question, again making explicit reference to the periodicity of the continued fraction expansion of quadratic surds, to Jacobi, who responded with his algorithm, an unnormalised three-dimensional version of that described above, and with examples of cases where it was and was not periodic. Jacobi's algorithm was published only posthumously in 1869 in an article, ATKA, completed by Heine from Jacobi's notes, and it was later generalised and studied intensively by Perron, who proved that it will converge, though it may not generate the best possible approximations, except for very special cases in low dimensions. And the fundamental issue of its periodicity still remains unresolved, despite much effort.[44]

---

[43] For a more sympathetic geometrical description of the Jacobi–Perron algorithm as a generalisation of the *Parmenides* algorithm, and related comments, see Pisot, SM.

[44] For references and details of the history and theory of the Jacobi–Perron algorithm, see Berstein, *JPA*. Higher dimensional algorithms seem to inspire a purple prose; Berstein refers to the attempts to generalise this periodic theory in terms of "Dantesque despair" of "abandoned hope", relieved only by the "great master" Jacobi and the nonagenarian Perron "still working feverishly on the periodicity question of the Jacobi algorithm", which, however, is "still waiting for the master-mind to decipher it completely" (see the historical sketch, pp. 1–10). One can also detect the excitement in the warm but sober letters of Minkowski to Hilbert: "I am nearly finished with the continued fractions for two real numbers; this final version of this examination will, I think, be pretty instructive. It is very similar to the Jacobi algorithm; however my algorithm has a few more frills (*Chicanen*) which allow for all sorts of questions to be asked which would be arbitrary for Jacobi's algorithm. Hermite has now scrutinised that part of my book which I have sent him. He wrote very delightedly to the translator about it: Je crois voir la terre promise, etc." (Minkowski to

Now we come to the final irony. By setting condition (iv) aside, so turning away from the fascination of the periodicity of the classical algorithm, and by concentrating on the subtractive formulation of the Euclidean algorithm, rather than division and a continued fraction algorithm, a new and successful kind of generalisation to all dimensions has been found by Ferguson and Forcade; see GEA and MEA.[45] These algorithms, though complicated, are geometrical and illuminating, and already seem to be more successful than the earlier attempts at periodic algorithms, as some of the results described in Ferguson, Bailey & Arno, APSLQ, demonstrate.

Hilbert, 20 August 1894). Alas, on 10 February 1896, he was to write: "The complete presentation of my investigations on continued fractions has reached almost a hundred printed pages but the all-satisfying conclusion is still missing: the vaguely conceived characteristic criterion for cubic irrational numbers." (Both extracts translated from Minkowski, *BDH*, 62 and 77.) Also see the next note.

[45] In an unpublished historical survey of higher dimensional algorithms Ferguson and Forcade call periodicity a "fluorescent red herring", and a "track laid by Lagrange down which many more trains were to roar impressively if fruitlessly". Of the Jacobi–Perron algorithm, they observe "there is undoubtedly a hoary counterexample of degree seven involving five digit coefficients". Minkowski's algorithm (see the previous note) does in some sense answer Lagrange and Hermite's question and also sets strict limitations on the scope of periodicity; but it is not the iterative kind of algorithm considered here. A modern account of Minkowski's algorithm and a report of extensive computer explorations using the Jacobi–Perron algorithm can be found in Ferguson, ACGLSE.

# APPENDIX:   

# NEW MATERIAL ADDED TO THE SECOND EDITION

## 10.1 A NEW INTRODUCTION: THE STORY OF THE DISCOVERY OF INCOMMENSURABILITY[1]

### 10.1(a) *The standard story*

Here is something that most mathematicians and historians of mathematics might say about the beginnings of Greek mathematics:

> The early Pythagoreans based their mathematics on commensurable magnitudes (or on rational numbers, or on common fractions *m/n*), but their discovery of the phenomenon of incommensurability (or the irrationality of $\sqrt{2}$) showed that this was inadequate. This provoked problems in the foundation of mathematics that were not resolved before the discovery of the proportion theory that we find in Book V of Euclid's *Elements*.

You must, at some time or another, have heard, or perhaps even have said, something like this; I have, but I shall try here to summarise why I now disagree with everything in this line of interpretation. This section will refer to many parts of the book where the assertions are developed in greater detail, and so will involve a certain amount of repetition. I arrange my comments under various headings.

### 10.1(b) *Some general remarks about our evidence*

Our evidence about Greek mathematics in general comes in very disparate forms, and almost all of it has been subject to an unknown amount of editing and interference. In particular, our late sources—editions, compilations, and commentaries dating from the second century AD onwards—are manifestly of very variable quality; it can often be difficult to separate what is reliable from what may not be, and these decisions may be subjective, depending on our own prior attitudes to meaning and relevance. So here, in the first phase of my reconstruction, I put almost all of this late material to one side as far as possible, and ignore it.[2] This is very drastic, and the reconstruction should go on to discuss, even to explain, this tendency to include unreliable opinions;

---

[1] This section is an adaptation of my self-contained article SDIR.

[2] At the outset, I must admit that one and only one piece of evidence of late provenance plays a crucial role, namely the material on 'side and diagonal' numbers and lines, found in Theon of Smyrna, Iamblichus, and Proclus. See the General Index, s.v. side and diagonal.

there will be an example of this later, in Section 10.1(e), on the changing role of arithmetic. Moreover, some of the relevant evidence in early sources, in particular in Euclid and Plato, comes in homogeneous slabs which often fit rather awkwardly in the various versions of the received interpretation. In some measure to redress the balance after my radical approach to the late texts, I endeavour to follow the principle that, if any one piece of such an early slab enters the reconstruction in a significant way, then the proposal should also engage with the whole of its context. For example, any important use of any aspect of the curriculum in Plato's *Republic* VII should also connect with the whole curriculum, especially since Plato insists on its unity (see Chapter 4); any significant application of a proposition in *Elements* II should eventually involve all 14 of them (see Chapter 3); anything about *Elements* XIII, the book which contains the construction of regular solids and a lot more, should say something about Book X, the classification of incommensurables, and perhaps also about Books II and IV (see Chapter 5).

Please note: I am not saying that all early evidence is all good, all late evidence is bad. I am saying that our evidence is a hotchpotch that we cannot sort out until later in the project, but the best evidence is likely to be found in the coherent but often seemingly marginal chunks of early provenance, so let us start from this material, taken all of a piece. It is a methodological principle, not a simple value judgement of the evidence.

But now to the opening assertion: my principal objection to it is that it is founded entirely on late evidence and speculation, uncorroborated to a remarkable extent by any of our earlier sources, so it forms a very insubstantial base on which to base a substantial part of our reconstructions of Greek mathematics. I shall spell this out in more detail, and then propose an alternative starting point.

### 10.1(c) *Our evidence concerning incommensurability and associated topics*

*On Pythagoras and the early Pythagoreans*, see Burkert, *LSAP*, Chapter VI. I subscribe fully to his general conclusion, that the only kind of scientific activities and discoveries we can attribute with confidence to the early Pythagoreans are some remarkable findings in music theory and acoustics, the most remarkable being that consonance is associated with some small integers. Most, if not all, of the mathematical stories have the ring of later legendising: "The beginnings of the development that is brought to a climax by Eudoxus was not far from Theodorus of Cyrene. . . . The thesis of the [early] Pythagorean foundation of Greek geometry cannot stand, any more than the legend of a great mathematics held secret" [*LSAP*, 464].[3] His explanation for

---

[3] I like Reviel Netz' formulation of Burkert's conclusions: "Pythagoras the mathematician finally died in 1962!" See Netz, *SDGM*.

the universality and tenacity of the general belief is succinct and convincing: "What is the origin of the firmly rooted conviction that Pythagoreanism was the source of Greek mathematics? This question is easy to answer: it came from the educational tradition. Everyone comes upon the name of Pythagoras for the first time in school mathematics; and this has been true from the earliest stages of the Western cultural tradition" (p. 406), and he goes on to catalogue examples from the early fifth century onwards. Plate 9 shows an earlier illustration of what Burkert is describing, though this is not referring to Pythagoras the mathematician. It is a standard kind of school exercise in grammar from the second century AD: a paradigm of Greek inflections, showing all of the different cases through variations of a simple sentence, this one about Pythagoras the philosopher.[4] Pupils who have worked through something like this will have his name branded in their minds for the rest of their lives.[5]

For these reasons, there are only isolated and very brief references to Pythagorean science in the book: see the General Index for a complete list. In particular, consider:

*The Pythagorean theory of proportion based on commensurable magnitudes.* I know of no explicit evidence for this, early or late. From what I can work out, the thinking goes something like this: we, since medieval times, have learned at school about common fractions, that is $\frac{m}{n}$ s, so that is what the Pythagoreans

---

[4] There are discussions of this tablet in Sedley, PGT and PGTDA. It may even tell us a bit about how Pythagoras was regarded in the third century AD. Compare today; how might the gap in the following sentence be filled: 'Pythagoras the *** was born in Samos and later went to Croton'? Various possibilities, arranged alphabetically, are: leader, mathematician, music theorist, mystic, philosopher, scientist, shaman, .... Thanks to several helpers, I was able to organise a simple and non-scientific survey over the Internet and, of around 190 replies, 40% said mathematician or some variant of that (geometer, mystic geometer, triangle theorist, ...), 28% said philosopher or some similar variant, 12% philosopher–mathematician, and the rest a very mixed bag of activities—number freak, bean-hater, vegetarian, polymath, new ager, .... One person said music-theorist, another father of acoustics, and these two were the only references to what may be the early Pythagoreans' most significant contribution to our scientific heritage. It is also perhaps worth noting that, from the first century up to the eighth and beyond, 'mathematician' (*mathēmatikos*) had very strong overtones of 'caster of horoscopes' and 'astronomer/astrologer'; and these last two words were interchangeable, in both Latin and Greek. And much later, in the sixteenth century, John Dee, who wrote the Preface and most probably participated in the editing of the first English translation of the *Elements*, made by Henry Billingsley and published in 1570, was an adept of these arts. I would like thank Lorenzo Smerillo for telling me about some of the detail of these very interesting developments of later antiquity.

[5] I speak from personal experience. Arthur Mee's *Children's Encyclopedia*, a standard reference for my generation, had a similar though less subtle set of variations on the first verse of Thomas Gray's *Elegy Written in a Country Churchyard*: "The curfew tolls the knell of parting day, / The lowing herd wind slowly o'er the lea, / The ploughman homeward plods his weary way, / And leaves the world to darkness and to me." I have never read the poem, but will never forget this opening verse.

must have used, especially since we find these fractions later in Greek mathematics and Greek accounting. Occasionally this opinion gets expressed in print; see, for example, van der Waerden, *SA*, p. 115:

It is probable that it was *calculation with fractions* which led to the setting up of [*Elements* Book VII, which van der Waerden attributes to the early Pythagoreans]. Fractions do not occur within the official Greek mathematics before Archimedes but, in practice, commercial calculations had to use them.

I discuss this topic at some considerable length, in Chapter 7 and elsewhere, summarise the argument in the article LFGM, and come to the following conclusions. Fractions do occur in 'official mathematics', even, in a limited way, in the *Elements*: see the use of the word *meros*, plural *merē*, translated as 'part' and 'parts', especially in Book VII. For example, VII 37 and 38 talk of 'homonymous parts', of three and the third, four and the quarter, etc., and these *merē*, 'unit fractions', are an ingredient of the way fractions are described in Greek. But I do not think there is any unambiguous evidence for '*common fractions*' $\frac{m}{n}$s, in any of our early texts including commercial calculations, and their apparent appearance in the late Byzantine copies which account for 99% or more of our evidence may instead be as scribal abbreviations. Our plentiful surviving explicit evidence is that Greek fractional practice was *exactly* the same as Egyptian practice: fractions were expressed as sums of different *merē*, as 'Egyptian fractions'. The details of this argument are long, painful, and contentious because, while we have many different kinds of evidence, most of it is of the wrong sort or it comes from the wrong place or the wrong time. And if you find the full conclusions of this overall thesis too much to swallow, we here need only a much weaker version of it, that the Greek mathematicians before Plato and Eudoxus used not common fractions, but Egyptian fractions. I return to fractions and arithmetic later in this section.

Note how, everywhere, 'Greek' as in 'Greek mathematics' simply means 'written in Greek', although almost all of our evidence has been transmitted via Egypt, and then via the whole eastern Mediterranean, and that, of course, complicates the argument; see Chapter 6. And, concerning the Greek use of fractions and division, I think that Diophantus needs a separate discussion, strongly influenced as he seems to have been by Mesopotamian mathematics, as also have the later Greek astronomers.

*The topic of incommensurability.*    In Section 8.3 I review all of the evidence on this topic for the thousand-year period up to Proclus and, there and here, I refer readers to another such review in Knorr, *EEE*, Chapter 2. Here are some conclusions of these investigations:

(i) In the first surviving explicit mention of incommensurability,[6] in Plato's *Theaetetus* (147a ff.), the topic is handled confidently as a source of interesting

---

[6] I here leave to one side the notorious 'nuptial number' at *Republic* 546c, with its talk of the 'rational and irrational diameters of five'.

mathematical research. Incidentally, as to the date of Theaetetus' death, 369 BC, which is generally regarded as one fixed point, perhaps the only secure fixed point, in the shifting sands of the incommensurability issue, there is now one specialist on pseudo-Pythagorean texts who dissents from this common view: "I find it essential to note that historians of mathematics who take it for granted that Theaitetos was still alive in the 370s must be wrong. He made some important discoveries as a young man, and Plato and his friends were deeply impressed by this. But he is likely to have died in 390 BC." (Thesleff, PC, 18, n. 47.)

(ii) There is a reference in the *Hippias Major* (whose authenticity and date have been vigorously debated, though it seems now accepted that it is an early dialogue by Plato) where, in a passage that revolves around properties of individuals and their combinations, Socrates says (303b): "Or is there nothing to prevent this, as in the case that when given things are both collectively even, they may perhaps individually be odd, or perhaps even, and again, when things are individually irrational [*arrhētos*, 'inexpressible' in the translation of Chapter 5] quantities they may perhaps both collectively be rational [*rhētos*, 'expressible'], or perhaps irrational [*alogos*] . . . ?" This latter remark is reminiscent of the kind of manipulation used extensively in *Elements* X (see Chapter 5); compare $(\sqrt{2} + 3) + (4 - \sqrt{2})$ and $\sqrt{2} + \sqrt{3}$.

(iii) The celebrated passage in Plato's *Laws* (817e ff.) where Plato talks of "ignorance . . . not worthy of human beings but pigs" may not be referring to incommensurability in our sense here, but something else, very possibly the kind of techniques used in land measurement, where again things are not what mathematicians and historians of mathematics seem to assume they ought to be when, for instance, they parade stories from commentators about Egyptian land measurement as the origin of mathematics; see Section 8.1.

(iv) Aristotle's favourite mathematical illustration is "the incommensurability of the diagonal" as something all mathematicians know; but he never suggests that it is or ever was a disaster for any mathematical theory, even though, in closely related passages, especially in the *Metaphysics*, he is highly critical of the Pythagoreans. Curiously, Aristotle never specifies that he is talking of the diagonal of a square. Twice, both times in *Prior Analytics*, at 41a 23 ff. and 50a 35 ff., he says something like: "from the assumption that the diagonal is commensurate, it follows that odd numbers are equal to evens", but gives no more details of what he means by this. I find Knorr's proposal (*EEE*, 228–32) completely convincing: that the two clumsy versions of the usual so-called Pythagorean proof, which revolve around this statement, may have been tacked on to the end of *Elements* X sometime after the time of Alexander of Aphrodisias in the third century AD, in response to the needs of Aristotelian commentators, and that Alexander himself cobbled together the further variant of this proof to be found in his own commentary. However, other natural mathematical explanations of this remark by Aristotle about odd and even numbers are possible, and our evidence for any interpretation of it is so tenuous

as to be unreliable as a basis for further reconstruction (see Section 8.3(c)). I will come back again to Aristotle later, in the next section.

(v) No surviving fragment or testimony of Eudoxus mentions incommensurability: the word-index to Eudoxus–Laserre, *FEK*, does not contain the words (a)*summetros*, (ar)*rhētos*, or *alogos*! To us, it may seem blindingly obvious that the principal achievement of the Eudoxan theory of *Elements* V must have been to accommodate ratios of incommensurable magnitudes, but no ancient source says that, and Section 4.4(b) tells this story of the antecedents to Book V Definition 5 differently; see, in particular, p. 121, at the end of a difficult passage in $B_{68}$: "Those [commensurable ratios] introduce complications, and I think incommensurable ratios may be easier to handle"![7]

(vi) Proclus, the source of most of our information about pre-Euclidean mathematics, never quotes anything from Eudemus on incommensurability, though he cites Eudemus by name several times, and also writes about incommensurability several times. The one apparent exception to this, the passage in the catalogue of geometers where Eudemus appears to refer to Pythagoras

Following upon these men [Thales, Mamercus, and Hippias of Ellis], Pythagoras transformed mathematical philosophy into a scheme of liberal education, surveying its principles from the highest downwards and investigating its theorems in an immaterial and intellectual manner. He it was who discovered the doctrine of proportionals and the structure of the cosmic figures [65.15–21]

is almost certainly an interpolation in Eudemus' text; and the remark there that Pythagoras discovered the "doctrine of proportionals" is a modern editorial emendation, $\dot{\alpha}\nu\alpha\lambda\acute{o}\gamma\omega\nu$ for $\dot{\alpha}\lambda\acute{o}\gamma\omega\nu$, of this interpolated text.[8] This is a good and clear illustration of how history can sometimes be manufactured by commentators, late and modern, who may adjust a text to fit with their own interpretations.

(vii) A proper discussion of this word *alogos*, often translated as 'incommensurable' and sometimes as 'rational' though neither word does justice to the range of its uses and meanings, will take us on a very long excursion into *Elements* Book X; see Chapter 5. Let us here just look very briefly at Euclid: as far as I am aware, the only time the topic of incommensurability appears anywhere in Euclid's works is in *Elements*, Books X and XIII, which in fact form our only coherent slab of early evidence on the topic of

---

[7] There is a shorter, simpler, and more straightforward account of the difficult arguments of this section in Fowler, EPP.

[8] Eudoxus–Laserre 65.16, translation from Proclus–Morrow, 52–3, which refers to Heath, *HGM* i, 84–5 (and also see 141) for a brief discussion of the passage and an (optimistic!) discussion of the textual emendation; there are many other translations, e.g. Thomas, *SIHGM*, 148–9. There is a thorough discussion of all aspects of the passage in Burkert, *LSAP*, 409–12, which concludes: "Thus its 'authority' is precisely reversed. . . . Nothing is left of the supposed testimony of Eudemus to the achievement of Pythagoras in the foundation of mathematics."

incommensurability, and a very massive, very coherent slab it is, in bulk and content well more than a quarter of the *Elements*, but most versions of the incommensurability story fail to do more than just mention it in passing. For proposals about the role of Book X in my story, see the beginning and end of Chapter 5.

(viii) The source of most of the stories about Pythagoras, Pythagoreanism, and incommensurability is Iamblichus, *On the Pythagorean Life*, so perhaps it is worth quoting the relevant passages in full (Iamblichus–Clark, *IPL*, 30 and 102–3):

§18 (88) . . . As for Hippasus, he was indeed a Pythagorean, but because he was the first to make public the sphere constructed from twelve pentagons he was lost at sea for his impiety: he got the reputation of having discovered it, but it all came from 'that man'— that is what they call Pythagoras: they do not use his name.

§34 (246) . . . The first man to reveal the nature of commensurability and incommensurability[9] to those unworthy to share his teachings was so much detested, they say, that not only was he excluded from their common life and meals, but they built him a tomb as if their former companion had left human life behind. (247) Some say the supernatural power took revenge on those who published Pythagoras' teachings. The man who revealed the construction of the 'twenty-angled shape' was drowned at sea like a blasphemer. (He told how to make a dodecahedron, one of the 'five solid figures', into a sphere.) Some say this fate befell the man who told about irrationality and incommensurability.

This farrago of mutually inconsistent stories, which appear for the first time in a source of doubtful reliability and relevance dating from some nine centuries after the time of Pythagoras, is the main evidential base for much of what has been written about the discovery of incommensurability! Compare O'Meara, *PR*, the most recent and authoritative study of the role of mathematics in neo-Pythagoranism, which does not even seem to mention incommensurability.

### 10.1(d) *The supposed effects of the discovery of incommensurability*

*The foundation crisis.* I find Freudenthal, CFMA, Knorr, *EEE* 306–12, and other writers convincing when they argue that, far from being a period of crisis and confusion, the early fourth century was an extraordinary period of creativity, especially in Plato's circle; we have no historical evidence for any of the postulated difficulties of a 'foundation crisis'. But I want to go further and explore the possibility that the discovery was no more than an incidental event in the early development of mathematics. So let us now look at the evidence concerning the effects of the discovery. This will involve tracking over some of the same material again.

---

[9] The translation here has "symmetry and asymmetry", but surely (in)commensurability is the appropriate translation here of (*a*)*summetros*, as in the phrase "irrationality (*alogos*) and incommensurability" at the end of the passage. The later "twenty-angled shape", *eikosagonon*, the regular polyhedron with twenty corners, is a non-standard name (perhaps found only here in Iamblichus) for *dodekahedron*, the regular polyhedron with twelve faces.

As far as I know, no Greek text, early or late, tells us clearly of the mathematical difficulties raised by incommensurability. Aristotle wrote of the innocent's surprise, and the way it then gives way to a more informed appreciation:

All men begin . . . by wondering that the matter is so (as in the case of automatic marionettes or the solstices or the incommensurability of the diagonal; for it seems wonderful to all men who have not yet perceived the explanation that there is a thing which cannot be measured even by the smallest unit). But we must end in the contrary and, according to the proverb, the better state, as is the case in these instances when men learn the cause; for there is nothing which would surprise a geometer so much as if the diagonal turned out to be measurable (*Metaphysics* 983a12–20, tr. Ross, with a slight adjustment; see p. 291).

Pappus wrote much later, in his *Commentary on Book X of Euclid's Elements*, I.2, of how thereafter:

the soul . . . wanders hither and thither on the sea of non-identity . . . immersed in the storm of the coming-to-be and the passing-away, where there is no standard of measurement

from which passage I shall grasp, below, the only mathematical straw, the last four words: "there is no standard of measurement". There is the similar Scholium 1 to Book X, quoted in part in the Introduction to Book X in Euclid–Heath, *TBEE* iii, 1–10, where there is a long discussion that is remarkable for its lack of any hard evidence.[10] And Proclus, *CFBEE*, 60, wrote:

The statement that every ratio is expressible (*rhētos*) belongs to arithmetic only, and not to geometry, for geometry contains inexpressible ratios (*arrhētos logos*),

but this is not, very much not, the terminology of *Elements* X, which is built on a completely different meaning for expressible incommensurable lines and ratios; see Chapter 5.

So, prompted by Aristotle, let us try to "perceive the explanation" and "learn the cause" of incommensurability. But first, may I persist a bit longer with a variation on the issue of the effects of the discovery of this section.

### 10.1(e)  *Objections to some proposed interpretations*

When we, today, look for reasons why the discovery of incommensurability would create serious difficulties for pre-Euclidean mathematics, a whole range of possibilities now seems to be on offer, some of which we now attribute to later commentators. With some overlap, and with my comments, there are the following:

(i) The Proclus objection:   Incommensurable ratios cannot be expressed in (or by) numbers. Surely that is nonsense! We express them in numbers, and

---

[10] Contrast this with the sparse sections in Thomas, *SIHGM* i, 110–11 and 214–17, on 'The Irrational', where this scholium and a passage from Aristotle, *Prior Analytics* 41a26–7, are the only texts excerpted.

much of mathematics is, ultimately, numbers. Early Greek mathematicians could have expressed them in numbers. We have no explicit evidence that they did this, let alone how they did it, but I am arguing for a speculative interpretation in which they were expressed in numbers in various different ways. This book is full of examples.

(ii) Incommensurable ratios could not be fitted within the pre-Eudoxan style and scope of mathematics. That is false! The definition for lines that $a:b::c:d$ means that the rectangle with sides $a$ and $d$ is equal to the rectangle with sides $b$ and $c$ uses only believed-to-be early ingredients in a believed-to-be early way, and will handle most of what we need for the *Elements*, perhaps everything except for the operation of compounding ratios, on which Euclid's own treatment is strange and unsatisfactory; see Section 4.5(c). But we have no evidence that this was an early definition.

(iii) Incommensurability showed the inadequacy of the Pythagorean doctrine that "all is number". Therefore, some add, it had to be concealed. Curiously, Aristotle, our principal early witness from whom we learn of this "all is number", never advances this criticism, even though all of this material comes together in the *Metaphysics*. We find there lots of mentions of incommensurability, we find summaries and harsh criticisms of Pythagorean philosophy, but we do not find this objection. (Zhmud, AN, has even proposed that this formula "all is number" was Aristotle's own invention.) And, I would add, incommensurability need not show the inadequacy of the doctrine that "all is number", *pace* Proclus; indeed, in my reconstruction, the behaviour of the anthyphairetic ratios of $\sqrt{n}:\sqrt{m}$ might even strongly reinforce such a doctrine, and give a satisfactory mathematical explanation, possibly the first, of the 'expressible' lines that underlie *Elements* X and XIII; see Chapters 3 and 5.

(iv) The discovery of incommensurability showed that the (Old Babylonian?) arithmetical basis of geometry was inadequate, so geometry had to be reformulated purely geometrically, for example as in *Elements* II. But we have absolutely no direct evidence of this supposed early, possibly Old Babylonian, arithmetical basis of early Greek mathematics, so this is pure speculation. Also, there are other explanations of the role of *Elements* II; see, for example, Chapter 3. I shall return to this issue below, when I discuss the role of the arithmetisation of geometry.

(v) Incommensurable ratios can only be approximated, while Greek mathematics aimed for precision. This assertion runs counter to our evidence: Archimedes is interested in approximation. Also Aristarchus and then, later, Hypsicles, Hero, Ptolemy, and so on. Moreover, the fraction 1/3 can only be approximated in the Greek system of land measurement which, by convention, only uses the parts $2', 4', 8', 16', \ldots$ (see Section 7.1(d)) and, for those who argue for the early influence of Old Babylonian mathematics (see the previous paragraph), the fraction 1/7 does not exist or (very rarely) has to he approximated in sexagesimal arithmetic. So Greek mathematics was involved

in approximation, and the necessity of approximating simple numerical ratios was a well-known phenomenon.

(vi) The discovery put into question the basic idea of ratio. I prefer to reformulate this, since no early text seems to express this concern: hints of it seem to surface first in later commentators like Iamblichus and Pappus, and then grow thereafter. So, in the next section, I shall take up the issue yet once more, this time from a different viewpoint.

### 10.1(f) *Our difficulties in defining ratio*

The previous two sections argued that only later commentators seem concerned with difficulties that incommensurability appears to pose when defining ratio. So let me rephrase the assertion just made in (vi), and ask: Why do none of the early testimonies seem concerned with our manifest problems arising from incommensurability? This emphasises that the difficulties may be *our* difficulties, not necessarily theirs; they may have had different ways of thinking about things. I finish this section by introducing this theme.

Incommensurability does present a problem to arithmetised mathematics, though ultimately I think it turns out that the problem was already lurking there even for commensurable manipulations. Arithmetised geometry is how we tend to think of geometry today: a line has a length, a number; a rectangle has an area, again a number which is equal to the product of the lengths of its sides; ratios are defined arithmetically, as quotients of numbers; and so on; see the General Index, s.v. arithmetised mathematics, for more examples and comments. So geometry becomes translated into the arithmetical manipulation of numbers—addition, subtraction, multiplication, division, taking roots, etc.—and then this arithmetic is later abstracted into algebra. For example, the so-called Pythagoras' theorem becomes $a^2 + b^2 = c^2$, where the usual interpretation of this statement involves the *lengths* (i.e. numbers) $a$, $b$, and $c$ of the sides of the triangle. It then seems that the numbers associated with things like $\sqrt{2}$ are rather complicated, so complicated that filling in all of these irrational numbers properly could not be done before the middle of the nineteenth century. Dedekind, the first, tells us that he succeeded on 24 November 1858.

But read Dedekind carefully, and you will see that there were equally serious problems with arithmetic. I go yet further than this, and argue that there need be no real problems in defining the numbers themselves, for example as decimal sequences, or sexagesimal sequences, or anthyphairetic sequences, or the astronomical sequences of Section 4.5, or other such descriptions, but a precise description of their arithmetic, even of the arithmetic of the rational numbers when they are not being conceived as common fractions, seems intractable.[11] Hence the crucial role of my proposal that early Greeks did not use common fractions, so they would not think of ratios in any way like our rational

---

[11] For discussions and examples, see my FHYDF and DT.

numbers, and so would not think that arithmetic was a natural and obvious basis on which to build their mathematics.

To pursue this theme further, early Greek geometry seems to me to be not arithmetised to a remarkable extent,[12] but then this changes as arithmetisation is gradually introduced. There are two different traditions in this later arithmetisation. The first is the astronomical one, using Old Babylonian sexagesimal numbers, which is not attested in Greece before Hypsicles and Hipparchus in the second century BC; it may have been transmitted much earlier, but that is pure speculation (see Section 7.1(a), note 7). The second is the Graeco-Egyptian unit fraction tradition which we find in our earliest testimonies like the Hibeh Papyrus (see Section 7.1(c)), then later in the Heronian Corpus, and in Ptolemy alongside the sexagesimal astronomical calculations. The arithmetic of both poses theoretical problems, and both are spectacularly absent from the geometry of the *Elements*: they seem to be irrelevant to the proportion theory of Book V and remote from the spirit of the treatment of incommensurability of Book X. No wonder commentators from antiquity onwards, who seem to work in a vaguely arithmetised geometry, have a hard time, and no wonder there seems to be some confusion.

So let us try to purge our mind of this arithmetised geometry. (I did not find this easy and it took me some years spent with some good alternatives.) One problem we then face is defining ratio or proportion, and note that this problem is not now directly concerned with incommensurability. I have already pointed out that we have no real difficulty in fitting proportionality into the pre-Eudoxan style of mathematics, but our *only* evidence on how early Greek mathematicians actually handled ratio or proportion is the celebrated passage in Aristotle's *Topics* 158b 29 ff. on antanairesis/anthyphairesis (quoted and discussed in Sections 1.2(d) and 2.1, and the basis of my first dialogue in Section 1.3), which suggests the use of the so-called Euclidean algorithm: given two numbers or two lines (or, with a bit of technique at our disposal, two more complicated geometrical objects), then count:

• how many times the second line can be subtracted from the first line;
• how many times the remainder can be subtracted from the second line;

---

[12] The only arithmetised passage I know, anywhere up to Archimedes and beyond, is in Plato's *Meno*, 82c ff., where Socrates says to the slaveboy: "Now if this side is two feet long, and this side the same, how many feet will the whole be"; and the passage continues in this arithmetised vein for the slaveboy's first two attempts. (I thank Wilbur Knorr for pointing this out to me in January 1991 in a train somewhere between Verona and Venice; I had used this passage for the introduction to the book without appreciating this feature!) The switch to geometry is then indicated by Socrates when he tells the slaveboy: "If you don't want to count it up, just show me on the diagram" (84a). And I think one can explain this singular exception by observing that the slaveboy is not a mathematician—that is the point of the episode. Note that an arithmetisation of aspects of everyday life occurs when a barter economy gives way to the use of money, which seems to have happened in Greece by the 7th or 6th century BC, though taxation and other documents are very often set out in terms of measures of wheat; see Section 7.3(c).

- how many times the next remainder can be subtracted from this remainder; etc.

and this gives a string of numbers, possibly infinite, that characterise the relationship of size between the two things, and which is here called the 'anthyphairetic ratio'.

For example, the ratio of 60 to 26 will be twice, three-times, four-times exactly. Do the process for two numbers, and you will quickly see that it must stop after a finite number of steps because, if not, "an infinite series of numbers will [arise], each of which is less than the other, which is impossible in numbers" (as formulated in *Elements* VII 31). Now do it with two lines. Early Greek geometers seem to take little notice of the philosopher's problems and their lines can be chopped up indefinitely, so the possibility of the anthyphairesis going on indefinitely presents itself. If you are a geometer, possessed of a bit of technique, and this question poses itself, you will soon find examples of it happening. Thus, as Aristotle might be saying, we could "learn the cause" and "perceive the explanation" of incommensurability in this way of thinking about the ratios of lines. But if, as Pappus might be saying, we involve geometry in "standards of measurement", which I take here as a hint of arithmetisation, then we encounter problems, as I have tried to explain.

### 10.1(g)  *Some examples of anthyphairetic geometry*

Here are some illustrations taken from Chapters 1–4. First, consider the problem of 'the diagonal and the side': draw any regular polygon, and evaluate the anthyphairetic ratio of one of its diagonals and its side. The easiest and best known example is the pentagon, so I leave that for the reader, and will consider here the square. We easily see that the side S of a square goes once, but not twice, into its diagonal D (this follows from Socrates' comments at Plato's *Meno*, 82a–85d), and so the ratio of diagonal to side is once, followed by the ratio of side to diagonal-minus-side. We are now faced with the evaluation of this ratio, and some may feel that, like Meno at 84a, we have made little progress: "It's no use Socrates, I just don't know". But, just as Socrates unblocks that impasse by conjuring a clever figure out of thin air, so I here draw Fig. 10.1. Starting from the small diagonally placed square in the left-hand corner, with side s and diagonal d, we construct a larger square whose side S is s+d, and check that its diagonal D is 2s+d.[13] (Note how the symbols, here and below, are pure shorthand for lines and contain no hint of arithmetisation.)

With one eye on Fig.10.1, and with the insight that the ratios are unaffected by the scale or orientation of the figures involved (but see Section 4.5(e)), we take up our problem again, and see that the ratio of big side to big diagonal-minus-side is the same as the ratio of little side-plus-diagonal to little side; and

[13] This is my proposed interpretation of the figure being described in the texts on side and diameter lines; see note 2, above.
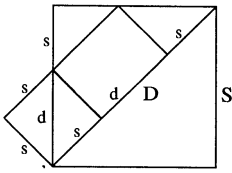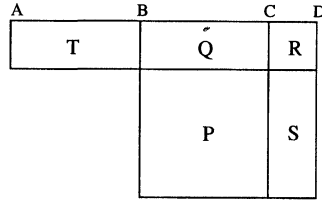
FIG. 10.1                    FIG. 10.2

we can now evaluate this as twice, followed by the ratio of little side to little diagonal-minus-side which is—scale and orientation aside—what we just started from. Hence the ratio of side to diagonal-minus-side is twice, twice, twice, twice, continuing thus indefinitely, and so the ratio of diagonal to side of a square is once, twice, twice, twice, . . . .

Let me now evaluate this same ratio another way; or, to be more exact, let me give another proof that the ratio of diagonal-plus-side to side is twice, twice, twice, twice, . . . Start with a square P and, by adding on a gnomon Q + R + S = P, as in Fig. 10.2, construct a larger square of size 2P, whose side will therefore be the diagonal of the original square; and then append a rectangle T equal to Q, as shown. Then the ratio of diagonal-plus-side to side will be the ratio of AD to BC, which we immediately see is twice, followed by the ratio of BC to CD. The required proof will now follow if we can show that the ratio of BC to CD is equal to the ratio of AD to BC; or, equivalently, by a standard result, the theory of geometrical proportions (see *Elements* VI 16; this manipulation was described earlier, in Section 10.1(e)), that the rectangle with sides AD and CD is equal to the square with sides BC, that is T + Q + R = P; but this underlies our construction, since T = Q = S.                    QED

The case of the ratio of the longer diagonal to the side of a hexagon gives rise to the ratio $\sqrt{3}$ to 1, where $\sqrt{3}$ denotes the side of the 3-fold square (which can be constructed using *Elements* II 14); and this example can be generalised to the investigation of the ratios like $\sqrt{n}$ to $\sqrt{m}$—the problem of 'the dimensions of squares' of Chapter 3. We first use numerical techniques very close to those found in *Elements* VII to explore and conjecture what the answer might be. The second kind of proof above then deals with the simpler cases; next, a comparison of the mechanisms of Fig. 10.2 and *Elements* II 11 gives us insight into the particular example of the ratios $n$-times, $n$-times, $n$-times, . . . ; and, finally, yet more heuristic exploration, followed by a generalisation of Fig. 10.1, gives a complete solution of this remarkable problem and a new interpretation of the whole of *Elements* II.

An analogous problem of 'the dimension of cubes' beckons, but proves to be of such redoubtable difficulty that in fact it remains unsolved today and Plato's remarks at *Republic* 528b–c are still perfectly applicable; see Section 4.3. We can also explore related problems such as 'the circumdiameter and side' of

polygons or polyhedra and see where they lead: they can provide explanations of the roles of *Elements* IV, X, and XIII; see Chapter 5. And the similar problem of 'the perimeter and the diameter' might be what lay behind Archimedes' original calculation in his *Measurement of a Circle*; see Section 2.2(b).

Developing these ideas takes us into a different world. The ingredients are all early Greek, but they are fitted together to create a completely different picture. It is mathematically appealing, amazingly coherent (too coherent, in fact, for my comfort), and consonant with a quite extraordinary breadth of our evidence. And the topic with which I started, the simple discovery of incommensurability, plays no significant part in it, which is why this book contains no discussion of it beyond the bald and sceptical catalogue of our evidence in Section 8.3.

As far as this topic of the discovery of incommensurability is concerned, the most significant change in the text of the second edition of this book is a little addition to the first of my dialogues, where my slaveboy describes this anthyphairetic definition of ratio under Socrates' prompting. They start doing it on numbers (heaps of stone, in fact) and the slaveboy realises that the anthyphairetic process must terminate. Then Socrates introduces the possibility of doing it with lines. At this point, at the end of $B_{36}$ in Section 1.3, a new short sentence has been added: "I wonder if it can then go on forever." That is my slaveboy realising one of the causes and explanations of incommensurability, and it is just a passing remark on the way to discovering much more remarkable things alluded to in $S_{37}$–$S_{43}$, alongside which this simple fact of incommensurability fades into insignificance.

## 10.2 RATIO AS THE EQUIVALENCE CLASS OF PROPORTIONALITY

By basing the discussion of ratio in Section 1.2(d) on *Elements* V, Definition 3 and Aristotle, *Topics* 158 f., I managed to sidestep the question: Does Euclid not here appeal to equivalence relations and classes? The reason for raising this issue here is that Book V, which deals with proportion theory, contains the remarkable Definitions 5 and 6:

[Four] magnitudes are said to be in the same ratio [or proportional], the first to the second and the third to the fourth when, if . . .

where the relation of proportionality that this goes on to define is clearly, at least to us, symmetric and reflexive. Then Proposition 11 proves that

Ratios which are the same with the same ratio are also the same with each other

which shows that it is also transitive. So is this not an explicit verification that proportionality is an equivalence relation, which then leads naturally to the definition of ratio as an equivalence class? Many mathematicians think so, but let us look at some other similar manipulations in the *Elements*.

First, there is the treatment of parallels in Book I. Proposition 30 reads:

Straight lines parallel to the same line are parallel to one another,

which expresses the transitivity of the relation of being parallel; but here we find no appeal to its equivalence class, nor any name for it such as 'direction' or 'point at infinity'. Indeed its very definition, I Definition 20:

Parallel straight lines are straight lines which, being in the same plane and being produced indefinitely in both directions, do not meet one another

shows that the relation is not even reflexive.

Many propositions with a similar structure to this occur in Book X, the massive classification of some kinds of incommensurable lines described here in Chapter 5. We have, for example, a straightforward instance of transitivity in X 12:

Magnitudes commensurable with the same magnitude are commensurable with one another,

and the relation is clearly reflexive and symmetric; so will we not find here an equally natural appeal to the equivalence class containing a given magnitude, corresponding to our important idea of the set of rational numbers with the size of that magnitude as unit? Here the answer is again a firm though complicated 'No'. Book X is based on the choice of an 'assigned line' in X Definition 3:

. . . Let then the assigned straight line be called *rational* (*rhētos*, better translated 'expressible'), and those lines that are commensurable with it, whether in length and square or in square only, *rational*, but those which are incommensurable with it *irrational* (*alogos*, 'without words', 'unreasonable').

Hence the side *and* the diagonal of a square with rational side are rational, and the underlying equivalence class in Book X, if we insist on imposing this point of view, is not the one that seems natural and familiar to us. For example, as we shall see in Chapter 5, the sum and the difference of two incommensurable rational lines is not rational, and this remark is the opening step of the classification of Book X.

Finally, Euclid always considers individuals, never sets. We see this most clearly again in Book X, which sets up a complicated set of relationships between lines—medial, binomial, major, apotome, minor, etc.—and the typical format, in X 23–70 and 66–103, reads:

A straight line commensurable with a medial/major/minor/ . . . straight line is a medial/major/minor/ . . . ,

while the propositions that show that all of these classes are disjoint are also phrased in terms of individuals, for example X 72/73, 111, and 111/112:

The binomial straight line and the irrational straight lines after it are neither the same with the medial nor with one another.

The apotome is not the same with the binomial straight line.
The apotome and the irrational straight lines following it are neither the same with the medial straight line nor with one another.

In conclusion, nowhere do we find any appeal to sets of objects, so nowhere could we find equivalence classes. This is an obstacle to the development of a general idea of the equivalence class that will hold for more than two thousand years.

Some early examples of equivalence-class-style arguments do occur in contexts where some specific well-defined concrete set underlies the subject—Gauss's number theory in his *Disquisitiones Arithmeticae* is an outstanding example—but the general technique of appealing to equivalence classes appeared only around the time of Dedekind, when set theory began to be introduced as a basis for mathematics in general, and it took some time to become established. This gives a good example of a piece of mathematics that has been popularised only recently, but which has already been retrospectively written back into the 'history' of the subject, even back to Euclid, and all this has happened almost within my own lifetime, in a process that is every bit as efficient and thorough as the work of the thought police in George Orwell's *Nineteen Eighty-Four*.

## 10.3 FURTHER REFLECTIONS ON THE METHOD OF GNOMONS, THE PROBLEM OF THE DIMENSION OF SQUARES, AND THEODORUS' LESSON IN *THEAETETUS*, 147c–158b

### 10.3(a) *Introduction*

Both of the following two sections fit best after Section 3.4, though they draw on attitudes and material developed elsewhere in the book. Both were in my mind as I was writing, but I did not include them for a variety of reasons: to be appreciated, they need some familiarity with and tolerance of my main argument, something that was not evident at the time; they would distract the main thrust of the argument by engaging too soon with some of the great and contentious issues of the subject; I did hope then that someone else would see and develop them, as there is no better way of generating acceptance of an idea than by letting others contribute to its construction in some significant way (though I regret that only Benno Artmann has briefly taken up this silent invitation, then to put it on one side in favour of his own attractive proposal[14]); my book already seemed too long; and the proposal of the first section, below, is speculative to a much greater extent than anything else in this book while, as to the second, the reader will see my feelings about it towards the end of the section.

[14] See note 20 below.

## 10.3(b) *From heuristic to deduction via algorithms*

Greek mathematics up to the time of Archimedes and Apollonius is unique among other ancient mathematical traditions in many respects: it was deductive, geometrical, non-arithmetised, and associated with an aristocracy rather than a scribal tradition, with all of these features to an extreme degree. Contrast this with neighbouring developments: setting aside all questions of chronology, Old Babylonian mathematics was algorithmic, profoundly arithmetised with only a modest and metrical involvement in geometry,[15] and associated with a governmental scribal class, while Egyptian mathematics was also arithmetised, though with a different way of representing numbers and doing arithmetic; it had a rather more geometrical involvement, though again metrical and non-deductive; and its scribes seemed to have some religious affiliations and a higher status in society.[16]

But very many details of the Greek tradition are lost. If they did work on some collection of problems, as Plato and Aristotle repeatedly tell us they did or ought to do,[17] then all that seem to survive of these are the three great problems of circle squaring, cube duplication, and angle trisection, and it is difficult to see how a tyro might cut his teeth on these, or even what a well-established practitioner might do with them on the Greek equivalent of a Monday morning. If they did have heuristic methods, then these have almost completely disappeared; we have the solitary exception of Archimedes' *Method*, and this text is so sophisticated that I know of no single result outside those described by Archimedes where it has been proposed as a preliminary exploratory technique. Moreover, Greek mathematics developed a deductive procedure, but our sources do not tell us how this came about. There are several strategies that we can take towards these issues. Consider deduction: we can proceed as if this was born *ex ovo*, or we can leave the topic aside as still being beyond our evidence and understanding (more or less the position I took when I wrote the book), or we can speculate about the origin of this unique feature. The current dominant speculation seems to be that Greek mathematics grew out of Old Babylonian mathematics, under the stimulus of problems raised by the Greek discovery of the phenomenon of incommensurability. But we have

---

[15] An extreme example of this dissociation from geometry can be found in later Babylonian astronomy: it is almost unbelievable to us that detailed computational astronomy can be pursued with not the slightest apparent involvement of some geometrical model.

[16] There is also the issue of our evidence: for Old Babylonian mathematics we have hundreds, if not thousands, of clay tablets, primary source material; almost all Greek mathematics is known through secondary sources, most of them closer in time to us than to their archetypes (see Chapter 6); and, for Egyptian mathematics, we have only a small handful of texts. Some of these textual issues are discussed in Chapters 6 and 7.

[17] For just one set of examples, see Plato's curriculum in *Republic* VII, quoted extensively in Chapter 4: "It is by means of problems, then, as in the study of geometry, that we will pursue astronomy too" (Section 4.4(a)), or "But [music theorists] do not ascend to generalised problems and the consideration which numbers are inherently concordant and which not and why in each case" (Section 4.5(a)).

absolutely no evidence whatsoever for this mathematical contact between the two cultures (if anything, our Greek sources emphasise contacts with Egypt, and this is corroborated to some extent by their joint way of handling fractions), and I believe that the story of the discovery of incommensurability owes more than is commonly realised to later invention and speculation than to reliable evidence, as I have just argued in Section 10.1. We can further add the Greek disdain for other languages and cultures, and the disdain of the early Greek aristocracy for trade and writing. Indeed, the hypothesis of the Old Babylonian roots of Greek mathematics seems to me to be something that flies in the face of much that we do know of early Greek society and mathematics.

I shall briefly propose here an alternative speculation concerning the origin of deduction and the development of other features of early Greek mathematics. It starts with a problem that already has a very characteristic Greek flavour; it pursues its exploration using an algorithmic procedure that leads to the appearance of a compelling pattern, followed by the realisation of the need for a proof (this distinction between true judgement and proof being the topic of Plato's *Theatetus*); it then moves on to the discovery of one such method of proof and some of its many consequences; then on to its limitations, and the discovery of extensions of all of this: better techniques for resolving this problem, and further problems, some eventually soluble, others not; and then, finally, the disappearance of the original motivating problem. Many of the details of this were already described and discussed in the first edition of this book; I now put the pieces together.

The problem is, of course, the dimension of squares; the algorithm is the procedure for evaluating anthyphaireses; the compelling pattern is the appearance of an apparent periodicity; the next stage is the realisation that, no matter how far this algorithmic procedure is continued, it will not constitute a proof; the next unrecorded and often unremarked episode may have been the slow and often painful realisation, typical everywhere of deductive mathematics, that the methods used so far seem insufficient to construct such a proof, so a complete change of technique, viewpoint, and language is needed; then follows the turn towards geometry which resolves a part of the investigation. Finally, geometry becomes a theory in its own right, and only vestiges remain of the original problems, which may be forgotten since they have been solved and then assimilated into the mainstream, or forgotten because they may seem unproductively insoluble, even with the newly developed techniques, or merely just forgotten. I have set my illustration of this process in the context of the problem of squares, but several others places within my reconstruction might equally plausibly follow the same kind of path, and there is no reason why these should not develop in parallel.

Of these consequences of this evolution, the most general and far reaching, apart from the introduction of proof into mathematics, possibly even for the first time, could be the need for some developed kind of geometry, and an understanding of just *why* geometry is 'true' and 'reliable', for it lies at the basis

of the proof. This deductive geometry may at first appear as a language needed for the solution of this and other such kinds of problems, but its investigation would be a topic of interest in its own right; then it could develop and becomes a mathematical topic of independent interest, and a new great mathematical theory would then have been born.[18]

I do not want to develop further what I have just outlined, as it is speculative to a much greater extent than the other aspects of the reconstruction. But I would like readers to bear it in mind as the story develops.

### 10.3(c)  *A compendium of examples*

Let us draw up a systematic list of the results we have now established for the expansions of $\sqrt{n}:1$, starting from $n = 2$:

$\sqrt{2}:1 = [1,\overline{2}]$: see Fig. 3.15 and also the cases of $\sqrt{5}:1$ (hereafter 5, etc.), 10, and 17.

$\sqrt{3}:1 = [1,\overline{1,2}]$: see Fig. 3.16, and 6, 8, 11, 12, 15, and 20, especially 15.

$\sqrt{4}:1 = [2]$, a commensurable ratio.

$\sqrt{5}:1 = [2,\overline{4}]$: see 2.

$\sqrt{6}:1 = [2,\overline{2,4}]$: see 3 and 15.

$\sqrt{7}:1 = [2,\overline{1,1,1,4}]$: see below.

$\sqrt{8}:1 = [2,\overline{1,4}]$: see 3 and 15.

$\sqrt{9}:1 = [3]$, a commensurable ratio.

$\sqrt{10}:1 = [3,\overline{6}]$: see 2.

$\sqrt{11}:1 = [3,\overline{3,6}]$: see 3 and 15.

$\sqrt{12}:1 = [3,\overline{2,6}]$: see 3 and 15.

$\sqrt{13}:1 = [3,\overline{1,1,1,1,6}]$: see below.

$\sqrt{14}:1 = [3,\overline{1,2,1,6}]$: see below.

$\sqrt{15}:1 = [3,\overline{1,6}]$: see 3 and Fig. 3.17.

$\sqrt{16}:1 = [4]$, a commensurable ratio.

$\sqrt{17}:1 = [4,\overline{8}]$: see 2.

$\sqrt{18}:1 = [4,\overline{4,8}]$: see 3.

$\sqrt{19}:1 = [4,\overline{2,1,3,1,2,8}]$: see below.

$\sqrt{20}:1 = [4,\overline{2,8}]$: see 3 and 15.

Nature has been very kind and helpful to us most of the time in this sequence of examples. There are the exceptional cases of the squares 4, 9, and 16, which

---

[18] This process of the shift from a specialised language to the establishment of a theory is now commonplace in mathematics. I here cite only one example; there will be more discussion of this in Section 10.5. The eighteenth- and nineteenth-century studies of vibrating strings and heat conduction led to new techniques which ultimately developed independently as our $\epsilon$–$\delta$ analysis; then problems within this analysis, about the behaviour of trigonometric series, led to transfinite arithmetic and a new branch of logic; further problems led to the beginnings of point set topology—and the list can be continued much further. How many of those who teach these theories, let alone of those who learn them, realise that, ultimately, they arise from the investigation of a vibrating string?

we ignore hereafter; then we start with the simplest cases $(\sqrt{2}+1):1 = [\overline{2}]$ and $(\sqrt{3}+1):1 = [\overline{2,1}]$ whose periods are short and contain the smallest possible numbers, and practise on these before meeting more complicated examples; for example $[1,\overline{1,2}]$ before $[2,\overline{2,4}]$, $[3,\overline{3,6}]$, etc. This allows some automatism to develop, and enables us to see generalities emerging; for example, it quickly becomes evident that $\sqrt{(n^2+1)}:1 = [n,\overline{2n}]$ and $\sqrt{(n^2+2)}:1 = [1,\overline{n,2n}]$. In this way, we build up our confidence before we tackle the more complicated cases of $\sqrt{7}:1$, $\sqrt{13}:1$, $\sqrt{14}:1$, and $\sqrt{19}:1$. By a fortunate accident of nature, these also present themselves roughly in order of increasing difficulty, though there is not here the nursery case of an example with three terms in its period, so a cautious reader should first try $\sqrt{5}:\sqrt{2} = [1,\overline{1,1,4}]$; see Fig. 3.20. Apart from these four cases, every expansion above has a figure associated with it, or is referred to a case or cases with a figure. For completeness, let us now deal with the figures and associated arguments for these final cases.

PROPOSITION    $\sqrt{7}:1 = [2,\overline{1,1,1,4}]$, i.e. $(\sqrt{7}+2) = [\overline{4,1,1,1}]$.

PROOF    See Fig. 10.3: start with the square P on the assigned line $AB = 1$, the four-fold square $P_4$ on $AC = 2$, and add a gnomon $Q + R + Q = 3P$ to give the square $P_7$ on the line $AD = \sqrt{7}$. Add $S + T + U + T + S = 2P$ to give the square $P_9$ on the three-fold line $AE = 3$, so $AB = BC = CE = 1$. Adjoin $FA = 2$. At this stage, ignore the dashed and dotted lines in the figure. Check
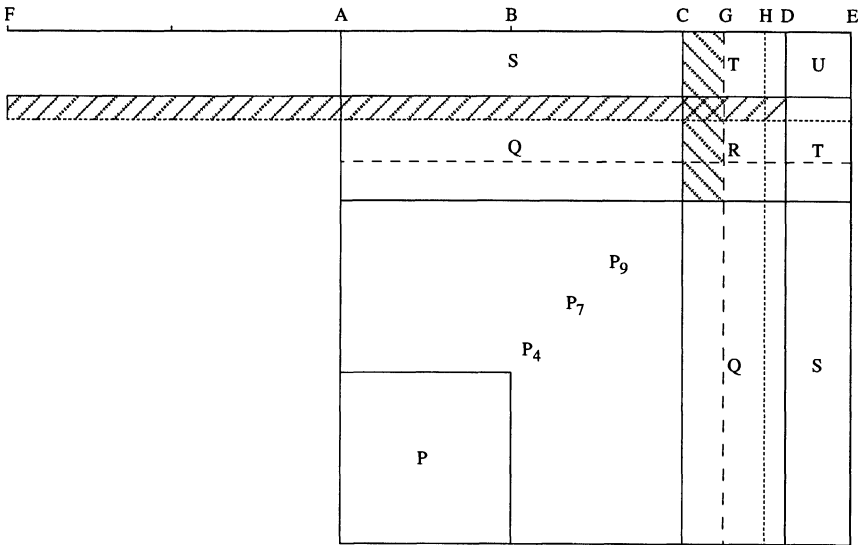


FIG. 10.3. $(\sqrt{7}+2):1 = [4,\overline{1,1,1}]$

that CD < CE < 2CD (if the gnomon Q + R + Q had width $\frac{1}{2}$AB, it would be too small; if width AB, too big). Start the anthyphairesis:

$$(\sqrt{7} + 2) : 1 = FD : AB = [4, AB : CD], \text{ since } CD < AB = CE,$$

$$AB : CD = CE : CD = [1, CD : DE], \text{ since } DE < CD.$$

Now let DG = DE, i.e. 'fold back' the last gnomon, generating the dashed lines in the diagram. We check CD < 2DE, i.e. CD < 2(AB − CD), i.e. 3CD < 2AB (since if Q + R + Q had width $\frac{2}{3}$AB it would be too big). So do the next subtraction:

$$CD : DE = [1, DE : CG].$$

Now let GH = GC, i.e. 'fold forward' the last remainder, generating the dotted lines in the diagram. Once again, check that GD = DE < 2GH, i.e. (AB − CD) < 2(CD − DE) = 2(CD − (AB − CD)), i.e. 3AB < 5CD, since if Q + R + Q had width $\frac{3}{5}$AB, it would again be too small. So the final subtraction step is:

$$DE : CG = DG : CG = [1, CG : HD].$$

Putting all these subtraction steps together, we get:

$$(\sqrt{7} + 2) : 1 = FD : AB = [4, 1, 1, 1, CG : HD],$$

so we finally prove that

$$CG : HD = FD : AB, \text{ i.e. rectangle } (CG, AB) = \text{rectangle } (FD, HD).$$

By a careful consideration of the figure, we see that this is the same as:[19]

$$(T + R) - (U + T) = (2S + T) - [(2Q + R) - (2S + T)],$$

i.e.

$$2Q + 2R = 4S + 2T + U,$$

so, adding 2T + U to each side,

$$(Q + R + Q) + (2T + U + R) = 2(S + T + U + T + S),$$

i.e.

$$3P + P = 4P. \hspace{3cm} \text{QED}$$

PROPOSITION $\sqrt{14} : 1 = [3, \overline{1, 2, 1, 6}]$, i.e. $(\sqrt{14} + 3) : 1 = [\overline{6, 1, 2, 1}]$.

This is similar to $\sqrt{7} : 1$ except that, at the penultimate stage where we checked that CD < 2DE, we now have to check that 2DE < CD < 3DE. This shows that, at this step of the anthyphairesis, the first 'folded-back' gnomon has to be subtracted twice. The reader is strongly encouraged to supply the details, to see just how terms larger than one in the anthyphairesis introduce complications into the argument, but that this case is still perfectly feasible.

---

[19] Note that these last steps are cutting up and rearranging rectangular areas, not doing algebra in our modern sense.
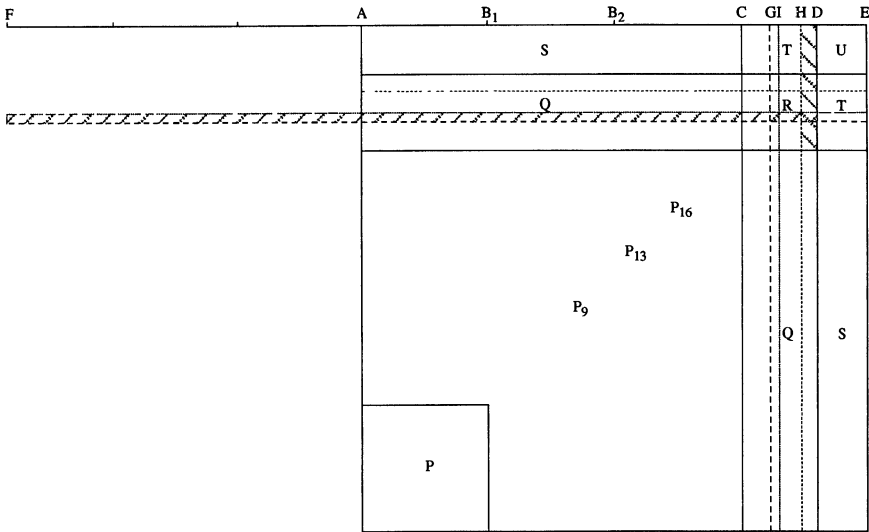
FIG. 10.4.   $(\sqrt{13}+3):1 = [6, \overline{1,1,1,1}]$

**PROPOSITION**   $\sqrt{13}:1 = [3, \overline{1,1,1,1,6}]$, i.e. $(\sqrt{13}+3):1 = [\overline{6,1,1,1,1}]$.

**PROOF**   See Fig. 10.4.

The basic construction: to the nine-fold square $P_9$, add a gnomon $Q + R + Q = 4P$, then a further gnomon $S + T + U + T + S = 3P$. Label the points as shown; then $AB_1 = B_1B_2 = B_2C = CE = 1$, and $AD = \sqrt{13}$. Adjoin $AF = 3$.

Preliminary checks, with the now routine steps elided: $CD < AB_1 = CD < 2CD$ (since . . . , etc.); $DE < CD < 2DE$, i.e. $(AB_1 - CD) < CD < 2(AB_1 - CD)$, i.e. etc. Fold $DE$ back to $DG$, and check that $DE = DG < CD < 2DG$, i.e. $(CD - DE) < CD < 2(CD - DE)$, etc. Fold $CG$ forward to $GH$ and check that $CG = GH < GD < 2GH$, i.e. etc. Finally—this is the new step—fold $HD$ back to $HI$ and check that $HD = HI < HG < 2HI$, i.e. $(CD - CH) < CG < 2(CD - CH)$, i.e. $(CD - 2(CD - DE)) < (CD - DE) < 2(\ldots)$, i.e. $3DE < 2CD$ and $3CD < 5DE$, i.e. $3(AB_1 - CD) < 2CD$ and $3CD < 5(AB_1 - CD)$, i.e. $3AB_1 < 5CD$ and $8CD < 5AB_1$, etc. Do not forget that $DG = DE$, $GH = GC$, and $HI = HD$!

Anthyphairesis:   $(\sqrt{13}+3):1 = FD:AB_1 = [6, AB_1:CD] = [6,1,CD:DE] = [6,1,1,DE(=GD):CG] = [6,1,1,1,CG(=GH):HD] = [6,1,1,1,1,HD:GI]$.

Periodicity: Now verify that $HD:GI = FD:AB_1$, i.e. rectangle $(HD, AB_1) = (GI, FD)$, i.e. (study the figure carefully!),

$$(U + T) - [(T + R) - (U + T)]$$
$$= (2S + T)] - 2[(2S + T) - \{(2Q + R) - (2S + T)\}];$$

i.e., rearranging,

$$6S + 4T + 2U = 4Q + 3R.$$

But, adding $2T + U$ to both sides, we confirm this, since

$$3(2S + 2T + U) = 2(2Q + R) + (2T + U + R), \text{ i.e. } 3 \times 3P = 2 \times 4P + P.$$

<div align="right">QED</div>

I think that anyone who works carefully through the details of this proof will confirm that it is intelligible. But, at this point I stop drawing diagrams and giving proofs, since nature's kindness is at an end: the next untreated case is $\sqrt{19}:1 = [4, \overline{2,1,3,1,2,8}]$, which presents the dual complications of a longer period containing several terms greater than 1. Though it is still just about possible to verify the much simpler examples $\sqrt{21}:\sqrt{8} = [1, \overline{1,1,1,1,1,2}]$ and $\sqrt{55}:\sqrt{8} = [2, \overline{1,1,1,1,1,4}]$ (though I do not know if such examples could be constructed without the aid of algebra), I do not think that I can draw the figure and prove this final remaining example of $\sqrt{19}:1$ by this method. Once again, we are at an impasse: it is not sufficient to assert airily that this method can, in principle, go on to deal with these cases. It is necessary to change tack and search for a new approach that will resolve this and other similar problems, and such a new method is described in Section 3.6; it involves completely new material and is rather sophisticated, but I think that it still lies well within the capacities of Greek mathematics. But let us set this new development to one side for the moment, and return to our original approach.

## 10.3(d) *The geometry lesson*

One of the most extensively debated of pre-Euclidean questions has been the interpretation of Theaetetus' description of a geometry lesson that he and his friend attended:

> Theodorus was proving for us via diagrams something about powers, in particular about the three-foot-power and the five-foot-power—demonstrating that these are not commensurable in length with the one foot power—and selecting each power individually in this way up to the seventeen-foot-power; but in this one for some reason he encountered difficulty [or: he stopped] [*Theaetetus* 147d–148b].

Whole books have been inspired by or written around this passage: the latest and greatest is Knorr's *EEE* (from which this translation has been taken), the book which, more than anything else, was the stimulus for my own venture into Greek mathematics. Knorr reviewed all significant prior explanations of the mathematics underlying this passage,[20] and its relevance to our understanding

---

[20] Subsequent ones are in Artmann, PTTDD, Caveing, DZV, and Kahane, TTCQR. (Artmann's attractive method is based, literally, on drawing diagrams and then producing anthyphairetic proofs based on information they suggest; my objection to it is that, like most other anthyphairetic proofs, it requires additional features to deal with the case of 13; I am almost, but not completely, persuaded that further such developments would not deal with

of early Greek mathematics, and he gave his own interpretation, a historically plausible procedure that grinds to a halt with the case 17. The final phrase in this passage is, however, still disputed: ἐν δὲ ταύτῃ πως ἐνεςχέτο, "in this one for some reason he encountered difficulty" in Knorr's very precise rendering, but translated by others with a rather more vague sense. Burnyeat sets out a range of possibilities:

The sentence could mean three things: (a) "at that point for some reason [*sc.* for no particular reason that Theaetetus knows of] he stopped", (b) "at that point for some reason [*sc.* for some particular reason] he stopped", (c) "at that point he somehow got tied up",

and

If the context implies anything on the matter, it rather implies that there was no difficulty at 17.[21]

I have just set out a sequence of proofs based on diagrams that snarls up at the case of $\sqrt{19}$, and so may be a possible candidate for anthyphairetic explanation of this passage.[22] Is it viable? Alas, I think not:

(i) It is susceptible to many of the objections to the whole class of anthyphairetic interpretations of the passage; see Knorr, *EEE*, 118–26. However, it has some very distinct advantages over them: all of its details are fully laid out here; all the cases are treated by a consistent geometrical technique; and the objection that it does not fit the style of fourth-century mathematics disappears within the context of my reconstruction. In truth, most other anthyphairetic techniques already run into difficulties with $\sqrt{13}$, or they change to a new method for this case, or they avoid it by invoking an unacceptable use of algebraic manipulation.

---

the case of 19. Caveing's proposal, based on a suggestion of Itard's, is most ingenious: he sets out an *arithmetical* proof which breaks down at 17 that $p^2 = Nq^2$ is imposible in integers; therefore "Theodorus stopped at 17 because he had shown that *his [geometrical]* method worked successfully for this number". He does not then go on to analyse what Theodorus' proof might have been, but this interpretation does open the door to the whole class of anthyphairetic proofs, provided—see above—they can realistically deal with the case of 13. Kahane's ingenious method is based heavily on subtle properties of the quadratic extensions of the rationals, and relies on arithmetico-algebraic techniques; it could perhaps be rephrased to remove some of these anachronistic features but would, I think, even then be a non-starter as a *historical* reconstruction.) Also see Høyrup, DBT, to which my subsequent note DMS brings us back to one of the considerations which opened this section: do we have any real evidence that Greek mathematics was influenced by the earlier Old Babylonian mathematics?

[21] From Burnyeat, PSTM: first quotation, 503–4 (translation (c) is taken from Plato, *Theaetetus*, tr. McDowell), and second quotation, 513. Also see the subsequent exchange in Knorr, MPP, with a reply by Burnyeat. There is a striking and important difference between Knorr and Burnyeat's approaches: Knorr is, as always, concerned with the technical mathematical details of the passage while Burnyeat leaves this aspect to one side in favour of philosophical issues, and the difference led to this regrettable altercation between them.

[22] There are nice summaries of typical supporting arguments for reconstructions in Artmann, PTTDD, 17–18 and 20–22.

(ii) It does not just prove incommensurability, as Theaetetus describes Theodorus' lesson. It does very much more, and actuallý evaluates the precise pattern of the anthyphairesis.

(iii) It does not encounter difficulty with $\sqrt{19}:1$ "for some (obscure) reason", as the passage may hint. It is obvious to someone who tries to press the procedure on this case why the method breaks down.

(iv) The case of $\sqrt{18}:1$ presents no difficulty, an objection to almost all other interpretations which never seems to me to be really satisfactorily explained.

(v) I am still emotionally attached to Knorr's interpretation and still find that it fits our evidence closer than any other interpretation so far offered, though I increasingly find its mathematical content unsatisfactory; I feel uncomfortable with it, though do not want to analyse it more precisely than that, as a piece of mathematics of its period. But I would have no objection to being persuaded that mine is a reasonable and plausible interpretation!

### 10.3(e)  *The overall structure of the Theaetetus*

Plato's *Theaetetus* presents us with many puzzles, and I would like to finish this section with a very brief discussion of some that are relevant to our approach here. Theaetetus was renowned as a mathematician but, while the dialogue is seen as an encomium to him, it only contains this explicit appeal to mathematics.[23] Moreover the bulk of the dialogue is given over to Socrates demonstrating that he is not a very acute philosopher. There seems to be an inconsistency: we would expect more mathematics and less philosophy. But it may indeed contain more mathematics than is generally realised.

First, there is the fine detail of its language, which seems in places to have strong mathematical overtones; there are some remarks on that in Section 5.2(a). Next, there is the question it is built around, of the difference between compelling belief, Plato's 'true judgement', and proof, Plato's '*logos*' or 'account', of which, as I have remarked, the problem of the dimension of squares provides a very striking illustration, the best that I know of within the plausible context of Greek mathematics. And, finally, there is the overall structure of the dialogue, which proceeds with Theaetetus making a series of four assertions that Socrates demonstrates to be wrong, and finishing with some progress towards an answer but, however, no actual correct answer. It has been suggested in Brown, TKCL, that this structure is itself a mathematical analogy, and Plato is here making reference to approximation by a sequence of harmonic and arithmetic means. The suggestion is very attractive, but this mechanism of implementing it rather clumsy, which may be why it does not seem to have been taken seriously. However, as a reference to anthyphairesis as

---

[23] There is some circularity in this assertion: An important piece of the evidence for Theaetetus as a mathematician is this eponymous dialogue, which is the most mathematical of Plato's dialogues, in part because it is dedicated to the mathematician Theaetetus.

a method of approximation, if anthyphairesis was then a well-known technique used by mathematicians and, in particular, a technique on which Theaetetus had made significant contributions, it looks spot on.[24]

## 10.4 *ELEMENTS*

### 10.4(a) *Pre-Euclidean Elements*

The topic of this section can be described by a minor variation of the opening of Section 6.2. Suppose we take a collection of texts relating to the early developments of Greek mathematics—the pre-Euclidean *Elements*, for example; what, precisely, are we considering? I shall look at aspects of our evidence about early elements from this point of view, and will only give my own position very briefly at the end, after a long review of the evidence.

### 10.4(b) *Lexica, dictionaries, and the scholarly literature*

Before starting, it is worth consulting the standard reference books.

Liddell, Scott, & Jones, *Greek–English Lexicon*, has a substantial entry on στοιχεῖον. The first and last senses and the beginning of II 3 are given verbatim below, the rest are summarised very briefly, and I have expanded the abbreviations:

I. In the form of a sun-dial, *the shadow of the gnomon, the length of which in feet indicated the time of day,* ὅταν ᾖ δεκάπουν τὸ στοιχεῖον when the *shadow* is ten feet long, Aristophanes, *The Ecclesiazusae* 652, vide Scholia; ὁπηνίκ ἂν εἴκοσι ποδῶν .. τὸ ϲτοιχεῖον ᾖ Eubulus 119.7, cf. Philemo 83.

II 1. A *simple sound* of speech . . . strictly different from letters (γράμματα) . . . though frequently not distinguished from them (as by Plato, *Theaetetus* 1c, *Cratylus* 426d, . . . ).

2. In Physics, the *ultimate components of matter,* reduced to four by Empedocles, who called then ῥιζώματα, the word ϲτοιχεῖα being first used (according to Eudemus, *apud* Simplicius, *In Physica* 7.13) by Plato, *Theaetetus* 201c, *Politicus* 278d, *Timaeus*, 48b, cf. . . . .

3. The *elements of proof,* e.g. in general reasoning the πρῶτοι ϲολλογιϲμοί, Aristotle *Metaphysics* 1014b 1; in Geometry the propositions whose proof is involved in the proofs of other propositions, *ib.* 998a 26, 1014b 36; title of geometrical works by Hippocrates of Chios, Leon, Theudios, and Euclid, Proclus, *in Euclid*, pp. 66, 67, 68f.: hence applied to whatever is one, small, and capable of many uses, Aristotle, *Metaphysics* 1014b 3; to whatever is most universal, e.g. the unit and the point, *ibidem* 6; the line and the circle, Idem *Topics* 158b 35; . . .

4. Generally, *elementary or fundamental principle* . . .

---

[24] Malcolm Brown has often urged me to develop this argument fully, but I have always been daunted by the prospect of a detailed struggle with this dialogue. There is an interesting suggestion in Artmann, PTTDD, p. 29, concerning Plato's substantial discussion of Protagoras' claim that "Man is the measure (*metron*) of all things": Theodorus has shown that there are many line segments with no common measure (again *metron*), so was Plato here, in a subtle way, subverting his teacher's thesis? Moreover Theodorus is himself declared to be the *metron* of geometrical diagrams, the judge of their validity (169a).

5. ἀςτέρων ςτοιχεῖα the stars . . .

6. ςτοιχεῖον = ἀριθμός, as etymology of Cτοιχαδεύς, Scholia Dionysius Thrax p. 192 H.

So we should not expect to have to look very widely in our pursuit of mathematical elements: the discussion will be dominated by Proclus, Plato, and Aristotle. But there are the curious first and last of these senses that we should not completely lose sight of, though they do look to be usages on the very fringes of the wide semantic shadow of the word.

We must also be aware of what our modern understanding of the translation of ςτοιχεῖον might bring to the discussion. The *Oxford English Dictionary* takes more than two pages over *element, elemental, elementary*, etc., far too much to begin to summarise here, and most of it not directly relevant to our purpose. But readers are encouraged to consult these items, if only to find how they will have to search through them closely to find what does interest us here. Only at the very end of the entries (this is a historical dictionary and the entries are as far as possible arranged chronologically as determined by the date of the first quotation supporting each sense, so these are the most recently introduced meanings) do we find:

*element* 14a. †[i.e. obsolete] The letters of the alphabet (obsolete). Hence, the rudiments of learning, 'the 'A.B.C.'; also the first principles of an art or science.

14b. Euclid's *Elements*: the title of a treatise on the rudiments of Geometry.

and

*elementary* 7a. Of the nature of elements or rudiments, rudimentary, introductory.

Hence we should not forget that these are relatively recent senses that have come to dominate our understanding of this cluster of English words, and try to approach the older and different Greek evidence with an open mind concerning the possibility of different connotations.

The standard, pre-eminent, long scholarly article on the topic, whose last quarter deals with the mathematical usages of the word, is Burkert, SSS. He demonstrates that ςτοιχεῖον did not originally mean 'letter' and then was generalised; rather, its original meaning was of a line of things in order: a column of soldiers, a row of stakes in the ground, etc., whence "arguments set in a line" (Burkert, *LSAP*, 402). But, much as I admire, respect, and have been influenced by Burkert's great book on Pythagoreanism, I find the sections of this earlier article that deal with mathematics are based more on assertion than evidence and argument, and feel that it represents a different scholarly tradition. I encourage others to read and judge it, but will not attempt a summary of it here.

### 10.4(c) *Proclus*

Our main source of evidence is, of course, the second Prologue of Proclus' *Commentary on the First Book of Euclid's Elements*, first in the passage where

he is believed to be quoting Eudemus, and then in the long passage where he discusses the nature of elements. Here are some extracts from the Eudemus passage, 66 ff. All passages will be quoted in the translation of Morrow:

Hippocrates wrote a book on elements, the first of whom we have any records who did so.

. . .

Younger than Leodamus were Neoclides and his pupil Leon, who added many discoveries to those of their predecessors, so that Leon was able to compile a book of elements more carefully designed to take account of the number of propositions that had been proved and of their utility.

. . .

Theudius of Magnesia . . . produced an admirable arrangement of the elements and made many partial theorems more general.

. . .

Hermotimus of Colophon pursued further the investigations already begun by Eudoxus and Theaetetus, discovered many propositions in the *Elements*,[25] and wrote some things about locus problems.

We know nothing more than is given here by Proclus about Leon, Theudius, and Hermotimus, though what is evident from these brief extracts is that he does say that Theudius and, perhaps, Hermotimus were associates of Plato, "making their inquiries in common". So Proclus–Eudemus are telling us—but can we trust them?—that the writing and perfecting of elements was a particular feature of the mathematical activity of the Academy. However, although it is generally agreed that Theaetetus and Eudoxus made very substantial contributions to Euclid's *Elements* (to Books X and XIII, and V and XII respectively, at least), and Archytas is believed to have contributed to the arithmetical Books VII–IX, no hint whatsoever of that is suggested here.

Proclus then goes on to give two substantial sections devoted to the purpose and arrangements of *Elements*, especially Euclid's *Elements*. Here he is clearly setting out much more in his own opinions, rather than quoting ancient sources; for example: "Looking at the subject matter, we assert that the whole of the geometer's discourse is obviously concerned with the cosmic figures" (70.23–25), a neo-Platonic opinion with which few who have made a close study of the whole work would agree. One should rather bear in mind Proclus' own relation to the material and the nature of his understanding of mathematics. There is a vivid description of this in Mueller, MPPC, 307–8:[26]

Whatever the precise character of his mathematical training, it seems reasonable to assume that Proclus came to mathematics late and that the mathematics he was given

---

[25] I have adopted Morrow's use here (but not always elsewhere) of 'elements' when we cannot be specific about just what is being considered, and '*Elements*' when we are talking about a particular book with that word in its title. But I do not know how he could justify the fine gradation of nuances here: "on elements", "of elements", "the elements", then "the *Elements*".

[26] Also see Mueller, IPEC, for a discussion of Proclus' sources in his First Prologue.

was elementary and dryly formalistic. This mathematics was presumably the best or close to the best available in his lifetime, and so was the mathematics he taught his own students. The contrast between Proclus' situation and Plato's is instructive. Plato lived at a time when mathematical knowledge was expanding rapidly, and technical advance mingled with philosophical speculation to create a sense of unlimited possibility. Not until the early modern period, when mathematics again enters a period of rapid expansion, do we find as convincing a proclamation of the broad powers of mathematical science as we find in Plato's *Republic*. . . . For Proclus Plato's word is authoritative. . . . But for Proclus and his predecessors and contemporaries mathematics was a closed-off subject to be formalised and reflected on, but not to be developed and advanced. Moreover, it seems likely that for many of them mathematics was closed off in another sense: namely incomprehensible, even at the level taught by Proclus and Theon. If Theon's description of his students isn't taken to be sufficient evidence for this claim, the way that he and Proclus present the subject surely is; one doesn't dot *i*'s and cross *t*'s for students who are following along well.

Notwithstanding this, modern commentators have made much of Proclus' comments here. Burkert, SSS, for example, discusses at length the passage:

The term 'element', however, can be used in two senses as Menaechmus tells us. For what proves is called an element of what is proved by it; . . . But in another sense 'element' means a simpler part into which a compound can be analysed [72.23–73.6],

and reasonably so, since the reference to Menaechmus may indicate that Proclus is here referring to some other source—but an unknown source of unknown reliability, transmitted by Proclus with an unknown degree of distortion. Moreover, this quotation is doing no more than referring to two of the well-established meanings of 'element', attested in other quotations. And the useful general discussion in Artmann, EEP, is carried out in the spirit of Proclus' description, possibly in Proclus' own words, of how:

Of those who have attempted [the proper selection and arrangement of elements,] some have brought together more theorems, some less; some have used rather short demonstrations, others have extended their treatment to great lengths; some have avoided the reduction to impossibility, others proportion; some have devised defences in advance against attacks upon the starting points; and in general many ways of constructing elementary expositions have been individually invented [73.17–25].

Here I shall take a description that I hope will embody the main features of the general opinion concerning pre-Euclidean elements and Euclid's *Elements*, which is strongly based on Proclus' *Commentary*, and analyse its presuppositions in detail. Here it is:

GENERAL OPINION:   Pre-Euclidean *Elements* were books on geometry and arithmetic which looked like Euclid's *Elements*, but with different contents, arrangements, and proofs, and sometimes arranged according to different principles; but Euclid's *Elements* is superior to all of its predecessors. Here "looked like" means with propositions arranged in logical order, and with each proposition structured in the distinctive style of the Euclidean proposition. However, some of these features may have taken some time to become established.

I shall concentrate here on those main features of Euclid's *Elements* that these early *Elements* may reasonably be expected to resemble: the structure of a Euclidean proposition, and the logical structure of the whole, but will not discuss their contents in any detail.

### 10.4(d) *The logical structure of Euclid's Elements*

Euclid's *Elements* has an almost universal reputation, even among historians of mathematics, for its relentlessly logical arrangement, and this reputation is justified for its first book, the only book that many look at, once it has got itself under way.[27] But thereafter things are not quite so straightforward. Mueller has tried to tease out the structure of the work in his massive *Philosophy of Mathematics and Deductive Structure in Euclid's Elements*, and here are a couple of his conclusions:

Propositions [of books III, IV, and VI] with no future use are proved . . . Needed propositions are not always supplied explicitly . . . In addition, in the first part of III and in VI, little attention is paid to logical sequence . . . Euclid clearly thinks of subject matter as a more important organising principle than deductive relevance—a fact which makes it difficult to read the *Elements* straight through [*PMDSEE*, 204].

The logical structure of books VIII and IX is puzzling, to say the least. Certain groups of propositions cohere, but the internal and external arrangement of the groups is not always perspicuous, and the details of the argumentation sometimes leaves much to be desired [*PMDSEE*, 83].

Elaborations of these remarks and further such comments can be found on pp. 158 (on book VI), 177–9 (on III), 197 (III again), 211 (on XI), and elsewhere.

So Euclid's *Elements* may aspire to a logical structure and achieve it for much of the time, but let us not exaggerate this too much. And, since no pre-Euclidean *Elements* survive, to say more about this feature of them would be to hang speculation on misrepresentation.

### 10.4(e) *The Euclidean proposition*

The most characteristic feature of Euclid's *Elements* is the structure of each proposition, and again it is Proclus the formalist who describes this:

Every problem and every theorem that is furnished with all its parts should contain the following elements: an enunciation (πρότασις), an exposition (ἔκθεσις), a specification (διορισμός), a construction (κατασκευή), a proof (ἀπόδειξις), and a conclusion (συμπέρασμα) . . . [Proclus–Morrow, *CFBEE*, 203–4].

Not all Euclidean propositions exhibit this structure,[28] and far from all of Greek mathematics begins to conform to these rules; for example, Archimedes has a much looser style. But, since we are looking here at the evidence for the predecessors of the Euclidean style, we shall look for pre-Euclidean evidence of

---

[27] There is, for example, a bit of trouble about the intersections in I 1, and the role of movement in I 4.

[28] For a simple example, the conclusion is an almost word-for-word repetition of the enunciation, but our manuscripts occasionally omit this, finishing with καὶ τὰ ἑξῆς, 'and the

this striking detail of the style.[29] And it will not be necessary to look for the whole shooting-match; it will be sufficient to look for its distinctive and useful opening feature, the enunciation or *protasis*.[30] But please be sure about my aim here. I do not deny that there was substantial pre-Euclidean mathematics; quite the reverse, in fact. What I want to explore is the much more narrow and limited question of whether there is any evidence that this mathematical activity was striving towards the style and content of Euclid's *Elements*.

### 10.4(f)  *Pre-Euclidean evidence on elements and mathematical style*

The first problem is that we know nothing about Euclid, not even the date of the *Elements*, so we adopt the conventional view, and look for evidence up to

rest'; see X 2, 6, 7–9, 11, 13–20 including the lemma to 17, 25, 91, 100, and XI 35 and 37. (This list is complete.) There is some variation among the manuscripts used by Heiberg, of which the following is a summary of the most important features. The Viennese ms. V shows the most variety, completing the enunciation in X 6–8 and 13, adding a bit more of the very long enunciations of 9, 11, 14, 15, 17, 91, and 100, and making no change or even omitting some of the abbreviated enunciation in 16, 19, and 20. Also V sometimes adds ὅπερ ἔδει δεῖξαι (X 2, 6, . . . , though I cannot completely understand Heiberg–Stamatis' apparatus at X 2) and sometimes doesn't (X 7, 8, . . . ). The 'Theonine' group of mss. BFVb make few systematic alterations: in X 9 they add ὅπερ ἔδει δεῖξαι; in X 91 and 100 they complete the enunciation but don't add ὅπερ ἔδει δεῖξαι; and in XI 35 they omit τῆς προτάςεως 'of the protasis' which is in P (and where Heiberg uniquely and inexplicably adds ὅπερ ἔδει δεῖξαι in brackets). P is the most important manuscript in Heiberg's edition, and he usually gives it as his main text except, for example, when all other texts are united against it; this also happens in X 25, where P has the complete enunciation but no ὅπερ ἔδει δεῖξαι, V simply adds ὅπερ ἔδει δεῖξαι after ἑξῆς, and there are minor variations in F. But what is remarkable is that none of these scribes and editors systematically restore this most evident and automatic feature of Euclidean style in all cases; this is perhaps an indication of the general stability of the transmission process, apart from scribal errors. Heath has a much more robust approach: he abbreviates the whole conclusion to "Therefore, etc." from I 4 onwards, and abbreviates ὅπερ ἔδει δεῖξαι or ποιῆσαι to QED. or F. from I 5; see his notes in *TBEE* i, 249 and 252.

[29] A study of the fine detail of Greek mathematical style is given by Reviel Netz in a forthcoming book *The Shaping of Deduction in Greek Mathematics*.

[30] This feature has been convenient from earliest times, in that a memorandum of the *Elements* can be given by just listing the enunciations, perhaps also including the diagrams. We find this in the papyrus P. Berol 17469 (see Plate 4), a fragment containing the diagrams for I 8–10 and the enunciations of 9 and 10, and possibly also in P. Oxy 1 29 (see Plate 2 and Section 6.2, above); the Bologna ms. b, which may derive from a pre-Theonine source (see Heath, *TBEE* i, 49), starts with all of the definitions and enunciations of the *Elements* and the *Data*; the early Latin version known as 'Boethius II' contains, amongst other things, the enunciations of almost all propositions of Books I–IV, but only the proofs of I 1–3; from the 12th century and the first Latin translation by Adelard of Bath up to the Renaissance, some believed that Euclid may have been the author of only the definitions, postulates, common notions, and enunciations, and that the proofs were added by later commentators (see Clagett, MLT, 19); and, for the latest example, Mueller's *PMDSEE* finishes with such a 62-page Appendix called 'The Contents of the *Elements*', though the author well knows that this is not the case, nor is it really a table of contents! There is a more recent German style of setting out mathematics, generally called 'Satz-Beweis' style, that has some affinities with *protasis*-style. Its premier exponent was Edmund Landau (see his *Grundlagen der Analysis* (1930), tr. F. Steinhardt, *Foundations of Analysis* (1951)), but Landau does occasionally permit himself a short remark!

300 BC. We shall find that the *only* evidence for *protasis*-style in this period will be in Aristotle, and most of that, and that part which is indubitably due to Aristotle, will not be in mathematics but in logic, in his *Prior Analytics*.[31] Outside Aristotle, the evidence will be in music, with Aristoxenus' *Elements of Harmony*, Book III; or in contexts which may not actually be pre-Euclidean, like Autolycus, *On Risings and Settings and On the Heavenly Spheres*. As to '*Elements*', the only book with a title using this name might actually have acquired this title later: Aristoxenus' *Elements of Harmony*. Here are the details:

*Plato and the Academy*: Notwithstanding Proclus–Eudemus' description of the element-writing activity practised in the Academy, there is no evidence of *protasis*-style in Plato's writings, nor even any parody of it, for example in the *Parmenides*. There is much talk of elements,[32] which can be roughly divided between the following periods:

- Early, in the *Cratylus*, in contexts such as "And the original elements are letters" [434b].
- Middle, as in Socrates' dream, at *Theaetetus* 201–2, an extract of which will be quoted later.
- Late, in the *Sophist, Philebus, Timaeus*, and *Laws*, for the four elements, or five in the *Epinomis*. For example: "And for these reasons, and out of such elements which are in number four, the body of the world was created, and it was harmonised by proportion . . ." [*Timaeus* 32c].

The mathematical passages at *Meno* 82b–85b (on doubling a square) and 86e–87b (on the use of hypotheses) are not cast in Euclidean style. And nowhere do we find a hint of any book which has *Elements* in its title. An even stronger indication of this can be found in the *Thirteenth Letter*, to Dionysius:

I am sending you some Pythagorean (treatises, $\Pi\upsilon\theta\alpha\gamma\acute{o}\rho\epsilon\iota\alpha$) and some classifications ($\delta\iota\alpha\iota\rho\acute{\epsilon}\varsigma\epsilon\iota\varsigma$). I am also sending you a man . . . who will perhaps be useful to you and Archytas. . . . His name is Helicon . . . a pupil of Eudoxus. . . . Take lessons of him, in addition to the rest of your philosophic training . . . and not only benefit yourself but add to your reputation ($\epsilon\mathring{\upsilon}\delta o\xi\mathring{\eta}\varsigma$[33]) [360b–e].

This letter may be a forgery, but that would not lower its evidential status since it would then have been composed to correspond as closely as possible in style and content to the period. Although the argument has been developed that Book VII was a (Pythagorean?) *Elements of Arithmetic* used by Archytas in his

---

[31] I would like to thank Ian Mueller for pointing this very important collection of examples out to me.

[32] I have used the useful, though surely far from complete, index in the Bollingen edition, Plato, *CD*. A much more thorough search could be made using Brandwood, *WIP*.

[33] Malcolm Brown (unpublished) sees this as a conscious pun on Eudoxus.

composition of the archetype of *Elements* VIII,[34] we do not find the name
'*Elements*' used here. Moreover the role of Helicon underlines what we might
now forget, with our efficient and user-friendly writing systems and our
written-, print-, and computer-based intellectual culture of the alphabet: that
writing in Plato's time was inefficient, requiring a lot of papyrus for its
generously laid-out majuscule script; unfriendly, with its unspaced script that
could be very difficult to decipher; fragile; lengthy to duplicate; and generally
denigrated by Plato (see the *Seventh Letter* 341 and *Phaedrus* 274–5; this topic
is discussed in Section 1.2(e) and Chapter 6). It was then probably more useful
as an *aide-mémoire* than the primary way of transmitting information. More-
over, during Plato's time, most mathematicians were either based around
Athens or, like Archytas, in direct contact with Athens, so for whom would
this writing be executed? This situation will change with Aristotle, with his
projects to collect and classify knowledge and practice of all kinds, and then
change even more dramatically with the foundation of Alexandria by his pupil.
So let us now turn to him.

*Aristotle and the Lyceum*:   Again, we find many references to 'elements'
throughout Aristotle,[35] especially in *Physics*, *Metaphysics*, and *Rhetoric*, and
these usages cover the whole spectrum of meanings of the word except as its use
in the title of a book. Here are just a few examples:[36]

If one entertains the same notions as Empedocles and has the same views about the
elements as he, is he unanimous with Empedocles? Surely not! [*Magna Moralia*
1212a15].

There is the argument that one who knows the letters ($\sigma\tau o\iota\chi\epsilon\hat{\iota}a$) knows the whole word,
since the word is the same thing as the letters which compose it . . . This whole
commonplace is fallacious [*Rhetoric* 1401a29].

Evidently, then, these thinkers [the Pythagoreans] also consider that number is the
principle both as matter for things and as forming their modifications and states, and
hold that the elements of number are the even and the odd . . . and the whole heaven, as
has been said, is numbers [*Metaphysics* 986a17–21].

The elements of geometrical proofs, and in general the elements of demonstrations, have
a similar character; for the primary demonstrations, each of which is implied in many
demonstrations, are called elements of demonstrations; and the primary deductions,
which have three terms and proceed by means of one middle, are of this nature
[*Metaphysics* 1014a35].

---

   [34] See Artmann, EEP, for more details.
   [35] Again, I have used the useful, though surely incomplete, index in the second, Bollingen,
edition, Aristotle–Barnes, *CW*.
   [36] It is perhaps of interest also to list the passages cited by Burkert, in his SSS:
*Categories* 14a38 f.; *Posterior Analytics* 84b19 f.; *Topics* 158b34 f.; *Metaphysics* 998a23 f.
and 1014a35 f.; and *Rhetoric* 1396b22 f. and 1403a18 f.

But, as I remarked above, the *Prior Analytics* does exhibit something close to a sustained Euclidean style of exposition, with enunciations, proofs with considerations of different cases, and alternative proofs. I choose just one example from scores, the figures known as 'disamis' and 'datisi':[37]

[Enunciation:]   If one term is related universally to the middle, the other in part only, when both are affirmative there must be a deduction, no matter which is universal.

[Setting-out. First Possibility:]   For if R belongs to every S, P to some S, P must belong to some R.

[Proof:]   For since the affirmative is convertible, S will belong to some P, consequently since R belongs to every S, and S to some P, R must also belong to some P; therefore P must belong to some R.

[Second Possibility:]   Again if R belongs to some S, and P to every S, P must belong to some R.

[Proof:]   This may be demonstrated in the same way as the preceding.

[Alternative Proof:]   And it is possible to demonstrate it also *per impossibile* and by exposition, as in the former cases [28b5f.].

There is no conclusion ("Therefore, etc.", *à la* Heath) and no diagram,[38] but most of the rest is there. And, what is more, we have a kind of symbolic manipulation such as we will not find in mathematics until the end of the sixteenth century.[39] Then there is the same kind of structure on a larger scale: each of Chapters 4 to 6 starts, for each figure, with a kind of enunciation, and then finishes with a kind of conclusion. There is also the same kind of use of *protasis*-style in the *Meteorologica* Book III, at 373a3–19:

. . . if the lines start from the same point and end at the same point and are equal, the points which form an angle will always lie on a circle.
Let ACB and AFB be lines each of which goes from the point A to the point B and forms an angle . . . ,

and in the locus problem at 375b16–376b12:

The rainbow can never be a circle nor a segment of a circle greater than a semicircle. The consideration of the diagram will show this and all the other properties of the rainbow. Let A be a hemisphere resting on the circle of the horizon, let its centre be K and let G be another point on the horizon . . . ,

---

[37] There is an explanation of these scholastic names in the introduction of the Loeb Classical Library edition of the *Prior Analytics*, vol. i of their Aristotle collection, p. 197.

[38] The diagrams are not mentioned by Proclus in his description of the parts of a proposition, but every proposition of Euclid's *Elements* has one, most of them make essential use of them, and the proofs in the *Meteorologica*, to be quoted below, explicitly refer to diagrams. The important role of diagrams is also analysed in Netz, *SDGM*; he thinks that the *Prior Analytics* may originally have had schematic diagrams.

[39] Why did these first steps towards a symbolic logic not develop into some rich and sophisticated language, as symbolic algebra did very quickly after its introduction at the end of the sixteenth century? I can offer no insights into this, and would welcome a discussion of the contrast by someone more competent than me.

but doubts have been cast over the authorship of this book of the *Meteorologica*.[40]

But all of this evidence seems to show that there is a possibility that Aristotle and the Lyceum may play a more important role in the development of the Euclidean style than is generally credited. Further reinforcement of this can be found in Aristoxenus, another member of the Lyceum, in his *Elements of Harmony* iii. Here is a typical example:

A tone is placed next to a ditone, only above it.
Suppose it is placed below. The consequence will be that . . . Since this is unmelodic, a tone placed immediately below a ditone must be unmelodic too,

where we also find a fluent proof by contradiction. However the evidence for Aristoxenus as the author of a book called '*Elements of . . .*' is less clear, since this name may well have been given later, when the tradition of using this word in technical treatises had been established; see Barker, *GMW* ii, 177 and 146, n. 121.

*Hippocrates and Archytas*:  We now turn to the only other pre-Euclidean authors of whom we may have any surviving mathematical texts. Once again, our ultimate source is Eudemus, who may or may not have modernised his sources—but if he did not, he is not only the first known historian of mathematics, but also the first for a long time who gives us his material straight, without rewriting it!—and again later commentators who quote Eudemus may also alter it in the transmission. Simplicius, the source of most of our fragments of Eudemus, is sometimes explicit about this:

I shall set out what Eudemus wrote [about Hippocrates' *Quadrature of Lunes*] word for word, adding only for the sake of clearness a few things taken from Euclid's *Elements* on account of the summary style of Eudemus, who set out his proofs in abridged form in conformity with the ancient practice.[41] He writes thus in the second book of his *History of Geometry* . . .

---

[40]  See, for example, the long discussion of this in Knorr, *ATGP*, 102–8, which refers back to earlier discussions; here are a few quotations from it: "This passage is unique in the Aristotelian corpus in that it provides a full geometric demonstration, and all the more remarkable in that the manner of proof is identical to that employed a century later by Apollonius in his second book *On Plane Loci*" (p. 103); "In speaking of the 'Aristotelian writer', I betray a reluctance to take Aristotle as author of this solution" (p. 107); "If these considerations do not quite rule out the authenticity of *Meteorologica* Book III, they surely deserve more serious attention than Aristotle scholars tend to afford them. . . . But the suggestion of its nonauthenticity at once raises the question of dating. It has already been observed that the *Meteorologica* appears to be a composite document, different sections reflecting times ranging from the mid-350's to the late 340's. Conceivably, then, some of its parts might be later still, and some might even be the work of disciples after Aristotle's death in 322. If so, how much later? If the passage on the locus problem derives from the 3rd century, near or even after the date of Apollonius, then its significance for the history of ancient geometry diminishes almost to nil" (p. 108). However, Knorr finally plumps for a date at the end of the fourth century, and its author a disciple of Aristotle, working from an unknown source.

[41]  This appears to mean that Eudemus is ancient for Simplicius (*c.* 800 years earlier),

so the text is corrupt, on its own admission. Much effort has been spent in trying to purge it of the additions, and there is no agreement over this; but it does seem that, taking the broad view of common agreement, the result is not in *protasis*-style. Here, for example, is the opening section, as presented in Knorr, IC 127–8, in which text that has been seriously proposed to be Simplicius' addition by some scholar, though not necessarily by all of them, has been set in italic:[42]

He made a beginning, and of the things useful for these [constructions] he first set out that [1] similar segments of circles have to each other the same ratio as their bases in power. This he proved from the proof that [2] the diameters have in power the same ratio as the circles, *which Euclid placed second in the tenth [twelfth!] book of the Elements, giving as its protasis thus: [2a] Circles are to each other as the squares on the diameters. [3] For as the circles are related to each other, so also the similar segments. [4] For similar segments are those which are the same part of the circle, for instance, the semicircle to the semicircle and the third part to the third part; because [5] also similar segments admit of equal angles. Now, [6] those [angles] of semicircles are right, and those of greater [segments] are less than right, and by as much as the segments are greater than the semicircles; and those [angles] of lesser [segments] are greater and by as much as the segments are less [than the semicircles].* Having proved this first he then proved that [7] if a lunule has a semicircle as its outer arc, its quadrature might be effected in a certain way. He did this by . . . [61.5–18]

This is indeed a sequence of *protases*, but most of them may be inserted by Simplicius, and even the remaining may be Simplicius–Eudemus' rendering of what Hippocrates wrote. The exposition then continues, not as *protasis* and proof, but in a loosely discursive style. And the same thing is found in our only other substantial pre-Euclidean fragment that we know, directly or indirectly, from Eudemus, namely the solution of the duplication of the cube by Archytas, which opens:

5. The solution of Archytas, according to Eudemus. Let the two given straight lines be *AΔ*, *Γ*; it is required to find two mean proportionals between *AΔ*, *Γ*. Let the circle *ABΔZ* be described . . .

In other words, no surviving fragment of mathematics transmitted by Eudemus is in *protasis*-style. Even our oldest primary Euclidean text, the very fragment-ary collection of ostraca from Elephantine Island, dated to *c.* 225 BC, which contains material cognate with *Elements* XIII 10 and 16, is not in *protasis*-

-

though how much more ancient these practices are than Eudemus seems undecidable. But this then also means that the 'abridged ancient practice' was still current in the last half of the fourth century, very close to the supposed time of Euclid, though it is not clear how much it refers to the style of philosophers rather than mathematicians.

[42] There is another translation and discussion of this extract in Mueller, *PMDSEE*, 199–202. The matter of purging the text of its additions has been discussed by Allman, Becker, Bulmer-Thomas, Diels, Tannery, Heiberg, Rudio, Heath, Knorr, etc. The question carries a certain charge; for instance, several of these items involve proportions, so what kind of evidence does this give us of the proportion theory that Hippocrates was using?

style.[43] What is more, none of these pieces of pre-Euclidean mathematics that we have—Plato's duplication of a square, Aristotle's rainbow locus, Hippocrates' quadrature locus, and Archytas' duplication of a cube—describe material to be found in Euclid's *Elements*. We might propose that the *Elements* constitutes material that was developed to deal with these three problems, but that also is speculative—no ancient author says so, and the *Elements* does not even contain a proposition setting out Hippocrates' first move, that of reducing the problem to finding two mean proportionals, nor about his *Quadrature of Lunes*, which looks like an associated problem of squaring; nor does it seem to explain Book X and cognate material. Yet more, Sense 6 of the *Greek–English Lexicon*, quoted at the beginning of this article, offers us some strange reflections on the word. This comes from the scholia to Dionysius Thrax (an influential second-century BC Alexandrine grammarian) of unknown authorship and date, which give a wide variety of explanations for the word $\text{c}\tau o\iota\chi\epsilon\hat{\iota}o\nu$: for example, the word came from $\dot{\alpha}\rho\iota\theta\mu\acute{o}\text{c}$, since there was a certain local Zeus called $\text{C}\tau o\iota\chi o\delta\epsilon\acute{\upsilon}\text{c}$ who numbered and ordered the people (so here the scholiast had to admit that the relevant sense is 'order', but his idea seems to be that the numerals were introduced first, and then they came to be used as letters); or there was an Athenian called $\text{C}\tau o\iota\chi\epsilon\hat{\iota}o\nu$, who may have invented letters, etc. But various things emerge from all these attempted, and not very convincing, explanations: already in antiquity the scholiast(s) were puzzled by the usages of this word and knew very little or nothing about their origins, and they did not bring the mathematical or other books called *Elements* into their discussion.[44]

10.4(g)  *When, where, and why was the Euclidean style introduced, and when were mathematics books first called Elements?*

My position on the question of the introduction of *protasis*-style in mathematics and the naming of mathematical treatises as *Elements* seems to me the only one permissible in the face of this lack of evidence: apart from Aristotle, there is no clear contemporary evidence concerning these, *even where we might expect to find it*, before Aristoxenus, Autolycus, and Euclid, so there is very little more that one can say on the topic. There are even problems with all of these authors: the name of *Elements of Harmonics* might have been given to Aristoxenus' treatise after the tradition of naming books *Elements of . . .* had been established; the evidence that Autolycus pre-dates Euclid's *Elements* is no more convincing than the other way round;[45] and we know nothing of the

---

[43] These are published in Mau & Müller, MOBS; see Section 6.2 and Dorandi, TPDEE.

[44] This last argument has been put together by Reviel Netz, whom I thank greatly.

[45] See Heath, *HGM* i, 348–53. The argument that Autolycus precedes Euclid is based on the fact that some of Autolycus' propositions appear in Euclid's *Phaenomena*, an argument that can just as easily be applied the other way round. Autolycus' dates are *c.* 315–*c.* 240 BC; and we do not know the relation between Euclid's *Elements* and *Phaenomena*, for example which of them came first, or even whether they were actually written by the same person or persons.

composition—place, date, and author or authors—of Euclid's *Elements*. And I do not know whether the use in mathematics of *protasis*-style could be an importation of its use in logic by Aristotle, and know of no discussion of this.

And yet there is a psychological need to think or say something more, so let me finish this section with some dreams. First Socrates' dream, *Theaetetus* 201c–202c:[46]

> If you have had a dream, let me tell you mine in return. I seem to have heard some people say that what might be called the first elements (*stoicheia*) of which we and all other things consist are such that no account (*aloga*) can be given of them. Each of them just by itself can only be named. . . . But when we come to things composed of these elements, then, just as these things are complex, so are the names combined to make an account (*logos*), a description being precisely a combination of names. Accordingly, elements (*stoicheia*) are inexplicable (*aloga*) and unknowable, but they are perceived, while complexes are knowable (*gnostai*) and explicable (*rhētai*), and you can have a true notion of them.

So the inexplicable elements, name-like, are combined into descriptive complexes, knowable in some deep sense, not unlike letters that can be combined into words, which form descriptions, explanations, and proofs.[47] Moreover Plato's *Theaetetus* seems to be one of his most mathematical dialogues, and Brown, TKCL, has even proposed that it proceeds by a process of successive approximation, never arriving at an end, on a conscious analogy with a mathematical process; see Section 10.3(c). Now my ratios are descriptions—principally anthyphairetic ratios, associated with Theaetetus, and the astronomical ratios, associated with Eudoxus—and these descriptions are all successions of numbers, set out in some form or another, as extended explanations of the ratio. And, in the Greek of this period, numbers were represented by the letters of the alphabet. And what then makes these descriptions not only true judgements, but also enables us to give an account of them and thus turn them into knowledge in the sense of the *Theaetetus*, is the way that we can construct mathematical proofs of some of their quite remarkable properties, and many of these proofs again proceed via approximations and are constructed around gnomon ('knowable') arguments; again see the first meaning of *stoicheia* in Section 10.4(b), above. We find here many of the well-attested meanings of 'element' coming together, and my dream is that these early elements were indeed the study of these kinds of ratios. A very substantial part of Euclid's *Elements* would then be the language and methods that were

---

[46] It is in this passage, rich in mathematical terms, that, according to Liddell, Scott, & Jones, *GEL*, citing Eudemus–Simplicius (see Section (b), above), that we first find the use of στοιχεῖα as the ultimate constituents of matter.

[47] Before embarking on my own dream, let me dream on behalf of Reviel Netz. In his *SDGM* he develops the argument that the study of lettered diagrams was crucial in the development of the Greek deductive process. So his candidate for elements as letters combined into descriptive complexes could be the lettered diagram. But dreaming on behalf of others is surely a chancy business!

developed to investigate and extend this study of ratios, as this book attempts to illustrate with its preoccupation with Books II, IV, X, and XII, plus a brief mention of Book V in connection with astronomical ratios; then, in a process that we will find time and time again in the future development of mathematics, the original motivating problems fall away and are forgotten, leaving this language as a 'theory' in its own right, to be developed and used in other contexts, as will be described in more detail in the following section.

## 10.5 . . . BUT WHY IS THERE NO EVIDENCE FOR THESE RATIO THEORIES?

At first sight, the single most serious objection to the reconstructions of early Greek mathematics developed here must be that if the exploration of ratio theories in general and anthyphairesis in particular did play a role such as I have proposed, then why do our sources, both early and late, not refer to these topics? Here, then, are some reactions to this issue.

First, our sources are not silent. Some of them may be explicitly referring to these topics, if only we could understand them; in other cases, we get a comprehensive and straightforward explanation of otherwise perplexing features when the material is explained from this point of view. For example, two perplexing fragments of Archytas concern *logistikē* and *logismos*; and Plato may tell us a great deal more about this 'ratio theory', in his many further references, if only we interpret him correctly.[48] Aristotle, in possibly our most explicit piece of information—perhaps even our only surviving explicit piece of information—about any early definition of ratio, tells us of *antanairesis*.[49] Euclid never gives us motivation for any of the mathematics he presents, but the only mathematically meaningful explanation of the basic distinction between the *rhētē* and *alogoi* lines of *Elements* X and XIII that I know is the one based on anthyphairesis described above in Chapter 5; and I believe that further relevant material can be found in *Elements* VII. The idiosyncratic features of Archimedes' *Measurement of a Circle* fit more naturally in an anthyphairetic context than an arithmetised one, as I have tried to illustrate in Sections 2.2(b) and 2.4(d). Nicomachus, one of our earliest commentators on *arithmētikē*, writing in the second century AD, gives a discussion of what he called *antaphairesis* and an elaboration of the classification into multiple, epimoric, and epimeric ratios.[50] The later Arab commentators, some of whom are beginning to be revealed as considerable mathematicians in their own right, do several times invoke anthyphairesis in their commentaries on

---

[48] See the catalogue in the Appendix to Chapter 4, and the proposals in that chapter.
[49] Aristotle, *Topics* VIII 3, 158b 29–35, quoted in Section 1.2(d), above.
[50] See Nicomachus–D'Ooge, *IA*, I xiii and xvii seqq., and Section 4.5 above, especially $A_{80}$–$B_{83}$.

*Elements* V.[51] The systematic way of calculating gear ratios such as are found in the Antikythera mechanism uses anthyphairesis.[52] And suggestive material on astronomical ratios is found in the fifth and fourth centuries BC, and on musical ratios throughout early and late antiquity.[53]

Nevertheless, the situation is curious. If anthyphairesis, for example, did provide an impetus to the development of mathematics, as in my reconstruction, then why does nobody tell us this more clearly? I shall approach this more refined version of the question from several different directions. I consider first the question in its historical context of early Greek mathematics; then consider the reactions of the later commentators; then introduce and develop an analogy from more recent times.

So, first, consider our early sources. My study here has been restricted to only a part of the evidence deriving from Plato and his friends, who met in their suburb of Athens and communicated amongst themselves by conversations and letters; from Archimedes, whose isolation from the mathematicians of Alexandria is testified by the prefaces to his works; and from the elusive figure of Euclid. We know the names and, sometimes, a few details of the work of many other mathematicians, working in other places, or at other times, or on other topics; and we also know that my small, selected group of mathematicians were not working exclusively or even primarily on mathematics, or exclusively on the topics in mathematics that I have been considering. So we should not think of the mathematicians of the fourth and third centuries BC as a homogeneous group, in close and constant contact; and the specialism of my study here should not be thought of as reflecting any exclusive specialism of its restricted cast of characters. And in this Greek world of intense intellectual, artistic, and political exploration, the excitement of the explosion of ideas across such a vast front might leave little time for, and less interest in, a clear and correct historical account by the participants.

Now turn to the later commentators, from whom the main structure of the received interpretation derives. If they had got main details of their story right, if they still had reliable sources and had been able to use them with insight, then there might not be the fundamental problems about understanding early Greek mathematics that still face us today. But we cannot have that confidence, since a long process of refinement and critical analysis of their writings has only led

---

[51] The Arabic commentaries on ratio are summarised and discussed in Plooij, *ECR*. Also see Omar Khayyam–Amir-Móez, *DDE*.

[52] See Price, *GG*, especially p. 58: "In this [now lost] planetarium Archimedes would have used, perhaps for the first time, sets of gears arranged to mesh in parallel planes, and he would have been led to the rather elegant number manipulation which is necessary to get a set of correct ratios for turning the various planetary markers". This is, of course, the same kind of investigation that we later find in Huygens, *DAP*; see the Historical Notes to Section 9.1(a), above.

[53] See the references in Sections 4.4 and 4.5.

to deeper and deeper problems and controversies.[54] If we did have a consistent and convincing knowledge of early Greek mathematics, there would be only the details to resolve and historians would have moved on to other fields. But this is manifestly not the case; it is still not possible to arrive at any general consensus even over some of the most fundamental issues.

The proposals set out here are only the beginning of a longer argument, since they have so far not taken into account the reactions of these later commentators whose possible misunderstandings can, I believe, be retrospectively understood and, to some extent, explained. To introduce my brief discussion here of this, which will be based on an analogy with more recent times, I first propose an experiment to my reader. I have tried, in the previous chapter, to indicate how the study of continued fractions, our version of anthyphairetic ratio theory, runs broad and deep through our mathematical heritage since the Renaissance, how some of the most influential and celebrated mathematicians have drawn inspiration from, and made contributions to, the theory, and how its properties lie behind some well-known and widely studied parts of mathematics today. But take down the fattest general history of mathematics from your shelf and look up 'continued fractions', 'Euclidean algorithm', or any other such synonym in the index; go to your nearest specialised mathematics library and search there; ask your neighbourhood mathematician for information about the subject; and so on. See if any of these sources give any hint of any of these developments. You will, I anticipate, be unlikely to find much beyond a general discussion of Pell's equation. Why do our commentators and compilers today make so little reference to this important theme in the modern history of the subject? There have been, since the seventeenth century, increasingly coherent communities of professional mathematicians and historians, drawn together through increasingly elaborate networks of personal and professional publications. So do we have any good reason to expect that the much more isolated ancient commentators and compilers could have done better?

It is a commonplace of the contemporary mathematical community that mathematical research fills out a broad spectrum. At one end are the informal and ephemeral activities of lectures, letters, and conversations, where underlying motivation is discussed, promising problems are formulated with suggestions for solution, blind alleys are described, and on, and where communication is carried out quickly and efficiently in what is often an inchoate, fragmentary, ambiguous, improvised language. The other end of the spectrum is dominated by the published tradition. Here much of the accompanying matter of the informal tradition is suppressed, and only successful investigations are presented, however irrelevant they may be to the original motivating problems.

---

[54] The discussion, in Section 3.6(b) and Chapter 8 above, gives an indication of some of the difficulties with this material, and the Epilogue, Chapter 11 below, reflects on some of these differences.

Few unsuccessful investigations, however relevant, get published because, quite simply, what can one say beyond that one has not succeeded?[55] The published works are set out in a very formal style, often involving a highly developed symbolic language that may be accessible only to those other mathematicians who specialise in the immediate topic. It is then only possible to get behind this published presentation and uncover with certainty the unexpected deeper motivations and the unpublished record of failed investigations when we can back up this published record with the conversations, letters, working papers, and rubbish baskets that belong to the informal tradition. Moreover, published mathematics can remain unread or unassimilated, even when it comes from the pen of a master. There are few who have the time, breadth of interest, and ability to work through the eighty-nine volumes of Euler's collected works, or the insight to understand Riemann's single volume, or the patience and persistence to read *Elements* X.

Such is the situation today and I see little reason to expect anything different at the dawn of this kind of mathematics; surely no one can believe that there was not some rich prehistory to what we find in Euclid's *Elements*. But nothing of what was then written has survived in its original form: still less can we, or could the later commentators, know of the conversations in the garden of Academe and elsewhere. In the case of early Greek mathematics, the important informal tradition has been completely lost.

Within the *Elements* we find ourselves facing a further barrier. No mathematician I can conceive talks like this:

If a straight line be bisected and a straight line be added to it in a straight line, the rectangle contained by the whole with the added straight line and the straight line together with the square on the half is equal to the square on the straight line made up of the half and the added straight line [*Elements* II 6],

though, unfortunately, many write thus, confident thereby of their place in an ancient tradition. The convoluted language is necessary here because formal Euclidean style does not allow any reference to any specific figure in its enunciation.[56] But many of the explorations and algorithms I have described in Part One simply cannot be expressed within the constraints of this formal

---

[55] Occasionally, however, an unsuccessful investigation may succeed in diverting mathematical activity away from the problem in hand. Poincaré's proof of the impossibility of one kind of solution of the three-body problem diverted attention away from the 2300-year-old mathematical study of the movement of sun and moon and towards topology; and I have proposed in Section 9.2(d), above, that Gauss's explorations of continued fractions in number theory, of which to my knowledge no trace survives, may have led him to develop further and promote the theory of quadratic forms.

[56] The formal style of the *Elements*, in particular the structure of a Euclidean proposition, apparently as strict as a piece of music in the strictest of classical styles and as often broken in practice, is described in Heath, *TBEE* i, Chapter 9, 114–51, and Mueller, *PMDSEE*, Chapter 1, 1–57. Here, again, our main source is Proclus, *Commentary on the First Book of Euclid's Elements*, 203–5. See the discussion of one aspect of it in the previous Section 10.4, and the drastic revision of the usual attitude to it in Netz, *PDMP*.

Euclidean style; see $S_{27}$, in Section 1.3. Moreover, while these algorithms and explorations play an essential role in the understanding óf anthyphairetic ratio theory, they have no part in the associated deductive theory. Observe, for example, how in Chapters 2 to 5 I have described and used in a fundamental way, but have not proved, any of the surprising and important behaviour that we observe in the *Parmenides* algorithm: that its run-lengths do in fact generate the anthyphairesis, that its behaviour will in fact be periodic when applied to $\sqrt{n} : \sqrt{m}$, that it will in fact always generate solutions to Pell's equation, and so on.[57] Here again, it is not easy to see how to give a deductive proof of these results without using the machinery of symbolic algebra and complete induction, let alone to see how this, also, could be expressed in formal Euclidean style. This then becomes the kind of material that belongs to the mathematician's craft, the everyday skills that are learnt orally, which are an essential part of his trade, but which often die along with the community that is investigating a particular corner of mathematics.

There is a further characteristic of continued fractions that makes them all the more prone to periods of fascination followed by neglect. On the one hand, the route via continued fractions is one of the quickest into the mathematical way of thinking. The *Parmenides* algorithm applied to $\sqrt{n} : \sqrt{m}$—or any of its equivalent manifestations—leads directly into subtle, surprising, and useful mathematics whose exploration will initiate and train a gifted novice. Further exploration will reveal rich further developments, and promise of continuing spectacular progress. But continued fractions are also the quickest route to a second mathematical phenomenon: the unsolved problem and, worse, the unfruitful unsolved problem. Mathematics can be a gratuitous activity and mathematicians are often wilful, as Plato complained at *Republic* VII, 528b–c.

---

[57] This is exactly the kind of criticism that Lagrange levelled against earlier treatments of Pell's equation and continued fractions. See, for example, the introduction to his SPA: "Given any non-square positive integer, find a square integer such that the product of these two numbers, increased by one, will be a square number. This is one of the problems that Mr. Fermat proposed, as a kind of challenge, to English mathematicians, in particular to Mr. Wallis who was the only one, to my knowledge, who resolved it or at least who published its solution. (See Chapter 98 of his *Algebra* and letters 17 and 19 of his *Commercium Epistolicum*.) But his method only consists of a kind of groping (*tâtonnement*) by which he arrives at the solution only in an uncertain manner, without knowing that he will actually ever get there. However, one should above all show that the solution of the problem is always possible, whatever the given number, a proposition that is generally regarded as true but which has to my knowledge not yet been firmly and rigorously proved. It is true that Mr. Wallis claims to have proved it, but by an argument that mathematicians find very unsatisfactory and which to me seems to beg the question (see Chapter 19 of his *Algebra* [and, in Section 9.2(d), above, Gauss's note *e* to the *Disquisitiones Arithmeticae*, §202]). Thus the problem has not yet been satisfactorily resolved in a way that leaves nothing to be desired. This is what determined me to make it the object of my study, the more so since the solution of this problem is like a key to all other problems of this type." (The original French is quoted in my BTEE2, 205 n. 18.) In fact many of the explorations before Lagrange were in the nature of what I have called 'heuristics' but which Lagrange characterised, here and elsewhere, by the very Lagrangian word '*tâtonnement*'

When the scent of an unsolved problem goes stale, when a new and promising point of view is revealed elsewhere, they will move off in another direction. The general ignorance of and little progress made in understanding the continued fraction expansion of higher-degree algebraic numbers, or higher-dimensional continued fractions, since the time of Gauss, despite the concentrated effort of a few maverick mathematicians, give us clear illustrations of stale problems and mathematicians' reactions to them. I have proposed here that analogous problems could also have arisen in the Greek context and they would, by the same token, have gone stale there too. In a process that is found many times throughout mathematics, the original problems themselves then disappear but the language that has developed in the pursuit of their solution lives on, transformed into the status of a theory. Many of the great theories of modern mathematics arose in this way, and I suggest that the same may have happened with early Greek mathematics, with geometry and *arithmētikē*.

So far in this final section I have considered only anthyphairetic ratio theory; but similar comments apply with even more force to the astronomical and musical ratio theories whose reconstruction I sketched in Chapter 4. Astronomical ratio theory, as I described it in Section 4.4, is so subtle that it scarcely gets beyond the stage of a definition and initial exploration before it peters out; though, as I suggested at $E_{57}$ and developed in my EPP, it may not have disappeared but survived, transformed, as the proportion theory of *Elements* V. And the bridge between musical ratio theory and the other ratio theories— namely, arithmetic with ratios—is so well hidden in the mathematical jungle that my proposed investigation could never have got beyond a few, faltering steps.

In addition to these general characteristics of the mathematical context, there is a further specific historical development that would compound the possibility of a later misunderstanding of any early Greek work on ratio theories. I have here throughout insisted on the non-arithmetised character of early Greek mathematics. We have abundant evidence of an earlier, profoundly arithmetical, Babylonian mathematics, but our historical record so far shows no explicit trace of any influence of Babylonian techniques on *early* Greek mathematics or astronomy.[58] Then, in the second century BC, with Hipparchus and Hypsicles, we find the first Greek usage of sexagesimal fractions; and thereafter Greek mathematics and science, as seen in the works of Ptolemy, Heron, Diophantus, and the later commentators, becomes a powerful mixture of the arithmetical and geometrical points of view. Now the attitudes and problems associated with different kinds of non-arithmetised and arithmetised mathematics can sometimes be very different, even incompatible; and this is particularly so for the case of ratio theories. Arithmetised ratio theory makes a sharp distinction between commensurable (our 'rational') and

---

[58] This is a typical example of a fundamental issue on which there is no general consensus. For a brief summary of the different positions, see Berggren, HGM, 397–8.

incommensurable (our 'irrational') ratios, the former being amenable to basic mathematical techniques, the latter requiring a much more subtle treatment; then when the ratios are conceived as sexagesimal or decimal numbers, there are other kinds of distinctions, such as whether the ratios terminate or not, or are periodic or not. To be sure, the dichotomy between commensurable and incommensurable also manifests itself in other ratio theories, for example as terminating or non-terminating anthyphairetic ratios, but this distinction is much less striking than the difference between anthyphairetic ratios of the form $\sqrt{n}:\sqrt{m}$ (Euclid's *rhētos*) and most of the rest (Euclid's *alogos*), a difference which lies at the heart of my reconstruction here. Reciprocally, the difference between $\sqrt{2}:1$ and $\sqrt[3]{2}:1$ is much less striking within an arithmetised ratio theory than an anthyphairetic theory. Or, for astronomical ratios, commensurability will manifest itself as a periodicity, with or without coincidences, but it is difficult, without very careful analysis, to comprehend the general commensurable astronomical ratio; here the understandable astronomical ratios are practically confined to multiple, epimoric, and multiple epimoric ratios.[59] Now arithmetised mathematics leads to a kind of algebraic reasoning that is absent from non-arithmetised mathematics, since the arithmetised operations which are then generalised and abstracted to give this algebra either are completely absent, or have a very special character, specific to each context.[60] So it may become difficult to appreciate the motivation, strengths, and weaknesses of one approach from within the context of another. For example, the ratio of the diagonal to side of a square poses a serious problem of understanding to arithmetised ratio theory; in the non-arithmetised anthyphairetic ratio theory, once an appropriate figure has been drawn, it becomes a basic and vivid example which marks an entrance into the theory. *Elements* V, Definition 5, posed a conundrum in arithmetised mathematics that was not really understood before Dedekind, but it corresponds to the first insight into the nature of astronomical ratios. Algebraic manipulations are abstractions from arithmetic; but the problem of describing the basic arithmetical operations within anthyphairetic ratio theory has only recently been solved, and I know of no such algorithms for astronomical ratio theory. Now extend these specific examples to a more global point of view: if, as I have proposed, early Greek mathematics was non-arithmetised, then we may not be able to understand or translate its motivations and some of its methods into our arithmetised point of view. So there would be a tendency for it to be misunderstood from the second century BC onwards.

So while I can understand a sceptical reader remaining worried by the lack of any clear and explicit anthyphairetic explanation of early Greek mathematics in the last two thousand and more years, I cannot agree that this absence is fatal to my reconstruction. Most of what has been said on the topic in this period has

---

[59] Witness the complications of the descriptions in $E_{67}$ and $B_{68}$ of Section 4.4(b).
[60] See Section 4.5(b), $A_{86-88}$.

been speculation, much of it has not been written by mathematicians, and most of it comes from within alien and incompatible mathematical traditions. However much it may tell us about the intellectual climate of the times at which it was written and of its author's ideas, most of it may be simply irrelevant to early Greek mathematics. I hope it is clear, too, that I am aware that however plausible and appealing my proposals here might be, they too may be irrelevant or wrong, like so many before them; though while I can see the possibility of quite drastic revisions to their details, I am not yet aware of any cogent arguments for rejecting them outright. And I do not think that one should confuse the exploration of the documented stages of some intellectual development, such as we have for some episodes in mathematics since the seventeenth century, with the kind of plausible rational reconstructions that have been attempted here. But, given the state of our evidence, no other approach to early Greek mathematics is possible.

I have dwelt very little on the historical and biographical aspects of early Greek mathematics here. Our sources give us very little reliable biographical information, even about matters as basic as chronology, and most of the details have to be reconstructed. In a subject with the logical coherence of mathematics, there is a tendency for the mathematical reconstruction to precede and determine the biographical reconstruction; for example, details about the discovery of incommensurability tend to affect reconstructions of pre-Socratic philosophy and mathematics, and the reconstructions of proportion theory have had an important part in the biographical reconstructions of the fifth and fourth centuries BC.[61] The drastic revision of the reconstruction of pre-Euclidean mathematics proposed here will provoke some revision of the historical picture, but I did not wish to complicate the issue by introducing these matters at this stage. However, I cannot resist finishing with a historical analogy. It is a convenient tendency, sometimes not without truth, to personify trends and locate movements in individuals. In this spirit, it is tempting to speculate that, just as I have suggested that Gauss rewrote number theory without continued fractions and thus may have prompted the decline in their study, so Eudoxus may have rewritten anthyphairetic and astronomical ratio theories as proportion theory with the same effect; and that both worked with a deep understanding of what they were omitting.

---

[61] See, for example, Knorr, *EEE*, Chapter 2, and de Santillana, EPSC.

# 11

## EPILOGUE: A BRIEF INTELLECTUAL AUTOBIOGRAPHY

> I've said it once, I've said it twice,
> What I tell you three times is true.
> *Lewis Carroll, The Taming of the Snark*[1]

Since the publication of 'Ratio in early Greek mathematics', my first article on Greek mathematics, the attitude towards these proposals among specialists has gradually changed from rank incredulity, from a refusal to begin even to take the first step towards considering the details of the interpretation, to a kind of distant respect.[2] People take anthyphairesis seriously,[3] and have now even begun to consider the possibility of the use of anthyphairesis as providing some way of thinking about ratio; but hardly any specialist has accepted that it could have played a role as precise and detailed as I have proposed. There is no real evidence for it, they say; it is far too elaborate a speculation to be acceptable, too much an importation of modern ideas into the past. Some of these objections have just been addressed, in Section 10.5; let me now, in a brief autobiographical sketch, describe my evolving attitude to these issues.

What changed my attitude towards early Greek mathematics from that of an amateur dilettante to someone obsessively gripped by the fascination of the topic was the material in the second of the two linked articles on *Elements* II: BTEE and BTEE2. We community of historians of Greek mathematics are still far from the beginnings of any agreement about the role of Book II, and some

---

[1] This is asserted twice in the poem, but not three times! This extract comes from the beginning, in Fit the First; then see Fit the Fifth: " 'Tis the note of the Jubjub! Keep count, I entreat. / You will find I have told you it twice. / 'Tis the song of the Jubjub! The proof is complete. / If only I've stated it thrice." It might also be said to describe the structure of a Euclidean proposition, with its three statements of what is to be proved and then has been proved, in its enunciation, its setting out, and its conclusion. (Lewis Carroll, under his own name Charles Dodgson, wrote several pamphlets on Euclid's *Elements* for the use of undergraduates; see Dodgson–Abeles, *MPCLD*.)

[2] One notable exception was Berggren, HGM, which gave a reasonable précis of the aims of the interpretation but ended with a tone of what I take to be a muted but qualified scepticism. It would be a very useful contribution if this general survey could now be brought up to date.

[3] The pioneer here was Becker, in his ESI.

of the disagreements have been acrimonious.[4] Many of those who have studied this topic have proposed some role or another for Book II, though usually for some but not all of the propositions, and this is just what I did in the first of these two articles. It gave some kind of context for everything except II 12 and 13, which, however, were easy to explain away glibly in the traditional way: "Only 12 and 13 do not appear to be relevant to anthyphairesis, and these propositions are generally accepted as generalisations of Pythagoras' theorem in their own right" (BTEE, 22). They were just two more instances within Book II of propositions that are unused elsewhere within and beyond the *Elements*, but these particular ones are closely related to this central and important result.[5] Here, from the most thorough modern study of the details of *Elements*, is Mueller, PMDSEE, 301, on the same issue:

. . . II 8 is a good example of the problematic character of Book II as a whole. II 9, of which the unused II 10 is a counterpart, is cited only in the problematic *lēmma* before X 60; and II 12, of which the unused II 13 is a counterpart, provides a possible explanation (given by a scholiast) for a problematic inference in XII 17. However the presence and position of propositions 12 and 13 may be sufficiently explained by the fact that they are completions of the Pythagorean theorem, but are proved by methods appropriate to Book II,

and his discussion goes on to consider and reject the main explanation of parts of Book II as the Greek counterpart to our algebra, the so-called 'geometric algebra'. And not only are some of these propositions of Book II unused within the *Elements* itself; the minutely detailed study of the role of the *Elements* and other early texts by Pappus, started in Saito, IPPC, has so far not unearthed any use by Pappus of II 7, 11, 12, and 13, though his published investigations have so far dealt only with Book 7 of the *Collection*.

The geometry needed in this investigation of an important collection of examples of anthyphairetic ratios was Pythagoras' theorem,[6] together with the figures associated with the other propositions of Book II apart from 12 and 13. Nothing more was needed: these figures contained all of the basic ingredients of the successively more complicated figures used in the proofs of the quite astonishing periodic behaviour of the anthyphairesis of a substantial subset of the corpus of incommensurable ratios $\sqrt{n}:\sqrt{m}$. But these proofs used two methods that got progressively more difficult to apply and then became unworkable in the cases where the expansions got more complicated; see Chapter 3, above. For example, the first method was defeated by $\sqrt{19}:1$. It was disappointing that the problem was left incomplete, but that is a common

[4] These rival interpretations are something that I have not mentioned in this book and do not intend to mention here. For sober discussions of some of the issues, see Mueller, *PMDSEE*, 41–5 & 301–2, Artmann, EEP, and, again, Berggren, HGM.

[5] In our terms today, they are equivalent to the generalisation of the so-called Pythagoras' theorem to the cosine rule for oblique- and acute-angled triangles.

[6] It can plausibly be argued that there was originally a version of this between Propositions 8 and 9; see the end of Section 3.2, above.

feature of mathematical explorations. But I then discovered, and described in BTEE2, how II 12 and 13 could play an absolutely central and essential role within a proposed interpretation of the successful continuation of this programme, in a generalisation of my interpretation of the geometrical side-and-diagonal configuration. To begin with, I treated this as a mathematical curiosity but, after further explorations, I became more and more persuaded that it could have some plausible historical relevance. This was yet one more example of the way that this programme of the exploration of early ratio theories[7] provided a coherence across a wide range of our evidence. It was also another example of my evolving interpretive principle that homogeneous slabs of evidence should be taken *as a whole* (see the new appendix, Section 10.1(b), for a description of this). Here, in this book, the main such slabs are Plato's curriculum, *Elements* II, and then the larger complex of *Elements* II, IV, X, and XIII (see Chapters 3, 4, and 5, which make up the central part of the interpretation). Another methodological principle is that of *difficilior lectio*, that the more obscure and awkward difficult versions of some bit of evidence should not be ignored or emended, but fitted if possible into the interpretation, for sometimes these carry the original intention that editors have been unable to understand and therefore have not smoothed out. For example, there is the awkward definition of synthesis to be found in Book XIII of the *Elements*, but consigned to an Appendix by Heiberg and so treated by Heath only in his notes; see *TBEE* iii, 442, where he comments: "There must apparently be some corruption in the text; it does not, in the case of synthesis, give what is wanted." This 'corrupt' Euclidean version of the text fits with the development of my anthyphairetic interpretation of Book II, and is discussed in BTEE, 19–22.[8]

Our knowledge of the details of pre-Euclidean mathematics is so sketchy that the various interpretations are more like religious beliefs than positions that can be defended in ways that go beyond faith; and the biggest influence on religious inclination tends to be the environment in which one grows up. To expect others to be persuaded into another interpretation is to expect something like a religious conversion, a tall order indeed. I have never in any serious way, beyond publications and talks at academic meetings, tried to create such an environment in which to nurture young and unformed talent to my point of

---

[7] I emphasise the plural as there are several examples of ratio theories described here; see, for example, the discussion in Chapter 4 of the different parts of Plato's curriculum. Some seem to think that anthyphairesis generates the only proposed ratio theory.

[8] I described briefly another way in which emendation can be used to create historical evidence in Section 10.1(c)(iv), but did not develop this beyond a bit of moralising. Another example of difficult text that is therefore often written off as spurious or misreported is Archytas Fragment B4 (see Section 4.6(b)), but there is no point in my discussing this particular text further until my interpretation of *logistikē* has had further examination. This topic has been discussed in Mueller, ME, who rejects my point of view; he sets the issue in the context of school education in mathematics, but approaches it from the opposite direction from that I take: he considers the literary evidence, I look at our material evidence from the schoolroom. Both are unlikely to give us more than a hint of what actually went on.

view; I have declined to take on graduate students on this topic, or have put almost impossible conditions on their doing so. And even a successful graduate of these studies is virtually unemployable in the late twentieth-century academic world. I grew up an innocent (but an innocent already in secure employment!), with little knowledge and no real beliefs concerning Greek mathematics—my historical taste originally lay in nineteenth- and early twentieth-century analysis—and it was reading Wilbur Knorr's *The Evolution of the Euclidean Elements*, jumping in at the deep end, that kindled my latent interest. My original project was to provide a more accessible account of the material in Knorr's book, directed towards mathematicians who were, as I had been, on the whole uninformed and uninterested in the subject beyond using it to spice their lectures with stereotyped and unreliable stories about Pythagoras and Euclid. It was while working on this expository project that I had my revelation early one morning in May 1978:[9] if anthyphairesis was used in the context of a pre-Euclidean proof of *Elements* VI 1, as Aristotle tells us at *Topics* 158b 29 that it was, in our *only* explicit evidence about this important issue, then was it used to define *proportion*, as seems almost universally believed by those who are inclined to assign any role to it,[10] or could it have been used to define *ratio*? That is, the pattern of the anthyphairesis might be worth exploring within the context of early Greek mathematics, since either interpretation fitted within Aristotle's testimony. This new point of view led to the first article, REGM, which was, in effect, a programme for further exploration; BTEE was the first of a projected series of articles on the details of the programme; and the programme was eventually extended and a substantial part of it carried through in the first edition of this book.

By the time of my heightened commitment to this interpretation that arose from the material of BTEE2, I had gained a little more historical maturity and had come to realise that, while historical red herrings may be easy to create, they may also be impossible to extirpate, especially if they are appealing, attractive, and fill some needed slot; the stories about Pythagoras and Euclid are fine examples of this. The anthyphairetic interpretation is so attractive and appealing that my responsibility was to investigate the historical context in much greater detail, and this involved exploring several other fields that were completely new to me.

Arithmetic, like religion, is also strongly determined by the environment in which one grows up. As soon as a society develops some kind of bureaucratic complexity, arithmetic plays an important role in everybody's life, as many commodities of everyday use and importance acquire a numerical value, taxes are calculated numerically, etc., and practical procedures for handling ratios

[9] I wish I kept a diary and so could locate the exact day. The time was around 5 am, and I was lying in bed!

[10] Many scholars, after initial enthusiasm, seem to distance themselves from this kind of interpretation as it fails to live up to its initial promise.

and proportions are developed.[11] So what was the arithmetical base of Greek society?—for that might influence the way mathematicians, even sophisticated mathematicians, might think about numbers and fractions.[12] The common, usually unconscious, assumption on this issue seems to be that the Greeks, being an intelligent society, calculated using our system of common fractions, our way of describing the rational numbers; by contrast the Egyptians, being less concerned with mathematical sophistication, used the cumbersome and unhelpful 'unit fractions'.[13] This view, though only occasionally articulated, is ubiquitous. For example, it lies behind the assertion that Pythagoras, or the Pythagoreans, or someone else, discovered that the 'rational' numbers were insufficient for geometry, that the diagonal of a square is 'irrational', and so directed mathematics away from arithmetic and towards geometry.

The conclusions of these investigations into school arithmetic are described

[11] The 'rule of three'—if eight cows cost a farmer twelve pounds, then how much would six cows cost?—a staple of the schoolroom since medieval times and earlier, is a good illustration of this.

[12] A more recent illustration of this kind of phenomenon may well be found in the introduction of decimal fractions into the West at the end of the sixteenth century. The first treatment setting out the details in a straightforward way was not addressed to specialists: Simon Stevin's *De Thiende* of 1585 was published in the vernacular (in Dutch; then later in the same year in French; and then in an English translation by Richard Norton in 1608) and was addressed to "Astronomers, Land-meters, Measurers of Tapestry, Gaugers, Stereometers in general, Money-Masters, and to all Merchants" (see Stevin–Struik, *PW*, vol. ii, or many other places). Decimal fractions gave a new fluency to arithmetic which permitted, perhaps for the first time, the feelings that all such calculations could now be taken for granted, and this paved the way to the next stage, their abstraction into symbolic algebra. There was an explosive development of these algebraic manipulations that was complete by the time of Descartes' *La Géométrie* in 1637, and it is a remarkable coincidence, if coincidence it be, that the people who contributed to this development—principally Stevin himself, Viete, and Harriot—were themselves calculators. (One can see the same kind of evolution later in the young Newton's annotations and arithmetical explorations of Wallis' *Arithmetica Infinitorum*, and the way his theory of infinite series is directly based on the procedures of decimal manipulations.) This confidence in decimal arithmetic still lies, I believe, behind our basic intuitions, even today, underlying the real numbers, even though it is a delusion: there is no algorithm for generating the digits of an arithmetical operation on two decimal numbers, even for the decimal expansions of rational numbers. For more comments, see the references to Dedekind in the Index.

[13] A most misleading name, for there is no 'unit' in their basic fractional sequence, which is more like 'the half', 'the third', 'the quarter', . . . , what I refer to, in translation of their Greek name *merē*, as 'parts'. And see van der Waerden, quoted in Section 4.5(c) above, for what I believe to be the kind of assumption described here. An even more explicit instance can be found in Toomer, MA, 45: "The truth is that Egyptian mathematics remained at too low a level to be able to contribute anything of value. The sheer difficulties of calculation with such a crude numerical system and primitive methods effectively prevented any advance or interest in developing the science for its own sake. It served the needs of everyday life (it is only a relatively advanced technology, such as was never achieved in the ancient world, which demands more than the most elementary mathematics), and that was enough. Its interest for us lies in its primitive character, and what it reveals about the minds of its creators and users, rather than its historical influence." This was written more than 25 years ago and may not represent its author's later point of view; it is given here only as an illustration of an attitude to the Egyptian fractional system.

in a series of articles (NFA, TP, HP, LFGM, FAT), and brought together in Chapter 7 of this book. Our evidence comes from the wrong time and place, but is unambiguous: Greek school arithmetic was exactly the same as Egyptian arithmetic. In this representation of fractional quantities, every basic arithmetical operation—addition, subtraction, multiplication, and division—is far from routine and automatic, and the abstraction of fractional quantities into something approaching our idea of the field of rational numbers is simply out of the question in that historical context. These assumptions about, and sometimes unconscious applications of, our own arithmetical intuitions to the past may be very misleading. Another topic involving the role of arithmetic is the postulated early influence of Old Babylonian arithmetical techniques on early Greek mathematics. There is little about this in this book since we have so far absolutely no direct evidence that this happened; it is another example of a purely speculative proposal, attractive to our point of view but implausible within the Greek context, such as we know it.

Another issue that needed exploration was the status of the material evidence that we possess—just what are our texts and how credible are our stories about the historical figures involved? This was fascinating; not only did I once again find myself in a totally new field but, once again, I enjoyed the unstinted and generous help of classicists and papyrologists. Here the conclusion seems to be that the texts, once they pass from the hands of their last editors, seem to be transmitted reasonably accurately, but we have almost no control over the period from their composition up to that last editorial stage, nor can we be confident that many relevant texts have not been lost in the winnowing entailed by the process of copying and recopying.[14] As to the stories about historical figures such as Pythagoras, Plato, Aristotle, Archimedes, etc., most of them belong to a tradition that grows and grows with time;[15] often a nugget of solid information may be carried within that tradition, though identifying that nugget can be problematic.

I returned from the exploration of these historical issues with even more confidence in the plausibility of the overall nature of my proposals, a realisation that their scope may be even wider than I had thought, and the feeling that they were no more speculative than many parts of the received interpretations—indeed that perhaps they were often less so. This gave me the resolve to bring a substantial part of them together into a whole. The first edition of this book was the outcome.

[14] One outstanding example of something that has disappeared at some unknown time is Eudemus' *History of Mathematics*. Even Proclus may not have had a copy, for it is not unlikely that his quotations came via Geminus, though it might be difficult to establish this with any confidence; in any case, no Byzantine copy is known, not even at second hand, and I do not think that any Arabic translation is known, though there are still very many unexplored Arabic texts.

[15] Detailed examinations of this 'expansion of the tradition' of Pythagoras are given in Burkert, *LSAP*, of Plato in Riginos, *PACLWP*, and of Archimedes in the first chapter of Dijksterhuis, *A*.

# BIBLIOGRAPHY

The general principles of the referencing system by acronym have been
described in the Introduction. To these can be added: authors are listed
alphabetically by first upper case letter of family name when abbreviations
have been expanded. (Thus V. De Falco under D; G. E. M. de Ste. Croix under
S, between Saffrey and Sandys; and D. de Solla Price under P.) Items that have
also appeared conveniently in collected works are also cross-referenced to these
works, which are identified by a keyword (*Works, Werke, Opera*, etc.), not an
acronym; and collected works precede any references to items they contain. In
general, the main reference is to the most convenient modern edition or reprint,
and the information about original publication may be restricted to a paren-
thetic date immediately after the title. Items within an author heading are
arranged chronologically by first publication; single authors come before joint
authors. Occasional parenthetic comments give references to reviews or other
information.

This bibliography is restricted to works that have been cited in the text. For more
comprehensive bibliographies see, for the history of mathematics in general: Dauben,
*HMAP*; for ancient mathematics, in particular: Knorr, *EEE* 353–68, Procissi, BMGA,
and Thomas, *SIHGM* I (repr. 1980) pp. xvii–lii; on scientific biography: the *DSB*; on
Plato: Cherniss, P50–57, Brisson, P58–75, Brisson & Ioannidi, P75–80, P80–85, . . . ; and
for editions of classical texts and general topics in classical antiquity: the Introduction to
Liddell, Scott, & Jones, *GEL*, and the *OCD*.

Alexander of Aphrodisias–Wallies, *IAP* = *In Aristotelis Analyticorum Priorum Librum 1
Commentarium*, ed. M. Wallies, Commentaria in Aristotelem Graeca $2_1$, Berlin:
Hayduck, 1883.
Alexander of Aphrodisias–Wallies, *IT* = *In Aristotelis Topicorum Libros Octo Com-
mentaria*, ed. M. Wallies, Commentaria in Aristotelem Graeca $2_2$, Berlin: Hayduck,
1891.
Angeli, A. & Dorandi, T., PMDL = Il pensiero matematico di Demetrio Lacone,
*Cronache Ercolanesi* 17 (1987), 89–103.
Archimedes–Heath, *WA* = *The Works of Archimedes Edited in Modern Notation* (1987)
with supplement *The Method of Archimedes* (1912), tr. T. L. Heath, Cambridge:
Cambridge University Press, 1912, repr. New York: Dover, n.d.
Archimedes–Heiberg, *Opera* = *Opera Omnia cum Commentariis Eutocii*, ed. J. L.
Heiberg, 2nd edn. 3 vols., Leipzig: Teubner, 1910–15, repr. Stuttgart: Teubner,
1972.
Aristarchus–Heath, *AS* = *Aristarchus of Samos, The Ancient Copernicus*, ed. & tr. T. L.
Heath, Oxford: Clarendon Press, 1913, repr. New York: Dover, 1981.
Aristotle, Individual works with the Greek text and parallel English translation in the
Loeb Classical Library, Cambridge, Massachusetts & London: Harvard University
Press & Heinemann.

Aristotle, Critical editions of the Greek text of individual works in the Scriptorum Classicorum Bibliotheca Oxoniensis (= The Oxford Classical Texts), Oxford: Clarendon Press.

Aristotle–Barnes, *CW = The Complete Works of Aristotle*, ed. J. Barnes, 2 vols., Bollingen Series 71₂, Princeton: Princeton University Press, 1984. (A revised version of Aristotle–Ross, *WA*.)

Aristotle–Bekker, *Opera = Aristotelis Opera*, ed. I. Bekker, 5 vols., Berlin: Königliche Akademie der Wissenschaften, 1831–70, often reprinted.

Aristotle–Ross, *WA = The Works of Aristotle Translated into English*, ed. W. D. Ross, various translators, 12 vols., Oxford: Clarendon Press, 1908–52. (Also see Aristotle–Barnes, *CW*.)

Artmann, B., EEP = Euclid's *Elements* and its prehistory, in *ΠΕΡΙ ΤΩΝ ΜΑΘΗΜΑΤΩΝ*, ed. I Mueller, *Aperion* 24 (1991), 1–47.

Artmann, B., PTTDD = A proof for Theodorus' theorem by drawing diagrams, *Journal of Geometry* 49 (1994), 3–35.

Baker, A., RACTAN = Rational approximation to $\sqrt[3]{2}$ and other algebraic numbers, *Quarterly Journal of Mathematics* (2nd series) 15 (1964), 375–83.

Barker, A., MPSA = Music and perception: A study in Aristoxenus, *Journal of Hellenic Studies* 98 (1978), 9–16.

Barker, A., MAESC = Methods and aims in the Euclidean *Sectio Canonis*, *Journal of Hellenic Studies* 101 (1981), 1–16.

Barker, A., *GMW = Greek Musical Writings*, 2 vols: vol. i, *The Musician and his Art*; vol. ii, *Harmonic and Acoustic Theory*, Cambridge: Cambridge University Press, 1984 and 1989.

Becker, O., ESI = Eudoxos–Studien I. Eine voreudoxische Proportionenlehre und ihre Spuren bei Aristotles und Euklid, *Quellen und Studien zur Geschichte der Mathematik, Astronomie und Physik*, Abteilung B: *Studien* 2 (1993), 311–33.

Berggren, J. L., HGM = History of Greek mathematics, A survey of recent research, *Historia Mathematica* 11 (1984), 394–410.

Berstein, L., *JPA = The Jacobi–Perron Algorithm, Its Theory and Application*, Lecture Notes in Mathematics 207, Berlin, Heidelberg, New York: Springer, 1971.

Bilabel, F., *SGUA = Sammelbuch Griechischer Urkunden aus Ägypten*, ed. F. Bilabel & others, 13 vols. to 1979, Berlin & Leipzig: de Gruyter. (Cited by papyrologists under the abbreviation SB.)

Blanchard, A., *SAPDG = Sigles et Abreviations dans les Papyrus Documentaires Grecs: Recherches de Paleographie*, London: Bulletin of the Institute of Classical Studies Supplement 30, 1974.

Boethius–Friedlein, *IM = De Institutione Arithmetica & De Institutione Musica*, ed. G. Friedlein, Leipzig: Teubner, 1867.

Bonitz, H., *IA = Index Aristotelicus*, Berlin: Reimer, 1831, often reprinted. (= vol. 5 of Aristotle–Bekker, *Opera*.)

Bos, H. J. M., AMRDMT = Arguments on motivation in the rise and decline of a mathematical theory; the "Construction of Equations", 1637—ca. 1750, *Archive for History of Exact Sciences* 30 (1984), 331–80.

Boter, G. M., PM = Plato, *Meno* 82c–3, *Phronesis* 33 (1988), 208–15.

Bowen, A. C., FEPHS = The foundations of early Pythagorean harmonic science: Archytas, fragment 1, *Ancient Philosophy* 2 (1982), 79–104.

Bowen, A. C. & Goldstein, B. R., MAALFC = Meton of Athens and astronomy in the late fifth century BC, pp. 39–81 in *A Scientific Humanist: Studies in Memory of Abraham Sachs*, ed. E. Leichty, M. de J. Ellis, & P Gerhardi; Philadelphia, 1988.

Brandwood, L., *WIP = A Word Index to Plato*, Leeds: W. S. Maney, 1976.

Brashear, W., CTAL = Corrections à des tablettes arithmétiques du Louvre, *Revue des Études Grecques* 97 (1984), 214–17.

Brashear, W., MS = The myrias-symbol in CPR vii 8, *Zeitschrift für Papyrologie und Epigraphik* 60 (1985), 239–42.

Brashear, W., VNTAB = Vier neue Texte zum antiken Bildungswesen, *Archiv für Papyrusforschung* 40 (1994), 29–35 and Plate 16.

Brezinski, C., LHCF = The long history of continued fractions and Padé approximants, pp. 1–27 in *Padé Approximation and its Applications, Amsterdam 1980*, ed. M. G. de Bruin & H. van Rossum, Lecture Notes in Mathematics 888, Berlin, Heidelberg, New York: Springer, 1981.

Brenzinski, C., *HCF = History of Continued Fractions and Padé Approximants*, Berlin, Heidelberg, New York: Springer, 1991.

Brisson, L., P58–75 = Platon 1958–1975, *Lustrum* 20 (1977), 5–304.

Brisson, L. & Ioannidi, H., P75–80 = Platon 1975–1980, *Lustrum* 25 (1983), 31–320.

Brisson, L. & Ioannidi, H., P80–85 = Platon 1980–1985, *Lustrum* 30 (1988), 11–294. Corrigenda, *ibid.*, 31(1989), 270–2.

Brisson, L. & Ioannidi, H., P85–90 = Platon 1985–1990, *Lustrum* 34 (1992), 7–330, and continuing.

Brown, M., PDSA = Plato disapproves of the slave-boy's answer. *The Review of Metaphysics* 20 (1967), 57–93, repr. in Brown, *PM*, 198–242.

Brown, M., TKCL = *Theaetetus:* Knowledge as continued learning, *Journal of the History of Philosophy* 7 (1969), 359–79.

Brown, M., *PM = Plato's Meno*, tr. W. K. C. Guthrie, with essays, ed. M. Brown, Indianapolis: Bobbs-Merrill, 1971.

Brun, V., GK = En generalisation av kjedebroken I & II, *Norske Videnkapsselskapets Skifter I, Matematisk-Naturvidenskapelig Klasse* 6 (1919), 1–29 & 6 (1920), 1–24.

Brun, V., AETQN = Algorithms euclidiens pur trois et quatre nombres, *Treizième congres des mathematiciens scandinaves, Helsinki* (1958), 46–64.

Bulmer-Thomas [= Thomas], I., PA = Plato's Astronomy, *Classical Quarterly* 34 (1984), 107–12.

Burkert, W., SSS = *ΣTOIXEION*, Eine semasiolgische Studie, *Philologus* 103 (1959), 167–97.

Burket, W., *LSAP = Lore and Science in Ancient Pythagoreanism*, tr. E. L. Minar of *Weisheit und Wissenschaft: Studien zu Pythagoras, Philoloas und Platon* (1962), Cambridge, Massachusetts: Harvard University Press, 1972.

Burnyeat, M. F., MSPD = The material and sources of Plato's dream, *Phronesis* 15 (1970), 101–22.

Burnyeat, M. F., PSTM = The philosophical sense of Theaetetus' mathematics, *Isis* 69 (1978), 489–513. (Also see Knorr, MPP.)

Butler, A. J., *ACE = The Arab Conquest of Egypt and the Last Thirty Years of the Roman Dominion* (1902), 2nd edn. ed. P. M. Fraser, Oxford: Clarendon Press, 1978.

Cameron, A., LDAA = The last days of the Academy in Athens, *Proceedings of the Cambridge Philogical Society* 95 (New Series 15) (1969), 7–29.

Cavallo, G., LSSE = Libri scritture scribe a Ercolano, *Cronache Ercolanesi* 13, Supplemento 1, 1983.

Caveing, M., DZV = The debate between H. G. Zeuthen and H. Vogt (1909–1925) on the historical sources of the knowledge of irrational quantities, *Centaurus* 38 (1996), 277–92.

Černý, J., *PBAE = Paper and Books in Ancient Egypt* (An inaugural lecture, May, 1947), London: University College & H. K. Lewis, 1952, repr. Chicago: Ares, 1977.

Chace, A. B., *RMP = The Rhind mathematical Papyrus*, 2 vols., Oberlin, Ohio: Mathematical Association of America, 1927–29, abridged repr. in 1 vol., Classics in Mathematics Education 8, Reston, Virginia: The National Council for Teachers of Mathematics, 1979.

Cherniss, H., *REA = The Riddle of the Early Academy*, Berkeley: University of California Press, 1945, repr. with index by L. Tarán, New York & London: Garland, 1980.

Cherniss, H., P50–57 = Plato 1950–1957, *Lustrum* 4 (1959), 5–308 & 5 (1960), 321–656.

Cherniss, H., *Papers = Selected Papers*, ed. L. Tarán, Leiden: Brill, 1977.

Cherniss, H., PM = Plato as mathematician, *The Review of Metaphysics 4* (1951), 395–425, repr. in *Papers*, 222–52.

Christoffel, E. B., OA = Observatio arithmetica, *Annali di Mathematica* (Series 2) 6 (1875), 148–52.

Chrystal, G., *TA = A Textbook of Algebra* (1886), 7th edn. 2 vols., repr. New York: Chelsea, 1964.

Chuquet–Flegg, Hay, & Moss, *NC = Nicholas Chuquet, Renaissance Mathematician, A Study with Extensive Translation of Chuquet's Mathematical Manuscript Completed in 1484* [= *Le Triparty en la Science des Nombres*], ed. G. Flegg, C. Hay, & B. Moss, Dordrecht: Reidel, 1985.

Churchhouse, R. F. & Muir, S. T. E., CFANMI = Continued fractions, algebraic numbers, and modular invariants, *Journal of the Institute of Applied Mathematics* 5 (1969), 318–28.

Clagett, M., MLT = The medieval Latin translations from the Arabic of the *Elements* of Euclid, with special emphasis on the versions of Adelard of Bath, *Isis* 44 (1953), 16–42.

Claggett, M., *AMA = Archimedes in the Middle Ages*, 5 vols. in 10 parts, Philadelphia: American Philosophical Society, 1964–84.

Cockle, W. E. H., RCP = Restoring and conserving papyri, *Bulletin of the Institute of Classical Studies* 30 (1983), 147–65.

Coldstream, J. N., *GG = Geometric Greece*, London: Benn, 1977.

Connelly, R., FS = Flexing surfaces, pp. 79–89 in *The Mathematical Gardner*, ed. D. A. Klarner, Boston: Prindle, Webster, & Schmidt and Belmont, California: Wadsworth, 1981.

*Corpus dei papiri filosofici greci e latini: testi e lessico nei papiri di cultura greca e latini,* 3 vols, Unione Accademica Nazionale, Accademia Toscana di Scienze e Lettere "La Colombaria", Florence: Olschki, Vol. 1 in 2 parts, 1989; Vol. 2, 1993; . . .

Cotes, R. L., L = Logometria, *Philosophical Transactions of the Royal Society* 29 (1714), 5–47.

Cotes, R., *HM = Harmonia mensurarum, sive Analysis et Synthesis per Rationum et Angulorum Mensuras Promotae: Accedent alia Opuscula Mathematica per Rogerum Cotesium*, London, 1722.

Coulton, J., TUGTD = Towards understanding Greek temple design: General considerations, *Annual of the British School at Athens* 70 (1975), 59–99; corrigenda, ibid. 71 (1976), 149–50.

Coxeter, H. S. M., *IG = Introduction to Geometry*, New York: Wiley, 1961.

Crawford, D. J., *KEVPD = Kerkeosiris: An Egyptian Village in the Ptolemaic Period*, Cambridge: Cambridge University Press, 1971.

Crawford, D. S., MT = A mathematical table [= P. Michael. 62], *Aegyptus* 33 (1953), 222–40.

Curchin, L. & Fischler, R., HANTDEMR = Hero of Alexandria's numerical treatment of division in extreme and mean ratio and its implications, *Phoenix* 35 (1981), 129–33.

Curchin, L. & Herz-Fischler, R., DQDPR = De quand date le premier rapprochement entre la suite de Fibonacci et la division en extrême et moyenne raison?, *Centaurus* 28 (1985), 129–38.

Dauben, J. W., *HMAP = The History of Mathematics from Antiquity to the Present, A Selective Bibliography*, New York & London: Garland, 1985. (An updated CD version of this is in preparation.)

Davenport, H., *HA = The Higher Arithmetic, An Introduction to the Theory of Numbers* (1952), 5th edn., Cambridge: Cambridge University Press, 1982.

Davenport, H. & Roth, K. F., RAAN = Rational approximations to algebraic numbers, *Mathematika* 2 (1955), 160–67.

Davy, H., SOEPFRH = Some observations and experiments on the papyri found in the ruins of Herculaneum, *Philosophical Transactions of the Royal Society* 111 (1821), 191–208.

Dedekind, R., *ETN = Essays on the Theory of Numbers: I. Continuity and Irrational Numbers [Stetigkeit und Irrationale Zahlen (1872)], II. The Nature and Meaning of Numbers [Was sind und was sollen die Zahlen (1888)]*, tr. W. W. Beman, Chicago: Open Court, 1901, repr. New York: Dover, 1963.

De Falco, V., *EDL = L'Epicureo Demetrio Lacone*, Naples: Achille Cimmarua, 1923.

Delone, B. N. & Faddeev, D. K., *TITD = The Theory of Irrationalities of the Third Degree*, tr. E. Lehmer & S. A. Walker of *Teoriya Irratsional'nostei Tretéi Stepeni* (1940), Translations of Mathematical Monographs 10, Providence: American Mathematical Society, 1964.

Dicks, D. R., *GFH = The Geographical Fragments of Hipparchus*, London: University of London Press, 1960.

Dickson, L. E., *HTN = History of the Theory of Numbers*, 3 vols., Washington, D.C.: Carnegie Insitute 1919–23, repr. New York: Chelsea, 1971.

Diels, H. & Kranz, W., *FV = Die Fragmente der Vorsokratiker*, 6th edn., Zurich: Weidmann, 1951–2. (Also see Freeman, *APSP*.)

Dijkstershuis, E. J., *A = Archimedes,* Dutch edn., 1938; English tr. C. Dikshoorn, Copenhagen: Munksgaard, 1956; paperback edition, with a bibliographic essay 'Archimedes after Dijksterhuis: A guide to recent studies' by W. R. Knorr, Princeton University Press, 1987.

Diogenes Laertius–Hicks, *LEP = Lives of the Eminent Philosophers*, tr. R. D. Hicks, Loeb Classical Library, Cambridge, Massachusetts & London: Harvard University Press & Heinemann, 1925, rev. 1938.

Diophantus–Heath, *DA = Diophantus of Alexandria: A Study in the History of Greek Algebra*, ed. & tr. T. L. Heath, 2nd edn., Cambridge: Cambridge University Press, 1910.

Diophantus–Sesiano, *BFSDA = Books IV to VII of Diophantus' Arithmetica in the Arabic Translation Attributed to Qusṭā ibn Lūqā*, ed. & tr. J. Sesiano, Sources in the History of Mathematics and Physical Sciences 3, Berlin, Heidelburg, New York: Springer, 1982.

Diophantus–Tannery, *Opera = Opera Omnia cum Graecia Commentariis*, ed. P. Tannery, 2 vols., Leipzig: Teubner, 1893–5.

Dodgson–Abeles, *MPCLD = The Mathematical Pamphlets of Charles Lutwidge Dodgson*, ed. F. F. Abeles, New York & London: Lewis Carroll Society of North America, 1994.

Dorandi, T., TPDEE = La tradizione papiracea degli "Elementi" di Euclide, pp. 306–11 of *Proceedings of the XX International Congress of Papyrologists*, ed. W. Bülow-Jacobsen, Copenhagen: Museum Tusculanum Press, 1994.

*DSB = Dictionary of Scientific Biography*, ed. C. C. Gillispie, 15 vols., New York: Charles Scribner's Sons, 1970–8.

Dutka, J., EHHF = The early history of the hypergeometric function, *Archive for History of Exact Sciences* 31 (1984), 15–34.

Edwards, H. M., *FLT = Fermat's Last Theorem*, Berlin, Heidelberg, New York: Springer, 1979; repr. 1997.

Eisenstein, E. L., *PREME = The Printing Revolution in Early Modern Europe*, Cambridge: Cambridge University Press, 1983.

Erichsen, W., *DG = Demotisches Glossar*, Copenhagen: Munksgaard, 1954.

Euclid. For an English translation of the *Elements*, see Heath, *TBEE*.

Euclid–Frajese, *EE = Gli Elementi di Euclide*, ed. & tr. A. Frajese & I. Maccioni, Turin: Unione Tipografica Editrice, 1970.

Euclid–Heiberg, *Opera = Opera Omnia*, ed. J. L. Heiberg, H. Menge, & M. Curtze, 9 vols., Leipzig: Teubner, 1883–1916.

Euclid–Peyrard, *OE = Les Oeuvres d'Euclide* (1804), traduites littéralement par F. Peyrard, Nouveau Tirage ed. J. Itard, Paris: Blanchard, 1966.

Euclid–Stamatis, *EE = Euclides Elementa*, ed. J. L. Heiberg, rev. E. S. Stamatis, 5 vols. in 6 parts, Leipzig: Teubner, 1969–77.

Eudemus–Wehrli, *ER = Eudemos von Rhodos, Die Schule des Aristoteles: Texte und Kommentar*, ed. F. Wehrili, Basel: Schwabe, 1955.

Eudoxus–Laserre, *FEK = Die Fragmente des Eudoxos von Knidos*, ed. F. Laserre, Berlin: de Gruyter, 1966. (See the review by G. J. Toomer, *Gnomon* 40 (1968), 334–7.)

Euler, L., *Opera Leonhardi Euleri Opera Omnia*, various editors, Series I (*Opera Mathematica*) 29 vols., Series II (*Opera mechanica et Asronomica*) 31 vols., Series III (*Opera Physica, Miscellanea*) 12 vols., Series IVA (*Epistolae*) 8 vols., in preparation, Series IVB (Unpublished Manuscripts) in preparation, Schweizerische Naturforschende Gesellschaft, Leipzig: Teubner, 1911–.

Euler, L., FCD = De fractionibus continuis dissertatio, *Commentarii academiae scientiarum Petropolitanae* 9 (1737), 98–137, repr. in *Opera* I xiv, 187–215.

Euler, L., *IAI = Introductio in Analysin Infinitorum* (1748), repr. as *Opera* I viii & ix.

Euler, L., *ICD = Institutiones Calculi Differentialis* (1775), repr. as Opera I x.

Euler, L., UNAPPS = De usu novi algorithmi in problemate Pelliano solvendo, *Novi commentarii academiae scientiarum Petropolitanae* 11 (1765), 28–66, repr. in *Opera* I iii, 73–111.

Euler, L., *VAA = Vollständige Anleitung zur Algebra*, (1770), repr. with Lagrange, *AEAE* as *Opera* I i.

Euler–Wyman & Wyman, ECF = An essay on continued fractions = Euler, FCD, tr. M. F. & B. F. Wyman, *Mathematical Systems Theory*, 18 (1985), 295–328.

Ferguson, H. R. P., ACGLSE = Arithmetic criteria, GL(n,Z[1/n!]) and some experiments, *1977 Annual Seminar of the Canadian Mathematical Congress, Queens Papers* 48 (1978), 398–433.

Ferguson, H. R. P., RRQN = Recognize rational and quadratic numbers with Euclid's and Brun's algorithms, now incorporated in the following paper, Ferguson, Bailey, & Arno, APSLQ.

Ferguson, H. R. P., Bailey, D. H., & Arno, S., APSLQ = Analysis of PSLQ, an integer relation finding algorithm, *Mathematics of Computation*, to appear.

Ferguson, H. R. P. & Forcade, R., GEA = Generalization of the Euclidean algorithm for real numbers to all dimensions higher than two, *Bulletin of the American Mathematical Society* (New Series) 1 (1979), 912–14.

Ferguson, H. R. P. & Forcade, R., MEA = Multidimensional Euclidean algorithms, *Journal für die reine und angewandte Mathematik* 334 (1982), 171–81.

Fischler [= Herz-Fischler], R., RE = A remark on Euclid II, 11, *Historia Mathematica* 6 (1979), 418–22.

Fletcher, T. J., AV = Approximating by vectors, *Mathematics Teaching* 63 (1973), 4–9 & 64 (1973), 42–4.

Fowler, D. H., REGM = Ratio in early Greek mathematics, *Bulletin of the American Mathematical Society* (New Series) 1 (1979), 807–46.

Fowler, D. H., BTEE = Book II of Euclid's *Elements* and a pre-Eudoxan theory of ratio, *Archive for History of Exact Sciences* 22 (1980), 5–36.

Fowler, D. H., ACPPCM = Archimedes' *Cattle Problem* and the pocket calculating machine, Coventry: University of Warwick, Mathematics Institute Preprint, 1980, supplemented 1981 & 1986.

Fowler, D. H., AREP = Anthyphairetic ratio and Eudoxan proportion, *Archive for History of Exact Sciences* 24 (1981), 69–72.

Fowler, D. H., GGS = A generalisation of the golden section, *Fibonacci Quarterly* 20 (1982), 146–58.

Fowler, D. H., BTEE2 = Book II of Euclid's *Elements* and a pre-Eudoxan theory of ratio, Part 2: Sides and diameters, *Archive for History of Exact Sciences* 26 (1982), 193–209.

Fowler, D. H., IEE = Investigating Euclid's *Elements, British Journal for the Philosophy of Science* 34 (1983), 57–70. (A review of Mueller, *PMDSEE*.)

Fowler, D. H., NFA = A note on fractions of the artaba, *Zeitschrift für Papyrologie und Epigraphick* 52 (1983), 273–4.

Fowler, D. H., TP = Tables of parts, *Zeitschrift für Papyrologie und Epigraphik* 53 (1983), 263–4.

Fowler, D. H., EROE = Eratosthenes' ratio for the obliquity of the ecliptic, with a reply by Dennis Rawlins, *Isis* 74 (1983), 556–62. (But also see the review of Thomas, *SIHGM*, by O. Neugebauer, *American Journal of Philology* 64 (1943), 452–7, on 453.)

Fowler, D. H., CQ = A review of Taisbak, *CQ, Mathematical Intelligencer* 5 (1983), 69–72.

Fowler, D. H., FHYDF = 400 years of decimal fractions, *Mathematics Teaching* 110 (1985), 20–1 & 111 (1985), 30–1.

Fowler, D. H., YMPM = Yet more on Plato, *Meno* 82c–85d, *Phronesis* 35 (1990), 175–81.

Fowler, D. H., ATUWT = An approximation technique, and its use by Wallis and Taylor, *Archive for History of Exact Sciences* 41 (1991), 189–233.

Fowler, D. H., DT = Dedekind's theorem: $\sqrt{2} \times \sqrt{3} = \sqrt{6}$, *The American Mathematical Monthly* 99 (1992), 725–33.

Fowler, D. H., DMS = *Dynamis, mithartum,* and *square, Historia Mathematica* 19 (1992), 418–19.

Fowler, D. H., IRBT = An invitation to read Book X of Euclid's *Elements, Historia Mathematica* 19 (1992), 233–64.

Fowler, D. H., RPEGM = Ratio and proportion in early Greek mathematics, pp. 98–118 in *Science and Philosophy in Classical Greece*, ed. A. C. Bowen, New York & London: Garland, 1992.

Fowler, D. H., LFGM = Logistic and fractions in Greek mathematics: a new interpretation, pp. 133–47 in P. Benoit, K. Chemla, & J. Ritter (eds.), *Histoire de Fractions, Fraction d'Histoire*, Stuttgart: Birkhaüser, 1992.

Fowler, D. H., NCRR2 = Newton, Cotes, and $\sqrt{\sqrt{2}}$: A footnote to Newton's theory of the resistence of fluids, pp. 355–68 in P. H. Harmann & A. E. Shapiro (eds), *An Investigation of Difficult Things: Essays on Newton and the Exact Sciences*, Cambridge: Cambridge University Press, 1992.

Fowler, D. H., CF = Article 'Continued Fractions', pp. 730–40, in I. Grattan-Guinness (ed), *Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences*, London: Routledge Companion Encyclopedias, 1993.

Fowler, D. H., SDIR = The story of the discovery of incommensurability, revisited, pp. 221–35 in K. Gavroglu, J. Christianidis, & E. Nicoliaidis (eds), *Trends in the Historiography of Science*, Boston Studies in the Philosophy of Science no. 151, Kluwer, 1994.

Fowler, D. H., FAT = Further arithmetical tables, *Zeitschrift für Papyrologie und Epigraphik* 105 (1995), 225–8.

Fowler, D. H., EPP = Eudoxus: parapegmata and proportionality, to appear.

Fowler, D. H. & Turner, E. G., HP = Hibeh Papyrus I 27: An early example of Greek arithmetical notation, *Historia Mathematica* 10 (1983), 344–59.

Fraser, P. M., PA = *Ptolemaic Alexandria*, 3 vols., Oxford: Clarendon Press, 1972.

Freeman, *APSP* = *Ancilla to the Pre-Socratic Philosophers*, Oxford: Blackwell, 1962. (English translations of the 'B fragments' of Diels & Kranz, *FV*.)

Freudenthal, H., CFMA = Y avait-il une crise de fondements des mathématiques dans l'antiquité, *Bulletin de la Société Mathématique de Belgique* 18 (1966), 43–55.

von Fritz, K., DIHM = The discovery of incommensurability by Hippasus of Metapontum, *Annals of Mathematics* 46 (1945), 242–64.

Furlani, G., SIBA = Sull'incendio della biblioteca di Alessandria, *Aegyptus* 5 (1924), 205–12.

Furlani, G., GFIBA = Giovanni il Filopono et l'incendio della biblioteca di Allessandria, *Bulletin de la Sociéta Archéologique d'Alexandria* 21 (New Series 6) (1925), 58–77.

Galen–Kühn, *Opera* = Galen, *Opera Quae Existant*, ed. C. G. Kühn, 20 vols., Leipzig: Officina Libraria Car. Cnoblochii, 1821–33.

Galois, E., *Oeuvres* = *Oeuvres mathématiques d'Évariste Galois*, ed. E. Picard, Société Mathématique de France, Paris: Gauthier-Villars, 1897.

Galois, E., DTFCP = Démonstration d'un théorème sur les fractions continues périodiques, *Annales de Mathématiques* 19 (1828–29), repr. in *Oeuvres* 1–8.

Gardies, J.-L., *HEEC* = *L'heritage epistémologique d'Eudoxe de Cnide*, Paris: Vrin, 1988.

Gardiner, A., *EG* = *Egyptian Grammar*, 3rd edn., London and Oxford: Oxford University Press for the Griffith Institute, Ashmolean Museum, 1957.

Gauss, C. F., *Werke* = *Werke*, 12 vols., Königlichen Gesellschaft der Wissenschaften zu Göttingen, Hildesheim: Olms, 1863–1933.

Gauss, C. F., *Tagebuch* = *Nachbildung und Abdruck des Tagebuchs (Notizenjournals) mit Erlanterungen*, in *Werke* $x_1$, 483–574.

Gauss, C. G., *DA* = *Disquisitiones Arithmeticae* (1801), repr. as *Werke* i.

Gauss, C. F., DGCSI = Disquisitiones generales circa seriem infinitam . . . , Pars prior, *Commentationes societatis regiae scientiarum Gottingensis recentiores* 2 (1813), repr. in *Werke* iii, 125–62.

Gauss, C. F., MNIVAI = Methodus nova integralim valores per approximationem inveniendi, *Commentationes societatis regiae scientiarum Gottingensis recentiores* 3 (1816), repr. in *Werke* iii 165–96.

Gauss–Clarke, *DA = Disquisitiones Arithmeticae*, tr. A. A. Clarke, New Haven: Yale University Press, 1966.

Geffcken, J., *ZGA = Zwei Griechische Apologeten [Aristides & Anathagoras]*, Leipzig: Teubner, 1907.

Gerstinger, H. & Vogel, K., *MNWPER = Mitteilungen aus der Nationalbibliothek in Wien: Papyrus Erzherzog Rainer*, Neue Series i [= M.P.E.R., N.S. i], Vienna: Osterreichischen Staatsdruckerei, 1932.

Gigante, M., *CPE = Catalogo dei Papiri Ercolanesi*, Naples: Bibliopolis, 1979.

Glucker, J., *ALA = Antiochus and the Late Academy*, Hypomnemata 56, Göttingen: Vandenhoeck & Ruprecht, 1978.

Goldstein, B. R., NMC = A note on the Metonic cycle, *Isis* 57 (1966), 115–16.

Goldstein, B. R., OEAGA = The obliquity of the ecliptic in ancient Greek astronomy, *Archives Internationales d'Histoire des Sciences* 33 (1983), 3–14.

Goldstein, B. R., & Bowen, A. C., NVEGA = A new view of early Greek astronomy, *Isis* 74 (1983), 330–40.

Gombrich, E. H., *SI = Symbolic Images, Studies in the Art of the Renaissance*, London: Phaidon, 1972.

Goody, J. *DSM = The Domestication of the Savage Mind*, Cambridge: Cambridge University Press, 1977.

Gosper, R. W., TSCFP = Table of the simple continued fraction of $\pi$ and the derived decimal approximation; see the review by J. W. Wrench, *Mathematics of Computation* 31 (1977), 1044.

Grabiner, J., CB = Cauchy and Bolzano, pp. 105–25 in *Transformation and Tradition in the Sciences*, ed. E. Mendelsohn, Cambridge: Cambridge University Press, 1984.

Grenfell, B. P., Hunt, A. S., & Hogarth, D. G., *FTP = Fayûm Towns and their Papyri* [= P. Fay.], London: Egypt Exploration Fund, 1900.

Grenfell, B. P., Hunt, A. S., & others, *HP = The Hibeh Papyri* [= P. Hib.], 2 vols., London: Egypt Exploration Society, 1906–55.

Grenfell, B. P., Hunt, A. S., & others, *OP = The Oxyrhynchus Papyri* [= P. Oxy.], 65 vols. to 1998, London: Egypt Exploration Society, 1898–.

Griffith, F. Ll., *HPKG = Hieratic Papyri from Kahun and Gurob* [= P. Kahun], 2 vols., London: Quaritch, 1897–8.

Guarducci, M., *EG = Epigrafia Greca*, 4 vols., Rome: Instituto Poligrafico, 1967–78.

Guthrie, W. K. C., *HGP = A History of Greek Philosophy*, 6 vols., Cambridge: Cambridge University Press, 1962–81.

Hall, F. W., *CCT = A Companion to Classical Texts*, Oxford: Clarendon Press, 1913.

Hall, M., SPCF = On the sum and product of continued fractions, *Mathematische Annalen* 48 (1947), 966–93.

Hardy, G. H. & Wright, E. M., *ITN = Introduction to the Theory of Numbers*, Oxford: Clarendon Press, 1938, 5th edn., 1979.

Harrauer, H. & Sijpesteijn, P. J., *NTAU = Neue Texte aus dem Antiken Unterricht* [= M.P.E.R., N.S. xv], 2 vols., Vienna: Hollinek, 1985.

Harvey, F. D., TKE = Two kinds of equality. *Classica et Mediaevalia* 26 (1965), 101–46; corrigenda, ibid. 27 (1966), 99–100.

Harvey, P. D. A., *HTM = The History of Topographical Maps: Symbols, Pictures, and Surveys*, London: Thames and Hudson, 1980.

Harvey, P. D. A. & Skelton, R. A., *LMPME = Local Maps and Plans from Medieval England*, Oxford: Clarendon Press, 1986.

Havelock, E. A., *LRG = The Literate Revolution in Greece and its Cultural Consequences*, Princeton: Princeton University Press, 1982.

Heath, T. L., *HGM* = *A History of Greek Mathematics*, 2 vols., Oxford: Clarendon Press, 1921, repr. New York: Dover, 1981.

Heath, T. L., *TBEE* = *The Thirteen Books of Euclid's Elements*, 2nd edn. 3 vols., Cambridge: Cambridge University Press, 1926, repr. New York: Dover, 1956.

Heath, T. L., *MA* = *Mathematics in Aristotle*, Oxford: Clarendon Press, 1949, repr. New York: Garland, 1980.

Heath, T. L., also see Aristarchus–Heath, *AS* & Diophantus–Heath, *DA*.

Heiberg, J. L., PEE = Ein palimpsest der Elemente Euklidis, *Philologus* 44 (1885), 353–66.

Heiberg, J. L., QPTM = Quelques papyrus traitant de mathématiques, *Oversigt over det Kgl. Danske Videnskabernes Selskabs Forhandlinger* [= *Proceedings of the Danish Academy of Sciences*] 2 (1900), 147–71.

Heiberg, J. L., ENA = Eine neue Archimedeschandschrift, *Hermes* 62 (1907), 235–303.

Heiberg, J. L., PE = Paralipomena zu Euklid, *Hermes* 39 (1930), 46–74, 161–201, & 321–56.

Heiberg, J. L., Also see Archimedes–Heiberg, *Opera*; Euclid–Heiberg, *Opera*; and Heron–Heiberg, Schöne, & others, *Opera*.

Henrici, P., *ACCA* ii = *Applied and Computational Complex Analysis* ii, New York: Wiley, 1977.

Herodotus–Godley, *H* = *Herodotus*, tr. A. D. Godley, 4 vols., Loeb Classical Library, Cambridge, Massachusetts & London: Harvard University Press & Heinemann, 1920–4.

Heron–Bruins, *CC* = *Codex Constantinopolitanus Palatii Veteris No. 1*, ed. E. M. Bruins, 3 vols. (i, Photographic reproduction of the manuscript; ii, Greek transcription; iii, Translation & commentary), Leiden: Brill, 1964. (A re-edition of our principal manuscript of Heron's mathematical writings and our sole source of his *Metrica*.)

Heron–Heiberg, Schöne, & others, *Opera* = *Opera Quae Supersunt Omnia*, 5 vols.: i, *Pneumatica et Automata*, ed. G. Schmidt; ii, *Mechanica et Catoprica*, ed. L. Nix & W. Schmidt; iii, *Rationes Dimetiendi* [= *Metrica*] & *Commentatio Dioptrica*, ed. H. Schöne; iv, *Definitones*, ed. J. L. Heiberg; v., *Stereometrica & De Mensuris*, ed. J. L. Heiberg, Leipzig: Teubner, 1895–1914, repr. 1976. (Also see Heron–Bruins, *CC*.)

Herz–Fischler [= Fischler], R., *MHDEMR* = *A Mathematical History of Division in Extreme and Mean Ratio*, Waterloo: Laurier, 1987.

Hogendijk, J., HTATGIG = How trisections of the angle were transmitted from Greek to Islamic geometry, *Historia Mathematica* 8 (1981), 417–38.

Horsfall, N., SB = Stesichorus at Bovillae?, *Journal of Hellenic Studies* 99 (1979), 26–48, with addendum: *Tabulae Iliacae* in the Collection Froehner, Paris, ibid. 103 (1983), 144–7.

Høyrup, J., DBT = *Dynamis*, the Babylonians, and Theaetetus 147c7–148d7, *Historia Mathematica* 17 (1990), 201–22; Erratum, *ibid.* 18 (1992), 89.

Huffman, C. A., AAF1 = The authenticity of Archytas Fr. 1, *Classical Quarterly* 35 (1985), 344–8.

Huffman, C. A., *PCPP* = *Philolaus of Croton, Pythagorean and Presocratic: A Commentary on the Fragments and Testimonia with Interpretative Essays*, New York: Cambridge University Press, 1993.

Hunt, A. S. & Edgar, C. C., *SP* i = *Select Papyri i, Non-Literary Papyri*, Loeb Classical Library, Cambridge, Massachusetts & London: Harvard University Press & Heinemann, 1932.

Hurwitz, A., *Werke = Mathematische Werke*, 2 vols., Eidgenössischen Technischen Hochschule in Zürich, Basel: Birkhäuser, 1932–3.

Hurwitz, A., KEZE = Über die Kettenbruch-Entwicklung der Zahl *e*, *Schriften der Physikalisch-Ökonomischen Gessellschaft zu Königsberg* 32 (1891), 59–62, repr. in *Werke* ii, 129–133.

Hurwitz, A., KTAR = Über die Kettenbrüche, deren Teilnenner arithmetische Reihen bilden, *Vierteljahrsschrift der Naturforschenden Gessellschaft in Zürich* 41 (1896), 34–64, repr. in *Werke* ii, 276–302.

Huygens, C., *Oeuvres = Oeuvres Complètes*, various editors, 22 vols., The Hague: Société Hollandaise des Sciences, 1888–1950.

Huygens, C., DAP = Descriptio automati planetarii, repr. in *Oeuvres* xxii, 579–652.

Hypsicles–De Falco, Krause, & Neugebauer, AG = Hypsikles, Die Auf-gangszeiten der Gestirne, ed. & tr. V. De Falco, M. Krause, & O. Neugebauer, *Abhandlungen der Akademie der Wissenschaften in Göttingen, Philologisch-Historische Klasse* 63 (1966).

Iamblichus–Clark, *IPL = Iamblichus: On the Pythagorean Life*, tr. G. Clark, Liverpool: Liverpool University Press, 1989.

Iamblichus–Pistelli, *NAI = In Nicomach Arithmeticam Introductionem Liber*, ed. H. Pistelli, Leipzig: Teubner, 1894.

Itard, J., *LAE = Les Livres Arithmétiques d'Euclide,* Histoire de la pense 10, Paris: Hermann, 1962.

Jacobi, C. G. J., *Werke = Gesammelte Werke*, ed. C. W. Borchardt & K. Weierstrass, 8 vols., Berlin, 1881–91, repr. New York: Chelsea, 1969.

Jacobi, C. G. J., ATKA = Allgemeine Theorie der Kettenbruchähnlichen Algorithmen, in welchen jede Zahl, aus drei Vorhergehenden gebildet wird (aus den hinterlassenen Papiren C. G. J. Jacobi's mitgetheilt durch E. Heine), *Journal für die reine und angewandte Mathematik* 69 (1868), 29–64, repr. in *Werke* vi, 385–426.

James, T. G. H., *EEEES = Excavating in Egypt: The Egypt Exploration Society 1882–1982*, ed. T. G. H. James, London: British Museum Publications, 1982.

Jan, K. von, *MSG = Musici Scriptores Graeci*, ed. K. von Jan, Leipzig: Teubner, 1895, repr. Hildesheim: Olms, 1962.

Johnston, A., *TGV = Trademarks on Greek Vases*, Warminster: Arts & Phillips, 1979.

Jones, A., LME = Land measurement in England 1150–1350, *Agricultural History Review* 27 (1979), 10–18.

Jones, W. B. & Thron, W. J., *CF = Continued Fractions: Analytic Theory and Applications*, Encyclopedia of Mathematics and its Applications 11, Reading, Massachusetts: Addison–Wesley, 1980.

Kahane, J.-P., TTCQR = La théorie de Théodore des corps quadratiques réels, *L'Enseignement Mathématique* 31 (1985), 85–92.

Kapsomenos, S. G., OPRT = The Orphic papyrus roll of Thessalonica, *Bulletin of the American Society of Papyrologists* 2 (1964–5), 3–31.

Keaney, J. J., ETTHP = The early tradition of Theophrastus' *Historia Plantarum*, *Hermes* 96 (1968), 293–8.

Kenyon, F. G., TGST = Two Greek school-tablets, *Journal of the Hellenic Society* 29 (1909), 29–40.

Kenyon, F. G., Bell, H. I., & others, *GPBM = Greek Papyri in the British Museum* [= P. Lond.], 7 vols., to 1974, London: British Museum, 1893–.

Khinchin, A. Ya., *CF = Continued Fractions*, tr. Scripta Technica Inc. of the 3rd Russian edn. (1961) of *Cepnye Drobi* (1936), Chicago: University of Chicago Press, 1964.

Klein, F., *EMAS* i = *Elementary Mathematics from an Advanced Standpoint*, vol. 1:

*Arithmetic, Algebra, Analysis* (1908), 3rd edn., tr. E. A. Hedrick & C. A. Noble, repr. New York: Dover, 1945.

Klein, J., *GMTOA = Greek Mathematical Thought and the Origin of Algebra*, tr. E Brann of Die griechische Logistik und die Entstehung der Algebra, *Quellen und Studien zu Geschichte der Mathematik, Astronomie und Physik*, Section B: *Studien*, $3_1$ (1934), 18–105 & $3_2$ (1936), 122–235, with Appendix containing Francois Viète's *Introduction to the Analytic Art*, tr. J. W. Smith of *Isagoge* (1591), Cambridge, Massachusetts: MIT Press, 1968.

Knorr, W. R., *EEE = The Evolution of the Euclidean Elements: A Study of the Theory of Incommensurable Magnitudes and Its Significance for Early Greek Geometry*, Dordrecht: Reidel, 1975.

Knorr, W. R., PIGNT = Problems in the interpretation of Greek number theory: Euclid and the 'Fundamental Theorem of Arithmetic', *Studies in the History and Philosophy of Science* 7 (1976), 353–68.

Knorr, W. R., AMC = Archimedes and the measurement of the circle: A new interpretation, *Archive for History of Exact Sciences* 15 (1976), 115–40.

Knorr, W. R., AE = Archimedes and the *Elements*: Proposal for a revised chronological ordering of the Archimedean corpus, *Archive for History of Exact Sciences* 19 (1978), 211–90.

Knorr, W. R., ANCSL = Archimedes' neusis-construction in spiral lines, *Centaurus* 22 (1978), 77–98.

Knorr, W. R., MPP = Methodology, philology, and philosophy, with a reply by M. Burnyeat, *Isis* 70 (1979), 565–70.

Knorr, W. R., APEPT = Archimedes and the pre-Euclidean proportion theory, *Archives Internationales d'Histoire des Sciences* 28 (1980), 183–244.

Knorr, W. R., TFAEG = Techniques of fractions in ancient Egypt and Greece, *Historia Mathematica* 9 (1982), 133–71.

Knorr, W. R., IC = Infinity and continuity: The interaction of mathematics and philosophy in antiquity, pp. 112–45 of *Infinity and Continuity in Ancient and Medieval Thought*, ed. N. Kretzmann, Ithaca: Cornell University Press, 1982.

Knorr, W. R., CM = "La Croix des Mathématiciens": The Euclidean theory of irrational lines, *Bulletin of the American Mathematical Society* (New Series) 9 (1983), 41–69.

Knorr, W. R., RTCQ = Review of Taisbak, *CQ*, *Centaurus* 27 (1984), 330–2.

Knorr, W. R., ETB = Euclid's tenth book: An analytic survey, *Historia Scientiarum* 29 (1985), 17–35.

Knorr, W. R., *ATGP = The Ancient Tradition of Geometric Problems*, Boston: Birkhäuser, 1986.

Knorr, W. R., AD = Archimedes after Dijksterhuis: A guide to recent studies, pp. 419–51 of Dijksterhuis, *Archimedes*, Princeton: Princeton University Press, 1987.

Knorr, W. R., *TSAMG = Textual Studies in Ancient and Medieval Geometry*, Boston: Birkhäuser, 1989.

Knorr, W. R., WEM = What Euclid meant: On the use of evidence in studying ancient mathematics, pp. 119–63, in *Science and Philosophy in Classical Greece*, ed. A. C. Bowen, New York and London: Garland Press, 1992.

Knorr, W. R., RDDI = "Rational diameters" and the discovery of incommensurability, *American Mathematical Monthly*, to appear.

Knorr, W. R., IM = In memoriam: An appreciation of Wilbur Knorr (1945–1998), with Bibliography, by David Fowler, *Historia Mathematica* 25 (1998), 123–32.

Knuth, D. E., *ACP* ii = *The Art of Computer Programming* ii, *Seminumerical Algorithms*, 2nd edn., Reading, Massachusetts: Addison-Wesley, 1981.

Lagrange, J. L., *Oeuvres = Oeuvres des Lagrange*, ed. J.-A. Serret, 14 vols., Le Ministre de l'Instruction Publique, Paris: Gauthier–Villars, 1867–92.

Lagrange, J. L., SPA = Solution d'un problème d'arithmétique, *Miscellanea Taurinensia* 4 (1766–9), repr. in *Oeuvres* i, 671–731.

Lagrange, J. L., REN = Sur la résolution des équations numériques, *Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin* 23 (1769), repr. in *Oeuvres* ii, 539–78.

Lagrange, L., AMREN = Additions au mémoire sur la résolution des équations numériques, *Memoires de l'Académie royale des Sciences et Belles-Lettres de Berlin* 24 (1770), repr. in *Oeuvres* ii, 581–652.

Lagrange, J. L., RRAE = Réflexions sur la résolution algébrique des équations, *Nouveaux Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin*, 1770 & 1771, repr. in *Oeuvres* iii, 205–421.

Lagrange, J. L., *AEAE = Additions aux Éléments d'Algèbre d'Euler: Analyse Indéterminée*, i.e. Appendix to the French translation (1774), by Jean III Bernoulli, of Euler, *VAA*; repr. in Lagrange, *Oeuvres* vii, 5–180 & Euler, *Opera* I i, 499–651.

Lagrange, J. L., RA = Recherches d'arithmétiques, *Nouveaux Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin*, 1775 and 1777, repr. in *Oeuvres* iii 697–793.

Lagrange, J. L., UFCCI = Sur l'usage des fractions continues dans le calcul intégral, *Nouveaux Mémoires de l'Académie royal des Sciences et Belles-Lettres de Berlin*, 1776, repr. in *Oeuvres* iv, 301–22.

Lagrange, J. L., *TREN = Traité de la résolution des équations numérique de tous les degrés*, 1st edn. 1798, revised & augmented 2nd edn. 1808, repr. in *Oeuvres* viii, 11–370.

Lagrange, J. L., EANTF = Essai d'analyse numérique sur la transformation des fractions, *Journal de l'École Polytechnique* 2 (prairial an vi [= 28 May–18 June, 1799]), repr. in *Oeuvres* vii, 291–313.

Landau, E., *FA = Foundations of Analysis*, tr. F. Steinhardt of *Grundlagen der Analysis* (1930), New York: Chelsea, 1951.

Lang, M., NNGV = Numerical notation on Greek vases, *Hesperia* 25 (1956), 1–24.

Lang, S. & Trotter, H., CFSAN = Continued fractions for some algebraic numbers, *Journal für die reine und angewandte Mathematik* 255 (1972), 112–34. Addendum, ibid. 267 (1974), 219–20.

Legendre, A. M., *ETN & TN = Essai sur la théorie des nombres*, 1st edn. 1798; supplemented 2nd edn. 1808; further supplemented 3rd edn. in 2 vols.: *Théorie des nombres*, 1830, repr. Paris: Blanchard, 1955.

Lehmer, D. H., EALN = Euclid's algorithm for large numbers, *American Mathematical Monthly* 45 (1938), 227–33.

Leibniz, G. G., NMMM = Nova methodvs pro maximis et minimis, itemque tangentibus, quae nec fractas, nec irrationales quantitates moratur, & singulare pro illis calculi genus, per G. G. L., *Acta Eruditorum* 3 (1684), 467–73. English translation in Raphson, *TF*.

Lemerle, P., *PHB = Le Premier Humanisme Byzantin*, Paris: Presses Universitaires, 1971.

Letronne, A. J. & Brunet de Presle, W., *NTPGML = Notices et textes des papyri grecs du Musée du Louvre et de la Bibliothèque Impériale* 18, Paris: Imprimerie Impériale, 1865.

Lewis, N., *PCA = Papyrus in Classical Antiquity*, Oxford: Clarendon Press, 1974.

Liddell, H. G., Scott, R., & Jones, H. S., *GEL = A Greek–English Lexicon*, 9th edn. (1940) with *Supplement* (1968), Oxford: Clarendon Press, often reprinted.

Lucas, A., *AEMI* = *Ancient Egyptian Material and Industries* (1926), 4th edn. rev. & enl. by J. R. Harris, London: Arnold, 1962.

Lynch, J. P., *AS* = *Aristotle's School: A Study of a Greek Educational Institution*, Berkeley: University of California Press, 1972.

McNamee, K., *AGLPO* = *Abbreviations in Greek Literary Papyri and Ostraca*, Bulletin of the American Society of Papyrologists Supplement 3, Chicago: Scholars Press, 1981.

Mattha, G., *DO* = *Demotic Ostraca from the Collections at Oxford, Paris, Berlin, Vienna, and Cairo*, Publications de la Societé Fouad I de Papyrologie, Cairo: Institute Française d'Archéologie Orientale, 1945.

Mau, J. & Müller, W., MOBS = Mathematische ostraka aus der Berliner Sammlung, *Archiv für Papyrusforschung* 17 (1982), 1–10.

Meyerhof, M., JGPAAM = Joannes Grammatikos (Philoponos) von Alexandrien und die Arabische Medizin, *Mitteilungen des Deutschen Instituts für Ägyptische Altertumskunde in Kairo* 2 (1932), 1–21.

Milne, H. J. M. & Skeat, T. C., *SCCS* = *Scribes and Correctors of the Codex Sinaiticus*, London: British Museum, 1938.

Minkowski, H., *Abhandlungen* = *Gesammelte Abhandlungen von Hermann Minkowski*, ed. A. Speiser, H. Weyl, & D. Hilbert, 2 vols., Leipzig: Teubner, 1911.

Minkowski, H., TK = Zur Theorie der Kettenbruche, *Abhandlungen* i, 278–92; French translation by L. Laugel, Généralisation de la théorie des fractions continues, *Annales de l'École Normale Supérieur*, (3$^e$ série) 13 (1896), 41–60.

Minkowski, H., KAZ = Ein Kriterium für die algebraischen Zahlen, *Nachrichten der K. Gesellschaft der Wissenschaften zu Göttingen, mathematische-physikalische Klasse* (1990), 64–88, repr. in *Abhandlungen* i, 293–315.

Minkowski, H., UP = Über Periodische Approximationen Algebraischen Zahlen, *Acta Mathematica* 26 (1902), 333–51, repr, in *Abhandlungen* i, 357–71.

Minkowski, H., *BDH* = *Briefe an David Hilbert*, ed. L. Rüdenberg & H. Zassenhaus, Berlin, Heidelberg, New York: Springer, 1973.

Most, G. W., RR = Reading Raphael: The *School of Athens* and its pre-text, *Critical Enquiry* 23 (1996), 145–82.

Mueller, I., *PMDSEE* = *Philosophy of Mathematics and Deductive Structure in Euclid's Elements*, Cambridge, Massachusetts: MIT Press, 1981.

Mueller, I., MPPC = Mathematics and philosophy in Proclus' Commentary on Book I of Euclid's *Elements*, pp. 305–18 in *Proclus—Lecture et interprète des anciens*, ed. J. Pépin & H. D. Saffrey, Paris: Editions du C.N.R.S., 1987.

Mueller, I., IPEC = Iamblichus and Proclus' Euclid Commentary, *Hermes* 115 (1987), 334–48.

Mueller, I., ME = Mathematics and education: Some notes on the Platonic programme, in *ΠΕΡΙ ΤΩΝ ΜΑΘΗΜΑΤΩΝ*, ed. I. Mueller, *Apeiron* 24 (1991), 85–104.

Muir, T., *EQSCF* = *The Expression of a Quadratic Surd as a Continued Fraction*, Glasgow: Maclehose, 1874.

Muir, T., *TD* = *The Theory of Determinants, in the Historical Order of Development*, 4 vols., London: St Martin's Press, 1906–23, repr. in 2 vols., New York: Dover, 1960; and *Contributions to the Theory of Determinants 1900–1920*, London: Blackie, 1930.

Nelson, H. L., SCAP = A Solution to Archimedes' Cattle Problem, *Journal of Recreational Mathematics* 13 (1980–1), 162–76.

Netz, R., *SDGM* = *The Shaping of Deduction in Greek Mathematics*, Cambridge: Cambridge University Press, to appear.

Netz, R., PDMP = Proclus' division of the mathematical proposition, *Classical Quarterly*, to appear.

Neuenschwander, E., EVBEE = Die ersten vier Bücher der Elemente Euklidis, *Archive for History of Exact Sciences* 9 (1973), 325–80.

Neuenschwander, E., SBEE = Die stereometrischen Bücher der Elemente Euklidis, *Archive for History of Exact Sciences* 14 (1975), 91–125.

Neugebauer, O., ESA = *The Exact Sciences in Antiquity*, 2nd edn., Providence: Brown University Press, 1957, repr. New York: Dover, 1969.

Neugebauer, O., APO = Astronomical papyri and ostraca: Biographical notes, *Proceedings of the American Philosophical Society* 106 (1962), 383–91.

Neugebauer, O., *HAMA* = *A History of Ancient Mathematical Astronomy*, 3 vols., Berlin, Heidelberg, New York: Springer, 1975.

Neugebauer, O. & Parker, R. A., *EAT* = *Egyptian Astronomical Texts*, 3 vols., Providence: Brown University Press, 1960–9.

von Neumann, J. & Tuckerman, B., CFE = Continued fraction expansion of $\sqrt[3]{2}$, *Mathematics of Computation* 9 (1955), 23–4.

Nicomachus–D'Ooge, *IA* = *Introduction to Arithmetic*, tr. M. L. D'Ooge, New York: Macmillan, 1926.

Nicomachus–Hoche, *IA* = *Introductionis Arithmeticae Libri II*, ed. R. Hoch, Leipzig: Teubner, 1866.

Niebel, E., *UBGKA* = *Untersuchungen über die Bedeutung der geometrischen Konstruktion in der Antike*, Kantstudien 76, Cologne: Kölner Universitäts-Verlag, 1959.

*OCD* = *The Oxford Classical Dictionary*, 2nd edn., ed. N. G. L. Hammond & H. H. Scullard, Oxford: Clarendon Press, 1970.

Odom, G., EP = Elementary Problem E 3007, *American Mathematical Monthly* 90 (1983), 482.

*OED* = *The Oxford English Dictionary*, Oxford: Clarendon Press, 1933, and often reprinted; and *Supplement*, 4 vols., ed. R. W. Burchfield, Oxford: Clarendon Press, 1972–86; and 2nd edn., 1989; 3rd edn., ed S. Hornblower & A. Spawforth, Oxford: Clarendon Press, 1996.

Omar Khayyam–Amir-Móez, DDE = Discussion of Difficulties in Euclid, by Omar ibn Abrahim al-Khayyami (Omar Khayyam), tr. A. R. Amir-Móez, *Scripta Mathematica* 24 (1959), 275–303.

O'Meara, D. J., *PR* = *Pythagoras Revived, Mathematics and Philosophy in Late Antiquity*, Oxford: Oxford University Press, 1989.

Pack, R. A., *GLLTGRE* = *The Greek and Latin Literary Texts from Greco-Roman Egypt*, 2nd edn., Ann Arbor: University of Michigan Press, 1965.

Pappus–Jones, BSC = *Book 7 of the Collection*, 2 vols, ed. & tr. A. Jones, Sources in the History of Mathematics and Physical Sciences, vol. 8, New York, Berlin, Heidelberg, & Tokyo: Springer, 1986.

Pappus–Thomson & Junge, *CPBXEE* = *The Commentary of Pappus on Book X of Euclid's Elements*, ed. & tr. W. Thomson & G. Junge, Harvard Semitic Series 8, Cambridge, Massachusetts: Harvard University Press, 1930, repr. New York: Johnson Reprint Company, 1968.

Parker, R. A., *CAE* = *The Calendars of Ancient Egypt*, Studies in Ancient Oriental Civilization 26, Chicago: University of Chicago Press, 1950.

Parker, R. A., DMPF = A demotic mathematical papyrus fragment, *Journal of Near Eastern Studies* 18 (1959), 275–9.

Parker, R. A., *DMP* = *Demotic mathematical Papyri*, Providence: Brown University Press, 1972.

Parker, R. A., ME = A mathematical exercise—P. dem. Heidelberg 663, *Journal of Egyptian Archaeology* 61 (1975), 189–96.

Perron, O., *LK = Die Lehre von den Kettenbrüchen*, 3rd edn. 2 vols., Stuttgart: Teubner, 1954.

Pestman, P. W., *GDTZA = Greek and Demotic Texts from the Zenon Archive* [= P.L. Bat. xx], Leiden: Brill, 1980.

Pestman, P. W., *GZA = A Guide to the Zenon Archive* [= P. L. Bat. xxi], 2 vols., Leiden: Brill, 1981.

Pfeiffer, R., *HCS = History of Classical Scholarship: From the Beginnings to the End of the Hellenistic Age*, Oxford: Clarendon Press, 1968.

Philoponus–Wallies, *IAP = In Aristotelis Analytica Posteriora Commentaria*, ed. M. Wallies, Commentaria in Aristotes Graeca 5₃, Berlin: Reimer, 1909.

Pintaudi, R. & Sijpestijn, P. J., *TLCVC = Tavolette Lignee e Cerate da Varie Collezioni* [= P. Florentina xviii], containing *Tavolette . . . della Biblioteca Apostolica Vaticana* (ed. R. Pintaudi & P. J. Sijpesteijn), *Quatre cahiers scholaires au Musée du Louvre* (ed. P. Cauderlier), and *Wax Tablets in the Pierpont Morgan Library* (ed. R. S. Bagnall), Florence: Gonelli, 1989.

Pisot, C., SM = Souvenirs mathématiques, *L'Enseignement Mathématique* 26 (1980), 185–92.

Plato, Individual works with the Greek text and parallel English translation in the Loeb Classical Library, Cambridge, Massachusetts & London: Harvard University Press & Heinemann.

Plato, Critical editions of the Greek text of individual works in the Scriptorum Classicorum Bibliotheca Oxoniensis (= The Oxford Classical Texts), Oxford: Clarendon Press.

Plato, *CD = The Collected Dialogues of Plato including the Letters*, ed. E. Hamilton & H. Cairns, Bollingen Series 71, Princeton: Princeton University Press, 1961.

Plato–Allen, *PP = Plato's Parmenides, Translation and Analysis*, ed. & tr. R. E. Allen, Oxford: Blackwell, 1983.

Plato–Bluck, *PM = Plato's Meno*, ed. R. S. Bluck, Cambridge University Press, 1961. (Critical edition of the Greek text, with translation and extensive commentary.)

Plato–Guthrie, *PM = Protagoras and Meno*, tr. W. K. C. Guthrie, Harmondsworth: Penguin, 1956. (Also repr. in Plato, *CD* and Brown, *PM*.)

Plato–McDowell, *Theaetetus* = Plato, *Theaetetus*, tr. J. McDowell, Oxford: Clarendon Press, 1963.

Plooij, E. B., *ECR = Euclid's Conception of Ratio and his Definition of Proportional Magnitudes as Criticised by Arabian Commentators*, Rotterdam: van Hengel, 1950.

Plutarch–Hubert, *M = Moralia* iv, ed. C. Hubert, Leipzig: Teubner, 1938, repr. 1971.

Poincaré, H., *Oeuvres = Oeuvres de Henri Poincaré*, 11 vols., L'Académie des Sciences, Paris: Gauthier-Villars, 1916–54.

Poincaré, H., GCF = Sur une généralisation des fractions continues, *Comptes rendus de l'Académie des Sciences* 99 (1884), 1014–16, repr. in *Oeuvres* v, 185–8.

Popper, K., *OSE* i = *The Open Society and its Enemies* i, *The Spell of Plato*, London: Routledge & Kegan Paul, 1945, 5th edn. 1966.

Price, D. de Solla, *GG = Gears from the Greeks: The Antikythera Mechanism—A Calendar Computer from ca. 80 B.C.*, Philadelphia: Transactions of the American Philosophical Society (New Series) 64 Part 7 (1974), repr. New York: Science History Publications, 1975.

Pritchard, P., MDEM = *Metaphysics Δ15 and pre-Euclidean mathematics*, *Apeiron* 30 (1987), 42–69.

Procissi, A., BMGA = Bibliografia della mathematica Greca antica, *Bollettino di Storia della Scienze Matematiche* 1 (1981), 7–149.

Proclus–Festugière, *CR = Commentaire sur la République*, tr. A. J. Festugière, 3 vols., Paris: Vrin, 1970.

Proculus–Friedlein, *PEEC = In Primum Euclidis Elementorum Commentarii*, ed. G. Friedlein, Leipzig: Teubner, 1873.

Proclus–Kroll, *IR = In Platonis Rem Publicam Commentarii*, ed. W. Kroll, 2 vols., Leipzig: Teubner, 1899–1901.

Proclus–Morrow, *CFBEE = A Commentary on the First Book of Euclid's Elements*, tr., with intro. and notes, G. R. Morrow, Princeton: Princeton University Press, 1970; repr. with foreword by I Mueller, Princeton: Princeton University Press, 1992. (All references to this book will use the pagination of Proclus–Friedlein.)

Ptolemy–Düring, *HKP = Die Harmonielehre des Klaudios Ptolemaios,* ed. I. Düring, Göteborg: Elander, 1930.

Ptolemy–Toomer, *PA = Ptolemy's Almagest*, tr. G. J. Toomer, London: Duckworth, 1984.

Raphson [Ralphson], J., *TF = The Theory of Fluxions, Shewing in a compedious manner The first Rise of, and various Improvements made in that Incomparable Method*, London, 1715.

Rawlins, D., EGU = Eratosthenes' geodesy unravelled: Was there a high-accuracy Hellenistic astronomy?, *Isis* 73 (1982), 259–62. (Also see Fowler, EROE.)

Raney, G. N., CFFA = On continued fractions and finite automata, *Mathematische Annalen* 206 (1973), 265–83.

Reid, C., *H = Hilbert*, Berlin, Heidelberg, New York: Springer, 1970.

Reynolds, L. D. & Wilson, N. G., *SS = Scribes and Scholars: A Guide to the Transmission of Greek and Latin Literature*, 2nd edn., Oxford: Clarendon Press, 1974.

Richtmyer, R. D., Deavney, M., & Metropolis, N., CFEAN = Continued fraction expansions of algebraic numbers, *Numerische Mathematik* 4 (1972), 64–84.

Riddell, R. C., EMES = Eudoxan mathematics and the Euoxan spheres, *Archive for History of Exact Sciences* 20 (1979), 1–19.

Riemann, B., *Werke = Gesammelte Mathematische Werke und Wissenschaftlicher Nachlass*, 2nd edn. (1892), ed. R. Dedekind & H. Weber, with *Nachträge* (1902), ed. M. Noether & W. Wirtinger, repr. New York: Dover, 1953.

Riemann, B., SQSI = Sullo svolgimento del quoziente di due serie ipergeometriche in frazione continus infinia (edited from notes of 1863 by H. A. Schwarz), in *Werke*, 424–30.

Riginos, A. S., *PACLWP = Platonica: The Anecdotes Concerning the Life and Writings of Plato*, Leiden: Brill, 1976.

Roberts, C. H. & Skeat, T. C., *BC = The Birth of the Codex*, Oxford: Oxford University Press, for the British Academy, 1983.

Rose, P. L., FHCAA = For the history of Codex A of Archimedes: Notes on the Estense, Carpi, and Ridolfi Libraries, *Manuscripta* 21 (1977), 180–3.

Rosen, D. & Shallit, J., CFA = A continued fraction algorithm for approximating all real polynomial roots, *Mathematics Magazine* 51 (1978), 112–16.

Rosenfeld, B. A., *HNEG = A History of Non-Euclidean Geometry*, Berlin, Heidelberg, & New York: Springer, 1988.

Rowland, R. J., BIW = The biggest island in the world, *The Classical World* 68 (1975), 438–9.

Sadie, S., *NGDMM = The New Grove Dictionary of Music and Musicians*, ed. S. Sadie, 20 vols., London: Macmillan, 1980.

Saffrey, H. D., AME = *ΑΓΕΩΜΕΤΡΗΤΟΣ ΜΗΔΕΙΣ ΕΙΣΙΤΩ*, une inscription légendaire, *Revue des Études Grecques* 81 (1968), 67–87.

Ste. Croix, G. E. M. de, GRA = Greek and Roman accounting, in *Studies in the History of Accounting*, ed. A. C. Littleton & B. S. Yamey, London: Sweet & Maxwell, 1956.

Ste. Croix, G. E. M. de, *CSAGW* = *The Class Struggle in the Ancient Greek World*, London: Duckworth, 1981.

Saito, K., BTEE = Book II of Euclid's *Elements* in the light of the theory of conic sections, *Historia Scientiarum* 28 (1985), 31–60.

Saito, K., CREA = Compounded ratios in Euclid and Apollonius, *Historia Scientiarum* 31 (1986), 25–59.

Saito, K., DRBS = Duplicate ratio in Book VI of Euclid's *Elements*, *Historia Scientiarum* 32 (1993), 115–35.

Saito, K., IPPC = Index of propositions used in Book VII of Pappus' *Collection*, *Jinbun Kanyku: The Journal of Humanities* 26 (1997), 155–88.

Sandys, J. E., *HCS* = *History of Classical Scholarship* (1903), 3rd edn. 3 vols., Cambridge: Cambridge University Press, 1921.

Santillana, G. de, EPSC = Eudoxus and Plato: A study in chronology, *Isis* 32 (1940/1947), 248–62.

Schwenter, D., *GP* = *Geometriae Practicae Novae et Auctae Tractatus*, Nurenberg, 1617(?).

Schwenter, D., *DPM* = *Deliciae Physico-Mathematicae*, Nurenberg, 1636.

Sedley, D., EA = The end of the Academy, *Phronesis* (1981), 67–75.

Sedley, D., PGT = Pythagoras the grammar teacher (PBrLibAdd Ms 37516), in *Papiri filosofici: miscellenea di studi* ii, Academia toscana di scienze e lettere "La Colombaria", Florence: Olschi, 1998.

Sedley, D., PGTDA = Pythagoras the grammar teacher and Didyrmon the adulterer, *Hyperboreus*, to appear.

Series, C., GMN = The geometry of Markoff numbers, *Mathematical Intelligencer* 7 (1983), 20–9.

Serret, J.-A., *CAS* = *Cours d'Algèbre Supérieure*, 3rd edn., 3 vols., Paris: Gauthier-Villars, 1866.

Sethe, K. H., *ZZAA* = *Von Zahlen und Zahlworten bei Alten Ägypter*, Strassburg: Trübner, 1916.

Sharples, R. W., NETDP = Nemesius of Emesa and some theories of divine providence, *Vigilae Christianae* 37 (1983), 141–56.

Sharples, R. W., MPM = More on Plato's *Meno* 82c–3, *Phronesis* 34 (1989), 220–6.

Sijpesteijn, P., WTMC = A wooden tablet in the Moen Collection, *Chronique d'Égypte* III (New Series 56) (1981), 99–101.

Smith, H. J. S., *Papers* = *The Collected Mathematical Papers of Henry John Stephen Smith*, ed. J. W. L. Glaisher, 2 vols., repr. New York: Chelsea, 1965.

Smith, H. J. S., NCF = Note on continued fractions, *Messenger of Mathematics* (Series 2) 6 (1876), 1–14, repr. in *Papers* ii, 135–47.

Smyly, J. G., EAGL = The employment of the alphabet in Greek logistic, pp. 515–30 in *Mélanges Nicole, Recueil des Memoirs de Philologie*, Geneva: Kundig, 1905.

Smyly, J. G., NTS = Notes on Theon of Smyrna, *Hermathena* 14 (1907), 261–79.

Stark, H. M., *INT* = *An Introduction to Number Theory*, Chicago: Markham, 1970, repr. Cambridge, Massachusetts: MIT Press, 1978.

Stark, H. M., EECFFB = An explanation of some exotic continued fractions found by

Brillhart, pp. 21–35 in *Computers in Number Theory, Proceedings of the Science Research Council Atlas Symposium 2, Oxford 1969*, ed. A. O. L. Atkin & B. J. Birch, London & New York: Academic Press, 1971.

Steele, A. D., URZLGM = Über die Rolle von Zirkel und Lineal in der griechischen Mathematik, *Quellen und Studien zur Geschichte der Mathematik, Astronomie und Physik*, Abteilung B: *Studien* 3 (1936), 287–369.

Steele, D. A., MRCP = A mathematical reappraisal of the Corpus Platonicum, *Scripta Mathematica* 17 (1951), 173–89.

Stevin–Struik, *PW = The Principal Works of Simon Stevin*, 5 vols: vol. i, *Mechanics*, ed. E. J. Dijksterhuis; vol. ii in two parts, *Mathematics*, ed. D. J. Struik, ..., Amsterdam: Swets & Zeitlinger, 1955–66.

Szegin, F., *GAS = Geschichte des arabischen Schrifttums*, vol. v: *Mathematik*, Leyden: Brill, 1974.

Taisbak, C. M., *DL = Division and Logos, A Theory of Equivalent Couples and Sets of Integers*, Odense: Odense University Press, 1971.

Taisbak, C. M., *CQ = Coloured Quadrangles: A Guide to the Tenth Book of Euclid's Elements*, Opuscula Graecolatina 24, Copenhagen: Museum Tusculanum Press, 1982. (Also see Knorr, RTCQ.)

Taisbak, C. M., EET = Eleven Eighty-Thirds. Ptolemy's reference to Eratosthenes in *Almagest* I.12, *Centaurus* 27 (1984), 165–7.

Tait, J. G., Préaux, C., & others, *GOBLO = Greek Ostraca in the Bodleian Library at Oxford and Various Other Collections* [= O. Bodl. or O. Tait], 3 vols., London: Egypt Exploration Society, 1930–64.

Tannery, P., *Memoires = Memoires Scientifiques*, 17 vols., Paris: Gauthier-Villars, 1912–50.

Tannery, P., SSCG = Sur la symbole de soustration chez le Grecs, *Bibliotheca Mathematica* 5 (1904–5), 342–9, repr. in *Memoires* xiii, 208–12.

Theon of Smyrna–Hiller, *ERMLPU = Expositio Rerum Mathematicarum ad Legendum Platonem Utilium*, ed. E. Hillier, Leipzig: Teubner, 1878.

Theon of Smyrna–Lawlor, *MUUP = Mathematics Useful for Understanding Plato*, tr. from the Greek/French edn. of J. Dupuis (1892) by R. & D. Lawlor, Secret Doctrine Reference Series, San Diego: Wizards Bookshelf, 1979. (See the review by R. D. McKirahan, *Historia Mathematica* 9 (1982), 100–4.)

Thesleff, H., *IPWHP = An Introduction to the Pythagorean Writings of the Hellenistic Period*, Åbo: Academie, 1961.

Thesleff, H., *PTHP = The Pythagorean Texts of the Hellenistic Period*, Åbo: Academie, 1965.

Thesleff, H., PC = Platonic chronology, *Phronesis* 34 (1989), 1–26.

Thomae, L. W., KGQ = Über die Kettenbrüchentwicklung der Gauss'chen Quotienten ..., *Journal für die reine und angewandte Mathematik* 67 (1867), 299–309.

Thomas [= Bulmer-Thomas], I., *SIHGM = Selections Illustrating the History of Greek Mathematics*, 2 vols., Loeb Classical Library, Cambridge, Massachusetts & London: Harvard University Press & Heinemann, vol. 1, 1939, supplemented repr. 1980; vol. 2, 1941. (See the review by O. Neugebauer, *American Journal of Philology* 64 (1943), 452–7.)

Tod, M. N., *AGNS = Ancient Greek Numerical Systems: Six Studies,* Chicago: Ares, 1979. (Six articles on the epigraphical evidence for numerical systems, originally published in the *Annual of the British School at Athens* vols. 18, 28, 37, 45, & 49 and the *Journal of Hellenic Studies* 33.)

Tod, M. N., GNN = The Greek numeral notation, *Annual of the British School of Athens* 18 (1911/12), 98–132, repr. in Tod, *AGNS*, 1–35.

Tod, M. N., LLGI = Letter-labels in Greek inscriptions, *Annual of the British School at Athens* 49 (1954), 1–8, repr. in Tod, *AGNS*, 98–105.

Toomer, G. J., MA = Mathematics and astronomy, pp. 27–54 in *The Legacy of Europe*, ed. J. R. Lewis, Oxford: Clarendon Press, 1971.

Toomer, G. J., LGMWAT = Lost Greek mathematical works in the Arabic tradition, *Mathematical Intelligencer* 6 (1984), 32–8.

Tóth, I., PCA = Das Parallelenproblem im Corpus Aristotelicum, *Archive for History of Exact Sciences* 3 (1967), 249–422.

Turner, E. G., ABFFCBC = *Athenian Books in the Fifth and Fourth Centuries B. C.* (An inaugural lecture, May 1951), London: University College & H. K. Lewis, 1952, 2nd edn. 1977.

Turner, E. G., FODMS = Four obols a day men at Saqqâra, pp. 573–7 & Plate X in *Hommages à Claire Préaux,* ed. J. Bingen, G. Cambier, & G. Nachtergael, Brussels: Editions de l'Université de Bruxelles, 1975.

Turner, E. G., GMAW = *Greek Manuscripts of the Ancient World*, Oxford: Clarendon Press, 1971.

Turner, E. G., PW = *The Papyrologist at Work,* Durham, North Carolina: Greek, Roman and Byzantine Monographs 6, 1973.

Turner, E. G., CCOS = A commander-in-chief's order from Saqqâra, *Journal of Egyptian Archaeology* 60 (1974), 239–42.

Turner, E. G., GP = *Greek Papyri: An Introduction*, 2nd edn., Oxford: Clarendon Press, 1980.

Turner, E. G., Fowler, D. H., Koenen, L., & Youtie, L. C., EEI = Euclid, *Elements*, I, Definitions 1–10 (P. Mich. iii 143), *Yale Classical Studies* 28 (1985), 13–24.

Turner, E. G. & Neugebauer, O., GDNM = Gymnasium debts and new moons [= P. Ryl. iv 589], *Bulletin of the John Rylands Library* 32 (1949), 80–96.

Turner, E. G., Tsantsanoglou, K., & Parássoglou, G. M., ODP = On the Derveni papyrus, *Gnomon* 54 (1982), 855–6.

Varberg, D. E., PTR = Pick's theorem revisited, *American Mathematical Monthly* 92 (1985), 584–7.

Vogel, K., BGL = Beiträge zur griechischen Logistik, *Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Abteilung* (1936), 357–472.

Vogel, K., BIZB = Buchstabenrechnung und indische Ziffern in Byzanz, pp. 660–4 in *Akten des XI. Internationalen Byzantinistenkongresses, München 1958*, Munich: Beck, 1960.

van der Waerden, B. L., SA = *Science Awakening*, tr. A. Dresden of *Ontwakende Wetenschap* (1950), Groningen: Noordhoff, 1954.

van der Waerden, B. L., GAC = Greek astronomical calendars: I, The parapegma of Euctemon; II, Callippos and his calendar; III, The calendar of Dionysios; IV, The Parapegma of the Egyptians and their "Perpetual Tables", *Archive for History of Exact Sciences* 29 (1984), 101–14, 115–24, 125–30 & 32 (1985), 95–104.

Wall, H. S., ATCF = *Analytic Theory of Continued Fractions*, New York: Van Nostrand, 1948, repr. New York: Chelsea, 1973.

Wallis, J., AI = *Arithmetica Infinitorum, sive nova methodus inquirendi in curvilineorum quadraturam aliaque difficiliora matheseos problemata*, Oxford, 1656.

Wallis, J., CE = *Commercium Epistolicum de quaestionbus quibusdam mathematicis nuper habitum inter nobilissimos viros . . .* , Oxford, 1658. (Correspondence concerning

Pell's equation between Brouncker, Digby, Fermat, Frenicle de Bessy, van Schooten, & Wallis.)

Wallis, J., *TA = A Treatise of Algebra, both Historical and Practical, shewing the Original, Progress, and Advancement thereof, from time to time, and by what Steps it hath attained to the Heighth at which now it is. . . .* , London, 1685.

Waschkies, H.-J., ENH = Eine neue Hypothese zur Entdeckung der inkommensurablen Grössen durch die Griechen, *Archive for History of Exact Sciences* 7 (1971), 325–53.

Wasserstein, A., THTN = Theaetetus and the history of the theory of numbers, *Classical Quarterly* 8 (1958), 165–79.

Waterhouse, W. C., DRS = The discovery of the regular solids, *Archive for History of Exact Sciences* 9 (1972), 212–21.

Watson, J. S., *LRP = The Life of Richard Porson, M.A.,* London: Longman, 1861.

Weil, A., *NT = Number Theory, An Approach through History from Hammurapi to Legendre*, Boston: Birkhäuser, 1984.

Weitzmann, K., *ABI = Ancient Book Illumination*, Cambridge, Massachusetts: Harvard University Press, 1959.

Westfall, R. S., *NR = Never at Rest, A Biography of Isaac Newton*, Cambridge: Cambridge University Press, 1980, repr. with additions 1983.

Whiteside, D. T., PMTLSC = Patterns of mathematical thought in the later seventeenth century, *Archive for History of Exact Sciences* 1 (1961), 179–388.

Whittaker, E. T., TCF = On the theory of continued-fractions, *Proceedings of the Royal Society of Edinburgh* 36 (1915–16), 243–55.

Whycherly, R. E., PAPS = Peripatos: The Athenian philosophical scene, *Greece and Rome* (2nd Series) 8 (1961), 152–63 & 9 (1962), 2–21.

Wilson, N. G., TBS = Three Byzantine scribes, *Greek, Roman and Byzantine Studies* 14 (1973), 223–8.

Wilson, N. G., MP = Miscellanea palaeographica, *Greek, Roman and Byzantine Studies* 22 (1981), 395–404.

Wilson, N. G., *SB = Scholars of Byzantium*, London: Duckworth, 1983.

Winter, J. G., Robbins, F. E., & others, *PUMC = Papyri from the University of Michigan Collection* [= P. Mich.], 15 vols to 1982, Ann Arbor: University of Michigan Press, 1931–.

Wolfskill, J., GBPQCN = A growth bound on the partial quotients of cubic numbers, *Journal für die reine und angewandte Mathematik* 346 (1984), 129–40.

Youtie, H. C., *TCDP = The Textual Criticism of Documentary Papyri: Prolegomena*, 2nd edn., London: Bulletin of the Institute of Classical Studies Supplement 33, 1974.

Zeeman, E. C., GG = Gears from the Greeks, *Proceedings of the Royal Institution* 58 (1986), 139–56.

Zeuthen, H. G., CLAEE = Sur la constitution des livres arithmétiques des Éléments d'Euclide et leur rapport à la question de l'irrationalité, *Oversigt over det Kgl. Danske Videnskabernes Selskabs Forhandlinger* [= *Proceedings of the Danish Academy of Sciences*] 5 (1910), 395–435.

Zhmud, L. Y., AN = "All is number"? "Basic doctrine" of Pythagoreanism reconsidered, *Phronesis* 34 (1989), 270–92.

This index covers only passages from ancient authors, and does not include incidental references. Papyri (Egyptian and Greek) are all listed together under 'Papyri', and minuscule manuscripts under 'Manuscripts'.

# INDEX OF NAMES

This index does not cover the following: the Preface and other front material; the Bibliography; the names of editors and translators of texts, unless there is some special reference to them; names that are only mentioned incidentally; and terms such as Farey series, the *Parminedes* proposition, Pythagoras' theorem, etc., which can be found in the General Index.

# GENERAL INDEX