The aim of this book is to explain the shape of Greek mathematical thinking. It can be read on three levels: first as a description of the practices of Greek mathematics; second as a theory of the emergence of the deductive method; and third as a case-study for a general view on the history of science. The starting point for the enquiry is geometry and the lettered diagram. Reviel Netz exploits the mathematicians' practices in the construction and lettering of their diagrams, and the continuing interaction between text and diagram in their proofs, to illuminate the underlying cognitive processes. A close examination of the mathematical use of language follows, especially mathematicans' use of repeated formulae. Two crucial chapters set out to show how mathematical proofs are structured and explain why Greek mathematical practice manages to be so satisfactory. A final chapter looks into the broader historical setting of Greek mathematical practice.

Reviel Netz is a Research Fellow at Gonville and Caius College, Cambridge, and an Affiliated Lecturer in the Faculty of Classics.

IDEAS IN CONTEXT 51

THE SHAPING OF DEDUCTION IN GREEK MATHEMATICS

IDEAS IN CONTEXT

The books in this series will discuss the emergence of intellectual traditions and of related new disciplines. The procedures, aims and vocabularies that were generated will be set in the context of the alternatives available within the contemporary frameworks of ideas and institutions. Through detailed studies of the evolution of such traditions, and their modification by different audiences, it is hoped that a new picture will form of the development of ideas in their concrete contexts. By this means, artificial distinctions between the history of philosophy, of the various sciences, of society and politics, and of literature may be seen to dissolve.

The series is published with the support of the Exxon Foundation.

A list of books in the series will be found at the end of the volume.

THE SHAPING OF DEDUCTION IN GREEK MATHEMATICS

A Study in Cognitive History

REVIEL NETZ

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge CB2 1RP, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge, CB2 2RU, United Kingdom http://www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011–4211, USA http://www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1999

Printed in the United Kingdom at the University Press, Cambridge

Typeset in 11/12½pt Baskerville No. 2 [GC]

A catalogue record for this book is available from the British Library

Library of Congress cataloguing in publication data

Netz, Reviel.
The shaping of deduction in Greek mathematics: a study in cognitive history / Reviel Netz.
p. cm. – (Ideas in context; 51)
Includes bibliographical references and index.
ISBN 0 521 62279 4 (hardback)
1. Mathematics, Greek. 2. Logic. I. Title. II. Series.
QA27.G8N47 1999
510′.938–dc21 98–20463 CIP

ISBN 0 521 62279 4 hardback

To Maya

Contents

Preface

This book was conceived in Tel Aviv University and written in the University of Cambridge. I enjoyed the difference between the two, and am grateful to both.

The question one is most often asked about Greek mathematics is: 'Is there anything left to say?' Indeed, much has been written. In the late nineteenth century, great scholars did a stupendous work in editing the texts and setting up the basic historical and mathematical framework. But although the materials for a historical understanding were there, almost all the interpretations of Greek mathematics offered before about 1975 were either wildly speculative or ahistorical. In the last two decades or so, the material has finally come to life. A small but highly productive international community of scholars has set up new standards of precision. The study of Greek mathematics today can be rigorous as well as exciting. I will not name here the individual scholars to whom I am indebted. But I can – I hope – name this small community of scholars as a third institution to which I belong, just as I belong to Tel Aviv and to Cambridge. Again I can only express my gratitude.

So I have had many teachers. Some were mathematicians, most were not. I am not a mathematician, and this book demands no knowledge of mathematics (and only rarely does it demand some knowledge of Greek). Readers may feel I do not stress sufficiently the value of Greek mathematics in terms of mathematical content. I must apologise – I owe this apology to the Greek mathematicians themselves. I study form rather than content, partly because I see the study of form as a way into understanding the content. But this content – those discoveries and proofs made by Greek mathematicians – are both beautiful and seminal. If I say less about these achievements, it is because I have looked elsewhere, not because my appreciation of them is not as keen as it should be. I have stood on the shoulders of giants –

to get a good look, from close quarters, at the giants themselves. And if I saw some things which others before me did not see, this may be because I am more short-sighted.

I will soon plunge into the alphabetical list. Three names must stand out – and they happen to represent the three communities mentioned above. Sabetai Unguru first made me read and understand Greek mathematics. Geoffrey Lloyd, my Ph.D. supervisor, shaped my view of Greek intellectual life, indeed of intellectual life in general. David Fowler gave innumerable suggestions on the various drafts leading up to this book – as well as giving his inspiration.

A British Council Scholarship made it possible to reach Cambridge prior to my Ph.D., as a visiting member at Darwin College. Awards granted by the ORS, by the Lessing Institute for European History and Civilization, by AVI and, most crucially, by the Harold Hyam Wingate Foundation made it possible to complete graduate studies at Christ's College, Cambridge. The book is a much extended and revised version of the Ph.D. thesis, prepared while I was a Research Fellow at Gonville and Caius College. It is a fact, not just a platitude, that without the generosity of all these bodies this book would have been impossible. My three Cambridge colleges, in particular, offered much more than can be measured.

I owe a lot to Cambridge University Press. Here, as elsewhere, I find it difficult to disentangle 'form' from 'content'. The Press has contributed greatly to both, and I wish to thank, in particular, Pauline Hire and Margaret Deith for their perseverance and their patience.

The following is the list – probably incomplete – of those whose comments influenced directly the text you now read (besides the three mentioned already). My gratitude is extended to them, as well as to many others: R. E. Aschcroft, Z. Bechler, M. F. Burnyeat, K. Chemla, S. Cuomo, A. E. L. Davis, G. Deutscher, R. P. Duncan-Jones, P. E. Easterling, M. Finkelberg, G. Freudental, C. Goldstein, I. Grattan-Guinness, S. J. Harrison, A. Herreman, J. Hoyrup, E. Hussey, P. Lipton, I. Malkin, J. Mansfeld, I. Mueller, J. Ritter, K. Saito, J. Saxl, D. N. Sedley, B. Sharples, L. Taub, K. Tybjerg, B. Vitrac, L. Wischik.[1]

[1] I have mentioned above the leap made in the study of Greek mathematics over the last two decades. This owes everything to the work of Wilbur Knorr, who died on 18 March 1997, at the age of 51. Sadly, he did not read this book – yet the book would have been impossible without him.

Abbreviations

GREEK AUTHORS

Abbreviation	Work (standard title)	Author
de Aedific.	*de Aedificiis*	Procopius
Amat.	*Amatores*	[Plato]
APo.	*Analytica Posteriora*	Aristotle
APr.	*Analytica Priora*	Aristotle
Av.	*Aves*	Aristophanes
Cat.	*Categoriae*	Aristotle
CF	*On Floating Bodies*	Archimedes
CS	*On Conoids and Spheroids*	Archimedes
DC	*Measurement of Circle*	Archimedes
D.L.	Lives of Philosophers	Diogenes Laertius
EE	*Ethica Eudemia*	Aristotle
El. Harm.	*Elementa Harmonica*	Aristoxenus
de Eloc.	*Demetrius on Style*	Demetrius
EN	*Ethica Nicomachea*	Aristotle
Epin.	*Epinomis*	[Plato] (Plato?)
Euthd.	*Euthydemus*	Plato
Euthyph.	*Euthyphro*	Plato
Grg.	*Gorgias*	Plato
HA	*Historia Animalium*	Aristotle
Hip. Mai.	*Hippias Maior*	Plato
Hip. Min.	*Hippias Minor*	Plato
In de Cael.	*In Aristotelis de Caelo Commentaria*	Simplicius
In Eucl.	*In Euclidem*	Proclus
In Phys.	*In Aristotelis Physica Commentaria*	Simplicius

Abbreviation	Work (standard title)	Author
de Int.	*de Interpretatione*	Aristotle
In SC	*In Archimedes' SC*	Eutocius
In Theaetet.	*Anonymi Commentarius In Platonis Theaetetum*	Anonymous
Lgs.	*de Legibus*	Plato
Mech.	*Mechanica*	[Aristotle]
Mem.	*Memorabilia*	Xenophon
Metaph.	*Metaphysica*	Aristotle
Meteor.	*Meteorologica*	Aristotle
Meth.	*The Method*	Archimedes
Nu.	*Nubes*	Aristophanes
Ort.	*Risings and Settings*	Autolycus
Parm.	*Parmenides*	Plato
de Part.	*de Partibus Animalium*	Aristotle
PE	*Plane Equilibria*	Archimedes
Phaedr.	*Phaedrus*	Plato
Phys.	*Physica*	Aristotle
QP	*Quadrature of the Parabola*	Archimedes
Rep.	*Republica*	Plato
SC	*On Sphere and Cylinder*	Archimedes
SE	*Sophistici Elenchi*	Aristotle
SL	*Spiral Lines*	Archimedes
Theaetet.	*Theaetetus*	Plato
Tim.	*Timaeus*	Plato
Vit. Alc.	*Vita Alcibiadis*	Plutarch
Vita Marc.	*Aristotelis Vita Marciana*	Anonymous
Vita Pyth.	*de Vita Pythagorica*	Iamblichus

ROMAN AUTHORS

Abbreviation	Work (standard title)	Author
Ann.	*Annales*	Tacitus
Nat. Hist.	*Naturalis Historia*	Pliny the Elder
ND	*de Natura Deorum*	Cicero
de Rep.	*de Republica*	Cicero
Tusc.	*Tusculanae Disputationes*	Cicero

DOCUMENTARY SOURCES

Abbreviation	Standard title
BGU	*Berliner griechische Urkunden*
FD	*Fouilles de Delphes*
ID	*Inscriptions Délos*
IG	*Inscriptionae Graecae*
IGChEg.	*Inscriptionae Graecae* (Christian Egypt)
IK	*Inschriften aus Kleinasien*
Ostras	*Ostraka* (Strasburg)
P. Berol.	*Berlin Papyri*
PCair.Zen.	*Zenon Papyri*
PFay.	*Fayum Papyri*
P. Herc.	*Herculaneum Papyri*
PHerm Landl.	*Landlisten aus Hermupolis*
POxy.	*Oxyrhynchus Papyri*
YBC	*Yale Babylonian Collection*

OTHER ABBREVIATIONS

Abbreviation	Reference (in bibliography)
CPF	*Corpus dei Papiri Filosofici*
DK	Diels–Kranz, *Fragmente der Vorsokratiker*
KRS	Kirk, Raven and Schofield (1983)
L&S	Long and Sedley (1987)
LSJ	Liddell, Scott and Jones (1968)
Lewis and Short	Lewis and Short (1966)
TLG	*Thesaurus Linguae Graecae*
Usener	Usener (1887)

NOTE ON GENDER

When an indefinite reference is made to ancient scholars – who were predominantly male – I use the masculine pronoun. The sexism was theirs, not mine.

The Greek alphabet

Capital *approximately the form used in ancient writing*	Lower case *a form used in modern texts*	Name of letter
Α	α	Alpha
Β	β	Bēta
Γ	γ	Gamma
Δ	δ	Delta
Ε	ε	Epsilon
Ζ	ζ	Zēta
Η	η	Ēta
Θ	θ	Thēta
Ι	ι	Iōta
Κ	κ	Kappa
Λ	λ	Lambda
Μ	μ	Mu
Ν	ν	Nu
Ξ	ξ	Xi
Ο	ο	Omicron
Π	π	Pi
Ρ	ρ	Rhō
Σ	σ ς[1]	Sigma
Τ	τ	Tau
Υ	υ	Upsilon
Φ	φ	Phi
Χ	χ	Chi
Ψ	ψ	Psi
Ω	ω	Ōmega

[1] A modern form for the letter in final position.

Note on the figures

As is explained in chapter 1, most of the diagrams in Greek mathematical works have not yet been edited from manuscripts. The figures in modern editions are reconstructions made by modern editors, based on their modern understanding of what a diagram should look like. However, as will be argued below, such an understanding is culturally variable. It is therefore better to keep, as far as possible, to the diagrams as they are found in Greek manuscripts (that is, generally speaking, in Byzantine manuscripts). While no attempt has been made to prepare a critical edition of the Greek mathematical diagrams produced here, almost all the figures have been based upon an inspection of at least some early manuscripts in which their originals appear, and I have tried to keep as close as possible to the visual code of those early diagrams. In particular, the reader should forgo any assumptions about the lengths of lines or the sizes of angles: unequal lines and angles may appear equal in the diagrams and vice versa.

In addition to the ancient diagrams (which are labelled with the original Greek letters), a few illustrative diagrams have been prepared for this book. These are distinguished from the ancient diagrams by being labelled with Latin letters or with numerals.

While avoiding painterly effects, ancient diagrams possess considerable aesthetic value in their austere systems of interconnected, labelled lines. I wish to take this opportunity to thank Cambridge University Press for their beautiful execution of the diagrams.

CHAPTER 1

The lettered diagram

PLAN OF THE CHAPTER

That diagrams play a crucial role in Greek mathematics is a fact often alluded to in the modern literature, but little discussed. The focus of the literature is on the verbal aspect of mathematics. What this has to do with the relative roles of the verbal and the visual in our culture, I do not claim to know. A description of the practices related to Greek mathematical diagrams is therefore called for. It will prove useful for our main task, the shaping of deduction.

The plan is: first, a brief discussion of the material implementation of diagrams, in section 1. Some practices will be described in section 2. My main claims will be that (a) the diagram is a necessary element in the reading of the text and (b) the diagram is the metonym of mathematics. I will conclude this section with a discussion of the semiotics of lettered diagrams. Section 3 will describe some of the historical contexts of the lettered diagram. Section 4 is a very brief summary.

This chapter performs a trick: I talk about a void, an absent object, for the diagrams of antiquity are not extant, and the medieval diagrams have never been studied as such.[1] However, not all hope is lost. The texts – whose transmission is relatively well understood – refer to diagrams in various ways. On the basis of these references, observations concerning the practices of diagrams can be made. I thus start from the text, and from that base study the diagrams.

[1] The critical edition most useful from the point of view of the ancient diagrams is Mogenet (1950). Some information is available elsewhere: the Teubner edition of the *Data*, for instance, is very complete on lettering; Jones's edition of Pappus and Clagett's edition of the Latin Archimedes are both exemplary, and Janus, in *Musici Graeci*, is brief but helpful. Generally, however, critical apparatuses do not offer substantial clues as to the state of diagrams in manuscripts.

1 THE MATERIAL IMPLEMENTATION OF DIAGRAMS

There are three questions related to the material implementation of diagrams: first, the contexts in which diagrams were used; second, the media available for drawing; finally, there is the question of the technique used for drawing diagrams – and, conversely, the technique required for looking at diagrams (for this is a technique which must be learned in its own right).

One should appreciate the distance lying between the original moment of inspiration, when a mathematician may simply have imagined a diagram, and our earliest extensive form of evidence, parchment codices. In between, moments of communication have occurred. What audience did they involve?

First, the 'solitaire' audience, the mathematician at work, like someone playing patience. Ancient images pictured him working with a diagram.[2] We shall see how diagrams were the hallmark of mathematical activity and, of course, a mathematician would prefer to have a diagram in front of him rather than playing the game out in his mind. It is very probable, then, that the process of discovery was aided by diagrams.

The contexts for communicating mathematical results must have been very variable, but a constant feature would have been the small numbers of people involved.[3] This entails that, very often, the written form of communication would be predominant, simply because fellow mathematicians were not close at hand. Many Greek mathematical works were originally set down within letters. This may be a trivial point concerning communicative styles, or, again, it may be significant. After all, the addressees of mathematical works, leaving aside the *Arenarius*,[4] are not the standard recipients of letters, like kings, friends or relations. They seem to have been genuinely interested mathematicians, and the inclusion of mathematics within a letter could therefore be an indication that works were first circulated as letters.[5]

[2] This is the kernel of the myth of Archimedes' death in its various forms (see Dijksterhuis (1938) 30ff.). Cicero's evocation of Archimedes 'from the dust and drawing-stick' (*Tusc.* v.64) is also relevant. Especially revealing is Archimedes' tomb, mentioned in the same context. What is Einstein's symbol? Probably '$E = MC^2$'. Archimedes' symbol was a diagram: '*sphaerae figura et cylindri*' (ibid. v.65).

[3] See the discussion in chapter 7, subsection 2.2 below (pp. 282–92).

[4] As well as Eratosthenes' fragment in Eutocius.

[5] Pappus' dedicatees are less easy to identify, but Pandrosion, dedicatee of book III, for instance, seems to have been a teacher of mathematics; see Cuomo (1994) for discussion.

Not much more is known, but the following observation may help to form some a priori conclusions. The lettered diagram is not only a feature of Greek mathematics; it is a predominant feature. Alternatives such as a non-lettered diagram are not hinted at in the manuscripts.[6] There is one exception to the use of diagrams – the *di' arithmōn*, 'the method using numbers'. While in general arithmetical problems are proved in Greek mathematics by geometrical means, using a diagram, sometimes arithmetical problems are tackled as arithmetical. Significantly, even this is explicitly set up as an exception to a well-defined rule, the *dia grammōn*, 'the method using lines'.[7] The diagram is seen as the rule from which deviations may (very rarely) occur.

It is therefore safe to conclude that Greek mathematical exchanges, as a rule, were accompanied by something like the lettered diagram. Thus an exclusively oral presentation (excluding, that is, even a diagram) is practically ruled out. Two methods of communication must have been used: the fully written form, for addressing mathematicians abroad, and (hypothetically) a semi-oral form, with some diagram, for presentation to a small group of fellow mathematicians in one's own city.

1.1 The media available for diagrams

It might be helpful to start by considering the media available to us. The most important are the pencil/paper, the chalk/blackboard and (gaining in importance) the computer/printer. All share these characteristics: simple manipulation, fine resolution, and ease of erasing and rewriting. Most of the media available to Greeks had none of these, and none had ease of erasing and rewriting.

The story often told about Greek mathematicians is that they drew their diagrams in sand.[8] A variation upon this theme is the dusted

6 I exclude the fragment of Hippocrates of Chios, which may of course reflect a very early, formative stage. I also ignore for the moment the papyrological evidence. I shall return to it in n. 31 below.

7 I shall return to this distinction below, n. 61.

8 Sand may be implied by the situation of the geometry lesson in the *Meno*, though nothing explicit is said; if the divided line in the *Republic* was drawn in sand, then Cephalus' house must have been fairly decrepit. Aristotle refers to drawing in γῇ – e.g. *Metaph.* 1078a20; it may well be that he has the *Meno* in mind. Cicero, *de Rep.* I.28–9 and Vitruvius VI.1.1, have the following tale: a shipwrecked philosopher deduces the existence of life on the island on whose shores he finds himself by (Vitruvius' phrase) *geometrica schemata descripta* – one can imagine the wet sand on the shore as a likely medium. The frontispiece to Halley's edition of the *Conics*, reproduced as the cover of Lloyd (1991), is a brilliant *reductio ad absurdum* of the story.

surface. This is documented very early, namely, in Aristophanes' *Clouds*;[9] Demetrius, a much later author, misremembered the joke and thought it was about a wax tablet[10] – a sign of what the typical writing medium was. Indeed, the sand or dusted surface is an extremely awkward solution. The ostrakon or wax tablet would be sufficient for the likely size of audience; a larger group would be limited by the horizontality of the sand surfaces. And one should not think of sand as directly usable. Sand must be wetted and tamped before use, a process involving some exertion (and mess).[11] Probably the hard work was done by Euclid's slaves, but still it is important to bear in mind the need for *preparation* before each drawing. Sand is a very cheap substitute for a drawing on wood (on which see below), but it is not essentially different. It requires a similar amount of preparation. It is nothing like the immediately usable, erasable blackboard.

The possibility of large-scale communication should be considered – and will shed more light on the more common small-scale communication. There is one set of evidence concerning forms of presentation to a relatively large audience: the evidence from Aristotle and his followers in the peripatetic school.

Aristotle used the lettered diagram in his lectures. The letters in the text would make sense if they refer to diagrams – which is asserted in a few places.[12] Further, Theophrastus' will mentions maps on *pinakes* (for which see below) as part of the school's property.[13] Finally, Aristotle refers to *anatomai*, books containing anatomical drawings, which students were supposed to consult as a necessary complement to the lecture.[14]

What medium did Aristotle use for his mathematical and semi-mathematical diagrams? He might have used some kind of prepared tablets whose medium is nowhere specified.[15] As such tablets were,

[9] Ashes, sprinkled upon a table: Aristophanes, *Nu.* 177. To this may be added later texts, e.g. Cicero, *Tusc.* v.64; *ND* II.48.

[10] Demetrius, *de Eloc.* 152.

[11] I owe the technical detail to T. Riehl. My own experiments with sand and ashes, wetted or not, were unmitigated disasters – this again shows that these surfaces are not as immediately usable as are most modern alternatives.

[12] E.g. *Meteor.* 363a25–6, *APr.* 41b14. Einarson (1936) offers the general thesis that the syllogism was cast in a mathematical form, diagrams included; while many of his individual arguments need revision, the hypothesis is sound.

[13] D.L. v.51–2. [14] See Heitz (1865) 70–6.

[15] Jackson (1920) 193 supplies the evidence, and a guess that Aristotle used a *leukoma*, which is indeed probable; but Jackson's authority should not obscure the fact that this is no more than a guess.

presumably, portable, they could not be just graffiti on the Lyceum's walls. Some kind of special surface is necessary, and the only practical option was wood, which is the natural implication of the word *pinax*. To make such writing more readable, the surface would be painted white, hence the name *leukōma*, 'whiteboard' – a misleading translation. Writings on the 'whiteboard', unlike the blackboard, were difficult to erase.[16]

Two centuries later than Aristotle, a set of mathematical – in this case astronomical – *leukōmata* were put up as a dedication in a temple in Delos.[17] This adds another tiny drop of probability to the thesis that wide communication of mathematical diagrams was mediated by these whiteboards.[18] On the other hand, the *anatomai* remind us how, in the very same peripatetic school, simple diagrams upon (presumably) papyrus were used instead of the large-scale *leukōma*.

Closer in nature to the astronomical tables in Delos, Eratosthenes, in the third century BC, set up a mathematical column: an instrument on top, below which was a résumé of a proof, then a diagram and finally an epigram.[19] This diagram was apparently inscribed in stone or marble. But this display may have been the only one of its kind in antiquity.[20]

The development envisaged earlier, from the individual mathematician thinking to himself to the parchment codex, thus collapses into small-scale acts of communication, limited by a small set of media, from the dusted surface, through wax tablets, ostraka and papyri, to the whiteboard. None of these is essentially different from a diagram as it appears in a book. Diagrams, as a rule, were not drawn on site. The limitations of the media available suggest, rather, the preparation of the diagram prior to the communicative act – a consequence of the inability to erase.

[16] See Gardthausen (1911) 32–9.

[17] *ID* 3. 1426 face B. col. II.50ff.; 3. 1442 face B. col. II.40ff.; 3. 1443 face B. col. II.108ff.

[18] It is also useful to see that, in general, wood was an important material in elementary mathematical education, as the archaeological evidence shows; Fowler (1987) 271–9 has 69 items, of which the following are wooden tablets: 14, 16, 18, 24, 25, 39, 42, 44, 45, 59.

[19] Eutocius, *In SC* II.94.8–14.

[20] Allow me a speculation. Archimedes' *Arenarius*, in the manuscript tradition, contains no diagrams. Of course the diagrams were present in some form in the original (which uses the lettered convention of reference to objects). So how were the diagrams lost? The work was addressed to a king, hence, no doubt, it was a luxury product. Perhaps, then, the diagrams were originally on separate *pinakes*, drawn as works of art in their own right?

1.2 Drawing and looking

In terms of optical complexity, there are four types of objects required in ancient mathematics.

1. Simple 2-dimensional configurations, made up entirely by straight lines and arcs;
2. 2-dimensional configurations, requiring more complex lines, the most important being conic sections (ellipse, parabola and hyperbola);
3. 3-dimensional objects, excluding:
4. Situations arising in the theory of spheres ('sphaerics').

Drawings of the first type were obviously mastered easily by the Greeks. There is relatively good papyrological evidence for the use of rulers for drawing diagrams.[21] The extrapolation, that compasses (used for vase-paintings, from early times)[22] were used as well, suggests itself.

On the other hand, the much later manuscripts do not show any technique for drawing non-circular curved lines, which are drawn as if they consist of circular arcs.[23] This use of arcs may well have been a feature of ancient diagrams as well.

Three-dimensional objects do not require perspective in the strict sense, but rather the practice of foreshortening individual objects.[24] This was mastered by some Greek painters in the fifth century BC;[25] an achievement not unnoticed by Greek mathematicians.[26]

Foreshortening, however, does little towards the elucidation of spherical situations. The symmetry of spheres allows the eye no hold on which to base a foreshortened 'reading'. In fact, some of the diagrams for spherical situations are radically different from other, 'normal' diagrams. Rather than providing a direct visual representation, they employ

[21] See Fowler (1987), plates between pp. 202 and 203 – an imperative one should repeat again and again. For this particular point, see especially Turner's personal communication on *PFay.* 9, p. 213.

[22] See, e.g. Noble (1988) 104–5 (with a fascinating reproduction on p. 105).

[23] Toomer (1990) lxxxv.

[24] In fact – as pointed out to me by M. Burnyeat – strictly perspectival diagrams would be less useful. A useful diagram is somewhat schematic, suggesting objective geometric relations rather than subjective optical impressions.

[25] White (1956), first part.

[26] Euclid's *Optics* 36 proved that wheels of chariots appear sometimes as circles, sometimes as elongated. As pointed out by White (1956: 20), Greek painters were especially interested in the foreshortened representation of chariots, sails and shields. Is it a fair assumption that the author of Euclid's theorem has in mind not so much wheels as representations of wheels? Knorr (1992) agrees, while insisting on how difficult the problem really is.

a quasi-symbolical system in which, for instance, instead of a circle whirling around a sphere, its 'hidden' part is shown *outside* the sphere.[27] I suspect that much of the visualisation work was done, in this special context, by watching planetaria, a subject to which I shall return below, in subsection 3.2.2. But the stress should be on the peculiarity of sphaerics. Most three-dimensional objects could have been drawn and 'read' from the drawing in a more direct, pictorial way.[28]

It should not be assumed, however, that, outside sphaerics, diagrams were 'pictures'. Kurt Weitzman offers a theory – of a scope much wider than mathematics – arguing for the opposite. Weitzman (1971, chapter 2) shows how original Greek schematic, rough diagrams (e.g. with little indication of depth and with little ornamentation) are transformed, in some Arabic traditions, into painterly representations. Weitzman's hypothesis is that technical Greek treatises used, in general, schematic, unpainterly diagrams.

The manuscript tradition for Greek mathematical diagrams, I repeat, has not been studied systematically. But superficial observations corroborate Weitzman's theory. Even if depth is sometimes indicated by some foreshortening effects, there is certainly no attempt at painterly effects such as shadowing.[29] The most significant question from a mathematical point of view is whether the diagram was meant to be *metrical*: whether quantitative relations inside the diagram were meant to correspond to such relations between the objects depicted. The alternative is a much more schematic diagram, representing only the qualitative relations of the geometrical configuration. Again, from my acquaintance with the manuscripts, they very often seem to be schematic in this respect as well.[30]

[27] Mogenet (1950). Thanks to Mogenet's work, we may – uniquely – form a hypothesis concerning the genesis of these diagrams. It is difficult to imagine such a system being invented by non-mathematical scribes. Even if it was not Autolycus' own scheme, it must reflect some ancient mathematical system.

[28] While foreshortening is irrelevant in the case of spheres, shading is relevant. In fact, in Roman paintings, shading is systematically used for the creation of the illusion of depth when columns, i.e. cylinders, are painted. The presence of 'strange' representations for spheres shows, therefore, a deliberate avoidance of the practice of shading. This, I think, is related to what I will argue later in the chapter, that Greek diagrams are – from a certain point of view – 'graphs' in the mathematical sense. They are not drawings.

[29] Effects which *do* occur in early editions – and indeed in some modern editions as well.

[30] Compare Jones (1986) 1.76 on the diagrams of Pappus: 'The most apparent . . . convention is a pronounced preference for symmetry and regularization . . . introducing [e.g.] equalities where quantities are not required to be equal.' Such practices (which I have often seen in manuscripts other than Pappus') point to the expectation that the diagram should not be read quantitatively.

To sum up, then: when mathematical results were presented in anything other than the most informal, private contexts, lettered diagrams were used. These would typically have been prepared prior to the mathematical reasoning.[31] Rulers and compasses may have been used. Generally speaking, a Greek viewer would have read into them, directly, the objects depicted, though this would have required some imagination (and, probably, what was seen then was just the schematic configuration); but then, any viewing demands imagination.

2 PRACTICES OF THE LETTERED DIAGRAM

2.1 *The mutual dependence of text and diagram*

There are several ways in which diagram and text are interdependent. The most important is what I call 'fixation of reference' or 'specification'.[32]

A Greek mathematical proposition is, at face value, a discussion of letters: *alpha*, *bēta*, etc. It says such things as 'AB is bisected at Γ'. There must be some process of fixation of reference, whereby these letters are related to objects. I argue that in this process the diagram is indispensable. This has the surprising result that the diagram is not directly recoverable from the text.

Other ways in which text and diagram are interdependent derive from this central property. First, there are assertions which are directly deduced from the diagram. This is a strong claim, as it seems to threaten the logical validity of the mathematical work. As I shall try to show, the threat is illusory. Then, there is a large and vague field of assertions which are, as it were, 'mediated' via the diagram. I shall try to clarify this concept, and then show how such 'mediations' occur.

[31] *P. Berol.* 17469, presented in Brashear (1994), is a proof of this claim. This papyrus – a second-century AD fragment of unknown provenance – covers *Elements* 1.9, with tiny remnants of 1.8 and 1.10. For each proposition, it has the enuncation together with an unlettered diagram, and nothing else. It is fair to assume that the original papyrus had more propositions, treated in the same way. My guess is that this was a memorandum, or an abridgement, covering the first book of Euclid's *Elements*. Had someone been interested in carrying out the proof, the lettering would have occurred on a copy on, e.g. a wax-tablet. (The same, following Fowler's suggestion (1987) 211–12, can be said of *POxy*. i.29.)

To anticipate: in chapter 2 I shall describe the practices related to the assigning of letters to points, and will argue for a semi-oral dress-rehearsal, during which letters were assigned to points. This is in agreement with the evidence from the papyri.

[32] The word 'specification' is useful, as long as it is clear that the sense is *not* that used by Morrow in his translation of Proclus (a translation of the Greek *diorismos*). I explain my sense below.

2.1.1 Fixation of reference

Suppose you say (fig. 1.1):

Let there be drawn a circle, whose centre is *A*.

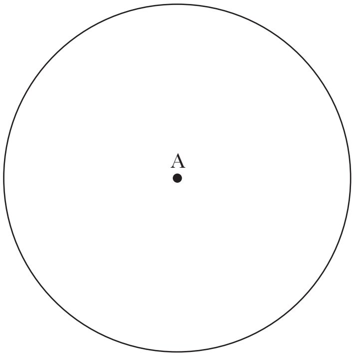

Figure 1.1.

A is thereby completely specified, since a circle can have only one centre.

Another possible case is (figs. 1.2a, 1.2b):

Let there be drawn a circle, whose radius is *BC*.

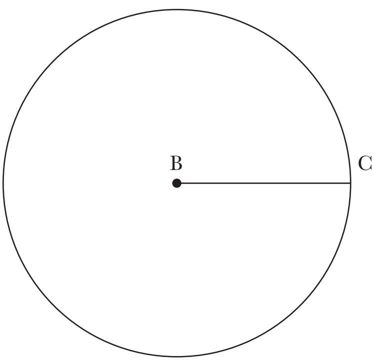

Figure 1.2a.

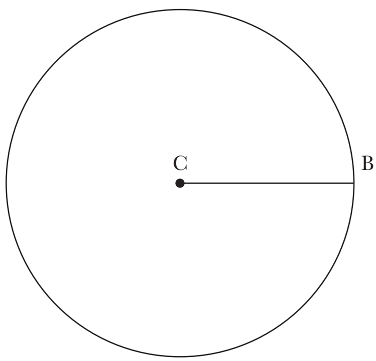

Figure 1.2b.

This is a more complicated case. I do not mean the fact that a circle may have many radii. It may well be that for the purposes of the proof it is immaterial which radius you take, so from this point of view saying 'a radius' may offer all the specification you need. What I mean by 'specification' is shorthand for 'specification for the purposes of the proof'.

But even granted this, a real indeterminacy remains here, for we cannot tell here *which of BC is which*: which is the centre and which touches the circumference. The text of the example is valid with both figures 1.2a and 1.2b. *B* and *C* are therefore underspecified by the text.

Finally, imagine that the example above continues in the following way (fig. 1.3):

> Let there be drawn a circle, whose radius is *BC*. I say that *DB* is twice *BC*.

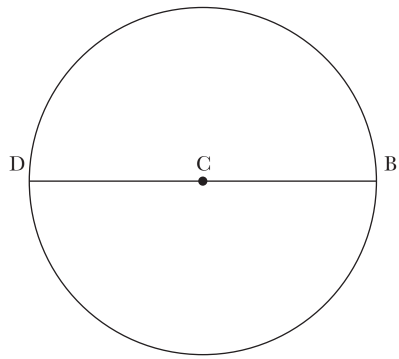

Figure 1.3.

D in this example is neither specified nor underspecified. Here is a letter which gets no specification at all in the text, which simply appears out of the blue. This is a completely unspecified letter.

We have seen three classes: completely specified, underspecified, and completely unspecified. Another and final class is that of letters which change their nature through the proposition. They may first appear as completely unspecified, and then become at least underspecified; or they may first appear as underspecified, and later get complete specification. This is the basic classification into four classes. I have surveyed all the letters in Apollonius' *Conics* I and Euclid's *Elements* XIII, counting how many belong to each class. But before presenting the results, there are a few logical complications.

First, what counts as a possible moment of specification? Consider the following case. Given the figure 1.4, the assertion is made: 'and therefore *AB* is equal to *BC*'. Suppose that nothing in the proposition so far specified *B* as the centre of the circle. Is this assertion then a specification of *B* as the centre? Of course not, because of the 'therefore' in the assertion. The assertion is meant to be a *derivation*, and

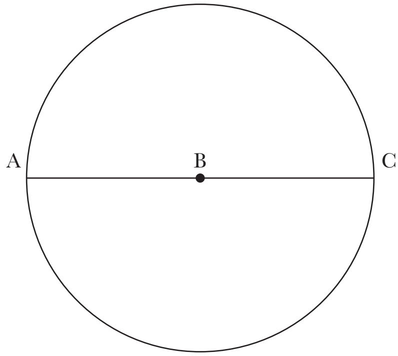

Figure 1.4.

making it into a specification would make it effectively a *definition*, and the derivation would become vacuous. Thus such assertions cannot constitute specifications. Roughly speaking, specifications occur in the imperative, not in the indicative. They are 'let the centre of the circle, *B*, be taken', etc.

Second, letters are specified by other letters. It may happen that those other letters are underspecified themselves. I have ignored this possibility. I have been like a very lenient teacher, who always gives his pupils a chance to reform. At any given moment, I have assumed that all the letters used in any act of specification were fully specified themselves. I have concentrated on *relative* specification, specification of a letter relative to the preceding letters. This has obvious advantages, mainly in that the statistical results are more interesting. Otherwise, practically all letters would turn out to be underspecified in some way.

Third and most important, a point which Grattan-Guinness put before me very forcefully: it must always be remembered, not only what the text specifies, but also what the mathematical sense demands. I have given such an example already, with 'taking a radius'. If the mathematical sense demands that we take *any* radius, then even if the text does not specify which radius we take, still this constitutes no underspecification. This is most clear with cases such as 'Let *some* point be taken on the circle, *A*'. Whenever a point is taken in this way, it is *necessarily* completely specified by the text. The text simply cannot give any better specification than this. So I stress: what I mean by 'underspecified letters' is not at all 'variable letters'. On the contrary: variable points have to be, in fact, completely specified. I mean letters which are left ambiguous by the text – which the text does not specify fully, *given the mathematical purposes*.

Now to the results.[33] In Euclid's *Elements* XIII, about 47% of the letters are completely specified, about 8% are underspecified, about 19% are completely unspecified, and about 25% begin as completely unspecified or underspecified, and get increased specification later. In Apollonius' *Conics* I, about 42% are completely specified, about 37% are underspecified, about 4% are completely unspecified, and about 16% begin as completely unspecified or underspecified, and get increased specification later. The total number of letters in both surveys is 838.

Very often – most often – letters are not completely specified. So how do we know what they stand for? Very simple: we see this in the diagram.

In fact the difficult thing is to 'unsee' the diagram, to teach oneself to disregard it and to imagine that the only information there is is that supplied by the text. Visual information is compelling itself in an unobtrusive way. Here the confessional mode may help to convert my readers. It took me a long time to realise how ubiquitous lack of specification is. The following example came to me as a shock. It is, in fact, a very typical case.

Look at Apollonius' *Conics* I.11 (fig. 1.5). The letter Λ is specified at 38.26, where it is asserted to be on a parallel to ΔΕ, which passes through Κ. Λ is thus on a definite line. But as far as the text is concerned, there is no way of knowing that Λ is a very specific point on that line, the one intersecting with the line ΖΗ. But I had never even thought about this insufficiency of the text: I always read the diagram into the text. This moment of shock started me on this survey. Having completed the survey, its implications should be considered.

First, why are there so many cases falling short of full specification? To begin to answer this question, it must be made clear that my results have little quantitative significance. It is clear that the way in which letters in Apollonius fail to get full specification is different from that in Euclid. I expect that there is a strong variability between works by the same author. The *way* in which letters are not fully specified depends upon mathematical situations. Euclid, for instance, in book XIII, may construct a circle, e.g. ΑΒΓΔΕ, and then construct a pentagon within

[33] The complete tables, with a more technical analysis of the semantics of specification, are to appear in Netz (forthcoming).

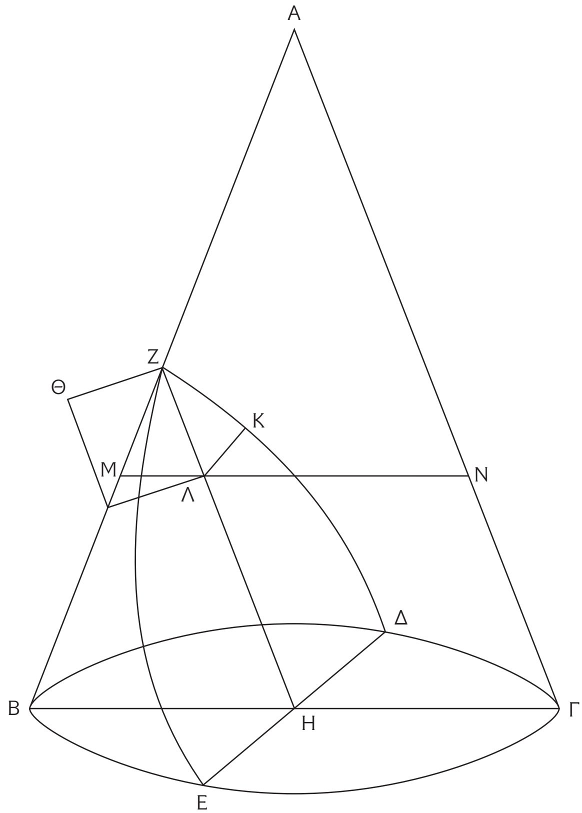

Figure 1.5. Apollonius' *Conics* 1.11.

the same circle, such that its vertices are the very same ΑΒΓΔΕ. This is moving from underspecification to complete specification, and is demanded by the subject matter dealt with in his book. In the *Conics*, parallel lines and ordinates are the common constructions, and letters on them are often underspecified (basically, they are similar to 'BC' in figs. 1.2a, 1.2b above).

What seems to be more stable is the percentage of fully specified letters. Less than half the letters are fully specified – but not much less than half. It is as if the authors were indifferent to the question of whether a letter were specified or not, full specification being left as a random result.

This, I claim, is the case. Nowhere in Greek mathematics do we find a moment of specification *per se*, a moment whose purpose is to make sure that the attribution of letters in the text is fixed. Such moments are very common in modern mathematics, at least since Descartes.[34] But specifications in Greek mathematics are done, literally, *ambulando*. The essence of the 'imperative' element in Greek mathematics – 'let a line be drawn . . .', etc. – is to do some job upon the geometric space, to get things moving there. When a line is drawn from one point to another, the letters corresponding to the start and end positions of movement ought to be mentioned. But they need not be carefully differentiated; one need not know precisely which is the start and which is the end – both would do the same job, produce the same line (hence underspecification); and points traversed through this movement may be left unmentioned (hence complete unspecification).

What we see, in short, is that while the text is being worked through, the diagram is assumed to exist. The text takes the diagram for granted. This reflects the material implementation discussed above. This, in fact, is the simple explanation for the use of *perfect* imperatives in the references to the setting out – 'let the point *A* have been taken'. It reflects nothing more than the fact that, by the time one comes to discuss the diagram, it has already been drawn.[35]

The next point is that, conversely, the text is not recoverable from the diagram. Of course, the diagram does not tell us what the proposition asserts. It could do so, theoretically, by the aid of some symbolic apparatus; it does not. Further, the diagram does not specify all the objects on its own. For one thing, at least in the case of sphaerics, it does not even look like its object. When the diagram is 'dense', saturated in detail, even the attribution of letters to points may not be obvious from the diagram, and modern readers, at least, reading modern diagrams, use the text, to some extent, in order to elucidate the diagram. The stress of this section is on *inter*-dependence. I have not merely tried to upset the traditional balance between text and

[34] In Descartes, the same thing is both geometric and algebraic: it is a line (called *AB*), and it is an algebraic variable (called *a*). When the geometrical configuration is being discussed, '*AB*' will be used; when the algebraic relation is being supplied, '*a*' is used. The square on the line is 'the square on *AB*' (if we look at it geometrically) or a^2 (if we look at it algebraically). To make this double-accounting system workable, Descartes must introduce explicit, *per se* specifications, identifying *symbols*. This happens first in Descartes (1637) 300. This may well be the first *per se* moment of specification in the history of mathematics.

[35] The suggestion of Lachterman (1989) 65–7, that past imperatives reflect a certain *horror operandi*, is therefore unmotivated, besides resting on the very unsound methodology of deducing a detailed philosophy, presumably shared by each and every ancient mathematician, from linguistic practices. The methodology adopted in my work is to explain shared linguistic practices by shared situations of communication.

diagram; I have tried to show that they cannot be taken apart, that neither makes sense in the absence of the other.

2.1.2 The role of text and diagram for derivations

In general, assertions may be derived from the text alone, from the diagram alone, or from a combination of the two. In chapter 5, I shall discuss grounds for assertions in more detail. What is offered here is an introduction.

First, some assertions do derive from the text alone. For instance, take the following:[36]

> As BE is to EΔ, so are four times the rectangle contained by BE, EA to four times the rectangle contained by AE, EΔ.

One brings to bear here all sorts of facts, for instance the relations between rectangles and sides, and indeed some basic arithmetic. One hardly brings to bear the diagram, for, in fact, 'rectangles' of this type often involve lines which do not stand at right angles to each other; the lines often do not actually have any point in common.[37]

So this is one type of assertion: assertions which may be viewed as verbal and *not* visual.[38] Another class is that of assertions which are based on the visual alone. To say that such assertions exist means that the text hides implicit assumptions that are contained in the diagram.

That such cases occur in Greek mathematics is of course at the heart of the Hilbertian geometric programme. Hilbert, one of the greatest mathematicians of the twentieth century, who repeatedly returned to foundational issues, attempted, in Hilbert (1899), to rewrite geometry without any unarticulated assumptions. Whatever the text assumes in Hilbert (1899), it either proves or explicitly sets as an axiom. This was never done before Hilbert, mainly because much information was taken from the diagram. As is well known, the very first proposition

[36] Apollonius' *Conics* 1.33, 100.7–8. The Greek text is more elliptic than my translation.

[37] Here the lines mentioned do share a point, but they are not at right angles to each other. See, for instance, *Conics* 1.34, 104.3, the rectangle contained by KB, AN – lines which do not share a point.

[38] This class is not exhausted by examples such as the above (so-called 'geometrical algebra'). For instance, any calculation, as e.g. in Aristarchus' *On Sizes and Distances*, owes nothing to the diagram. It should be noted that even 'geometrical algebra' is still 'geometrical': the text does not speak about multiplications, but about rectangles. This of course testifies to the primacy of the visual over the verbal. In general, see Unguru (1975, 1979), Unguru and Rowe (1981–2), Unguru and Fried (forthcoming), Hoyrup (1990a), for a detailed criticism of any interpretation of 'geometrical algebra' which misses its visual motivation. The term itself is misleading, but helps to identify a well-recognised group of propositions, and I therefore use it, quotation marks and all.

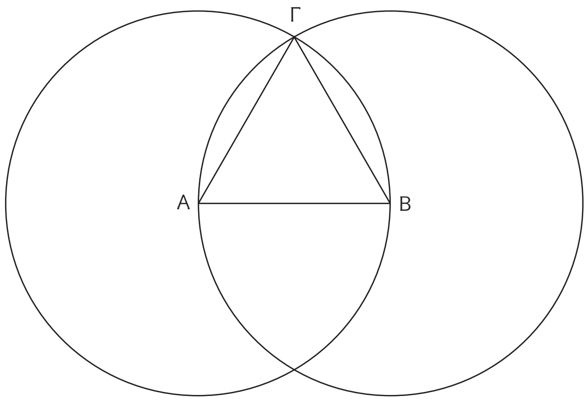

Figure 1.6. Euclid's *Elements* 1.1.

of Euclid's *Elements* contains an implicit assumption based on the diagram – that the circles drawn in the proposition meet (fig. 1.6).[39]

There is a whole set of assumptions of this kind, sometimes called 'Pasch axioms'.[40] 'A line touching a triangle and passing inside it touches that triangle at two points' – such assumptions were generally, prior to the nineteenth century, taken to be diagrammatically obvious.

Many assertions are dependent on the diagram alone, and yet involve nothing as high-powered as 'Pasch axioms'. For instance, Apollonius' *Conics* III.1 (fig. 1.7):[41] the argument is that ΑΔΒΖ is equal to ΑΓΖ and, therefore, subtracting the common ΑΕΒΖ, the remaining ΑΔΕ is equal to ΓΒΕ. Adopting a very grand view, one may say that this involves assumptions of additivity, or the like. This is part of the story, but the essential ground for the assertion is identifying the objects in the diagram.

My argument, that text and diagram are interdependent, means that many assertions derive from the combination of text and diagram. Naturally, such cases, while ubiquitous, are difficult to pin down precisely. For example, take Apollonius' *Conics* 1.45 (fig. 1.8). It is asserted – no special grounds are given – that ΜΚ:ΚΓ::ΓΔ:ΔΛ.[42] The implicit ground for this is the similarity of the triangles ΜΚΓ, ΓΔΛ. Now diagrams cannot, in themselves, show satisfactorily the similarity of triangles. But the diagram may be helpful in other ways, for, in fact, the similarity

[39] Most recommended is Russell (1903) 404ff., viciously and in a sense justly criticising Euclid for such logical omissions.

[40] For a discussion of the absence of Pasch axioms from Greek mathematics, see Klein (1939) 201–2.

[41] 318.15–18. [42] 138.10–11.

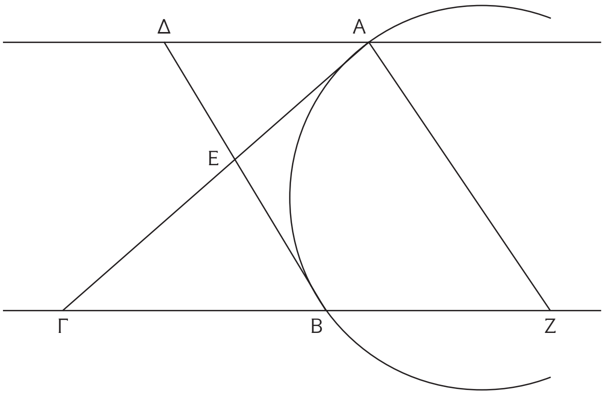

Figure 1.7. Apollonius' *Conics* III.1 (Parabola Case).

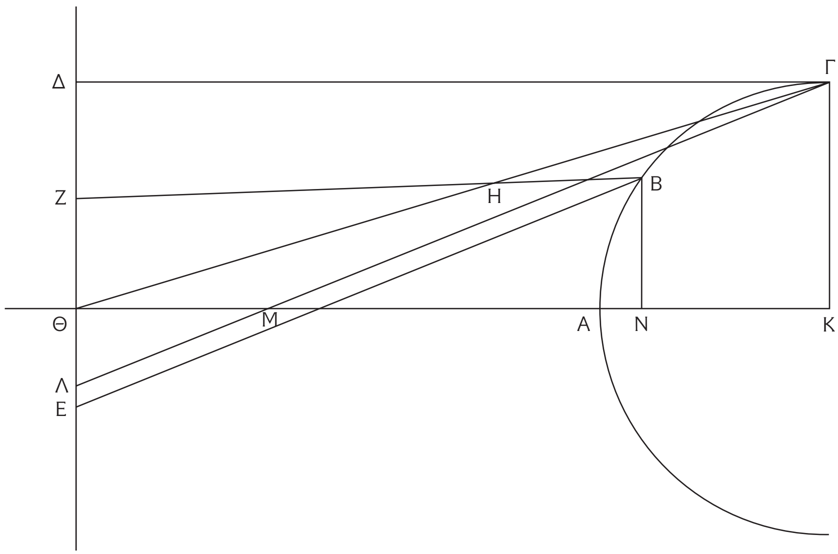

Figure 1.8. Apollonius' *Conics* I.45 (Hyperbola Case).

of the relevant triangles is not asserted in this proposition. To see this similarity, one must piece together a few hints: ΓΔ is parallel to ΚΘ (136.27); Μ lies on ΚΘ (underspecified by the text); ΓΚ is parallel to ΔΘ (138.6); Λ lies on ΔΘ (underspecified by the text); Μ lies on ΓΛ (136.26). Putting all of these together, it is possible to prove that the two triangles are similar. In a sense we do piece together those hints. But we are supposed to be able to do so at a glance (a significant phrase!). How do we do it then? We coordinate the various facts involved, and we coordinate them at great ease, because they are all simultaneously available on the diagram. The diagram is synoptic.

Note carefully: it is not the case that the diagram asserts information such as 'ΓΚ is parallel to ΔΘ'. Such assertions cannot be shown to be true in a diagram. But once the *text* secures that the lines are parallel, this piece of knowledge may be encoded into the reader's representation of the diagram. When necessary, such pieces of knowledge may be mobilised to yield, as an ensemble, further results.

2.1.3 The diagram organises the text

Even at the strictly linguistic level, it is possible to identify the presence of the diagram. A striking example is the following (fig. 1.9):[43]

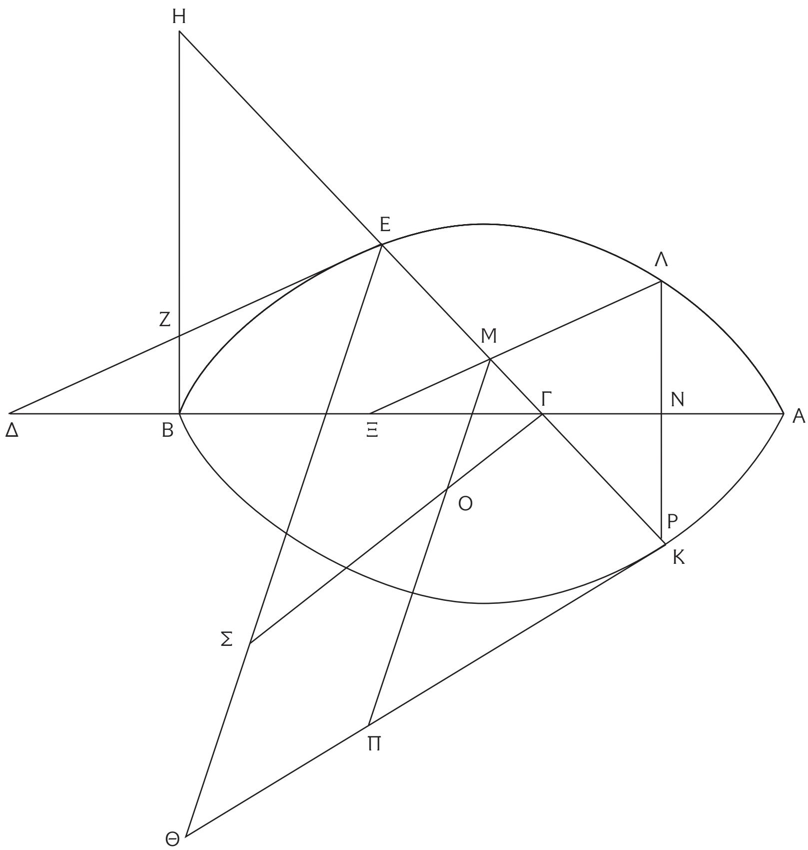

Figure 1.9. Apollonius' *Conics* 1.50 (Ellipse Case).

[43] Apollonius, *Conics* 1.50, 150.23–5: καὶ εἰλήφθω τι ἐπὶ τῆς τομῆς σημεῖον τὸ Λ, καὶ δἰ αὐτοῦ τῇ ΕΔ παράλληλος ἤχθω ἡ ΛΜΞ, τῇ δὲ ΒΗ ἡ ΛΡΝ, τῇ δὲ ΕΘ ἡ ΜΠ.

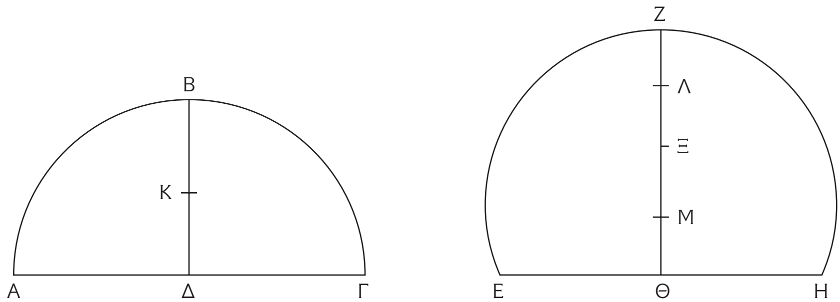

Figure 1.10. Archimedes' *PE* II.7.

And let some point be taken on the section, Λ, and, through it, let ΛΜΞ be drawn parallel to ΕΔ, ΛΡΝ to ΒΗ, ΜΠ to ΕΘ.

Syntactically, the sentence means that ΜΠ passes through Λ – which ΜΠ does not. The diagram forces one to carry Λ over to a part of the sentence, and to stop carrying it over to another part.[44] The pragmatics of the text is provided by the diagram. The diagram is the framework, the set of presuppositions governing the discourse.

A specific, important way in which the diagram organises the text is the setting of cases. This is a result of the diagrammatic fixation of reference. Consider Archimedes' *PE* II.7: ΕΖΗ, ΑΒΓ are two similar sections; ΖΘ, ΒΔ are, respectively, their diameters; Λ, Κ, respectively, their centres of gravity (fig. 1.10). The proposition proves, through a *reductio*, that ΖΛ:ΛΘ::ΒΚ:ΚΔ. How? By assuming that a different point, Μ, satisfies ΖΜ:ΜΘ::ΒΚ:ΚΔ. Μ could be put either above or below Λ. The cases are asymmetrical. Therefore these are two distinct cases. Archimedes, however, does not distinguish the cases in the text. Only the diagram can settle the question of which case he preferred to discuss.

There are many ways in which it can be seen that the guiding principle in the development of the proof is spatial rather than logical. Take, for instance, Apollonius' *Conics* I.15 (fig. 1.11): the proposition deals with a construction based on an ellipse. This construction has two 'wings', as it were. The development of the proof is the following: first, some work is done on the lower wing; next, the results are reworked on the ellipse itself; finally, the results are transferred to the

[44] Compare also the same work, proposition 31, 94.2–3: the syntax seems to imply that ΔΘ passes through Ε; it does not. In the same proposition, 92.23–4: is Γ on the hyperbola or on the diameter? The syntax, if anything, favours the hyperbola; the diagram makes it stand on the diameter: two chance examples from a chance proposition.

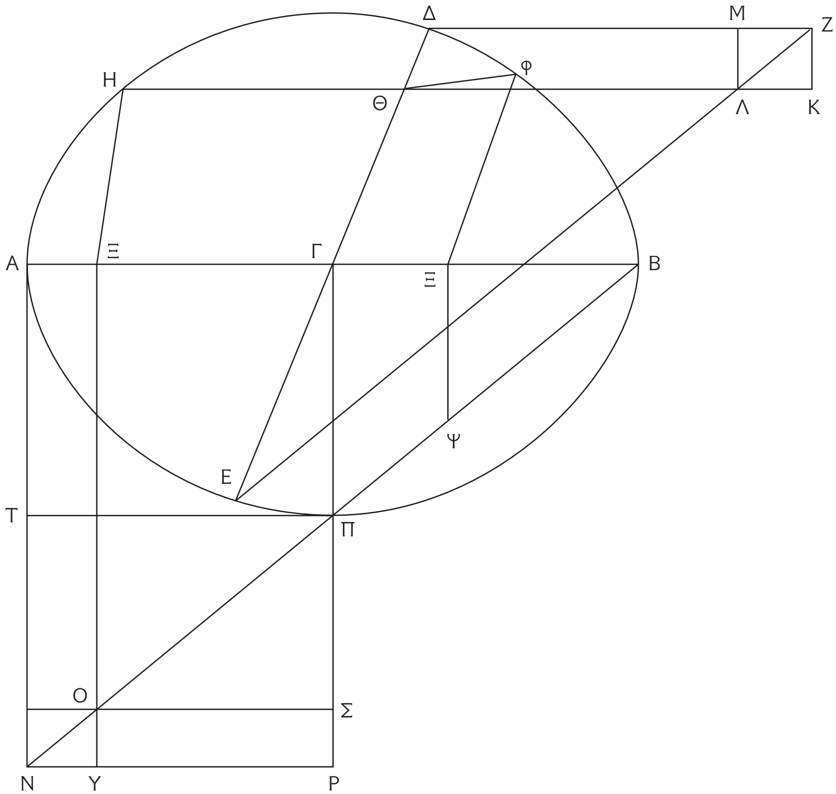

Figure 1.11. Apollonius' *Conics* 1.15.

upper wing. One could, theoretically, proceed otherwise, collecting results from all over the figure simultaneously. Apollonius chose to proceed spatially.[45] There are a number of contexts where the role of spatial visualisation can be shown, on the basis of the practices connected with the assignment of letters to objects, and I shall return to this issue in detail in chapter 2 below. The important general observation is that the diagram sets up a world of reference, which delimits the text. Again, this is a result of the role of the diagram for the process of fixation of reference. Consider a very typical case: Λ in Apollonius' *Conics* 1.6. It is specified in the following way (fig. 1.12):[46] 'From K, let a

[45] The first part is 60.5–19, the second is 60.19–29, the third is 60.29–62.13. That the second part casts a brief glance – seven words – back at the lower wing serves to show the contingency of this spatial organisation.

[46] 22.3.

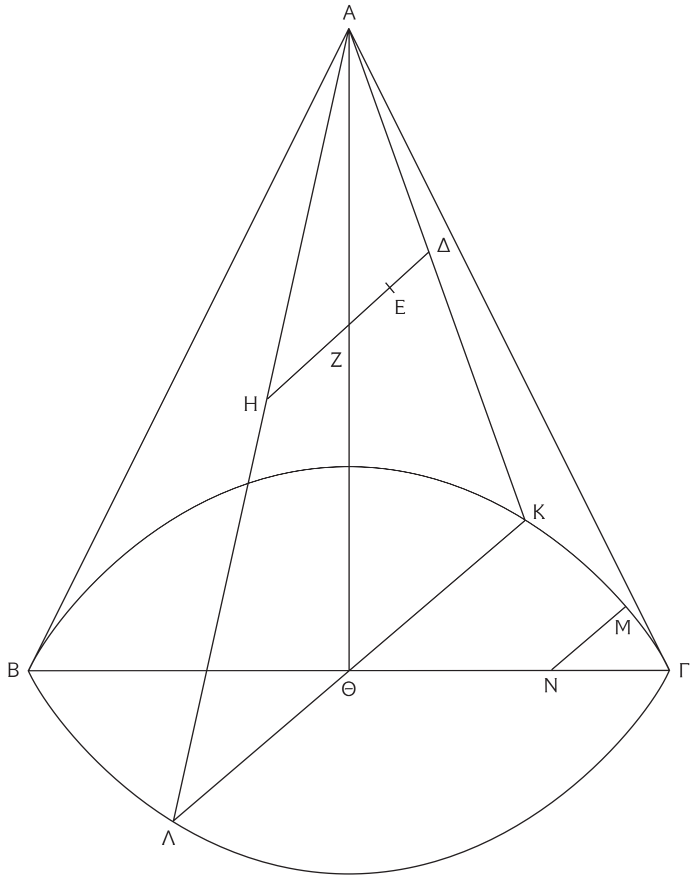

Figure 1.12. Apollonius' *Conics* 1.6 (One of the Cases).

perpendicular, to ΒΓ be drawn (namely) ΚΘΛ.' The locus set up for Λ is a line. How do we know that it is at the limit of that line, on the circle ΓΚΒ? Because Λ is the end point of the action of drawing the line ΚΘΛ – and because this action must terminate on this circle *for this circle is the limit of the universe of this proposition*. There are simply no points outside this circle.

Greek geometrical propositions are not about universal, infinite space. As is well known, lines and planes in Greek mathematics are always finite sections of the infinite line and plane which *we* project. They are, it is true, indefinitely extendable, yet they are finite. Each geometrical proposition sets up its own universe – which is its diagram.

2.1.4 The mutual dependence of text and diagram: a summary

Subsections 2.1.1–2.1.3, taken together, show the use of the diagram as a vehicle for logic. This might be considered a miracle. Are diagrams not essentially misleading aids, to be used with caution?

Mueller, after remarking on Greek implicit assumptions, went on to add that these did not invalidate Greek mathematics, for they were *true*.[47] This is a startling claim to be made by someone who, like Mueller, is versed in modern philosophy of mathematics, where truth is often seen as relative to a body of assumptions. Yet Mueller's claim is correct.

To begin with, a diagram may always be 'true', in the sense that it is there. The most ultra-abstract modern algebra often uses diagrams as representations of logical relations.[48] Diagrams, just like words, may be a way of encoding information. If, then, diagrams are seen in this way, to ask 'how can diagrams be true?' is like asking 'how can language be true?' – not a meaningless question, but clearly a different question from that we started from.

But there is more to this. The problem, of course, is that the diagram, *qua* physical object, does *not* model the assertions made concerning it. The physical diagram and the written text often clash: in one, the text, the lines are parallel; in the other, the diagram, they are not. It is only the diagram perceived in a certain way which may function alongside the text. But this caveat is in fact much less significant than it sounds, since whatever is perceived is perceived in a certain way, not in the totality of its physical presence. Thus the logical usefulness of the diagram *as a psychological object* is unproblematic – the important requirement is that the diagram would be perceived in an intersubjectively consistent way.

Poincaré – having his own axe to grind, no doubt – offered the following interpretation of the diagram:[49] 'It has often been said that geometry is the art of reasoning correctly about figures which are poorly constructed. This is not a quip but a truth which deserves reflection. But what is a poorly constructed figure? It is the type which can be drawn by the clumsy craftsman.'

Immediately following this, Poincaré goes on to characterise the useful diagram: 'He [the clumsy craftsman] distorts proportions more or less flagrantly . . . But [he] must not represent a closed curve by an

[47] Mueller (1981) 5.
[48] See e.g. Maclane and Birkhoff (1968), *passim* (explanation on the diagrammatic technique is found in 5ff.).
[49] I quote from the English translation, Poincaré (1963) 26.

open curve . . . Intuition would not have been impeded by defects in drawing which are of interest only in metric or projective geometry. But intuition will become impossible as soon as these defects involve *analysis situs*.'

The *analysis situs*[50] is Poincaré's hobby-horse, and should be approached with caution. The diagram is not just a graph, in the sense of graph theory. It contains at least one other type of information, namely the straightness of straight lines; that points stand 'on a line' is constantly assumed on the basis of the diagram.[51] This fact is worth a detour.

How can the diagram be relied upon for the distinction between straight and non-straight? The technology of drawing, described in section 1 above, showed that diagrams were drawn, probably, with no other tools than the ruler and compasses. Technology represented no more than the distinction between straight and non-straight. The man-made diagram, unlike nature's shapes, was governed by the distinction between straight and non-straight alone. The infinite range of angles was reduced by technology into a binary distinction.[52] This is hypothetical, of course, but it may serve as an introduction to the following suggestion.

There is an important element of truth in Poincaré's vision of the diagram. The diagram is relied upon as a *finite* system of relations. I have described above the proposition as referring to the finite universe of the diagram. This universe is finite in two ways. It is limited in space, by the boundaries of the figure; and it is discrete. Each geometrical proposition refers to an infinite, continuous set of points. Yet only a limited number of points are referred to, and these are almost always (some of) the points standing at the intersections of lines.[53] The great multitude of proletarian points, which in their combined efforts construct together the mathematical objects, is forgotten. All attention is fixed upon the few intersecting points, which alone are named. This,

[50] Corresponding – as far as it is legitimate to make such correspondences – to our notion of 'topology'.

[51] That the full phrase of the form ἡ εὐθεῖα γραμμὴ AB is almost always contracted to the minimum ἡ AB, even though this may equally well stand for ἡ γραμμὴ AB *simpliciter* – i.e. for a curved rather than a straight line – reflects the fact that this basic distinction, between curved and straight, could generally be *seen* in the diagram.

[52] So far, the technology is not confined to Greece; and Babylonian 'structural diagrams', described by Hoyrup (1990a: 287–8), are useful in this context.

[53] In Archimedes' *SL*, which includes 22 geometrical propositions (i.e. a few hundred letters), there are 24 which do not stand in extremes, or intersections, of lines, namely proposition 14: Β, Γ, Κ; 15: Β, Κ; 16: Β, Γ, Κ, Ν; 17: Β, Κ, Ν; 18: Β, Γ, Κ, Λ; 19: Β, Ε, Κ, Λ; 20: Β, Λ; 21: Δ; 28: Β. I choose this example as a case where there are relatively many such points, the reason being Archimedes' way of naming spirals by many letters, more letters than he can affix to extremes and intersections alone – essentially a reflection of the peculiarity of the spiral.

finally, is the crucial point. The diagram is named – more precisely, it is lettered. It is the lettering of the diagram which turns it into a system of intersections, into a finite, manageable system.

To sum up, there are two elements to the technology of diagrams: the use of ruler and compasses, and the use of letters. Each element redefines the infinite, continuous mass of geometrical figures into a man-made, finite, discrete perception. Of course, this does not mean that the object of Greek mathematics is finite and discrete. The perceived diagram does not exhaust the geometrical object. This object is partly defined by the text, e.g. metric properties are textually defined. But the properties of the perceived diagram form a true subset of the real properties of the mathematical object. This is why diagrams are good to think with.[54]

2.2 *Diagrams as metonyms of propositions*

A natural question to ask here is whether the practices described so far are reflected in the Greek conceptualisation of the role of diagrams. The claim of the title is that this is the case, in a strong sense. Diagrams are considered by the Greeks not as appendages to propositions, but as the core of a proposition.[55]

2.2.1 *Speaking about diagrams*[56]

Our 'diagram' derives from Greek *diagramma* whose principal meaning LSJ define as a 'figure marked out by lines', which is certainly etymologically correct. The word *diagramma* is sandwiched, as it were, between its anterior and posterior etymologies, both referring simply to drawn figures. Actual Greek usage is more complex.

Diagramma is a term often used by Plato – one of the first, among extant authors, to have used it – either as standing for mathematical

[54] A disclaimer: I am not making the philosophical or cognitive claim that the only way in which diagrams can be deductively useful is by being reconceptualised via letters. As always, I am a historian, and I make the historical claim that diagrams came to be useful as deductive tools in Greek mathematics through this reconceptualisation.

[55] That they *put* diagrams as 'appendages' – i.e. at the end of propositions rather than at their beginning or middle – shows something about the relative role of beginning and end, not about the role of the diagram. It should be remembered that the titles of Greek books are also often put at the *end* of treatises. My guess is that, reading a Greek proposition, the user would unroll some of the papyrus to have the entire text of the proposition (presumably a few columns long) ending conveniently with the diagram. It was the advent of the codex which led to today's nightmare of constant backwards-and-forwards glancing, from text to diagram, whenever the text spills from one page to the next.

[56] Part of the argument of this subsection derives from Knorr (1975) 69–75.

proofs or as the *de rigueur* accompaniment of mathematics.[57] With Aristotle, *diagrammata* (the plural of *diagramma*) can practically mean 'mathematics', while *diagramma* itself certainly means 'a mathematical proposition'.[58] Xenophon tells us that Socrates used to advise young friends to study geometry, but not as far as the unintelligible *diagrammata*,[59] and we begin to think that this may mean more than just very intricate diagrams in the modern sense. Further, Knorr has shown that the cognates of *graphein*, 'to draw', must often be taken to carry a logical import.[60] He translates this verb by 'prove by means of diagrams'. Certainly this phrase is the correct translation; however, we should remember that the phrase stands for what, for the Greeks, was a single concept.

Complementary to this, the terminology for 'diagram' in the modern sense is complex. The word *diagramma* is never used by Greek mathematicians in the sense of 'diagram'. When they want to emphasise that a proposition relies upon a diagram, they characterise it as done *dia grammōn* – 'through lines', in various contexts opposed to the only other option, *di' arithmōn* – 'through numbers'.[61]

A word mathematicians may use when referring to diagrams present within a proof is *katagraphē* – best translated as 'drawing'.[62] The verb *katagraphein* is regularly used in the sense of 'completing a figure', when the figure itself is not specified in the text. The verb is always used within this formula, and with a specific figure: a parallelogram (often rectangle) with a diagonal and parallel lines inside it.[63]

57 As in *Euthd.* 290c; *Phaedr.* 73b; *Theaetet.* 169a; the [pseudo?]-Platonic *Epin.* 991e; and, of course, *Rep.* 510c.

58 E.g. *APr.* 41b14; *Meteor.* 375b18; *Cat.* 14b1; *Metaph.* 998a25, 1051a22; *SE* 175a27.

59 *Mem.* IV.7.3. 60 Knorr (1975) 69–75.

61 See, e.g. Heron: *Metrica* II.10.3; Ptolemy: *Almagest* I.10, 32.1, VIII.5, 193.19, *Harmonics* I.5, 12.8; Pappus VI.600.9–13. Proclus, *In Rem Publicam* II.23. The treatment of book II by Hero, as preserved in the Codex Leidenensis (Besthorn and Heiberg (1900: 8ff.), is especially curious: it appears that Hero set out to prove various results with as few lines as possible, preferably with none at all, but with a single line if the complete avoidance of lines was impossible (one is reminded of children's puzzles – 'by moving one match only, the train changes into a balloon'). Hero's practice is comparable to the way a modern mathematician would be interested in proving the result X on the basis of fewer axioms than his predecessors. Modern mathematicians prove with axioms; Greek mathematicians proved with lines.

62 See e.g. Euclid's *Elements* III.33, IV.5, XII.4; Apollonius, *Conics* IV.27. Archimedes usually refers simply to σχήματα (*CF* II.394.6, 406.2, 410.24; *SC* II.224.3). This is 'figure' in the full sense of the word, best understood as a *continuous* system of lines; a single diagram – especially an Archimedean one! – may include more than a single σχῆμα. Finally, Archimedes uses once the verb ὑπογράφειν (*PE* I.5 Cor. 2, 132.12), a relative of καταγράφειν.

63 The first five propositions of Euclid's *Elements* XIII, and also: II.7, 8; VI.27–9; X.91–6. The formula is a feature of the Euclidean style – though the fact that Apollonius and Archimedes do not use it should be attributed, I think, to the fact that they do not discuss this rectangle.

Aristotle's references to diagrams are even more varied. On several occasions he refers to his own diagrams as *hupographai*, yet another relative of the same etymological family.[64] *Diagraphai* – a large family – are mentioned as well.[65] None of these diagrams are *mathematical* diagrams; when referring to a proof where a mathematical diagram occurs, Aristotle uses the word *diagramma*, and we are left in the dark as to whether this refers to the diagram or to the proof as a whole. What does emerge in Aristotle's case is a certain discrepancy between the standard talk *about* mathematics and the talk *of* mathematics. We will become better acquainted with this discrepancy in chapter 3.

Mathematical commentators may combine the two discourses, of mathematics and about mathematics. What is their usage? Pappus uses *diagramma* as a simple equivalent of our 'proposition'.[66] In several cases, when referring to a diagram inside a proposition, he uses *hupographē*.[67] Proclus never uses *diagramma* when referring to an actual present diagram, to which he refers by using the term *katagraphē* or, once, *hupogegrammenē*.[68] Eutocius uses *katagraphē* quite often.[69] *Schēma*, in the sense of *one* of the diagrams referred to in a proposition, is used as well. It is interesting that one of these uses derives directly from Archimedes,[70] while all the rest occur in – what I believe is a genuine – Eratosthenes fragment.[71]

The evidence is spread over a very long period indeed, but it is coherent. Alongside more technical words signifying a 'diagram' in the modern sense – words which never crystallised into a systematic terminology – the word *diagramma* is the one reserved for signifying that which a mathematical proposition is. Should we simply scrap, then, the notion that *diagramma* had anything at all to do with a 'diagram'? Certainly not. The etymology is too strong, and the semantic situation can be easily understood. *Diagramma* is the metonym of the proposition.

[64] *de Int.* 22a22; *Meteor.* 346a32, 363a26; *HA* 510a30; *EE* 1220b37.

[65] *EE* 1228a28; *EN* 1107a33; *HA* 497a32, 525a9. The γεγραμμέναι of *de Part.* 642a12 is probably relevant as well; I guess that the last mentioned are ἀνατομαί-type diagrams, included in a book, and that diagrams set out in front of an audience (e.g. on wooden tablets) are called ὑπογραφαί; but this is strictly a guess.

[66] E.g. VII.638.17, 670.1–2. When counting propositions in books, Pappus often counts θεωρήματα ἤτοι διαγράμματα, 'theorems, or diagrams' – a nice proof that 'diagrams' may function as metonyms of propositions.

[67] Several cognate expressions occur in IV.200.26, 272.14, 298.6; VI.542.11, 544.19 and, perhaps, III.134.22.

[68] *In Eucl.*: καταγραφή: 340.11, 358.11, 370.14, 400.9–15; ὑπογεγραμμένη: 286.22.

[69] Seventeen times in the commentary to Archimedes, for which see index II to Archimedes vol. III.

[70] 216.24.

[71] 88.15, 92.7, 94.13, 19.

It is so strongly entrenched in this role that when one wants to make quite clear that one refers to the diagram and *not* to the proposition – which happens very rarely – one has to use other, more specialised terms.[72]

2.2.2 *Diagrams and the individuation of propositions*

That diagrams may be the metonyms of propositions is surprising for the following reason. The natural candidate from our point of view would be the 'proposition', the enunciation of the content of the proposition – because this enunciation *individuates* the proposition. The hallmark of Euclid's *Elements* I.47 is that it proves 'Pythagoras' theorem' – which no other proposition does. On the other hand, nothing, logically, impedes one from using the same diagram for different propositions.

Even if this were true, it would show not that diagrams cannot be metonyms, but just that they are awkward metonyms. But interestingly this is wrong. The overwhelming rule in Greek mathematics is that propositions *are* individuated by their diagrams. Thus, diagrams are convenient metonyms.[73]

The test for this is the following. It often happens that two separate lines of reasoning employ the same basic geometrical configuration. This may happen either within propositions or between propositions.[74] Identity of configuration need not, however, imply identity of diagram, since the lettering may change while the configuration remains. My claim is that identity of configuration implies identity of diagram *within* propositions, and does not imply such identity *between* propositions.

What is an 'identity' between diagrams? This is a matter of degree – one can give grades, as it were:

1. 'Identity *simpliciter*' – the diagrams may be literally identical.
2.1. 'Inclusion' – the diagrams may not be identical, because the second has some geometrical elements which did not occur in the

[72] Note that I am speaking here not of diachronic evolution, but of a synchronic situation. It is thus useful to note that in contexts which are not strictly mathematical διάγραμμα has clearly the sense 'diagram' – e.g. Bacchius, in *Musici Graeci* ed. Janus, 305.16–17: Διάγραμμα . . . τί εστι; – Συστήματος ὑπόδειγμα. ἤτοι οὕτως, διάγραμμά ἐστι σχῆμα ἐπίπεδον . . .

[73] Here it should be clarified that the 'diagram' of a single proposition may be composed of a number of 'figures', i.e. continuous configurations of lines. When these different figures are not simply different objects discussed by a single proof, but are the same object with different cases (e.g. Euclid's *Elements* I.35), the problem of transmission becomes acute. Given our current level of knowledge on the transmission of diagrams, nothing can be said on such diagrams.

[74] Such continuities may be singled out in the text by the formulae τῶν αὐτῶν ὑποκειμένων/κατασκευασθέντων, καὶ τὰ ἄλλα τὰ αὐτὰ προκείσθω/κατασκευάσθω – see e.g. Euclid's *Elements* III.3, 14; VI.2, 3; Archimedes, *SC* II.6; Apollonius, *Conics* III.6. I will argue below that such continuities do not imply identities. Whether the continuity is explicitly noted or not does not change this.

first (or vice versa). However, the basic configuration remains. Furthermore, all the letters which appear in both diagrams stand next to identical objects (some letters would occur in this diagram but not in the other; but they would stand next to *objects* which occur in this diagram but not in the other). Hence, wherever the two diagrams describe a similar situation they may be used interchangeably.

2.2. 'Defective inclusion' – diagrams may have a shared configuration, but some letters change their objects between the two diagrams. Thus, it is no longer possible to interchange the diagrams, even for a limited domain.

3. 'Similarity' – the configuration is not identical, and letters switch objects, but there is a certain continuity between the two diagrams.

'F'. No identity at all – although the two propositions refer to a mathematical situation which is basically similar, the diagrams are flagrantly different.

Conics III offers many cases of interpropositional continuity of subject matter. I have graded them all.[75] The results are: a single first, seven 2.1, four 2.2, six thirds and four fails. Disappointing; in fact, the results are very heterogeneous and should not be used as a quantitative guide. The important point is the great rarity of the first – which makes it look like a fluke.

To put this evidence in a wider context, it should be noted that *Conics* III is remarkable in having so many cases of continuities. More often, subject matters change between propositions, ruling out identical diagrams. An interesting case in the Archimedean corpus is *CF* 4/5: a 2.2 by my marking system, but the manuscripts are problematic. Euclid's *schēma*, used in the formula 'and let the figure be drawn' to which I have referred in n. 63 above, is usually in the range 3–F.[76] There are no relevant cases in Autolycus; I shall now mention a case from Aristarchus (and, in n. 79, Ptolemy).

The best way to understand the Greek practice in this respect is to compare it with Heath's editions of Archimedes and Apollonius. One of the ways in which Heath mutilated their spirit is by making diagrams as identical as possible. This makes the individuated unit larger

[75] 1: 46 (identical to 45); 2.1: 2 (compared with 1), 14 (13), 29 (28), 47–50 (15); 2.2: 7–8 (6), 10 (8), 21 (20); 3: 3 (1), 6 (5), 9 (6), 25 (24), 35 (34), 38 (37); F: 11 (8), 12 (11), 36 (34), 40 (39).

[76] In this I ignore *Elements* X.91–6, which is a specimen from a strange context. In general, book X works in hexads, units of six propositions proving more or less the same thing. It is difficult to pronounce exactly on the principle of individuation in this book: are propositions individuated, or are hexads?

than a given proposition: it is something like a 'mathematical idea'. But such identities ranging over propositions are Heath's, not Archimedes' nor Apollonius'.

The complementary part of my hypothesis has to do with internal relations. It is not at all rare for a proposition to use the same configuration twice. For instance, this is very common in some versions of the method of exhaustion, where the figure is approached from 'above' and from 'below'. The significance of the diagram changes; yet, there is no evidence that it has been redrawn.[77]

The following case appears very strange at first glance: the construction of Aristarchus 14 begins with ἔστω τὸ αὐτὸ σχῆμα τῷ πρότερον – 'let there be the same figure as before'.[78] Having said that, Aristarchus proceeds to draw a diagram which I would mark 2.2 – not at all the identity suggested by his own words (figs. 1.13a and 1.13b)! How can we account for this? I suggest the following: Aristarchus' motivation is to save space; that is, he does not want to give the entire construction from scratch – that would be tedious. But then, saying 'let *A* and *B* be the same, *C* and *D* be different, and so on' is just as tedious. So he simply says 'let it be the same', knowing that his readers would not be misled, for no reader would expect two diagrams to be literally identical. When you are told somebody's face is 'the same as Woody Allen's', you do not accept this as literally true – the pragmatics of the situation rule this out. Faces are just too individual. Greek diagrams are, as it were, the faces of propositions, their metonyms.[79]

2.2.3 Diagrams as metonyms of propositions: summary

I have claimed that diagrams are the metonyms of propositions; in effect, the metonyms of mathematics (as mentioned in n. 58 above).

[77] See, e.g. Archimedes, *CS* 21 352.9, 25 380.16, 26 388.10, 27 402.7–8, 29 420.15; *SC* II.6 204.14; *QP* 16 296.26. For examples from outside the method of exhaustion, see Apollonius' *Conics* I.26 82.20–1; 32 96.23–6; Euclid's *Elements* III.3 172.17, 14 204.11.

[78] Aristarchus 14 398.23. Incidentally, this is another mathematical use of σχῆμα for 'diagram'.

[79] I have not discussed Ptolemy's diagrams in this subsection. Ptolemy often uses expressions like 'using the same diagram'. Often the diagrams involved are very dissimilar (e.g. the first diagram of *Syntaxis* V.6, in 380.18–19, referring to the last diagram of V.5). Sometimes Ptolemy registers the difference between the diagrams by using expressions such as 'using a *similar* diagram' (e.g. the first diagram of XI.5, in 393.1–2, referring to the first diagram of XI.1). Rarely, diagrams are said to be 'the same' and are indeed practically identical (e.g. the fourth diagram of III.5, in 245.6–7, referring to the third diagram of III.5). But this is related to another fact: Ptolemy uses in the *Syntaxis* a limited *type* of diagram. Almost always, whether he does trigonometry or astronomy, Ptolemy works with a diagram based on a single circle with some lines passing through it. A typical Greek mathematical work has a wide range of diagrams; each page looks different. Ptolemy is more repetitive, more schematic. L. Taub suggested to me that this should be related to Ptolemy's wider programme – that of preparing a 'syntaxis', *organised* knowledge.

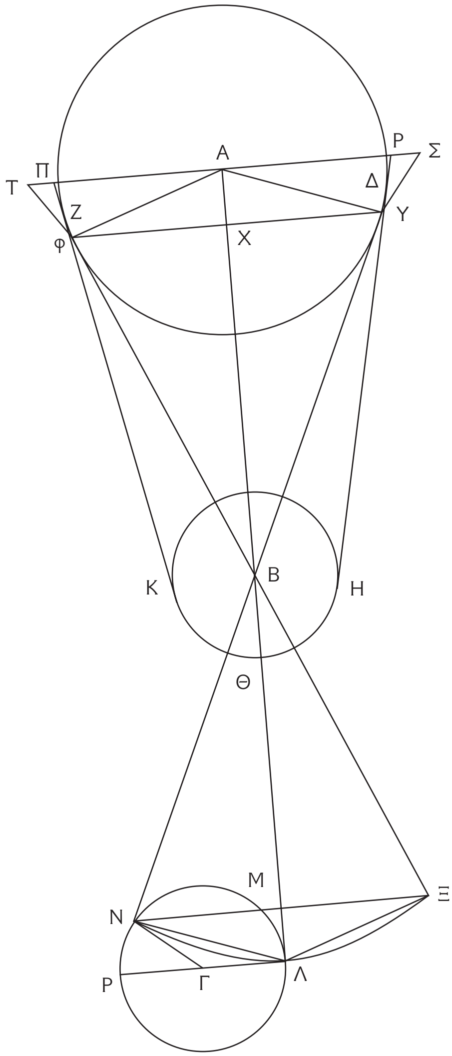

Figure 1.13a. Aristarchus 13.

That diagrams were considered essential for mathematics is proved by books V, VII–IX of Euclid's *Elements*. There, all the propositions are accompanied by diagrams, as individual and – as far as the situations allow – as elaborate as any geometrical diagram. Yet, in a sense, they are redundant, for they no longer represent the situations discussed. As Mueller points out, these diagrams may be helpful in various ways.[80]

[80] Mueller (1981) 67.

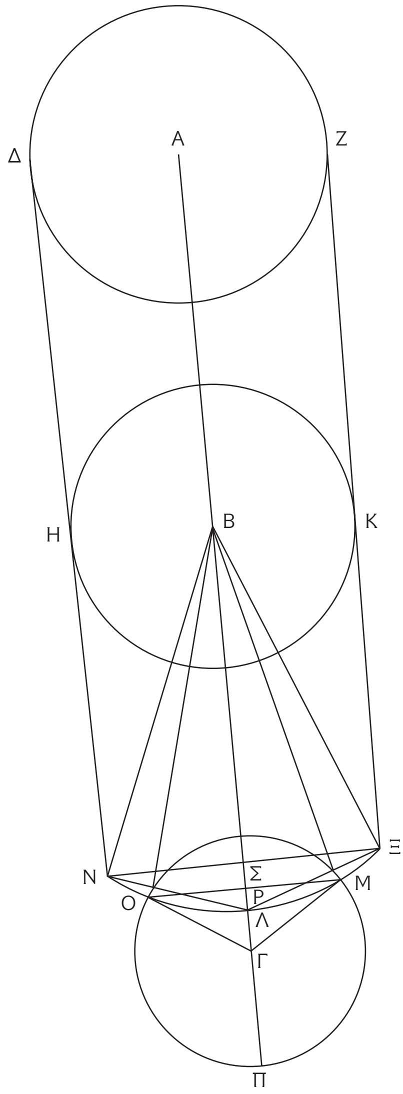

Figure 1.13b. Aristarchus 14.

Yet, as he asserts, they no longer have the same function. They reflect a cultural assumption, that mathematics *ought* to be accompanied by diagrams. Probably line diagrams are not the best way to organise proportion theory and arithmetic. Certainly symbolic conventions such as '=', for instance, may be more useful. The lettered diagram functions here as an obstacle: by demanding one kind of representation, it obstructs the development of other, perhaps more efficient representations.[81] An obstacle or an aid: the diagram was essential.

2.3 The semiotic situation

So far I have used neutral expressions such as 'the point *represented* by the letter'. Clearly, however, the cognitive contribution of the diagram cannot be understood without some account of what is involved in those 'representations' being given. This may lead to problems. The semiotic question is a tangent to a central philosophical controversy: what is the object of mathematics? In the following I shall try not to address such general questions. I am interested in the semiotic relation which Greek mathematicians have used, not in the semiotic relations which mathematicians in general ought to use. I shall first discuss the semiotic relations concerning letters, and then the semiotic relations concerning diagrams.

2.3.1 The semiotics of letters

Our task is to interpret expressions such as ἔστω τὸ μὲν δοθὲν σημεῖον τὸ Α[82] – 'let the given point be the Α'. To begin with, expressions such as τὸ Α, 'the Α', are not shorthand for 'the *letter* Α'; Α is not a letter here, but a point.[83] The letter in the text refers not to the letter in the diagram, but to a certain point.

Related to this is the following. Consider this example, one of many:[84]

ἔστω εὐθεῖα ἡ ΑΒ

(I will give a translation shortly).

This is translated by Heath as 'Let *AB* be a straight line.'[85] This creates the impression that the statement asserts a correlation between a symbol and an object – what I would call 'a moment of specification *per se*'.

[81] By a process which eludes our knowledge, manuscripts for Diophantus developed a limited system of shorthand, very roughly comparable to an abstract symbolic apparatus. Whether this happened in ancient times we can't tell; at any rate, Diophantus requires a separate study.

[82] Euclid's *Elements* I.2, 12.21.

[83] This can be shown through the wider practice of such abbreviations, which I discuss in chapter 4.

[84] Euclid's *Elements* XIII.4, 256.26; Heath's version is vol. III.447.

[85] Heath probably preferred, in this case, a slight unfaithfulness in the translation to a certain stylistic awkwardness. It so happens that this slight unfaithfulness is of great semiotic significance. It should be added that I know of no translation of Euclid which does not commit – what I think is – Heath's mistake. Federspiel (1992), in a context very different from the present one, was the first to suggest the correct translation.

In fact, this translation is untenable, since the article before AB can only be interpreted as standing for the elided phrase 'straight line',[86] so Heath's version reads as 'let the straight line *AB* be a straight line', which is preposterous. In fact the word order facilitates the following translation:

'Let there be a straight line, [viz.] *AB*.'

First, what such clauses do not assert: they do not assert a relation between a symbol and an object. Rather, they assert an action – in the case above, the taking for granted of a certain line – and they proceed to localise that action in the diagram, on the basis of an independently established reference of the letters. The identity of 'the *AB*' as a certain line in the diagram is assumed by Euclid, rather than asserted by him.

So far, expressions use the bare article and a combination of one or more letters. This is the typical group of expressions. There is another, rarer, group of expressions, which may shed some light on the more common one. Take the Hippocratic fragment, our evidence for earliest Greek mathematics[87] (fig. 1.14):

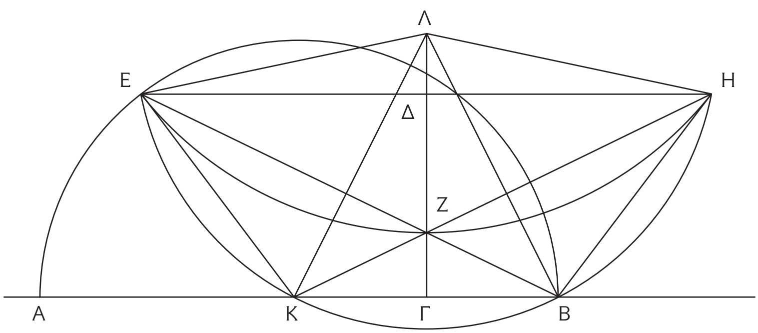

Figure 1.14. Hippocrates' Third Quadrature.

ἔστω κύκλος οὗ διάμετρος ἐφ' ᾗ AB κέντρον δε αὐτοῦ ἐφ' ᾧ K

'Let there be a circle whose diameter [is that] on which AB, its centre [that] on which K'.

[86] While the feminine gender, in itself, does not imply a straight line, the overall practice demands that one reads the bare feminine article, *ceteris paribus*, as referring to a *straight* line.

[87] Becker (1936b) 418.32.

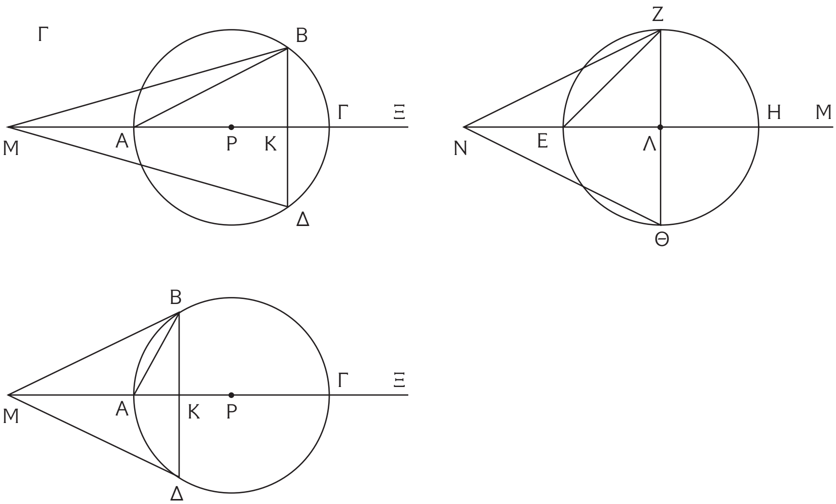

Figure 1.15. Archimedes' *SC* II.9.

I translate by 'on which' a phrase which in the Greek uses the preposition *epi* with the dative (which is interchangeable with the genitive).[88] Our task is to interpret this usage.

Expressions such as that of the Hippocratic fragment are characteristic of the earliest Greek texts which use the lettered diagram, that is, besides the Hippocratic fragment itself, the mathematical texts of Aristotle.[89] However, Aristotle – as ever – has his own, non-mathematical project, which makes him a difficult guide. I shall first try to elucidate this practice out of later, well-understood *mathematical* practice, and then I shall return to Aristotle.

The Archimedean corpus contains several expressions similar to the *epi* + dative. First, at *SC* II.9 Archimedes[90] draws several *schēmata*, and in order to distinguish between them, a Γ (or a special sign, according to another manuscript)[91] is written next to that *schēma* (fig. 1.15). Later

[88] For the genitive in the Hippocratic fragment, see Simplicius, *In Phys.* 65.9, 16; 67.21–2; 68.14. It is interesting to see that in a number of cases the manuscripts have either genitive or dative, and Diels, the editor, *always* chooses the dative: 64.13, 15; 67.29, 37 – which gives the text a dative-oriented aspect stronger than it would have otherwise (though Diels, of course, may be right).

[89] E.g. *Meteor.* 375b22, 376a5, 15, b5, 13, etc.; as well as many examples in contexts which are not strictly mathematical, e.g. *Meteor.* 363a34; *HA* 510a31, 550a25; *Metaph.* 1092b34. The presence of a diagram cannot always be proved, and probably is not the universal case.

[90] Or some ancient mathematical reader; for our immediate purposes, the identification is not so important.

[91] The same sign (astronomical sun) is used to indicate a scholion, in *PE* II.7, 188.18.

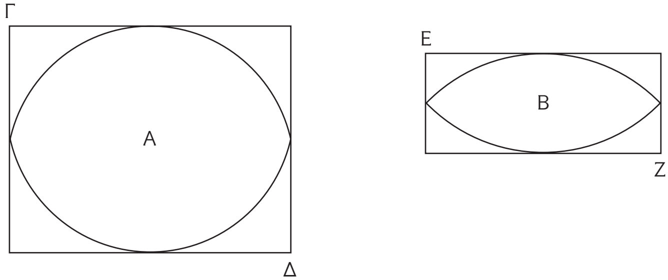

Figure 1.16. Archimedes' *CS* 6.

Λ	Λ	Λ	Λ	Λ	Λ
K	K	K	K	K	K
I	I	I	I	I	I
Θ	Θ	Θ	Θ	Θ	Θ

Figure 1.17. Archimedes' *CS* 2.

in the same proposition, at 250.6–7, when referring to that *schēma*, the expression used is πρὸς ᾧ το Γ σημεῖον – 'that, next to which is the sign[92] Γ'. This uses the preposition *pros* with the dative. I shall take *CS* 6 282.17–18 next. In order to refer to areas bounded by ellipses, in turn surrounded by rectangles, Archimedes writes the letters A, B inside the ellipses (fig. 1.16), then describes them in the following way: ἔστω περιεχόμενα χωρία ὑπὸ ὀξυγωνίου κώνου τομᾶς, ἐν οἷς τὰ A, B – 'let there be areas bounded by ellipses, in which are A, B'. This uses the preposition *en* with the dative. Proposition 2 in the same work refers, first, to signs which stand near lines and, consequently, within rectangles (fig. 1.17). It comes as no surprise now that the rectangles are mentioned at 268.1 as ἐν οἷς τὰ Θ, I, K, Λ – 'in which the Θ, I, K, Λ'. More interestingly, the lines in question are referred to at, e.g. 266.22–3 as ἐφ' ὧν . . . Θ, I, K, Λ – 'on which Θ, I, K, Λ' – where we finally get as far as the *epi* + genitive.[93] A certain order begins to emerge.

[92] Undoubtedly this is the sense of σημεῖον here. That the word becomes homonymous is not surprising: we shall see in chapter 3 that, in the border between first-order and second-order language, many such homonyms occur.

[93] For further examples of prepositions with letters, see Archimedes, *SL* 46.27, 48.20, 52.22, 72.18, 76.21, 80.24, 84.5, 86.11, 92.24, 102.13, 24 116.22; *CS* 266.17, 268.7, 19, 270.2, 276.7–9, 276.14, 280.23–5, 282.18, 24, 372.18, 276.3, 12–13, 378.13, 19, 390.10, 394.9, 396.12, 400.12, 410.27, 414.5, 418.10, 426.20; Apollonius, *Conics* III.13, 338.12; Pappus, book II, *passim* (in the context 'ἀριθμοὶ ἐφ' ὧν τὰ A . . .').

When Archimedes deviates from the normal letter-per-point convention, he often has to clarify what he refers to. A fuller expression is needed, and this is made up of prepositions, relatives and letters. Now the important fact is that the prepositions are used in a spatial sense – as is shown by their structured diversity. Different prepositions and cases are used in different spatial configurations. They describe various *spatial* relationships between the letters in the diagram and the objects referred to.

There is a well-known distinction, offered by Peirce, between three types of signs. Some signs are *indices*, signifying by virtue of some deictic relation with their object: an index finger is a good example. Other signs are *icons*, signifying by virtue of a similarity with their object: a portrait is a good example. Finally, some signs are *symbols*, signifying by virtue of arbitrary conventions: most words are symbols. We have gradually acquired evidence that in some contexts the letters in Greek diagrams may be seen as indices rather than symbols. My theory is that this is the case generally, i.e. the letter *alpha* signifies the point next to which it stands, not by virtue of its being a symbol for it, but simply because it stands next to it. The letters in the diagram are useful signposts. They do not stand *for* objects, they stand *on* them.

There are two different questions here. First, is this the correct interpretation of *epi* + dative/genitive in the earliest sources? Second, should this interpretation be universally extended?

The answer to the first question should, I think, be relatively straightforward. The most natural reading of *epi* is spatial, so, given the presence of a diagram which makes a spatial reading possible, I think such a reading cannot be avoided. It is true that many spatial terms are used metaphorically (if this is the right word), probably in all languages. In English, one can debate whether 'Britain should be inside the European Union', and it is clear that no spatial reading is intended: 'European Union' is (in a sense) an abstract, non-spatial object. The debate can be understood only in terms of inclusion in a wide, non-spatial sense. But if you ask whether 'the plate should be inside the cupboard', it is very difficult to interpret this in non-spatial terms. When a spatial reading suggests itself at all, it is irresistible. I have argued that the mathematical text is focused on the strictly spatial object of the diagram. It is as spatial as the world of plates and cupboards; and a spatial reading of the expressions relating to it is therefore the natural reading.

The case of Aristotle is difficult.[94] Setting aside cases where a reference to a diagram is clear, the main body of evidence is from the *Analytics*. There, letters are used very often.[95] When the use of those letters is of the form '*A* applies to all *B*', etc., the bare article + letter is used, i.e. the *epi* + dative/genitive is never used in such contexts. From time to time, Aristotle establishes a relation between such letters and 'real' objects – *A* becomes man, *B* becomes animal, etc. Usually, when this happens, the *epi* + dative/genitive is used at least with one of the correlations, and should probably be assumed to govern all the rest.[96] A typical example is 30a30:

> ἐφ᾽ ᾧ δὲ τὸ Γ ἄνθρωπος
>
> 'And [if that] on which Γ [is] man' / 'and [if that] which Γ stands for [is] man'.

I have offered two alternative translations, but the second should probably be preferred, for after all Γ does not, spatially speaking, stand on the class of all human beings. It's true that the antecedent of the relative clause need not be taken here to be 'man'. Indeed, often it cannot, when the genders of the relative pronoun and the signified object clash.[97] But there are other cases, where the gender, or more often the number of the relative pronoun do change according to the signified object.[98] The most consistent feature of this Aristotelian usage is its inconsistency – not a paradox, but a helpful hint on the nature of the usage. Aristotle, I suggest, uses language in a strange, forced way. That his usage of letters is borrowed from mathematics is extremely likely. That in such contexts the sense of the *epi* + dative/genitive would have been spatial is as probable. In a very definite context – that of establishing external references to letters of the syllogism – Aristotle uses this expression in a non-spatial sense. Remember that Aristotle had to start logic from scratch, the notions of referentiality included. I suggest that the use of the *epi* + dative/genitive in the *Analytics* is a bold *metaphor*, departing from the spatial mathematical

[94] Readers unfamiliar with Greek or Aristotle may prefer to skip the following discussion, which is relatively technical.

[95] The letter A is used more than 1,200 times; generally, the density of letters is almost comparable to a mathematical treatise.

[96] There are about – very roughly – a hundred such examples in the *Analytics*, which I will not list here. In pages 30–49 of *APr.* the examples are: 30a30, 31b5, 28, 34a7–8, b34, 39, 37b1, 38a31, b20, 44a13–17, 26, b3, 46b3–5, 13, 34, 47b21–2, 30–1, 48a3, 33–4, 48b6, 49a15–16, 32, 34, 39, b1.

[97] E.g. *APr.* 64a24: ἰατρικὴ δ᾽ ἐφ᾽ οὗ Δ.

[98] E.g. *APr.* 44a13: ἑπόμενα τῷ Α ἐφ᾽ ὧν Β; *APo.* 94a29: ἡμισεῖα δυοῖν ὀρθαῖν ἐφ᾽ ἧς Β.

usage. Aristotle says, ‘let A stand on “man” ’, implying ‘as mathematical letters stand on their objects and thus signify them’, meaning ‘let A signify “man” ’. The index is the metaphor through which the general concept of the sign is broached.[99] This, I admit, is a hypothesis. At any rate, the *contents* referred to by Aristotle are like ‘Britain’ and ‘European Union’, not like ‘plates’ and ‘cupboards’; hence a non-spatial reading becomes more natural.

Moving now to the next question: should the mathematical letters be seen as indices even in the absence of the *epi* + dative/genitive and its relatives?

The first and most important general argument in favour of this theory is the correction offered above to Heath’s translation of expressions such as ἔστω εὐθεῖα ἡ AB, ‘let there be a line, <namely> AB’. If the signification of the ‘AB’ is settled independently, and antecedently to the text, then it could be settled only via the letters as indices. The setting of symbols requires speech; indices are visual. The whole line of argument, according to which specification of objects in Greek mathematics is visual rather than verbal, supports, therefore, the indices theory.

Next, consider the following. In the first proposition of the *Conics* – any other example with a similar combination of genders will do – a point is specified in the following way:[100]

ἔστω κωνικὴ ἐπιφάνεια, ἧς κορυφὴ τὸ Α σημεῖον

‘Let there be a conic surface, whose vertex is the point A’.

The point A has been defined as a vertex, and it will function in the proposition *qua* vertex, not *qua* point. Yet it will always be called, as in the specification itself, τὸ Α, in the neuter (‘point’ in Greek is neuter, while ‘vertex’ is feminine). This is the general rule: points, even when acquiring a special significance, are always called simply ‘points’, never, e.g. ‘vertices’. The reason is simple: the expression τὸ Α is a periphrastic reference to an object, using the letter in the diagram, A, as a signpost useful for its spatial relations. This letter in the diagram, the actual shape of ink, stands in a spatial relation to a *point*, not to a vertex – the point is spatial, while the vertex is conceptual.

[99] Another argument for the ‘metaphor’ hypothesis is the fact that the *epi* + dative/genitive is not used freely by Aristotle, but only within a definite formula: he never uses more direct expressions such as καὶ Γ ἐπ’ ἀνθρώπῳ – ‘and [if] Γ stands for man’ – instead he sticks to the cumbersome relative phrase. Could this reflect the fact that the expression is a metaphor, and thus cannot be used outside the context which makes the metaphor work?

[100] 8.25.

Third, an index (but not a symbol) can represent simultaneously several objects; all it needs to do so is to point to all of them. Some mathematical letters are polyvalent in exactly this way: e.g. in Archimedes' *SC* 32, the letters Ο, Ξ, stand for both the circles and for the cones whose bases those circles are.[101]

Fourth, my interpretation would predict that the letters in the text would be considered as radically different from other items, whereas otherwise they should be considered as names, as good as any. There is some palaeographic evidence for this.[102]

Fifth, a central thesis concerning Greek mathematics is that offered by Klein (1934–6), according to which Greek mathematics does not employ variables. I quote:[103] 'The Euclidean presentation is not symbolic. It always intends determinate numbers of units of measurements, and it does this without any detour through a "general notion" or a concept of a "general magnitude".'

This is by no means unanimously accepted. Klein's argument is philosophical, having to do with fine conceptual issues.[104] He takes it for granted that *A* is, in the Peircean sense, a symbol, and insists that it is a symbol of something determinate. Quite rightly, the opposition cannot see why (the symbolhood of *A* taken for granted) it cannot refer to whatever it applies to. My semiotic hypothesis shows why *A* must be determinate: because it was never a symbol to begin with. It is a signpost, and signposts are tied to their immediate objects.

Finally, my interpretation is the 'natural' interpretation – as soon as one rids oneself of twentieth-century philosophy of mathematics. My proof is simple, namely that Peirce actually took letters in diagrams as *examples* of what he meant by 'indices':[105] '[W]e find that indices are

[101] Or a somewhat different case: Archimedes' *PE* I.3, where Α, Β are simultaneously planes, and the planes' centres of gravity.

[102] It should be remembered that, as a rule, Greek papyri do not space words. *P. Berol.* 12609, from *c.*350–325 BC (Mau and Mueller 1962, table II): the continuous text is, as usual, unspaced. Letters referring to the diagram are spaced from the rest of the text. *P. Herc.* 1061, from the last century BC, contains no marking off of letters, but the context is non-mathematical. *PFay.* 9, later still, marks letters by superscribed lines, as does the *In Theaetet.* (early AD? *CPF* III 341, n. ad XXIX.42–XXXI.28). This practice can often be seen in manuscripts. Generally, letters are comparable to *nomina sacra*. Perhaps it all boils down to the fact that letters, just as *nomina sacra*, are not read phonetically (i.e. 'ΑΒ' was read *'alpha-bēta'*, not 'ab')?

[103] The quotation is from the English translation (Klein 1968: 123). Klein has predominantly arithmetic in mind, but if this is true of arithmetic, it must *a fortiori* be true of geometry.

[104] Unguru and Rowe (1981–2: the synthetic nature of so-called 'geometric algebra'), Unguru (1991: the absence of mathematical induction; I shall comment on this in chapter 6, subsection 2.6) and Unguru and Fried (forthcoming: the synthetic nature of Apollonius' *Conics*), taken together, afford a picture of Greek mathematics where the absence of variables can be shown to affect mathematical contents.

[105] Peirce (1932) 171.

absolutely indispensable in mathematics . . . So are the letters *A*, *B*, *C* etc., attached to a geometrical figure.'

The context from which the quotation is taken is richer, and one need not subscribe to all aspects of Peirce's philosophy of mathematics there. But I ask a descriptive, not a prescriptive question. What sense did people make of letters in diagrams? Peirce, at least, understood them as indices. I consider this a helpful piece of evidence. After all, why not take Peirce himself as our guide in semiotics?

2.3.2 The semiotics of diagrams

So far, I have argued that letters are primarily indices, so that representations employing them cannot but refer to the concrete diagram. A further question is the semiotics of the diagram itself: does it refer to anything else, or is it the ultimate subject matter?

First, the option that the diagram points towards an ideal mathematical object can be disposed of. Greek mathematics cannot be about squares-as-such, that is, objects which have no other property except squareness, simply because many of the properties of squares are not properties of squares-as-such; e.g. the square on the diagonal of the square-as-such is the square-as-such, not its double.[106] It is not that speaking about objects-as-such is fundamentally wrong. It is simply not the same as speaking about objects. The case is clearer in algebra. One can speak about the even-as-such and the odd-as-such: this is a version of Boolean algebra.[107] Modern mathematics (that is, roughly, that of the last century or so) is characterised by an interest in the theories of objects-as-such; Greek mathematics was not.[108]

So what *is* the object of the proof? As usual, I look to the practices for a guide. We take off from the following. The proposition contains imperatives describing various geometrically defined operations, e.g.: κύκλος γεγράφθω – 'let a circle have been drawn'.[109] This is a certain action, the drawing of a circle. A different verb is 'to be', as in the

[106] The impossibility of Greek mathematics being about Platonic objects has been argued by Lear (1982), Burnyeat (1987).

[107] As the above may seem cryptic to a non-mathematician, I explain briefly. What is 'the essence' of the odd and the even? One good answer may be, for instance, to provide their table of addition: Odd + Odd = Even, O + E = O, E + O = O, E + E = E. One may then assume the existence of objects which are characterised by this feature only. One would thus 'abstract' odd-as-such and even-as-such from numbers. Such abstractions are typical of modern mathematics.

[108] Of course, the import of Greek proofs is general. This, however, need not mean that the proof itself is about a universal object. This issue will form the subject of chapter 6.

[109] Euclid's *Elements* 1.1, 10.19–12.1.

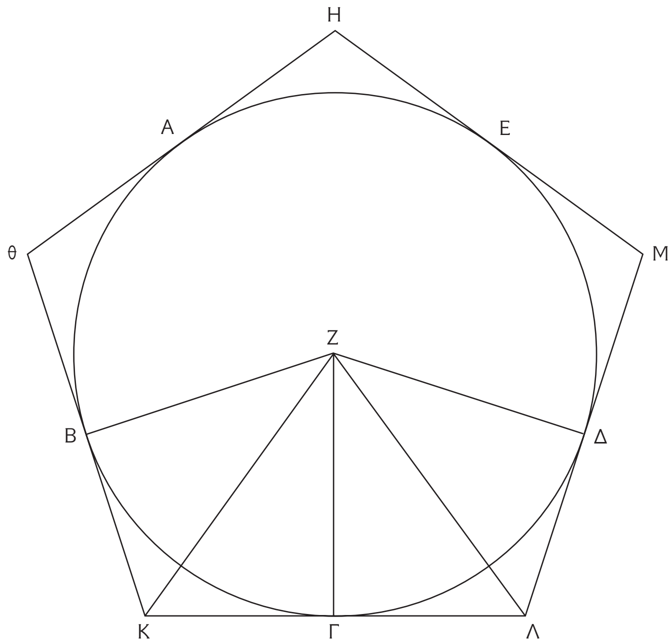

Figure 1.18. Euclid's *Elements* IV.12.

following:[110] ἔστω ἡ δοθεῖσα εὐθεῖα πεπερασμένη ἡ AB – 'let the given bounded straight line be AB'. The sense is that you identify the bounded given straight line (demanded earlier in the proposition) as AB. So this is another action, though here the activity is that of visually identifying an object instead of constructing it.

A verb which does not fit into this system of actions is *noein*, which may be translated here as 'to imagine', as in the following:[111]

> νενοήσθω τοῦ ἐγγεγραμμένου πενταγώνου τῶν γωνίων σημεῖα τὰ A, κτλ
>
> 'Let the points A, etc. be imagined as the points of the angles of the inscribed pentagon'.

What is the point of imagination here? The one noticeable thing is that the inscribed pentagon does not occur in the diagram, which for once should, with all the difficulties involved, be taken to reflect Euclid's diagram (fig. 1.18). On the logical plane, this means that

[110] Ibid. 10.16. [111] Euclid's *Elements* IV.12, 302.10–11.

the pentagon was taken for granted rather than constructed (its constructability, however, has been proved, so no falsity results).

Though not as common as some other verbs, *noein* is used quite often in Greek mathematics.[112] It is used when objects are either not drawn at all, as in the example above, or when the diagram, for some reason, fails to evoke them properly. The verb is relatively rare because such cases, of under-representation by the diagram, are relatively rare. It is most common with three-dimensional objects (especially the sphere, whose Greek representation is indeed indistinguishable from a circle). Another set of cases is in 'applied' mathematics, e.g. when a line is meant to be identified as a balance. Obviously the line is not a balance, it is a line, and therefore the verb *noein* is used.[113]

However, if the diagram is meant as a representation of some ideal mathematical object, then one should have said that any object whatever was 'imagined'. By delegating some, but not all, action to 'imagination', the mathematicians imply that, in the ordinary run of things, they literally mean what they say: the circle of the proof is drawn, not imagined to be drawn. It will not do to say that the circle was drawn in some ideal geometrical space; for in that geometrical space one might as easily draw a sphere. Thus, the action of the proof is literal, and the object of the proof must be the diagram itself, for it is only in the diagram that the acts of the construction literally can be said to have taken place.

This was one line of argument, showing that the diagram is the object of the proposition. In true Greek fashion, I shall now show that it is not the object of the proposition.

An obvious point, perhaps, is that the diagram must be false to some extent. This is indeed obvious for many moderns,[114] but at bottom this

[112] There are at least ten occurrences in Euclid's *Elements*, namely IV.12 302.10, XI.12 34.22, XII.4 lemma 162.21, 13 216.20, 14 220.7, 15 222.22 (that's a nice page and line reference!), 17 228.12, 228.20, 18 242.16, 244.6. There are three occurrences in Apollonius' *Conics* I, namely 52 160.18, 54 168.14, 56 178.12. Archimedes' works contain 38 occurrences of the verb in geometrical contexts, which may be hunted down through Heiberg's index. The verb is regularly used in Ptolemy's *Harmonics*. Lachterman (1989) claims on p. 89 that the verb is used by Euclid in book XII alone (the existence of Greek mathematicians other than Euclid is not registered), to mitigate, by its noetic function, the operationality involved in the generation of the sphere and the cylinder. We all make mistakes, and mine are probably worse than Lachterman's; but, as I disagree with Lachterman's picture of Greek mathematics as non-operational, I find it useful to note that this argument of his is false.

[113] E.g. Archimedes, *Meth.* 434.23 – one of many examples. The use of the verb in Ptolemy's *Harmonics* belongs to this class.

[114] E.g. Mill (1973), vol. VII 225: 'Their [sc. geometrical lines'] existence, as far as we can form any judgement, would seem to be inconsistent with the physical constitution of our planet at least, if not of the universe.' For this claim, Mill offers no argument.

is an empirical question. I imagine our own conviction may reflect some deeply held atomistic vision of the world; there is some reason to believe that atomism was already seen as inimical to mathematics in antiquity.[115] An ancient continuum theorist could well believe in the physical constructability of geometrical objects, and Lear (1982) thinks Aristotle did. This, however, does not alter the fact that the actual diagrams in front of the mathematician are not instantiations of the mathematical situation.

That diagrams were not considered as exact instantiations of the object constructed in the proposition can, I think, be proved. The argument is that 'construction' corresponds, in Greek mathematics, to a precise practice. The first proposition of Euclid's *Elements*, for instance, shows how to construct an equilateral triangle. This is mediated by the construction of two auxiliary circles. Now there simply is no way, if one is given only proposition I.1 of the *Elements*, to construct this triangle without the auxiliary circles. So, in the second proposition, when an equilateral triangle is constructed in the course of the proposition,[116] one is faced with a dilemma. Either one assumes that the two auxiliary circles have been constructed as well – but how many steps further can this be carried, as one goes on to ever more complex constructions? Or, alternatively, one must conclude that the so-called equilateral triangle of the diagram is a fake. Thus the equilateral triangle of proposition I.2 is a token gesture, a make-believe. It acknowledges the shadow of a possible construction without actually performing it.

We seem to have reached a certain impasse. On the one hand, the Greeks speak as if the object of the proposition is the diagram. Verbs signifying spatial action must be taken literally. On the other hand, Greeks act in a way which precludes this possibility (quite regardless of what their ontology may have been!), and the verbs signifying spatial action must, therefore, be counted as metaphors.

To resolve this impasse, the 'make-believe' element should be stressed. Take Euclid's *Elements* III.10. This proves that a circle does not cut a circle at more than two points. This is proved – as is the regular

[115] Plato's peculiar atomism involved, apparently, some anti-geometrical attitudes (surprisingly enough), for which see Aristotle, *Metaph.* 992a20ff. Somewhat more clear is the Epicurean case, discussed in Mueller (1982) 92–5. The evidence is thin, but Mueller's educated guess is that Epicureans, as a rule, did assume that mathematics is false.

[116] Euclid's *Elements* I.2, 12.25–6. Needless to say, the text simply says 'let an equilateral triangle have been set up on [the line]', no hint being made of the problem I raise.

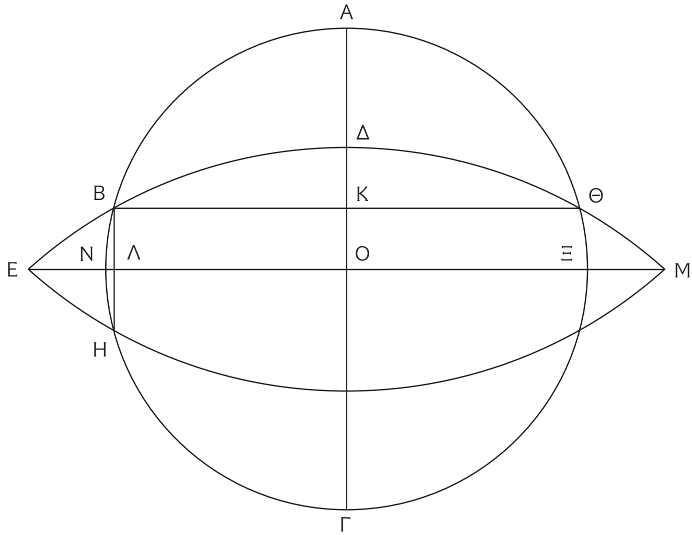

Figure 1.19. Euclid's *Elements* III.10.

practice in propositions of this nature – through a *reductio ad absurdum*: Euclid assumes that two circles cut each other at more than two points (more precisely, at four points), and then derives an absurdity. The proof, of course, proceeds with the aid of a diagram. But this is a strange diagram (fig. 1.19): for good geometrical reasons, proved *in this very proposition*, such a diagram is impossible. Euclid draws what is impossible; worse, what is patently impossible. For, let us remember, there is reason to believe a circle is one of the few geometrical objects a Greek diagram could represent in a satisfying manner. The diagram cannot be; it can only survive thanks to the make-believe which calls a 'circle' something which is similar to the oval figure in fig. 1.19. By the force of the make-believe, this oval shape is invested with circlehood for the course of the *reductio* argument. The make-believe is discarded at the end of the argument, the bells of midnight toll and the circle reverts to a pumpkin. With the *reductio* diagrams, the illusion is dropped already at the end of the *reductio* move. Elsewhere, the illusion is maintained for the duration of the proof.

Take Pünktchen for instance.[117] Her dog is lying in her bed, and she stands next to it, addressing it: 'But grandmother, why have you got such large teeth?' What is the semiotic role of 'grandmother'? It is not

[117] Kästner (1959), beginning of chapter 2 (and elsewhere for similar phenomena, very ably described. See also the general discussion following chapter 3).

metaphorical – Pünktchen is not trying to insinuate anything about the grandmother-like (or wolf-like) characteristics of her dog. But neither is it literal, and Pünktchen knows this. Make-believe is a *tertium* between literality and metaphor: it is literality, but an as-if kind of literality. My theory is that the Greek diagram is an instantiation of its object in the sense in which Pünktchen's dog is the wolf – that the diagram is a make-believe object. It shares with Pünktchen's dog the following characteristics: it is similar to the intended object; it is functionally identical to it; what is perhaps most important, it is never questioned.

2.4 *The practices of the lettered diagram: a summary*

What we have seen so far is a series of procedures through which the text maintains a certain implicitness. It does not identify its objects, and leaves the identification to the visual imagination (the argument of 2.2). It does not *name* its objects – it simply points to them, via indices (the argument of 2.3.1). Finally, it does not even hint *what*, ultimately, its objects are; it simply works with an ersatz, as if it were the real thing (the argument of 2.3.2). Obviously there is a certain vague assumption that some of the properties of the 'real thing' are somehow captured by the diagram, otherwise the mediation of the proposition via the diagram would collapse. But my argument explaining why the diagram is useful (because it is redefined, especially through its letters, as a discrete object, and therefore a manageable one) did not deal with the ontological question of why it is assumed that the diagram could in principle correspond to the geometrical object. Undoubtedly, many mathematicians would simply assume that geometry is about spatial, physical objects, the *sort* of a thing a diagram is. Others could have assumed the existence of mathematicals. The centrality of the diagram, however, and the roundabout way in which it was referred to, meant that the Greek mathematician would not have to speak up for his ontology.[118]

[118] Let me explain briefly why the indexical nature of letters is significant. This is because indices signify references, not senses. Suppose you watch a production of *Hamlet*, with the cast wearing soccer shirts. John, let's say, is the name of the actor who plays Hamlet, and he is wearing shirt number 5. Then asking 'what's your opinion of John?' would refer, probably, to his acting; asking 'what's your opinion of Hamlet?' would refer, probably, to his indecision; but asking 'what's your opinion of no. 5?' would refer ambiguously to both. Greek letters are like numbers on soccer shirts, points in diagrams are like actors and mathematical objects are like Hamlet.

Plato, in the seventh book of the *Republic*, prized the ontological ambiguity of mathematics, especially of its diagrams. An ontological borderline, it could confuse the philosophically minded, and lead from one side of the border to the other. He was right. However, this very ambiguity meant also that the mathematicians could choose not to engage in the philosophical argument, to stick with their proofs and mutual agreements – a point (as claimed above) conceded by Plato.

To conclude, then: there are two main ways in which the lettered diagram takes part in the shaping of deduction. First, there is the whole set of procedures for argumentation based on the diagram. No other single source of evidence is comparable in importance to the diagram. Essentially, this centrality reverts to the fact that the specification of objects is done visually. I shall return to this subject in detail in chapter 5. Second, and more complex, is this. The lettered diagram supplies a universe of discourse. Speaking of their diagrams, Greek mathematicians need not speak about their ontological principles. This is a characteristic feature of Greek mathematics. Proofs were done at an object-level, other questions being pushed aside. One went directly to diagrams, did the dirty work, and, when asked what the ontology behind it was, one mumbled something about the weather and went back to work. This is not meant as a sociological picture, of course. I am speaking not of the mathematician, but of the mathematical proposition. And this proposition acts in complete isolation, hermetically sealed off from any second-order discourse.[119] There is a certain single-mindedness about Greek mathematics, a deliberate choice to do mathematics and nothing else. That this was at all possible is partly explicable through the role of the diagram, which acted, effectively, as a *substitute* for ontology.[120]

It is the essence of cognitive tools to carve a more specialised niche within general cognitive processes. Within that niche, much is automatised, much is elided. The lettered diagram, specifically, contributed to both elision (of the semiotic problems involved with mathematical discourse) and automatisation (of the obtaining of a model through which problems are processed).

119 I will discuss this in chapter 3 below.

120 I am not saying, of course, that the *only* reason why Greek mathematics became sealed off from philosophy is the existence of the lettered diagram. The lettered diagram is not a cause for sealing mathematics off from philosophy; it is an important explanation of how such a sealing off was possible. I shall return to discussing the single-mindedness of Greek mathematics in the final chapter.

3 CONTEXTS FOR THE EMERGENCE OF THE LETTERED DIAGRAM

The lettered diagram is a distinctive mark of Greek mathematics, partly because no other culture developed it independently.[121] Indeed, it would have been impossible in a pre-literate society and, obvious as this may sound, this is an important truth.[122] An explanatory strategy may suggest itself, then: to explain the originality of the lettered diagram by the originality of the Phoenician script. The suggestion might be that alphabetic letters are more suitable, for the purpose of the lettered diagram, than pictograms, since pictograms suggest their symbolic content. The coloured constituents of some Chinese figures may be relevant here.[123]

But of course such technological reductionism – everything the result of a single tool! – is unconvincing. The important question is *how* the tool is used. This is obvious in our case, since the technology involved the combination of two different tools. Minimally, the contexts of diagrams and of letters had to intersect.

The plan of this section is therefore as follows. First, the contexts of diagrams and letters outside mathematics are described. Next, I discuss two other mathematical tools, abaci and planetaria. These, too, are 'contexts' within which the lettered diagram emerged, and understanding their limitations will help to explain the ascendancy of the lettered diagram.

3.1 Non-mathematical contexts for the lettered diagram

3.1.1 Contexts of the diagram

As Beard puts it,[124] 'It is difficult now to recapture the sheer profusion of visual images that surrounded the inhabitants of most Greek cities.' Greeks were prepared for the visual.

[121] Babylonian and Chinese diagrams exist, of course – though Babylonian diagrams are less central for Babylonian mathematics, or at least for Babylonian mathematical texts (Hoyrup 1990a), while Chinese diagrams belong to a different context altogether, of representations endowed with rich symbolic significance (Lackner 1992). Neither refers to the diagram with a system similar to the Greek use of letters. Typically, in the Babylonian case, the figure is referred to through its geometric elements (e.g. breadth and width of rectangles), or it is inscribed with numbers giving measurements of some of its elements (e.g. *YBC* 7289, 8633: Neugebauer 1945).

[122] Also, while this point may sound obvious, it would have been impossible to make without Goody (1977), Goody and Watt (1963) on the role of writing for the historical development of cognition and, generally, Goody's *œuvre*; this debt applies to my work as a whole.

[123] See Chemla (1994), however, for an analysis of this practice: what is important is not the *individual* colours, but their existence as a *system*. In fact, one can say that the Chinese took colours as a convenient metaphor for a system.

[124] Beard (1991) 14.

This is true, however, only in a limited sense. Greek elite education included literacy, numeracy, music and gymnastics, but not drawing or indeed any other specialised art.[125] The educated Greek was experienced in looking, not in drawing. Furthermore, the profusion of the visual was limited to the visual as an aesthetic object, not as an informative medium. There is an important difference between the two. The visual as an aesthetic object sets a barrier between craftsman and client: the passive and active processes may be different in kind. But in the visual as a medium of information, the coding and decoding principles are reciprocal and related. To the extent that I can do anything at all with maps I must understand some of the principles underlying them. On the other hand, while the 'readers' of art who know nothing about its production may be deemed philistines, they are possible. The visual as information demands some exchange between craftsmen and clients, which art does not.

Two areas where the use of the visual *qua* information is expected are maps and architectural designs. Herodotus gives evidence for world maps, designed for intellectual (IV.36ff.) and practical (V.49–51) purposes. Such maps could go as far back as Anaximander.[126] Herodotus' maps were exotic items, but we are told by Plutarch that average Athenians had a sufficiently clear grasp of maps to be able to draw them during the euphoric stage of the expedition to Sicily, in 415 BC.[127] Earlier, in 427, a passage in Aristophanes' comedy *The Clouds* shows an understanding of what a map is: schematic rather than pictorial,[128] preserving shapes, but not distances.[129] The main point of Aristophanes' passage is clear: though diagrammatic representations were understood by at least some members of the audience, they were a technical, specialised form. It may be significant that the passage follows immediately upon astronomy and geometry.

Our later evidence remains thin. There is a map in Aristotle's *Meteorology*,[130] and *periodoi gēs* – apparently world maps – are included, as

[125] Excluding mathematics itself – to the extent that it actually gained a foothold in education (see chapter 7).

[126] Agathemerus I.1; D.L. II.1–2; Herodotus II.109. Anaxagoras may have added some visual element to his book (D.L. II.11 – the first to do so? See also DK 59A18 (Plutarch), A35 (Clement)). I guess – and I can do no more – that this was a cosmological map (both Plutarch's and Clement's reference come from a cosmological context).

[127] *Vit. Alc.* XVII.3. The context is historically worthless, but the next piece of evidence could give it a shade of plausibility.

[128] 208–9: a viewer of the map is surprised to see Athens without juries!

[129] Shapes: 212, the 'stretched' island Euboea leads to a pun. Distances: 215–17, the naive viewer is worried about Sparta, which is too near.

[130] 363a26ff.

mentioned already, in Theophrastus' will.[131] There is also some – very little – epigraphic and numismatic evidence, discussed by Dilke.[132] Most interestingly, it seems that certain coins, struck in a military campaign, showed a relief-map of its terrain.[133] All these maps come from either intellectual or propaganda contexts. As early as Herodotus, the drawing of a map in pragmatic contexts was meant to impress rather than to inform. Otherwise, much of the evidence comes from sources influenced by mathematics.

Surprisingly, the same may be true of architectural designs.[134] The main tools of such design in classical times were either verbal descriptions (*sungraphai*), or actual three-dimensional and sometimes full-scale models of repeated elements in the design (*paradeigmata*). Rules of trade, especially a modifiable system of accepted proportions, allowed the transition from the verbal to the physical. There is a strong *e silentio* argument against any common use of plans in early times. From Hellenistic times onwards, these began to be more common, especially – once again – in the contexts of persuasion rather than of information. This happened when competition between architects forced them to evolve some method of conveying their intentions beforehand, in an impressive manner. Interestingly, the use of visual representations in architecture is earliest attested in mechanics, which may show a mathematical influence.

What is made clear by this brief survey is that Greek geometry did not evolve as a reflection upon, say, architecture. The mathematical diagram did not evolve as a modification of other practical diagrams, becoming more and more theoretical until finally the abstract geometrical diagram was drawn. Mathematical diagrams may well have been the first diagrams. The diagram is not a representation of something else; it is the thing itself. It is not like a representation of a building, it is like a building, acted upon and constructed. Greek geometry is the study of spatial action, not of visual representation.

However speculative the following point may be, it must be made. The first Greeks who used diagrams had, according to the argument above, to do something similar to building rather than to reflect upon building. As mentioned above, the actual drawing involved a practical skill, not an obvious part of a Greek education. Later, of course, the lettered diagram would be just the symbol of mathematics, firmly

[131] D.L. v.51–2. [132] Dilke (1985) chapter 2.
[133] Johnston (1967). [134] The following is based on Coulton (1977) chapter 3.

situated there; but at first, some contamination with the craftsman-like, the 'banausic', must be hypothesised. I am not saying that the first Greek mathematicians were, e.g. carpenters. I am quite certain they were not. But they may have felt uneasily close to the banausic, a point to which I shall return in the final chapter.

3.1.2 *Contexts of letters as used in the lettered diagram*

Our earliest direct evidence for the lettered diagram comes from outside mathematics proper, namely, from Aristotle. There are no obvious antecedents to Aristotle's practice. Furthermore, he remained an isolated phenomenon, even within the peripatetic school which he founded. Of course, logical treatises in the Aristotelian tradition employed letters, as did a few quasi-mathematical works, such as the pseudo-Aristotelian *Mechanics*. But otherwise (excluding the mathematically inclined Eudemus) the use of letters disappeared. The great musician Aristoxenus, just like the great mechanician Strato – both in some sense followers of Aristotle – do not seem to have used letters. The same is true more generally: the Aristotelian phenomenon does not recur. And, of course, nothing similar to our common language use of '*X*' and '*Y*' ever emerged in the Greek language.[135]

Otherwise, few cases of special sign systems occur. At some date between the fifth and the third centuries BC someone inserted an acrophonic shorthand into the Hippocratic *Epidemics* III.[136] Galen tells us about another shorthand designed for pharmaceutical purposes, this time based, in part, upon iconic principles (e.g. *omicron* for 'rounded').[137] A refined symbolic system was developed for the purposes of textual criticism. Referring as it did to letters, the system employed *ad hoc* symbols.[138] This system evolved in third-century Alexandria. Another case of a special symbolic system is musical notation, attested from the third century BC but probably invented earlier.[139] Letters, grouped and repeated in various ways, are among other symbols considered to have magical significance.[140] Finally, many systems

[135] Which should not surprise us: the Greek letters as used in diagrams, being indices, were inseparable from specific situations, unlike the modern symbolic '*X*'.

[136] This is not a feature of the manuscripts alone – which might have suggested a Byzantine origin – since Galen reports the system, XVII 600ff.

[137] Galen XIII 995–6. The system is due to Menecrates, of an early AD provenance.

[138] See Turner (1968) esp. 112–18.

[139] West (1992) chapter 9.

[140] See Betz (1992) for many examples, e.g. 3, 17 (letters), 191 (other symbols). For a discussion, see Dornseiff (1925).

of abbreviation are attested in our manuscripts, and while the vast majority are Byzantine, 'shorthand' was known already in antiquity.[141] The common characteristic of all the above is their reflective, written context. These are all second-order signs: signs used to refer to other signs. Being indices to diagrams, the letters of Greek mathematics form part of the same pattern.

What we learn is that the introduction of a special sign-system is a highly literate act – this indeed should have been obvious to start with. The introduction of letters as tools is a reflective use of literacy. Certainly the social context within which such an introduction could take place was the literate elite.

3.2 Mathematical non-verbal contexts

Generally speaking, mathematical tools are among the most widespread cultural phenomena of all, beginning with the numerical system itself and going through finger-reckoning, abaci, etc., up to the computer.[142] Many of these tools have to do with calculation rather than proof and are thus less important for my purposes here. Two tools used by Greek mathematics, besides the lettered diagram, may have been of some relevance to proof, and are therefore discussed in the following subsections: these are abaci and planetaria.

It is natural to assume that not all tools can lead equally well to the elaboration of scientific theories. To make a simple point, science demands a certain intersubjectivity, which is probably best assisted through language. A completely non-verbalised tool is thus unlikely to lead to science.[143] On the other hand, intersubjectivity may be aided by the presence of a material object around which communication is organised. Both grounds for intersubjectivity operate with the lettered diagram; I shall now try to consider the case for other tools.

[141] See, e.g. Milne (1934). The compendia used in mathematical manuscripts are usually restricted to the scholia. It doesn't seem that abbreviations were important in Greek mathematics, as, indeed, is shown by the survival of Archimedes in Doric.

[142] See, e.g. Dantzig (1967). Schmandt-Besserat (1992, vol. I: 184ff.) is very useful.

[143] I am thinking of the Inca *quipu* (where strings represent arithmetical operations) as a tool where verbalisation is not represented at all (as shown by the problematic deciphering) (Ascher 1981).

3.2.1 The abacus in Greek mathematics

The evidence is:

(a) Greeks used pebbles for calculations on abaci.[144]
(b) Some very few hints suggest that something more theoretical in nature was done with the aid of pebbles.[145]
(c) It has been argued that a certain strand in early Greek arithmetic becomes natural if viewed as employing pebbles. According to this theory, some Greeks represented numbers by configurations of pebbles or (when written) configurations of dots on the page: three dots represent the number three, etc.[146] However:
(d) Not a single arithmetical BC text refers to pebbles or assumes a dot representation of an arithmetical situation.

Philip[147] argued that we should not pass too quickly from (b) to (c). Certainly, Eurytus' pebbles need not be associated with anything the Greeks themselves would deem arithmetical. I shall argue below[148] that what is sometimes brought as evidence, Epicharmus' fragment 2, belongs to (a) and not to (b), let alone (c). Similarly, Plato's analogy of mathematical arts and *petteutikē* – pebble games[149] – need not involve any high-powered notion of mathematics.

This leaves us with two Aristotelian passages:

'Like those who arrange numbers in shapes [such as] triangle and square';[150]

'For putting gnomons around the unit, and without it, in this [case] the figure will always become different, in the other it [will be] unity'.[151]

[144] Lang (1957).

[145] The only substantial early hints are the two passages from Aristotle quoted below (which can be somewhat amplified for Eurytus by *DK* 45A2: he somehow related animals(?) to numbers, via pebble-representations).

[146] Becker (1936a). Knorr (1975) goes much further, and Lefevre (1988) adds the vital operational dimension.

[147] Philip (1951), appendix II, esp. 202–3. [148] Chapter 7, subsection 1.1 272–5.

[149] *Grg.* 450cd; *Lgs.* 819d–820d; also relevant is *Euthyph.* 12d.

[150] *Metaph.* 1092b11–12: ὥσπερ οἱ τοὺς ἀριθμοὺς ἄγοντες εἰς τὰ σχήματα τρίγωνον καὶ τετράγωνον.

[151] *Phys.* 203a13–15: περιτιθεμένων γὰρ τῶν γνωμόνων περὶ τὸ ἓν καὶ χωρὶς ὁτὲ μὲν ἄλλο ἀεὶ γίγνεσθαι τὸ εἶδος, ὁτὲ δὲ ἕν. Both passages are mere clauses within larger contexts, and are very difficult to translate.

Philip maintained that, however arithmetical these passages may sound, they are relatively late fourth-century and therefore might be due to the great mathematical progress of that century, and so need have nothing to do with the late fifth century. Knorr[152] quite rightly objected that this makes no evolutionary sense: could that progress lead to mathematics at the pebbles level? Knorr must be right, but he does not come to terms with the fact that our evidence is indeed late fourth-century. Moreover, the texts refer to Pythagoreans, in connection with Plato, and the natural reading would be that Aristotle refers to someone roughly contemporary with Plato. Thus, our only evidence for an arithmetical use of pebbles comes from a time when we know that mathematically stronger types of arithmetic were available.

I certainly would not deny the role of the abacus for Greek arithmetical concept-formation.[153] The question is different: whether any arithmetical proof, oral or written, was ever conducted with the aid of pebbles. The evidence suggests, perhaps, oral proofs. Aristotle talks about people *doing* things, not about anything he has *read*. Why this should be the case is immediately obvious. Pebble manipulations admit a transference to a written medium, as is amply attested in modern discussions. However, the special advantage of pebbles over other types of arithmetical representations is a result of their direct, physical manipulations, which are essentially tied up with actual operations. It is not the mere passive looking at pebbles which our sources mention: they mention pebbles being moved and added. This must be lost in the written medium, which is divorced from specific actions. Thus, it is only natural that pebbles would lose their significance as the written mode gained in centrality. They would stay, but in a marginal role, emerging in a few asides by Plato and Aristotle, never as the centre of mathematical activity.[154]

[152] Knorr (1975) 135–7.

[153] Lefevre (1988) offers a theory of such concept formation, with a stress on the general role of operations for concept-formation.

[154] An important comparison is the following, which, however, being no Assyriologist, I will express tentatively and in this footnote alone. The geometrical reconstructions offered by Hoyrup (1990a) for Babylonian 'algebra' take the shape of *operations* upon spatial objects, moved, torn and appended – following the verbs of the Akkadian text. I would say:

1. The loss of (most) diagrams from Babylonian mathematics is related to this manner in which Babylonian mathematics was visualised. The texts refer to objects which were actually moved, not to inscribed diagrams.
2. The visualisation was operational because the role of the text was different from what it is in the Greek case. Babylonian mathematical texts are not context-independent; they are

3.2.2 Planetaria in Greek mathematics

The earliest and most extensive piece of evidence on planetaria in Greek astronomy is Epicurus' – biased – description of astronomical practices, in *On Nature* XI.[155] The description is of a school in Cyzicus, where astronomers are portrayed as using *organa*, 'instruments', while *sullogizesthai*, *dialegesthai* (i.e. reasoning in various ways), having *dianoia* (translated by Sedley in context by 'a mental model') and *epinoēsis* ('thought-process') and referring to a *legomenon* (something 'said' or 'asserted'). What is the exact relation between these two aspects of their practice, the instrument and the thought? One clue is the fact that Epicurus claims that the aspects are irreconcilable because, according to him, the assumption of a heavens/model analogy is indefensible. This assumes that some dependence of the verbal upon the mechanical is necessary. This dependence might be merely the thesis that 'the heavens are a mechanism identical to the one in front of us', or it might be more like 'setting the model going, we see [e.g.] that some stars are never visible, QED'. Where in the spectrum between these options should we place the mathematicians of Cyzicus?[156]

My following guess starts from Autolycus, a mathematician contemporary with this Epicurean text. Two of his astronomical works survive – *The Moving Sphere* and *The Risings and Settings*. He never mentions any apparatus, or even hints at such, even though *The Risings and Settings* are practical astronomy rather than pure spherical mathematics. Neither, however, does he give many definitions or, generally, conceptual hints.[157] Furthermore, as mentioned above, his diagrams – belonging

the internal working documents of scribes, who know the operational context in which these texts are meant to be used.

3. The different contexts and technologies of writing meant that in one case (Mesopotamia) we have lost the visualisations alone, while in the other (Greek pebble arithmetic) we have lost both visualisations and text.
4. Babylonian mathematics is limited, compared to Greek mathematics, by being tied to the particular operation upon the particular case; which reflects the difference mentioned above.

155 Sedley (1976) 32–4. The text survives only on papyrus.

156 And not only them: the evidence for the use of planetaria (and related star-modelling mechanisms) in antiquity goes beyond any other archaeological evidence for mathematics. A truly remarkable piece of evidence is the Antikytheran 'planetarium', described in Price (1974). See there the evidence for sundials (51), and for other planetaria (55–60).

157 That the definitions of *The Moving Sphere* are spurious is probable, though not certain. See Aujac (1979) 40 (in the edition of Autolycus used in this study: see Appendix, p. 314), who rejects them. If they are spurious, then they are the result of a perplexity similar to that which the modern reader must feel. The definitions of *The Risings and Settings* explain the terminology of observation, not the spatial objects discussed.

as they do to the theory of spheres – are sometimes only very roughly iconic. The reader – who may be assumed to be a beginner – is immediately plunged into a text where there is a very serious difficulty in visualising, in conceptualising. No doubt much of the difficulty would have been solved by the Greek acquaintance with the sky. But a model would certainly be helpful as well, at such a stage. After all, you cannot turn the sky in your hands and trace lines on its surface. An object which can be manipulated would contribute to concept-formation.[158] This acquaintance is more than the mere analogy claim – the model is used to understand the heavens – yet this is weaker than actually using the model for the sake of proof.[159]

Timaeus excuses himself from astronomy by claiming that τὸ λέγειν ἄνευ δι᾽ ὄψεως τούτων αὖ τῶν μιμημάτων μάταιος ἂν εἴη πόνος – 'again, explaining this without watching models would be a pointless task'.[160] This, written by the staunch defender of mathematical astronomy! It seems that models were almost indispensable for the pedagogic level of astronomy. The actual setting out in writing of mathematical astronomy, however, does not register planetaria. Again, just as in the case of the abacus, the tool may have played a part in concept-formation. And a further parallelism with the abacus is clear. Why is it difficult for Timaeus to explain his astronomy? Why indeed could he not have brought his planetaria? The answer is clear: the written text filtered out the physical model.

In Plato's case, of course, not only physical models were out of the question: so were diagrams, since the text was not merely written, but also the (supposed) reflection of conversation, so that diagrams used by the speakers must be reconstructed from their speeches (as is well known, e.g. for the *Meno*). Plato's text is double-filtered. More generally, however, we see that the centrality of the written form functions as a filter. The lettered diagram is the tool which, instead of being filtered out by the written mode, was made more central and, with the marginalisation of other tools, became the metonym of mathematics.

158 For whatever its worth, it should be pointed out that Epicurus' criticisms fasten upon the concept-formation stage.

159 This is certainly not the only purpose of building planetaria. Planetaria could do what maps did: impress. Epicurus is setting out to persuade students away from Cyzicus. The planetarium seems to have been set up in order to persuade them to come.

160 Plato, *Tim.* 40d2–3.

4 SUMMARY

Much of the argument of this chapter can be set out as a list of ways in which the lettered diagram is a combination of different elements, in different planes.

(a) On the logical plane, it is a combination of the continuous (diagram) and the discrete (letters), which implies,
(b) On the cognitive plane, a combination of visual resources (diagram) and finite, manageable models (letters).
(c) On the semiotic plane, the lettered diagram is a combination of an icon (diagram) and indices (letters), allowing the – constructive – ambiguity characteristic of Greek mathematical ontology.
(d) On the historical plane, it is a combination of an art, almost perhaps a banausic art (diagram) and a hyper-literate reflexivity (letters).

The line of thought suggested here, that it is the fertile intersection of different, almost antagonistic elements which is responsible for the shaping of deduction, will be pursued in the rest of the book.