Lucio Russo

The Forgotten Revolution

How Science Was Born in 300 BC and Why It Had to Be Reborn

With the Collaboration of the Translator, Silvio Levy

Springer
2
Hellenistic Mathematics

2.1 Precursors of Mathematical Science

The term “mathematics” is seldom defined by historians of the subject: for instance, Boyer’s History of Mathematics explicitly avoids the task, saying merely that “much of the subject […] is an outgrowth of thought that originally centered on the concepts of number, magnitude and form.”\(^1\) Were we to take this as the basis for a definition, mathematics would not only go back to paleolithic times — and indeed long before, since one can talk about the “mathematical abilities” of various animals, and research has been done on the issue — but would encompass even the Neapolitan smorfia, a series of rules for extracting from dreams information supposed to be helpful in predicting winning lottery numbers. This too, one must admit, deals with questions centered on the concept of number.

But mathematical science, in the sense we have given the word, arises in the Hellenistic period. Of course Hellenistic mathematics does not come out of nothing. Earlier mathematics can be divided, roughly speaking, into two phases. The first, extremely long, phase includes the mathematics of Old Babylonia and of Egypt under the Pharaohs. The second consists of a period of approximately two and a half centuries in which classical Greece created what we will call Hellenic mathematics, to distinguish it from Hellenistic mathematics.

The first period started in the paleolithic, with the ability to count,\(^3\) and saw the accumulation of a remarkable body of knowledge, as in Egypt.

\(^1\)[Boyer], p. 1.
during the Pharaonic era; there, for the first time, appeared specialized writings with mathematical content. These writings can be called mathematical only in that their object is solving problems that we would call arithmetical or geometric; they completely lack the rational structure that we associate with mathematics today. They contain recipes for solving problems—for example, calculating the volume of a truncated pyramid or the area of a circle (the latter being, of course, unintentionally approximate)—but there is no sign of anything like a justification for the rules given. At that stage, then, fairly elaborate notions beyond the integers had already been developed, including many plane and solid figures, area, and volume; problem-solving methods were passed down the generations; but the correctness of the solutions was based solely on experience and tradition. This was very far from being a science in the sense we have given the word. It was simply a part of that enormous store of empirical knowledge that enabled the Egyptians to achieve their famous technological feats; it was methodologically homogeneous with the rest of such knowledge, and transmitted in the same way.

Otto Neugebauer, one of the twentieth century’s most accomplished scholars of ancient exact science, wryly remarked:

Modern authors have often referred to the marvels of Egyptian architecture [in connection with their mathematics], though without ever mentioning a concrete problem of statics solvable by known Egyptian arithmetical procedures.

There is nothing surprising about the lack of applications of Egyptian mathematics to statics or other theories with technological interest. Since mathematical and technological knowledge alike were purely empirical, either could be applied only to directly related, concrete, specific problems; there was no scientific theory within which technological planning could be carried out, so there could not have been what we have called scientific technology. The quantities considered in mathematical problems known from Pharaoh-era Egypt are not internal to a theory, but instead had immediate and concrete interest: the number of bricks needed for a building of a given shape and size, or the area of a field.

3 Animal bones have been found in today’s Lebanon, dating from 15000 to 12000 B.C., with series of notches arranged into groups of equal cardinality. Thus the recording of tallies far predates the invention of writing, which for that matter seems to have arisen precisely from the evolution of a bookkeeping system (see Section 7.2).

4 For a review of sources, see [Gillings].

5 The papyrus Anastasi I gives some perspective on the role of “mathematical” knowledge in the context of the competencies required of a scribe. See [Gardiner].

6 [Neugebauer: ESA], p. 151.
Similar considerations apply to Old Babylonian mathematics, though it had reached a higher level.

The bridge between the empirical knowledge of Pharaoh-era Egypt and ancient Mesopotamia, on the one hand, and the sophisticated mathematical science of the Hellenistic period, on the other, is Hellenic mathematics. Without it, the transition would have been unthinkable. During these two and a half centuries Greek culture assimilated the Egyptian and Mesopotamian results and subjected them to a sharp rational analysis, closely linked to philosophical inquiry. Greek tradition names two pioneers in these investigations: Thales, who supposedly started developing, at the beginning of the sixth century B.C., the geometry that he had learned in Egypt, and Pythagoras, who founded his famous political and religious association during the second half of the same century.

An ancient tradition, attested by (among other things) the name “Pythagorean theorem” given to the famous theorem of geometry, holds that the deductive method arose at the very beginning of Hellenic mathematics. This belief goes back at least to the History of geometry written by Aristotle’s disciple, Eudemus of Rhodes, according to whom Thales proved that a diameter divides a circle into two equal parts, and that opposite angles at a vertex are equal. But it is not possible that statements so apparently obvious could have been among the first to be demonstrated. The usefulness of the deductive method must have been noticed first in proving nonobvious statements. Only when a well-developed deductive system is attained can the demand arise for demonstrations of such apparently obvious statements as the ones attributed to Thales.

In fact Eudemus systematically backdated mathematical results, in a process made explicit by Proclus in at least one case:

Eudemus, in his History of geometry, attributes to Thales this theorem [that triangles having one side and two adjacent angles equal are congruent], because, in his opinion, the method with which Thales is said to have determined the distance of ships in the sea depends on the use of this theorem.

It is clear from this passage how Eudemus confused the logical order, which requires a theorem to be proved before it can be applied, with the historical order. In fact, for the application he mentions, it is not necessary
to know the theorem: it is enough to be convinced (and only in the particular case in question) of the truth of the statement of the theorem, without knowing its proof; and without the proof one cannot, of course, talk about its being a theorem. The error made by Eudemus is still made today. One often reads that the "Pythagorean theorem" was known in Mesopotamia in the Old Babylonian period; actually what was known was its empirical basis, the fact that the square on the hypotenuse has the same area as the sum of the squares on the sides. The idea of proofs and theorems had not been invented in Old Babylonia, nor yet in Pythagorean times.

Without getting into the history of Hellenic mathematics,¹⁰ we will show with some examples how it was not a science—not only at the time of Thales and Pythagoras, but even much later.

A remarkable example of the state of Hellenic mathematics in mid-fifth century is afforded by Zeno's paradoxes, which are so famous (particularly the one with Achilles and the turtle) that we need not repeat them here.¹¹ Why are they thought of today primarily as a philosophical subject, although they deal with the concept of a continuous quantity, which is essential in mathematics? Above all because Zeno talks about space and time and not about their mathematical model, which had not been constructed then. The instruments used to analyze the notions of space and time are ordinary language and philosophy; the structure of a scientific theory, in the sense defined in Section 1.3, is still lacking. Zeno's paradoxes certainly influenced significantly the evolution of the concept of a continuous magnitude, which eventually resulted in the sophisticated theory expounded in Book V of Euclid's Elements, but once the building had been erected there was no place in it for that type of reasoning.

Another important example, traditionally attributed to the Pythagorean school, is the discovery of the incommensurability of the side and diagonal of a square. This is often quoted as a demonstration of the irrationality of the square root of 2, but the original argument should not be blurred with its later development. One reconstruction of the early state of affairs is the following.¹² We know that the early Pythagoreans thought that

---

¹⁰Perhaps because of the complete absence of primary sources, Hellenic mathematics has induced much more writing that its Hellenistic heir. Among the books devoted wholly or mostly to Hellenic mathematics, we mention [Lasserre], [Szabó], [Knorr: EEE], and [Fowler]. For Greek mathematics of both periods, the standard reference is [Heath: HGM], while [Loria] can also be useful. An anthology of original texts can be found in [GMW], whose two little volumes are all the Loeb Classical Library spares for Greek mathematics.

¹¹The main source for Zeno's paradoxes is Aristotle, Physica, VI, ix, 239b–240a. This passage and all other relevant sources are reported in [FVJ, L, 247–258. The paradoxes are discussed in [Heath: HGM], vol. 1, pp. 271–283.

¹²Reconstructions of this episode are based primarily on two sources. The older one is a passage of Aristotle (Analytica priora, I, xxiii, 41a:26–27), which says that the diagonal is not commensurable
every segment was made up of finitely many points.\(^{13}\) If we construct a square whose side is made up of an odd number of points, say \(k\), we can ask whether the number \(n\) of points of the diagonal is even or odd. Since the square of \(n\), by the Pythagorean theorem, is \(2k^2\), and therefore even, and only an even number can have an even square, it can easily be concluded that \(n\) is even. But someone must have noticed that if \(n\) were even its square would be a multiple of 4, whereas \(2k^2\) is not such a multiple if \(k\) is odd; therefore \(n\) must be odd. Since both reasonings appear convincing, but an odd number cannot be even, they did not know what to conclude.\(^{14}\)

The result is an impasse, analogous (from the viewpoint of the culture of the time) to Zeno's paradoxes. Because the Pythagoreans attached great importance to the opposition between even and odd numbers,\(^{15}\) it is reasonable to assume that the difficulty just described arose among them, as tradition maintains. But if the reconstruction above is correct, the Pythagoreans had not proved anything by contradiction; they had simply reached a contradiction (to their chagrin!), while trying to find the parity of the diagonal. Note that to get to this point it is not necessary to know the Pythagorean theorem in full generality, but only in the case of an isosceles right triangle, and the validity of this case can easily be verified by counting half-squares in a square array.

Unlike Zeno's paradoxes, the argument just discussed, which most likely dates from the last quarter of the fifth century,\(^{16}\) was later incorporated into mathematical science, where it provides the base for the proof of a theorem. But to reach a theorem there must be a qualitative leap allowing

\[^{13}\text{This can reasonably be deduced from several elements: the fact, reported by many sources, that the Pythagoreans based geometry on the integers; the Pythagorean theories of "figurative numbers" (for which see [Heath: HGM], vol. I, pp. 76-84, and [Knorr: EEE], Chapter 5); and above all Aristotle's assertion that the Pythagoreans attributed a magnitude to the units that made up material bodies (Aristotle, \textit{Metaphysica}, XIII, vi, 1080b:16-21 + 1083b:8-18 = [FV], I, 453:39 - 454:9, Pythagoreans B9, B10). Sextus Empiricus seems to still be thinking in Pythagorean mode when he says that it is impossible to bisect a segment formed by an odd number of points: \textit{Adversus physicos} I (= \textit{Adv. dogmaticos} III = \textit{Adv. mathematicos} IX), \$283; \textit{Adversus geometras} (= \textit{Adv. math.} III), \$110-111.}\]

\[^{14}\text{The reconstruction given here, apart from the modernized notation (the use of letters to denote numbers is not part of the Greek tradition) and the use of "points" as in the Pythagorean tradition, follows in essence the sources that report the theorem on the incommensurability of a square's side and diagonal. It is generally accepted that incommensurability was first discovered in this case, particularly because Plato and Aristotle always talk about it in connection with this example. A different conjecture about the origin of the idea of incommensurability is argued for in [von Fritz].}\]

\[^{15}\text{Philolaus, as quoted by Stobaeus (\textit{Eclogae}, I, xxi §7c, 1889-12 = [FV], I, 408:7-10, Philolaus B5); Aristotle, \textit{Metaphysica}, I, v, 986a:18+23-24. Evenness and oddness were still at the basis of arithmetic for Plato (\textit{Gorgias}, 451a-b). For a discussion of the Pythagorean ideas about even and odd numbers, see [Knorr: EEE], pp. 134-142.}\]
the circumvention of the impasse. Only by abandoning the Pythagorean notion that a line segment is made up of a succession of points — and therefore abandoning the Pythagorean program of basing explanations about real objects on the concept of an integer — can one attain the idea that two segments may not admit a common subdivision — may be incommensurable. The impasse can then be transformed into the proof by contradiction, known to Aristotle, that the side and diagonal of a square are incommensurable. There is no evidence for classifying this step as being Pythagorean. That neither Plato nor Aristotle, in any of several passages where they discuss the problems posed by incommensurability, ever attribute its discovery to the Pythagoreans is a strong indication that neither should we; that neo-Pythagoreans did so attribute it is sufficiently explained by the likelihood that the realization of the difficulty dated to Pythagoreans.

Even after the essential step of transforming the impasse into a proof by contradiction, the result remains purely negative: the statement of an impossibility, insufficient to serve as the basis for a theory of continuous magnitudes. The mathematicians of the beginning of the fourth century knew several pairs of incommensurable magnitudes, as we know from Plato, but they did not have a “theory of irrationalities”. As attested by the very word “irrational”, they did not say that the ratio between the side and the diagonal of a square is irrational, but rather that there is no such ratio.

Because the Pythagoreans probably thought that statements about the parity of the number of points on the diagonal referred to something in
2.1 Precursors of Mathematical Science

the real world, there may be some truth to the tradition that they regarded
the discovery of incommensurability (or rather, of the impasse that led to
the notion of incommensurability) as a dramatic event. If we discover that
a scientific theory is contradictory, it's no big deal: we change theories. But
what can we do if we discover, or think we have discovered, a contradic­
tion in reality itself?

The widespread idea that the discovery of incommensurability shook
the foundations of mathematics is based on the assumption that in the
fifth century B.C. mathematics already existed in our sense; it must have
arisen by analogy with the shaking of the foundations of mathematics
at the turn of the twentieth century. What could have been shaken at
that time was the original Pythagorean philosophical framework, which
wished to be the foundation (in a sense very different from the one used
by today's scholars) of much more than our mathematics. The studies that
we call mathematical today would have continued without much trouble,
precisely because they did not have a monolithic foundation.

Our third example of Hellenic mathematics comes from Plato, who was
very interested in the methods used by geometers of his time, and who
presents in his works mathematical arguments of great interest. As an
example of a demonstration expounded by Plato, we recall the famous
passage in the Meno where it is shown that, given a square, the square
built on the diagonal is double. The proof consists in observing that the

![Diagram of geometric shapes]

second square is formed of four triangles, each of which is equal to half the
initial square. This presupposes assumptions that are not made explicit
(among them: a square can be built on a given side, and the four triangles
into which a square is divided by the two diagonals are equal). In other
words, the truth of a geometric statement is deduced logically from other
statements chosen ad hoc among those that are visibly true. Precisely be­
cause a proof such as the one in the Meno is not embedded in a theory, but
stands on its own, independent of any other geometric line of argument,

---

30 This point is emphasized in [Knorr: EEE], p. 307.
31 Fabio Acerbi has shown that one can recognize in Plato (Parmenides, 149a.7–c:3) an example
of a proof by complete induction, a method that is usually considered to have been introduced in
modern mathematics. See [Acerbi: Plato].
32 Plato, Meno, 84e–85b.
it can, as in this case, be understood by a young slave completely ignorant of geometry.

In the *Timaeus* Plato explains the growth of a young body as follows:

[Consider] the young constitution of the whole animal, which has the triangles of the elements new.... Since the triangles coming in from outside, which make up food and drink, are older and weaker than its own triangles, it overpowers them and cuts them up with its new triangles, making the animal grow by nourishing it with many similar elements.  

What does Plato mean by "triangle"? It is clearly something real, not a theoretical entity such as we, after twenty-three hundred years of Euclidean geometry, naturally think it should be. Indeed, Plato states elsewhere that mathematical objects are endowed with a higher level of reality than that of their perceptible images; this was certainly an important thought in the process leading to the conscious construction of theoretical entities.

The Hellenic period — or at least most of it — can be considered as a long gestation of mathematical science, in which ever more refined logical instruments were being accumulated, but mathematics had not yet reached the stage of a science in the sense we have given the word, since there was not a single logically coherent and connected corpus of knowledge inside which any student whatsoever could solve an unlimited number of "exercises".

Probably a key person in the transition from the Hellenic to the Hellenistic period was Eudoxus of Cnidus, who according to the traditional boundary falls at the end of the Hellenic period. Because all his works have perished, it is hard to ascertain whether he was a precursor or the founder of mathematical science of the sort that appears in the *Elements*. We will revisit the relationship between Euclid and his fourth-century predecessors in Section 2.6.

In the next few sections, without intending to outline a history of Hellenistic mathematics, we will illustrate some of its fundamental aspects by means of examples, taken primarily from Euclid and Archimedes.

---

25 The critical edition of his works is [Euclid: OO]. An English translation of the *Elements* and a rich historical and critical apparatus can be found in [Euclid/Heath].
26 All his surviving works can be found in [Archimedes/Mugler]. Another very useful book is [Dijksterhuis: Archimedes], which contains a detailed exposition of the contents of his existing works.
2.2 Euclid’s Hypothetico-Deductive Method

Three types of problems came up in Hellenic mathematics. First, it was noticed how certain apparently obvious statements about geometric figures could imply others, much less obvious. This revealed the usefulness of the deductive method; but of course, as already remarked in Aristotle,\(^{27}\) one could not prove everything without causing the proof of any statement to involve an infinite regression. Second, paradoxes such as Zeno’s and the impasse found by the Pythagoreans had made apparent the high degree of subtlety of concepts such as space, time, and infinity, and of the relations between discrete and continuous magnitudes; it had also shown how inadequate everyday language is for dealing with such questions. Finally, there was the question of the unclear relationship between the concepts of mathematics and the real world.

In Euclid’s *Elements* we can see for the first time the solution to these problems, which was reached by establishing mathematics as a scientific theory—more precisely, by explicitly defining the theory’s entities (circles, right angles, parallel lines, and so on) in terms of a few fundamental entities (such as points, lines, and planes)\(^{28}\) and by listing the statements about such entities that must be accepted without proof.

In the *Elements* there are five statements of this type, called “postulates” (αἵτιματα):\(^{29}\)

1. [One can] draw a segment from any point to any point.
2. [One can] continuously extend a segment to a line.
3. [One can] draw a circumference with arbitrary center and radius.
4. All right angles are equal to one another.
5. If a line transversal to two lines forms with them in the same half-plane internal angles whose sum is less than two right angles, the two lines meet in that half-plane.

Any other statement regarding geometric entities can and should be accepted as true only if it can be supported by a proof (ἀπόδειξις), that is, a chain of logical implications that starts from the postulates (and the “common notions”) and leads to the given statement. This method is known to anyone who has studied mathematics in high school (at least that was the


\(^{28}\) In the *Elements* even these fundamental entities are “defined”, and the presence of these “definitions” (which are mere tautologies or purely illustrative statements) appears to contradict the thesis of the present discussion. This important question will be the subject of Section 10.14, where we will be able to study it in light of the material contained in intervening chapters.

\(^{29}\) There are also five “common notions”, that is, statements that are not about the specific entities of geometry. However, the authenticity of the “common notions” has often been contested. See, for example, [Euclid/Heath], vol. I, pp. 221 ff.
case until a while ago), because it was inherited by modern mathematics. Note the privileged role played, from the postulates on, by lines and circles. The reason for this choice is clear: these two entities play a special role because they are the mathematical models of what can be drawn with ruler and compass. Euclidean geometry arises explicitly as the scientific theory of the objects that can be drawn with ruler and compass. Euclid's first three postulates are nothing but the straightforward transposition, into the context of mathematical theory, of the usual operations carried out with these basic instruments. Of course, there is a tremendous difference between mathematics and drawing. A compass cannot draw circumferences of arbitrary radius — in fact it cannot draw a true circumference at all. Mathematical science arises from the replacement of the ruler and compass by an ideal ruler and compass, theoretical models of the real instruments, capable of the operations described by the first three postulates; for this theoretical model, both its origin and the correspondence rules that allow its application are perfectly clear.

The difference between the first three postulates, which affirm the constructibility of lines and circumferences, and the last two, whose nature is more theoretical, is reflected in the propositions that make up the treatise, which are of two types: "problems" (προβλήματα) and "theorems" (θεωρήματα). The first type consists in the description of a geometric figure with specified properties, followed by the figure's construction and a proof that the figure constructed does satisfy the desired properties. The first proposition of the Elements, for instance, solves the problem of constructing an equilateral triangle. The theorems, by contrast, consist in the statement that certain properties imply others, and can be followed by the demonstration alone. One famous theorem, for example, states that the square built on the hypotenuse of a right triangle is equivalent to the sum of the squares built on the other two sides. In the Elements this theorem is immediately preceded by the problem of building the squares (and proving that the construction works). In fact, Euclid never uses a geometric figure unless he has given its construction and demonstrated the validity thereof.

30 The distinction between problems and theorems is discussed at length by Proclus (in primum Euclidis elementorum librum commentarii, 77–81, ed. Friedlein), and appears twice in Pappus (Collectio, III, 30:3–24; VII, 650:16–20). Euclid does not differentiate between the two types of propositions in these terms, but the distinction is clear from the formula that closes the demonstration, which is either "as was to be shown" (ὅπερ ἔδει δεῖξαι) or "as was to be done" (ὅπερ ἔδει ποιήσαι).
31 Euclid, Elements, I, proposition 47.
32 Euclid, Elements, I, proposition 46.
2.3 Geometry and Computational Aids

Mathematics has always been used to obtain quantitative results, and its theoretical structure has always been influenced, if often unconsciously, by the way in which these results are obtained. Today we have digital computers. What were the computational aids of classical and Hellenistic Antiquity? For calculations with integers abaci of various types were used, but we know very little about the classical versions. Our ignorance of, and the usual attitude toward, the subject is well illustrated by this quotation from an authoritative history of technology (italics mine):

The form of the Greek abacus is obscure, but the more developed Roman type is well known...\(^33\)

The other computational aid, used above all for noninteger quantities, was geometry. Every problem about continuous magnitudes was cast into geometric language, the data being represented by lengths of segments. Knowing how to solve the problem meant knowing how to construct geometrically a segment whose length represented the solution; this length was then measured. The instruments used in geometric constructions were primarily the ruler and the compass, which thus became not just drawing instruments but analog computational tools.\(^34\) The use of analog computational tools may seem strange to us, accustomed as we are to digital computers, but remember that until a few decades ago engineers did a large part of their calculations with slide rules, whose precision is less than that attainable with the ruler and compass of Hellenistic mathematics. Two features of ruler and compass solutions made them particularly useful. First, their relative error was very small (of the order of the ratio between thickness and length of the lines drawn): no technical application could want better. Second, the construction was easily reproducible in solving an equal problem with different numerical data. Today we consider independent three activities that were indissolubly linked in the practice of Hellenistic mathematics: deductive reasoning, calculation, and drawing.

\(^33\)[HT], vol. III, p. 501.

\(^34\)The problems that can be solved in this way are those that we would express in terms of algebraic equations of the first or second degree. For example, the determination of the fourth proportional of three given segments (Euclid, *Elements*, V, proposition 12) is equivalent to the calculation of a ratio, once one segment is chosen as the unit of measurement. The determination of the proportional mean of two given segments (Euclid, *Elements*, V, proposition 13) amounts to taking a square root. Obviously the algebraic formulation is not necessary for applications: every problem solvable by taking a square root can be solved equivalently by finding a proportional mean.
Forgetting that ruler and compass were the main computational aids of Hellenistic mathematics can lead one badly astray. Boyer writes in his *History of Mathematics*:

The Greek definition and study of curves compare quite unfavorably with the flexibility and the extent of the modern treatment. Indeed, the Ancients overlooked almost entirely the part that curves of various sorts played in the world about them. Aesthetically one of the most gifted people of all times, the only curves that they found in the heavens and on the earth were combinations of circles and straight lines.\(^{35}\)

Finding combinations of circles and lines in heaven and on earth meant successfully reducing the solution of problems both earthly and astronomical to calculations that could be performed with elementary instruments such as ruler and compass.\(^{36}\) Boyer might as well have charged contemporary scientists with infinite intellectual poverty because, in using digital computers, they are unable to imagine anything other than combinations of zeros and ones.

Moreover the Greeks, from the Hellenistic period on, did study curves that cannot be drawn with ruler and compass. They knew, for instance, that the quadratrix of Hippias (the curve described by the intersection of a segment in uniform translational motion with one in uniform rotational motion) allows one to square circles and trisect angles. However, they considered this a pseudo-solution, or "sophistic solution", to these problems. Why? Clearly because it transferred the difficulty from the original problem to that of building a machine that could carry out in practice the two required synchronized motions, tracing the intersection point. The task was certainly feasible, but not with the same precision with which segments and circles could be traced with ruler and compass, and above all not in such an easily reproducible way and with a precision so easy to check.

The preference on the part of Hellenistic mathematicians for ruler and compass solutions has often been considered an intellectual prejudice. That misses the point; geometers who proposed "sophistic" solutions such as the one just described were much in the same position as someone today who might propose to solve a physics problem not by finding a theoretical method translatable into an algorithm that can be implemented on digital computers, but by using an analog "computer" that measures


\(^{36}\)How far one can get with "combinations of circles" was clear in Hellenistic times, and is even clearer today to anyone who has studied Fourier series expansions. This point will be taken up again in Section 3.8.
the desired physical magnitude by reproducing the phenomenon under study. Such a procedure can certainly be useful, but no one would say that it provides a true solution for the problem.

Archimedes introduced in his *Arenarius* a numbering system whose expressive power equals not only that of our positional system, but even that of today's exponential notation. Despite the creation of an analogous system by Apollonius and the introduction of zero, there was no consequent spread of the positional notation system; its use remains largely limited to the base sixty system used in astronomy and trigonometry. It can be conjectured that the efficiency of geometric algorithms contributed to the slow rate of diffusion of "algebraic" computation methods. This impression is supported by the fact that both Archimedes and Apollonius developed their (essentially equivalent) methods in close connection with the problem of representing very large numbers, and so invented exponential representation before positional representation pure and simple. Clearly the geometric method was so efficient that the need for improving it came up chiefly when very large ratios were involved, a case which geometry does not serve well (how can one represent with segments two numbers that differ by several orders of magnitude?).

The efficiency of algorithms based on ruler and compass was closely tied to the possibility of making accurate drawings on papyrus. The link between theoretical structures and material instruments is illustrated by the very different route taken by mathematics in Mesopotamia. There, as we have said, clay tablets were used all the way down to the Hellenistic period, and they do not allow accurate drawings. This meant that Mesopotamian mathematics had to be based on numerical, rather than geometric, methods. The scarcity of sources, already mentioned in Section 1.1, prevents us from following Hellenistic-era Mesopotamian mathematical

---

37 The first versions of positional notation arose in Old Babylonia; they were the result of millennia of unplanned evolution (as discussed further in Section 7.2) and had not, until Hellenistic times, led to a completely coherent and ambiguity-free system. Archimedes' creation, by contrast, was consciously designed, with full knowledge of the conventional nature of such systems.

38 A précis of Apollonius' system can be found in Pappus, *Collectio*, II, 6–28.

39 Zero was being systematically used in Mesopotamia with sexagesimal notation around 300 B.C. Its possible earlier history in Babylonian mathematics is unclear; see [Neugebauer: ESA], p. 29. Its present symbol, transmitted by Indians and Arabs, appears in Ptolemy's trigonometric tables and in papyri from the Ptolemaic era (where it is modified by a line above or other decorations); see [Neugebauer: ESA], p. 13–14. A late mention of the role played by the number zero in arithmetic can be found in Lamblichus (*In Nicomachi arithmeticae introductionem*, 17–19). Lamblichus' word for zero is οὐδέν, from whose first letter may have derived the symbol we use.

40 Because of its usefulness in astronomical calculations, the positional system was imported into India together with astronomy (see Section 2.8).

41 Another case where geometric algorithms fell short was that of trigonometric tables; see p. 53.

42 Among the clay tablets dating from the Seleucid era many have been found to have mathematical content, but again the great majority of these have never been published or translated.
developments as well as we’d like, but the few cuneiform texts that have been translated make it clear that a transition from prescientific mathematics to mathematical science took place during the Hellenistic period in the Seleucid kingdom, just as it did in the Mediterranean world whose main center was Alexandria.\textsuperscript{43}

Although in the rest of this chapter we will continue to deal mostly with mathematics characterized by the use of geometric methods, we should keep in mind that this is not all there was to mathematical science in the Hellenistic period. The different strand of mathematics pursued in Mesopotamia led not only to the introduction of zero but also to certain algebraic methods which, after being taken up to some extent in Alexandria during the imperial age (first by Heron and especially by Diophantus), reappeared in different guise after centuries of dormancy, as part of new developments in India and China.

2.4 Discrete Mathematics and the Notion of Infinity

Two classes of objects of study in the \textit{Elements} are the integer numbers (\textit{\varphi\rho\theta\mu\omicron\omicron\omicron}) and the magnitudes, or continuous quantities. As an example of a theorem about integers, we recall the famous proof that there are infinitely many primes. The statement of Euclid’s theorem (IX, proposition 20) is:

There are more prime numbers than any preset multitude of prime numbers.

The proof is the following. Given any “multitude” (finite set) of prime numbers, let \( k \) be the number obtained by multiplying them all together\textsuperscript{44} and adding 1 to the result. Clearly, \( k \) cannot be a multiple of any of the given prime numbers (which are assumed to be different from 1). Thus, if \( m \) is a prime factor of \( k \) distinct from 1, \( m \) cannot be one of the given prime numbers. Thus we have found a prime number not included among those originally preset.

It is often said that in Antiquity the concept of infinity was not used in mathematics. For example, Morris Kline writes:

In Greek science the concept of the infinite is scarcely understood and frankly avoided.\ldots\ The concept of a limitless process fright-\textsuperscript{43}[Neugebauer: ESA, p. 48.\textsuperscript{44}Actually Euclid considers not the product, but the least common multiple, of the numbers. Since the numbers are prime, the result is the same. Euclid’s choice makes possible the geometric interpretation provided by the illustrators of our manuscripts, who, representing each prime number by a segment, represent in the same way also their least common multiple.
2.5 Continuous Mathematics

The use of "magnitudes", or continuous quantities, in addition to integers, gave rise to a difficult problem. Consider segments. To operate with these "magnitudes" one must know how to carry out the basic arithmetic operations. Addition poses no problem: if \(a\) and \(b\) are two segments, the sum \(a + b\) is the segment obtained in a natural way by extending the first segment a length equal to that of the second. Differences are defined analogously. These rules correspond to what one effectively does in order to add or subtract noninteger quantities using the geometric method. For multiplication things were also simple: the product \(ab\) was thought of as a rectangle whose sides were represented by \(a\) and \(b\). But what meaning could be assigned to the ratio \(a:b\)? Of course, the operation of addition

---

45 [Kline], p. 57.
46 See, for instance, Conica, II, proposition 44, where Apollonius, after showing how to construct a diameter of a conic, concludes: "In this way we will find infinitely many diameters" (\(\varphi\sigma\iota\alpha\iota\varsigma\ ο\acute{\iota}\mu\epsilon\tau\rho\varsigma\)). This use of \(\varphi\sigma\iota\alpha\iota\varsigma\) in the sense of actual infinity in a mathematical context appears already in Plato’s Theaetetus, 147d. Theaetetus reports a conversation between the mathematician Theodorus and his students (of whom he was one), dealing with squares that are multiples of the unit square but whose sides are not multiples of the unit length (and therefore are incommensurable with it). They remark that such sides are infinite in number (\(\varphi\sigma\iota\alpha\iota\varsigma\ ο\acute{\iota}\mu\epsilon\tau\rho\varsigma\) to \(\pi\lambda\rho\beta\varsigma\iota\)).
47 This way of regarding products returns magnitudes nonhomogeneous with the factors, so it makes expressions such as \(a + ab\), where \(a\) and \(b\) are lengths, nonsensical. This introduces a kind of automatic "dimension control".
between magnitudes induces in an obvious way the operation of multiplication of a magnitude by a natural number: if \( k \) is a natural number, the product \( ka \) can be defined as the sum of \( k \) magnitudes, each equal to \( a \). If two integers, \( k \) and \( h \), can be found such that \( ka = hb \), the ratio \( a:b \) can be defined as the ratio between integers \( h:k \). In other words, the ratio \( a:b \) can be defined as a fraction. When there are no two integers \( h \) and \( k \) satisfying \( ka = hb \), the magnitudes \( a \) and \( b \) are called incommensurable. In this case (which happens, for example, when \( a \) and \( b \) represent the side and diagonal of a square, as we saw in Section 2.1), it is not clear what can be understood by a "ratio" between the two magnitudes. Yet the theory of similarities — which we would not want to restrict to the case of commensurable magnitudes — requires that we assign a meaning to proportions such as \( a:b = c:d \) even when the magnitudes are incommensurable.

This problem is solved by Euclid with his definition of proportion (Elements, Book V, definition 5, Heath translation):

Magnitudes are said to be in the same ratio, the first to the second and the third to the fourth, when, if any equimultiples whatever be taken of the first and third, and any equimultiples whatever of the second and fourth, the former equimultiples alike exceed, are alike equal to, or alike fall short of, the latter equimultiples respectively taken in corresponding order.

Expressing this in an algebraic language more familiar to us, we say that the proportion \( a:b = c:d \) holds when, for any chosen natural numbers \( h \) and \( k \), one of the following statements is true:

(A) \( ka > hb \) and simultaneously \( kc > hd \);
(B) \( ka = hb \) and simultaneously \( kc = hd \);
(C) \( ka < hb \) and simultaneously \( kc < hd \):

In this way we manage to define equality between ratios using only multiplication by natural numbers, even in the case of incommensurable magnitudes.

For a long time, this definition of a ratio was criticized by modern mathematicians, who, for reasons to be explained later,\(^47\) did not realize the need for such complexity. The idea was finally grasped by Weierstrass and Dedekind, who founded the modern theory of real numbers essentially by translating Euclid's definition into the language used nowadays. The translation,\(^48\) into the terms used by Dedekind, is essentially this: If we define a real number as any possible "Euclidean ratio",\(^49\) the Euclidean

---

\(^47\) See Section 11.4, especially page 350.

\(^48\) Obviously, this translation into modern language led to significant modifications. In particular, the modern notion of the algebraic structure of the set of real numbers is new and comes from the
definition boils down to saying that a real number is uniquely determined by its behavior regarding every pair \((h, k)\) of integers; that is, it is identified by the set of pairs for which cases A, B, or C obtain. The sets corresponding to cases A and C are called by Dedekind “contiguous classes of rational numbers”, and are clearly sufficient to identify the ratio \(a:b\), that is, the real number. The first works on the “modern” theory of real numbers date from 1872.\(^{50}\)

On the subject of Euclid’s proportions, Heath wrote:

The greatness of the new theory itself needs no further argument when it is remembered that the definition of equal ratios in Eucl. V, Def. 5 corresponds exactly to the modern theory of irrationals due to Dedekind, and that it is word for word the same as Weierstrass’s definition of equal [real] numbers.\(^{51}\)

As other authors whom we will encounter have done in similar cases, Heath—one of the foremost modern historians of ancient science—regards it as the greatest glory of Greek mathematicians that they managed to anticipate modern theories. Here he seems almost to suggest that the “word for word” correspondence he notes may have to do with Euclid’s ability to anticipate results that would come two thousand years later, rather than with Weierstrass’s “word for word” use of Euclid’s definition, which he knew well. (Let’s not forget that the Elements were the textbook at the foundation of Weierstrass’s and Dedekind’s early mathematical education.)

It may seem that Euclid’s definition of proportions, like the modern definition of real numbers that derives from it, is impossible to apply, since it requires the consideration of infinitely many integers. The question can be clarified by examining an application of Euclid’s definition. Consider his proof that the ratio of two circles equals the ratio of the squares on their diameters.\(^{52}\) The existence—in the sense of constructibility with ruler and

modern primacy of algebraic over geometric aspects. But these changes, though important, do not contradict the fundamental fact that the “modern” theory appeared with the recovery of Euclid’s definition.

\(^{49}\)Physicists and engineers know very well that even today a real number is just a ratio of homogeneous magnitudes (and so they know that in their formulas the arguments of functions such as sines and exponentials must be ratios of homogeneous magnitudes). This awareness seems sometimes to elude mathematics students.

\(^{50}\)They are an article by Dedekind and one by Heine, both based on the ideas of their teacher, Weierstrass.


\(^{52}\)Euclid, Elements, XII, proposition 2. “The ratio of two circles” may sound odd to modern ears, but the notion of area arose precisely as the ratio between a given figure and another chosen as a unit of measurement. Euclid always talks about ratios between plane figures (or segments, or solids) rather than about areas.
compass — of the mathematical objects involved has already been established: that of circles is the content of the third postulate, and that of squares was proved earlier.\textsuperscript{53} The definition of proportions is used to state a relationship between geometric objects already constructed, one whose validity can be demonstrated in a finite number of logical steps, as can be checked by reading Euclid’s proof. Thus there is an important difference between Euclid’s ratio between magnitudes and today’s real numbers: whereas modern mathematicians have introduced axioms about the set of all real numbers and have often considered real numbers whose existence is proved thanks to these axioms and is not supported by constructive procedures, Euclid considers only ratios of constructible magnitudes.

2.6 Euclid and His Predecessors

According to a widespread opinion, the contents of the \textit{Elements} had appeared before Euclid in similar treatises, since lost. This belief rests largely on the fact that most of the statements of theorems contained in the \textit{Elements} had indeed been known before Euclid, and that many proofs had been accomplished, including complex ones.\textsuperscript{54} Much effort has been expended, often fruitfully, on the task of tracing the origins of the material contained in the various books of the \textit{Elements}. But from our point of view the main feature of Euclid’s work is not the set of results presented, but the way in which these results connect together, forming infinitely extensible “networks” of theorems, drawn out from a small number of distinguished statements. To judge the originality of the \textit{Elements}, therefore, one must ask whether a similar structure (without which one cannot extend the theory by doing “exercises”: that is the whole point!) had been achieved prior to Euclid.

In the surviving fragments on pre-Euclidean mathematics there is no evidence for sets of postulates similar to Euclid’s. The works of Plato and Aristotle, moreover, offer an explicit description of what the “principles” accepted by mathematicians as the initial assumptions of their science were like at the time. Plato writes that “those who work with geometry, arithmetic, and the like lay as ‘hypotheses’ evenness and oddness, figures, the three kinds of angles and similar things.”\textsuperscript{55} Aristotle, in a passage where he discusses the role of principles in the deductive sciences, makes

\textsuperscript{53}Euclid, \textit{Elements}, I, proposition 46.

\textsuperscript{54}For example, Archytas gave a construction for two proportional means (which amounts to the extraction of a cube root), and Eudoxus proved the formulas for the volume of the pyramid and cone. Both proofs date from the first half of the fourth century.

\textsuperscript{55}Plato, \textit{Republic}, VI, 510c.
a distinction between the principles common to all sciences and those particular to each. As an example of the first type he mentions the assertion “Subtract equals from equals and equals remain”, which appears in the *Elements* exactly as one of the “common notions”. Immediately before that he had written: “Particular [principles] are ‘The line is such-and-such’, and likewise for straightness.”

There is an obvious difference between the type of “geometric principles” exemplified by Plato and Aristotle, which surely could not serve as the basis for proving theorems, and the postulates contained in the *Elements*.

As to the premises actually used in the demonstration of geometric theorems, several passages from Plato and Aristotle attest to a deductive method much more fluid in the choice of initial assumptions than that transmitted by the *Elements* and later works.

The logical unity of the *Elements*, or of a large portion of it, is clearly not due to chance; it is the result of conscious work on the part of the same mathematician to whom we owe the postulates. There is no reason to suppose that this unity is not an innovation due to Euclid, and a very important one at that.

### 2.7 An Application of the “Method of Exhaustion”

As an example of an application of what in modern times was named the “method of exhaustion”—that is, of Hellenistic mathematical analysis—we recall how Archimedes computed the area of a segment of parabola, in his *Quadrature of the parabola*. This is probably the simplest of the surviving proofs of Archimedes (and therefore the most popular), but it will suffice for giving an idea, if not of Archimedes’ ability to solve difficult problems, at least of his level of rigor. Readers who dislike detailed mathematical arguments may proceed to the first full paragraph of page 52.

Let a parabola be given. If $A$ and $B$ are points on it, the part of the plane comprised between the segment $AB$ and the arc of the parabola joining $A$ and $B$ is called the segment of parabola with base $AB$. The point $C$ of the arc of parabola that lies farthest from the line $AB$ is called the vertex of the segment of parabola.

Archimedes’ proof is based on a postulate, fundamental in nature and discussed at the beginning of the work, and on three technical lemmas.

---

57 The vertex of a segment of parabola depends on the base $AB$, and should not be confused with the vertex of the parabola, which is usually a different point.
58 Our exposition differs slightly from the original one, which contained more propositions.
Postulate. If two areas are unequal, there is some multiple of the difference that exceeds any previously fixed area.

Lemma 1. If C is the vertex of the segment of parabola of base AB, the area of the triangle ABC is more than half the area of the segment of parabola.

From the technical point of view, the fundamental ingredient in the proof lies in the following lemma, whose proof, together with that of the preceding lemma, is given in the Appendix.

Lemma 2. If C is the vertex of the segment of parabola of base AB and D is the vertex of the segment of parabola of base CB, the area of triangle CBD is one-eighth that of triangle ABC.

Archimedes' basic idea is to cover the segment of parabola with infinitely many triangles. He starts by inscribing the triangle ABC into the segment of parabola of base AB, thereby dividing the latter into three parts: the triangle ABC and two segments of parabola, in each of which we can inscribe another triangle following the same procedure (like triangle CBD in the figure). The procedure can of course be iterated, leading to ever smaller triangles. Let S be the area of the initial segment of parabola, $A_0$ the area of the triangle $ABC$ inscribed in it, and $A_1, A_2, A_3, \ldots$ the total area of the triangles inscribed at each successive iteration. Since at each iteration the number of triangles doubles, whereas the area of each, by Lemma 2, becomes 8 times smaller, it is clear that $A_1 = \frac{1}{8}A_0$, $A_2 = \frac{1}{8}A_1$, and so on. At this point Archimedes proves another lemma:

Lemma 3. If $A_0, A_1, A_2, \ldots, A_n$ form a finite sequence of magnitudes, each of which is four times the next, we have

\[ A_0 + A_1 + A_2 + \cdots + A_n + \frac{1}{3}A_n = \frac{4}{3}A_0. \]

We will not reproduce the proof of this lemma; we merely note that it is a particular case of a property of geometric progressions that is very well known today.

Archimedes can now compute the area $S$ of the segment of parabola by proving the following theorem:

Theorem. $S = \frac{4}{3}A_0$.

The theorem is proved by contradiction, by showing that $S$ cannot be either more or less than $\frac{4}{3}A_0$. Suppose that $S > \frac{4}{3}A_0$, and call $E$ the difference
2.7 An Application of the "Method of Exhaustion"

Let $S - \frac{4}{3}A_0$, so that

$$S = A_0 + A_1 + A_2 + \cdots + A_n + \varepsilon_n = \frac{4}{3}A_0 + E,$$

where we have indicated with $\varepsilon_n$ the area of the part of the segment of parabola not covered by triangles after $n$ iterations. If $n$ is large enough, the area $\varepsilon_n$, which at each step gets smaller by a factor greater than 2 (by Lemma 1), will end up smaller than $E$ (by the postulate). Therefore

$$A_0 + A_1 + A_2 + \cdots + A_n > \frac{4}{3}A_0.$$ 

But this inequality is false, because it contradicts (*). Thus we have excluded the case $S > \frac{4}{3}A_0$.

Suppose instead that $S < \frac{4}{3}A_0$. Using the postulate again, we see that if $n$ is a sufficiently large integer the area $\frac{1}{3}A_n$ must be less than the difference $\frac{4}{3}A_0 - S$. Using (*) we deduce that

$$\frac{4}{3}A_0 < A_0 + A_1 + A_2 + \cdots + A_n + \frac{4}{3}A_0 - S,$$

that is,

$$S < A_0 + A_1 + A_2 + \cdots + A_n.$$ 

This inequality, too, is clearly false, since the right-hand side represents the area of a portion of the segment of parabola of area $S$. This concludes the proof of the theorem.

We note (and it will be clearer to those who read the Appendix) that the proof depends crucially on the study of triangles, which don’t appear at all in the formulation of the problem. They are used merely as a tool. This example makes it clear why Hellenistic mathematicians laid out with great care such simple theories as that of triangles, presented in the *Elements*: they were useful tools for tackling even problems whose original statement had no connection whatsoever with the auxiliary theory. Triangles were studied so that figures could be triangulated. We will encounter a similar use of circles as a tool in the study of planetary orbits.

Every real number different from zero has a multiple that is greater than an arbitrary fixed real number. In modern axiomatizations of the reals, this property is assumed true and is called the "Archimedean postulate". The postulate that Archimedes actually stated is different: the magnitudes that he (and Hellenistic mathematicians in general) considered in fact form a non-Archimedean set (in the modern terminology), in that the magnitude of a segment, though nonzero, has no multiple that exceeds the magnitude of a square. In the parlance of Hellenistic mathematicians, two magnitudes have a ratio, and are called homogeneous, if each has a multiple that
Archimedes’ surviving writings may give the impression that the level of scientific works transmitted through late Antiquity and the Middle Ages was not as low as we claimed on page 8. In fact, the selection criteria we mentioned are confirmed even in this case, because some of Archimedes’ works have reached us only through exceptionally lucky circumstances. In spite of their author’s fame, some of his writings (such as the Quadrature of the parabola) apparently hung on for several centuries in a single copy, a codex prepared in Byzantium in the ninth century, at the initiative of Leo the Mathematician. This manuscript, now lost, found its way to the Norman court in Sicily in the twelfth century and thence to the hands of Frederick II, Holy Roman Emperor; after the battle of Benevento (1266) it ended up in the Vatican Library. It still existed in the fifteenth century, when copies were made in France and Italy, but there the trail ends. Another manuscript, which contained different works and was probably given to the pope at the same time, is lost track of earlier, in the fourteenth century. From this second manuscript was derived a Latin version of the treatise On floating bodies. Our only other source for the works of Archimedes is the already mentioned palimpsest (page 8) discovered by Heiberg in 1906, subsequently lost, and recently found again.

If we had none of his works, our knowledge about Archimedes would be limited to remarks transmitted by authors such as Plutarch, Athenaeus, Vitruvius and Heron. We would be exactly in the same situation we are with respect to, say, Ctesibius: a scientist who, to judge from the same sources, appears no less interesting than Archimedes. Circumstances such as the preservation for six centuries of a codex owned successively by Byzantines, Normans, German emperors, Angevins, popes and Florentine humanists are hardly replicable. In how many other cases have we been less fortunate?

2.8 Trigonometry and Spherical Geometry

We conclude this chapter on Hellenistic mathematics with a brief mention of plane and spherical trigonometry. We make this choice not because of

---

59 Euclid, Elements, V, definition 4.
60 A more general version of the postulate, applying not only to surfaces but also to lines and solids, appears in Archimedes, De sphaera et cylindro, 11:16–20 (ed. Mugler, vol. I).
61 See [Dijksterhuis: Archimedes], Chapter 2. For a full discussion of the transmission of Archimedes in the Middle Ages, see [Clagett: AMA].