The Sources and Limits of Geometric Rigor from Euclid Through Descartes

Ryan C. Thompson

May 7, 2008

Of all the systems of thinking that have made their way through the ages, Euclidean geometry remains one of the most appealing and intuitive. It is also one of the most successful, having been practiced continuously from Greek antiquity through modern-day schools. We will examine the aspects of geometry that account for this intuitiveness, as well as several innovations that allowed it to tackle new and more difficult problems. We begin with an analysis of Euclid's *Elements*, and then we will consider the contributions of two ancient authors, Archimedes and Apollonius. Lastly, we will see how two authors during the scientific revolution, Galileo Galilei and René Descartes, pushed geometry into new areas, namely the realistic and the algebraic.

The focus of each proposition in Euclid's *Elements* is what we call a "construction," and we must consider the significance of this term – that is, what separates a construction from a mere drawing. Herein lies one source of the Greeks' confidence in the truth of their propositions: their proofs are constructed, rather than simply drawn, by a method based on two kinds of objects, the circle and the straight line. These were the two simplest shapes known to the Greeks, and also two of the most fundamental to Greek philosophy. In Plato's *Timaeus*, he describes the shape of the world:[1]

> "Wherefore [the Creator] made the world in the form of a globe, round as from a lathe, having its extremes in every direction equidistant from the centre, the most perfect and the most like itself of all figures ... the movement suited to his spherical form was assigned to him, being of all the seven that which is most appropriate to mind and intelligence; and he was made to move in the same manner and on the same spot, within his own limits revolving in a circle. All the other six motions [rectilinear motions up, down, left, right, forward, and backward] were taken away from him, and he was made not to partake of their deviations."

In Aristotle's *De Caelo*, too, we see a similar conception of motion: "But all movement ... is either straight or circular or a combination of these two, which

[1]Plato, *Timaeus.* (The Internet Classics Archive);

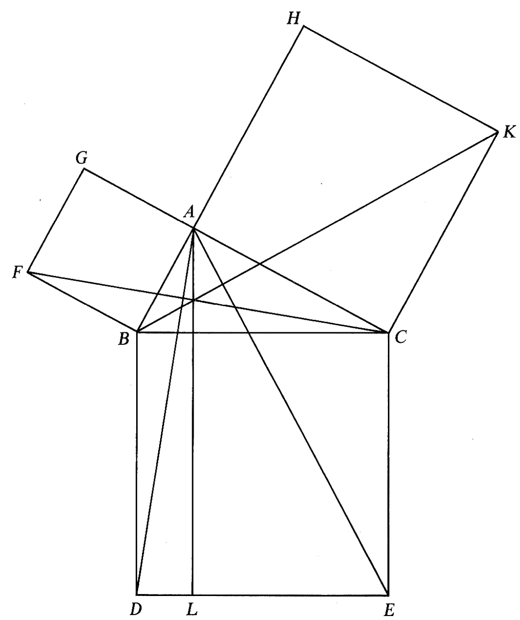

Figure 1: *Elements* I.47 – Fauvel and Gray, p. 115.

are the only simple movements."[2] Greek geometers, then, clearly had a solid philosophical basis for choosing the figures generated by these two fundamental motions as the starting points for their geometry. Furthermore, using "natural" motions as mathematical starting points must certainly have appealed to the Pythagorean and Platonic ideal of understanding nature through mathematics.

In every extant copy of Euclid's *Elements*, a diagram accompanies each proposition. We may rightly ask whether these diagrams (or similar ones) were included in the original, or whether later editors added them in an effort to make the constructions easier to visualize. Several arguments attest to the conclusion that the original Greek texts must have included diagrams, but no piece of evidence shows this more clearly than the fact that many of the propositions fail to specify essential lettered points in the text yet use those points later in the proof. As an example, consider Prop. I.47, "Pythagoras's Theorem," in which three squares are constructed on the sides of a right triangle ABC: "For let there be described on BC the square BDEC, and on BA, AC the squares GB, HC."[3] The first square, BDEC, is fully specified, with all four vertices named in an order that uniquely determines their positions. However, the other two

[2]Aristotle, *The Works of Aristotle Vol. 1* (Chicago: Encyclopedia Britannica, Inc., 1952) p. 359.;

[3]John Fauvel and Jeremy Gray, *The History of Mathematics: A Reader* (New York: Palgrave Macmil- lan, 1987), p. 115.[6]

squares are nam ed in the usual Greek fashion, by two diagonally opposite vertices, GB and HC. Point A is implicitly specified as third vertex of both of these two squares since it is an endpoint of the line segment on which each square is constructed. The fourth vertices of each these two squares, points F and K respectively, are never specified in the text of this construction, though both points appear in the diagram (see Figure 1). However, Euclid uses these points later in the proof as though they *were* specified: "And, since the angle DBC is equal to the angle FBA: for each is right: let the angle ABC be added to each; therefore the whole angle DBA is equal to the whole angle FBC."[4] He also uses the unspecified point K in the same fashion. From this and other more subtle indications, we see that Greek texts not only included diagrams, they exhibited a strong interdependence between the text and the diagram.[5]

In modern mathematics, "proof by picture" is considered a logical fallacy and is especially common among new students of geometry. Furthermore, ancient Greek philosophers and mathematicians understood this fallacy – Proclus criticizes those who would accept without proof Euclid's 5th postulate for exactly this reason:[6]

> To them Geminus has given the proper answer when he said that we have learned from the very founders of this science not to pay attention to plausible imaginings in determining what propositions are to be accepted in geometry ... And Simmias is made by Plato to say, 'I am aware that those who make proofs out of probabilities are impostors.' ... the conclusion that because [the lines] converge more as they are extended farther they will meet at some time is plausible, but not necessary, in the absence of an argument proving that this is true of straight lines.

However, as we have seen above, Greek proofs depend intimately on their diagrams. We may fairly ask then, what difference Greek geometers saw between their diagrams and "plausible imaginings." That is, how could each of Euclid's proofs rely on an accompanying diagram without relying on implicit and unjustified assumptions contained in that diagram?

By our modern standards, of course, Euclid's proofs fail this criterion; certain implicit axioms are indeed supported purely by the diagram. For instance, the construction in Prop. I.1 does not prove that the two circles intersect. The circles in the diagram intersect a point C (Figure 2), yet the text gives no justification for this intersection at all. Rather, the text simply uses the point of intersection with no prior introduction, in the same way that the text of Prop. I.47 uses the points F and K.[7] This suggests that, to Euclid, the existence of the intersection point was no less obvious than the existence of the implicitly

[4] *Ibid,* p. 116

[5] Reviel Netz, *The Shaping of Deduction in Greek Mathematics* (New York: Cambridge University Press, 1999), pp. 19-26.

[6] Proclus, A Commentary on the First Book of Euclid's Elements (New Jersey: Princeton University Press, 1970), pp. 150-151.

[7] Euclid, Elements (Chicago: Encyclopedia Britannica, Inc., 1952) p. 2.

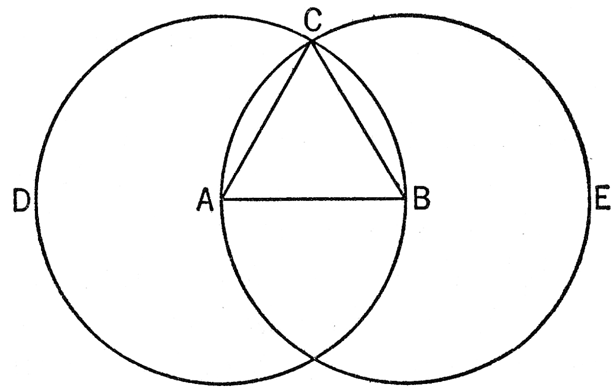

Figure 2: *Elements* I.1 – *Great Books of the Western World*, Vol. 11., p. 2.

specified points F and K.

Since modern geometry is so different from the geometry of antiquity, let us seek an analogy in the learning of arithmetic. Children learn arithmetic initially through concrete examples – start with one apple, add two more, and the result is three apples. Absent is any talk of the commutativity of addition, the reflexivity of equality, or other concepts that we normally associate with arithmetic. This does not mean that children learn a noncommutative form of addition, or a nonreflexive version of equality. On the contrary, addition simply is commutative; equality simply is reflexive. This knowledge is so basic that an explanation would be meaningless, until we introduce a noncommutative example, such as matrix multiplication.

In the same way, any person with only a few minutes' experience using a compass and straightedge may convince herself that the circles in Prop. I.1 must intersect, even if she cannot explain how she is so certain. The conclusion is even more clear in this particular case, because any construction of Prop.I.1 must look exactly like Figure 2, the only difference being size, since all equilateral triangles are similar. The knowledge of this intersection is at least as basic to an understanding of a circle as commutativity is to addition. Indeed, until modern students begin to learn abstract algebra, teachers rarely demand an explicit justification for commutative transformations of expressions, and mathematics textbooks rarely call attention to such details. Similarly, if Euclid's Elements is a geometrical reference book, then we might reasonably expect the same degree of trivial omission, even if Euclid was aware of the issue.

On a similar note, we may observe that the numbered list of definitions, postulates, and axioms at the beginning are not as they seem. We have seen above that the geometer's fundamental understanding of circles and straight lines governs certain implicit conclusions about the diagrams. Consistent with this, Reviel Netz argues that the definitions in Greek mathematical texts have little bearing on the actual mathematics, and that working definitions of important terms are instead established by their usage throughout the text.[8] The list of "definitions," "postulates," and "common notions" is really nothing more than Euclid's prose introduction to his work, which later editors have transformed into an ordered list. As such, the definitions are not rigorous, but upon closer inspection read rather like general guidelines. Many definitions that we would consider vital, such as "section of a circle" and "tangent," are clearly ambiguous. Indeed, in the original Greek, Euclid defines "tangent" circularly in terms of another form of the same word, and then he goes on to use both forms of in the propositions as though the definition did notexist.[9] Later authors avoided this confusion by substituting another unrelated word, for which no author bothered to give any definition at all. To a modern mathematician who requires a precise definition for every term, such behavior is unforgivable. However, we have seen that Euclid relied on properties inherent in the various figures, so what we call "definitions" are most likely simply an attempt to assign names to previously understood concepts. If this is the case, vague, imprecise definitions would suffice, and the substitution of words for "tangent" is simply the renaming of an already well-defined property of lines.

We have repeatedly seen that the roots of Euclidean geometrical rigor lie in the intuitive appeal to simplicity. Both the simplest motions described by philosophers and the clearest intuitions about geometrical figures form the basis for the Greek reasoning about geometry. We now examine how later authors built on this base.

Archimedes, who lived approximately a century after Euclid, repeatedly pushed the limits of Greek construction by proving propositions about the areas or volumes of figures bounded by curved lines and surfaces. Very few such results were known before Archimedes, and for good reason: with only a compass and straightedge, one can only draw circular arcs and straight line segments. Using these two tools (or the shapes that they produce) to construct any other type of curve is difficult at best, and generally impossible. Archimedes proved his propositions through the use of potential infinity in the method of "indirect passage to the limit."[10] Unlike the simpler Euclidean proofs, the proofs of Archimedes did not furnish a construction of the objects in question, nor did they hint at how he first arrived at his results. This resulted in considerable speculation (and envy) through the ages until 1906, when a codex surfaced containing Archimedes's mechanical method for investigating geometrical questions. The mechanical method treats geometric figures as physical objects with masses and centers of gravity, and it also goes further than the method of indirect passage

[8]The argument here partially follows that of Reviel Netz, *The Shaping of Deduction in Greek Mathematics* (New York: Cambridge University Press, 1999), pp. 94-102.

[9]This particular example is consistent with the hypothesis that ancient Greek writers worked as many authors do today and wrote the main text before the introduction. In this case, we could not expect the definitions to have any real bearing on the text of the propositions themselves. Of course, we have no easy way to determine in what order Euclid wrote the original text, so this is idle speculation.

[10]I follow Dijksterhuis in avoiding the name "method of exhaustion" for "a mode of reasoning which has arisen from the conception of the inexhaustibility of the infinite." E. J. Dijksterhuis, *Archimedes* (Princeton, NJ: Princeton University Press, 1987), p. 130.

to the limit in that it makes direct use of actual infinity. Archimedes himself could not seem to decide whether he considered this method rigorous, but he clearly realized that his peers would not accept them, or else he would not have bothered to reformulate each proof using the more accepted indirect passage to the limit. In order to understand the difference between these two kinds of infinity (potential and actual) and why one but not the other was accepted in Greek geometry, we must look to prevailing attitudes on the infinite in ancient Greece – namely, the writings of Aristotle on infinity.

Aristotle's teachings held sway from his time (ca. 300 BCE) until well into the scientific revolution in Europe,[11] and with them came Aristotle's writings on the infinite. In his Physics, Aristotle concludes that the infinite must exist in some sense, but that in another sense it cannot exist.[12] He resolves this dispute by considering that "being" can mean having the potential to exist, or it can mean actually existing. Furthermore, one can encounter infinity either by repeated addition or by repeated division. He concludes that, regarding magnitudes, in the case of additive infinity, potential existence must imply actual existence, since a magnitude that exceeds every determinate magnitude is by definition infinite. Infinitely large magnitudes may not exist in the ancient Greek conception of the world, which has a finite size and nothing outside of it. On the other hand, Aristotle sees "no difficulty in refuting the theory of indivisible lines,"[13] since Greek proportion theory requires that any magnitude is divisible into smaller magnitudes – that is, any line may be divided at any point along its length, but such division can never yield an infinitely small line. In summary, the infinite "by way of division" exists potentially in Aristotle's philosophy, while the infinite "by way of addition" cannot exist potentially, and no form of infinity can exist actually.

Despite the theoretical acceptabiliyt of potential infinty, in practice, Greek geometers tended to avoid arguments involving *any* form of infinity, partly because infinity is manifestly impossible to construct, and partly because if they did use it, they would have to resolve paradoxes like the famous "Zeno's paradox" associated with infinite division. However, they had one tool, the aforementioned "indirect passage to the limit," that they had applied to relatively simple cases, such as the ratio of the volumes of a cylinder and a cone of equal base and height. Each example of this method in Euclid's *Elements* invariably included a construction that resembles Figure 3: a regular 4n-gon inscribed inside a circle (and in the text, a similar one circumscribed outside the circle).[14] The method worked by a double *reductio ad absurdum*; in order to prove that the required equality holds, one assumes that either one of the two equated magnitudes is greater than the other. In either case, the assumption results in the conclusion that an inscribed polygon with a sufficiently large number of sides can be constructed with area larger than the circle, or that a circumscribed

[11] David C. Lindberg, *The Beginnings of Western Science* (Chicago: The University of Chicago Press, 1992), pp. 67-68.

[12] Aristotle, *The Works of Aristotle Vol. 1*, pp. 284-285.

[13] *Ibid.*, p. 284.

[14] Euclid, *Elements* (Chicago: Encyclopedia Britannica, Inc., 1952) pp. 351-354. 9

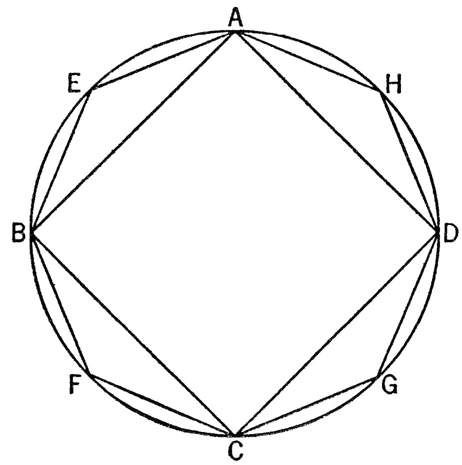

Figure 3: Elements XII.10 – *Great Books of the Western World*, Vol. 11., p. 351.

polygon can be constructed with smaller area than the circle, in both cases a clear contradiction. Finally, since neither of the two quantities considered is greater than the other, they must be equal.

This argument differs in one important respect from most of the other propositions in Greek geometry: because the construction invokes the potentially infinite divisibility of an area (an acceptable practice, according to Aristotle), the required polygon is not actually constructed. The argument merely proves that such a polygon (and therefore a contradiction) *could* be constructed if the two quantities differ by any finite magnitude. Presumably, the Greeks felt uneasy about this logic, because none that we know of before Archimedes applied it to any construction other than the exact one above – *every* indirect passage to the limit in the *Elements* involves a regular 4n-gon inscribed in a circle; nothing else. This is precisely where Archimedes surpassed his predecessors: he generalized a method that had previously been applied only to a specific diagram and used it to prove propositions about numerous other curved figures, both planar and solid.

Archimedes's argument in *Quadrature of the Parabola*, Proposition 24 follows exactly the same structure as the indirect passage propositions in *Elements*, except that the ratio between successive triangles is different than for a circle, with the result that area is commensurable with unity (in modern terms, a rational number).[15] This was certainly a startling result at the time, as the parabola must have seemed, if anything, more complicated than the circle, which had resisted all attempts at quadrature by compass and straightedge. Inscribing triangles in a parabolic segment might not seem to be much of a generaliza-

[15] Archimedes, The Works of Archimedes (Chicago: Encyclopedia Britannica, Inc., 1952), p. 537.

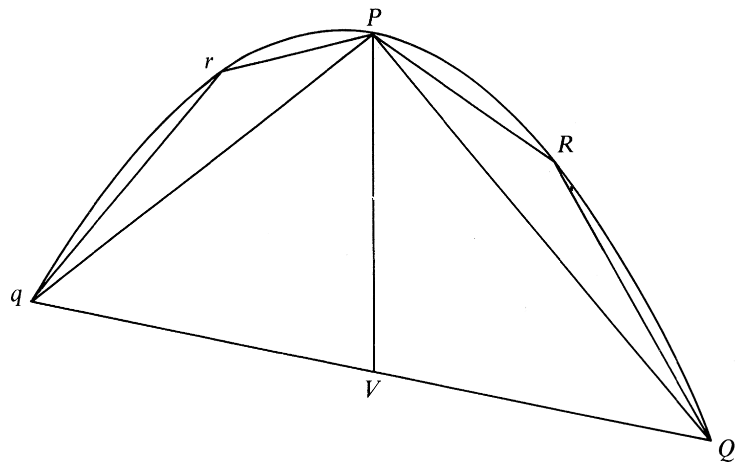

Figure 4: Archimedes's quadrature of the parabola, prop. 24. Fauvel and Gray, p. 154.

tion, since Archimedes's diagram (Figure 4)[16] closely resembles the diagram that would result from cutting away the bottom half of Euclid's.[17] However, Archimedes showed a much more fundamental understanding of the processes involved. After the parabola, he moved on to three-dimensional solids of revolution, such as spheres, cylinders, paraboloids, and spheroids. In each case, we see the same general technique: an assumed finite difference between two magnitudes leads to the potential construction of a contradictory figure (typically, a figure that is both less than and greater than some other figure), allowing Archimedes to prove the desired equality by double *reductio ad absurdum*. The particular figures that Archimedes constructed depended on the situation; he did not limit himself to regular polygons in a circle, as did his predecessors. Yet, since his logic is essentially the same, his results are as rigorous as those in Euclid's *Elements*. Archimedes did not, in these proofs, come any closer to infinity than earlier geometers, but he certainly demonstrated far more confidence in coming as close as he did.

However, the caveat, mentioned above, that indirect passage to the limit does not actually construct the required figure, left a mystery for Archimedes's peers and his successors: how did Archimedes arrive at the results that he then set out to prove? The answer was finally discovered in 1906, with the revelation of the *Method of Treating Mechanical Problems*. In this, Archimedes reveals the method by which he came to know the results whose proofs he had previously published. The text describes a series of decidedly atypical constructions in

[16]This figure and some others have been rotated to save vertical space. Other than rotating the letters to keep them upright, no further modifications have been made.

[17]Indeed, if the Greeks drew all curved lines, circular or not, using their compasses, as Netz and others indicate, then the diagrams might be identical! Netz, *The Shaping of Deduction in Greek Mathematics*, p. 17.

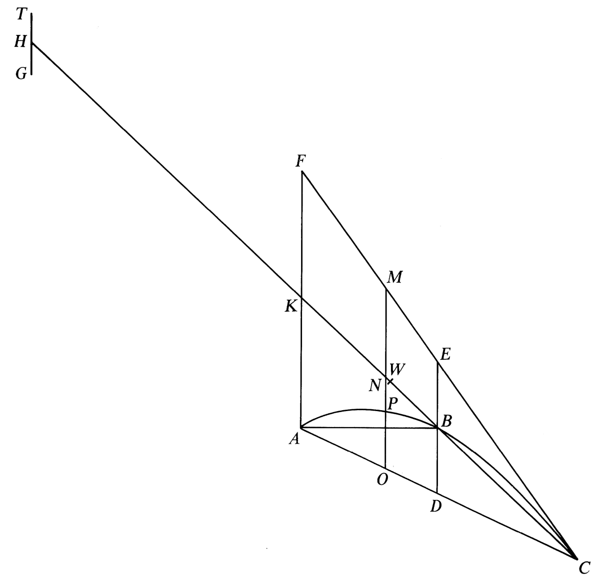

Figure 5: Archimedes's *Method*, Prop. 1. Fauvel and Gray, p. 170.

which an object whose center of gravity is moved to a known position is shown to balance another static object whose center of gravity is already known. The law of the lever (proved in Archimedes's *On Balancing Planes I*) then gives the ratio of the magnitudes (which may be areas or volumes, depending on whether plane or solid objects are under consideration) as the inverse ratio of their distances from the balance's fulcrum. Though any type of movement was rare in Greek constructions, it is not completely absent, especially from Archimedes's own work, for example *On Spiral Lines*. While the movement described in the Method is not of the same nature, one could easily describe a Greek construction using the Prop. I.2 of Euclid's *Elements* to translate all the points in a figure using a given line as a vector. Hence, the motion described in the proof, while unusual, is by no means novel in the context of Greek geometry. The real curiosity is the method by which Archimedes may take a object whose center of gravity is unknown and move that unknown center of gravity to a known position, from which he may then derive the desired ratio.

Let us take the first proposition as an example. In this proposition, Archimedes exhibits his method on the quadrature of the parabola.[18] In the diagram (Figure 5), Archimedes's goal is to place the center of gravity of parabola ABC at the point H, which he has already proven stands in a known proportion with the center of gravity W of the triangle AFC. However, there is no parabola visible around H, because the parabola does not arrive there in one piece; instead, Archimedes takes an arbitrary cross-section MO (i.e. a line segment) of

[18] Archimedes, "The method of treating mechanical problems" in Fauvel & Gray, pp. 169-171.

both the parabola and the triangle, and shows that if the parabola's segment is placed at TG, it will balance the triangle's segment. Since any arbitrary pair of such lines will balance, it follows that any multitude of such pairs of lines will balance. Now comes the critical step:[19]

> And, since the triangle CFA is made up of all the parallel lines like MO, and the [parabolic] segment CBA is made up of all the straight lines like PO within the curve, it follows that the triangle, placed where it is in the figure, is in equilibrium about K with the segment CBA placed with its centre of gravity at H.

In the context of Greek thought, Archimedes has committed several sins in this step: two against Aristotelian philosophy, and one against geometric dimensionality. First, he has taken, implicitly, an *actually* infinite number of line segments, since a plane figure cannot possibly contain a finite number of line segments. Second, he appears to have added these lines and taken the result to be a plane figure. Furthermore, as noted, this number is infinite, so he has claimed to reach the infinite by way of addition, another slight against Aristotle. On a closer inspection, though, the situation is even more ambiguous, because Archimedes did not say "the triangle CFA is the *sum* of all the parallel lines;" he said it is "made up of" all the lines. Clearly, this ambiguity cannot be resolved without a careful study of the original Greek text, so for now we must leave the logic of this claim in its current shaky state: all we know is that Archimedes has claimed that a plane figure is somehow "made up of" an infinite number of lines.

After this "proof," Archimedes quickly admits that "the fact here stated is not actually demonstrated by the argument used."[20] Yet, this seems to contradict his statement from the introduction: "This procedure is, I am persuaded, no less useful even for the proof of the theorems themselves."[21] Again, without a study of the original Greek, we may only speculate as to Archimedes's true opinion of his method and the rigor thereof. Whatever his own thoughts, though, Archimedes clearly understands that the method is fundamentally different and probably incompatible with contemporary geometry and natural philosophy of his time. Still, the care with which he prepared the remainder of the argument according to geometrical standards shows in what high regard he held his method, and reflects the hope of its usefulness that he expressed in the introduction. In summary, Archimedes's approaches to infinity, both potential and actual, did much to expand the scope of Greek mathematical thought.

Approximately a century after Archimedes, another geometer expanded the limits of Greek geometry in a different direction. Apollonius of Perga composed a comprehensive study of the three types of conic sections: the ellipse, the parabola, and the hyperbola. The conic sections were well-known before Apollonius's treatment of them; we have already seen how Archimedes squared the

[19] *Ibid.*, p. 170.
[20] Archimedes, *The Works of Archimedes*, p. 572. 21
[21] *Ibid.*, 569.

parabola and considered conoids and spheroids. However, Apollonius claims, quite truthfully, to have developed the geometry of the conic sections "more fully and universally than in the writings of others."[22] Such reworkings of previously covered material were common in the Greek mathematical tradition.[23] At least in this case, Apollonius backs up his criticisms of earlier authors with a well-developed and intuitive treatment of the conic sections. Indeed, the means by which Apollonius generates his sections and the names that he gives to them are the same that we use today.

Before Apollonius's *Conica* (for example, in Archimedes's treatises), the ellipse, parabola, and hyperbola were referred to, respectively, as sections of an acute-, right-, and obtuse-angled cone. These names reflected the manner of their generation. The angle referred to is the one inside the vertex of the cone, and the section is always taken in a plane perpendicular to one of the lines on the surface on the cone.[24] If one wished to generate one of each conic section, one would require three separate cones, one with each kind of angle at its vertex. Apollonius simplified the generation of the sections in such a way that any cone could generate any kind of conic section. A slice whose axis is parallel to a line in the surface of the cone generates a parabola; a slice whose axis intersects the surface line outside the cone itself generates a hyperbola; a slice whose axis intersects the surface line on the surface of the cone generates an ellipse.[25] The ability to generate all the conic sections from a single cone must certainly have appealed to the Greek philosophical ideal of simplicity.

Furthermore, Apollonius gives much more attention to the cone itself – indeed, he does not even define the three conic sections until propositions 11-13. [26] The cone and its partner, the conic surface, have the first ten propositions to themselves, a sort of "mini-treatise" on cones and their sections, in preparation for the next three propositions, in which the cones may be sectioned in any which way, as described above.[27] To begin, Apollonius gave a more general definition of a cone than Euclid, who defined a cone as the surface of revolution generated by a right triangle about one of its shorter legs. Euclid's definition necessarily generates what Apollonius calls a right cone, in which the axis and the base are perpendicular. Apollonius himself also allows oblique cones, whose bases and axes are not perpendicular. He defines the cone as the surface generated by a straight line segment with one end (the vertex) fixed, and the other end revolving around the circumference of the circular base. From this, he also derives the conic surface, which is generated by the same line segment, when it is extended indefinitely. He thereby creates an unbounded surface "which

[22] Apollonius, *Conica* (Chicago: Encyclopedia Britannica, Inc., 1952), p. 603.

[23] This is something of an understatement. See Reviel Netz, *The Transformation of Mathematics in the Early Mediterranean World* (New York: Cambridge University Press, 2004), pp. 60-63.

[24] Michael N. Fried and Sabetai Unguru, *Apollonius of Perga's Conica* (Leiden, Netherlands: Koninklijke Brill NV, 2001), p. 75.

[25] Apollonius, Conica, p. 615-620.

[26] Fried and Unguru, *Apollonius of Perga's Conica*, p. 74. 27.

[27] *Ibid.*, p. 76.

increases indefinitely as the generating straight line is produced indefinitely,"[28] and in Proposition I.4, he provides the means to cut off a cone of any required size and similar to the original. He thereby satisfies the need for arbitrarily large conic sections while skirting around any issues relating to infinity and providing the intuitiveness of a bounded object. Elegant indeed.

From these solid foundations, Apollonius developed a comprehensive and rigorous treatment of all the geometrical aspects of the conic sections, including their symptomata – the proportional relations that hold for the points of a conic section. Both before and especially after Apollonius, Greek geometers had realized that these symptomata could give the solutions to some open problems that no one had solved using a compass and straightedge – namely the trisection of the angle and the doubling of the cube. However, such solutions encountered widespread resistance among the mathematical community of the time, precisely because they were not reproducible with only a compass and a straightedge. This seems to be an apparent contradiction, as Apollonius had developed the symptomata of the conics with only a compass and straightedge. The problem lies, once again, in infinity. Apollonius proved propositions that allow one to use the symptomata to construct points on the curve of a conic section using a compass and straightedge. However, one cannot construct the entire curve in this manner, unless one were to construct an infinite number of points and add them together, which is impossible in Aristotelian philosophy. Without the entire curve, one cannot find the intersection of that curve with another curve. Therefore, to solve the aforementioned problems required the construction of complete, non-circular, curved lines, an impossible feat by Greek standards.

Thus, at least part of the mathematical community never truly accepted the conic sections as acceptable for constructions. However, as it became clear that trisecting the angle and doubling the cube could not be done without them, pragmatic concerns largely prevailed, and many geometers accepted the conics, and even more complicated curves, into geometrical constructions, with the compromise that geometry was partitioned into three kind of problems.[29] Plane problems were those whose solutions required only a compass and straightedge; solid problems required the conic sections, and linear problems required some other type of curve. In this way, Euclidean geometry embraced a larger range of problems and constructions, at the cost of instituting a class hierarchy of constructions.

We will next consider the innovations in Euclidean geometry that two more modern mathematicians have contributed: Galileo Galilei and René Descartes. First, though, we must briefly note that Euclidean geometry underwent several important changes as it worked its way though the ages between classical Greece and Renaissance Europe. In one of the most marked changes, later authors regarded the starting points of geometry, that is, the definitions, postulates, and axioms, as being much more important than Euclid and his contemporaries. We have already seen a few hints of this trend: the contention over which lines are

[28] 28Apollonius, *Conica*, p. 604.

[29] "Pappus on the three types of geometrical problem," in Fauvel & Gray, pp. 209-210.

acceptable as starting points for constructions, and Proclus's lengthy discourse on all the starting points of Euclid's *Elements* Book I (from which a small sample was excerpted above). We observed earlier that the starting points in the Elements and other texts were not part of the main body of the texts, but rather comprised a more informal introduction. Later authors, like Proclus, seem not to have been aware of this, and in any case they found numerous deficiencies in the postulates and axioms, and sought to correct them. The ultimate result was that the axiomatic foundation of Euclidean geometry grew to become regarded as an essential contributor to its rigor.

The second change relevant to the current work was the coupling of geometry and algebra in medieval Islamic mathematics. Beginning at least with al-Khwarizmi, Islamic mathematicians connected the solutions of equations involving "numbers," "roots," and "squares" (that is, what we call quadratic equations) to the geometrical constructions in *Elements* Book II. Al-Khwarizmi's goal in making this connection was to justify the well-known algebraic and numerical algorithms by lending the rigor of geometry to their justification – that is, to justify the solution by geometrical construction. In his text, he first worked through numerical examples of each type of equation, and then gave a geometrical construction to justify his calculations: "Now, however, it is necessary that we should demonstrate geometrically the truth of the same problems which we have explained in numbers."[30] This, then is the second important development between antiquity and the scientific revolution: the use of Euclidean geometry to justify algebraic solutions to problems.

The first of these changes, the increased importance of axioms as starting points for geometry, allowed Galileo Galilei convincingly to extend Euclidean geometry into the real world. None of the Greek authors had done anything approaching what Galileo did in this direction. Even in Archimedes's treatises On Balancing Planes and On Floating Bodies, in which he explores the physical phenomena of balance and buoyancy, he considers these properties only with respect to the same ideal figures (though Archimedes must certainly have tested with real objects). However, by establishing the appropriate axioms, Galileo could apply geometry and proportion theory to material objects, capable of breaking and moving.

Like the Platonists, Galileo believed that nature was "written in the language of mathematics."[31] However, he conceived differently of the distinction between the abstract and the concrete. In Plato's philosophy, there exists an imperfect world of the senses and the perfect world of the intellect. Galileo claims, though, that one may conceive of an abstract object, such as a sphere, that is not perfect, so that it may touch an abstract imperfect plane in more than one point, even though a perfect sphere and plane may be tangent at only a single point.[32] Logically, then, he decoupled abstractness from perfection, and claimed that abstract thought could indeed correspond as closely as one wished to imperfect, concrete phenomena. In his *Dialogues Concerning the Two New*

[30]"al-Khwarizmi on the algebraic method," Fauvel & Gray, p. 230.
[31]Galileo Galilei, "On mathematics and the world," in Fauvel & Gray, p. 328.
[32]Ibid.

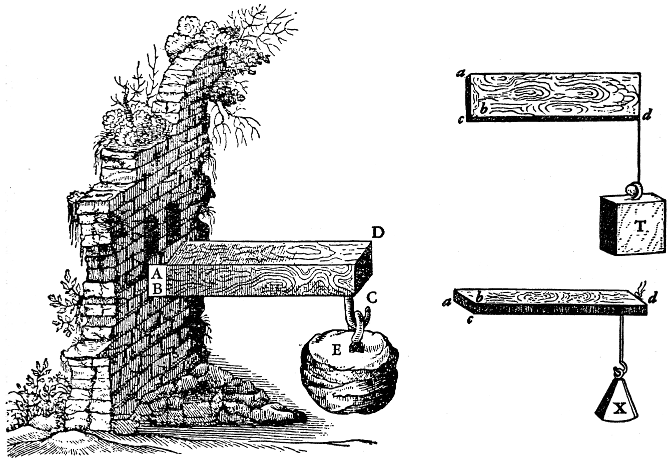

Figure 6: Typical Galilean diagrams. Galileo Galilei, *Two New Sciences*, p. 181.

Sciences, Galileo sought to follow through on his claim by applying mathematics to real, physical problems. We see this already in the diagrams, many of which appear at first sight to be mere drawings rather than geometrical constructions (see Figure 6). A closer inspection reveals that the object of interest within the diagram is in fact a rectangular prism, and that it has indeed been drawn using a straightedge. However, the realistic style of the remainder of the illustration serves to reinforce Galileo's point that he is proving propositions about the behavior of real, physical objects.

At the beginning of the first day's dialogue, Salviati, the representative of Galileo's New Sciences, takes the following axiom:[33]

> Since I assume matter to be unchangeable and always the same, it is clear that we are no less able to treat this constant and invariable property in a rigid manner than if it belonged to simple and pure mathematics.

In this statement, Galileo presents a proportional relation, though he has not worded it as such. The assumed proportion, clarified by subsequent discourse, is that if two objects are composed of the same material, the ratio of their weights is the same as that of their volumes. By this axiom along with another kind of proportional relation, Archimedes's law of the lever, Galileo can reduce the question of the weight required to break a solid object to a problem involving nothing but proportions. At that point, solving the problem by geometrical means is simple, as Euclidean geometry is well-equipped to manipulate proportions. Galileo's strategy, then, is to express clear physical relations as axiomatic

[33]Galileo Galilei, Dialogues Concerning the Two New Sciences (Chicago: Encyclopedia Britannica, Inc., 1952), pp. 131-132.

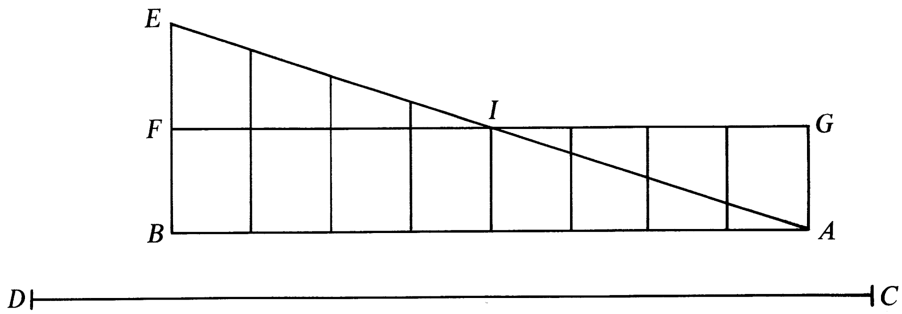

Figure 7: Diagram for Prop. 1 of naturally accelerated motion. Fauvel & Gray, p. 332.

proportions, and from these, to derive a geometrical problem from a physical one.

We see this strategy more clearly in the third day's dialogue, in which the subject of discourse transitions moves on to the second New Science, that of motion. Again, Galileo begins with four axioms regarding motion, all of which follow easily from the Aristotelian definition of uniform motion. As expected, each of these axioms expresses a proportional relation. For example:[34]

> Axiom I: In the case of one and the same uniform motion, the distance traversed during a longer interval of time is greater than the distance traversed during a shorter interval of time.

As in the first two days' dialogues, Galileo uses these axioms once again to transform problems involving motion and acceleration into problems involving proportions, for which he then gives geometrical solutions. Perhaps equally important, the same axioms provide a means to effect the reverse transformation on the solutions. Let us take as an example the well-known first proposition on naturally accelerated motion:[35]

> The time in which any space is traversed by a body starting from rest and uniformly accelerated is equal to the time in which that same space would be traversed by the same body moving at a uniform speed whose value is the mean of the highest speed and the speed just before acceleration began.

Using his axioms as a guide, Galileo proceeds to construct the diagram shown in Figure 7, in which CD represents the interval of time and the areas of the triangle ABE and the rectangle ABFG represent the distances traversed by the first (accelerated) and second (uniformly moving) objects. At the end of the proof, the areas are shown equal, and this then allows Galilei to work backwards and conclude that the two objects will traverse the same distance, because he has

[34] *Ibid.*, p. 197.
[35] *Ibid.*, p. 205.

established a bidirectional correspondence between the areas and the distances traversed. Thus, Galileo has proved by geometry a proposition that describes how real objects move under real forces.

In some sense, then, Galileo did not extend Euclidean geometry. Rather than truly apply geometry to the real world, Galileo brought real problems into the logical space occupied by geometry, using proportions as the critical link between the real and the abstract. However, by using the same proportions to translate the geometrical conclusion into a physical one, he extended the domain of geometrical rigor into new, previously unsolved problems, even if he did not truly extend the geometry itself.

Finally, we consider René Descartes, with respect to his contributions to Euclidean geometry. We saw earlier that medieval Islamic mathematicians had established that one could use geometric constructions to lend rigor to algebraically derived solutions to arithmetic problems. The new algebra gradually filtered from the Islamic world into Europe throughout the late medieval period and the early parts of the scientific revolution. Descartes strongly supported the complementary idea to this one; that is, that one may convert a geometric problem to an algebraic one, solve it algebraically, and then translate the algebraic results back to draw geometric conclusions. In short, Descartes's innovation was to justify geometry using algebra. Just as Galileo established a bidirectional link from ideal geometry to real-world problems, Descartes did the same for geometry and algebra. However, the situation here is somewhat reversed. Galileo used his link to bring problems into the realm of geometry, while Descartes used his to bring them from geometry into algebra, and then accepted the algebraic conclusions as geometrically proven. Naturally, such a move met with considerably more criticism from his contemporaries.

In order to regard any algebraic equation as geometrically rigorous, Descartes necessarily had to regard all algebraic curves as geometrically constructable,[36] for the geometric equivalent of an algebraic manipulation would require the construction of the curves described by the equations. However, in order to uphold the geometric standard of rigor, he needed a geometrical criterion to determine which curves were constructable. He claimed that any curve that "can be conceived of as described by a continuous motion or by several successive motions, each motion being completely determined by those which precede" ought to be acceptable for the purposes of geometrical construction, because for such curves "an exact knowledge of the magnitude of each is always obtainable."[37] He might have said, using the terminology that Apollonius and others had applied to the conics, that such curves have defined symptomata that may be used to determine ratios and proportions. Indeed, he used an analogy to the generations of the conic sections by the intersection of a cone and a plane, saying that any of his curves may be traced by the intersection of two or more moving lines.[38]

To demonstrate his meaning, Descartes described a mechanical instrument

[36] "H. J. M. Bos on Descartes's Geometry," Fauvel & Gray, pp. 349-350.

[37] René Descartes, *The Geometry of René Descartes* (Chicago: The Open Court Publishing Company, 1925), p. 43.

[38] *Ibid.*

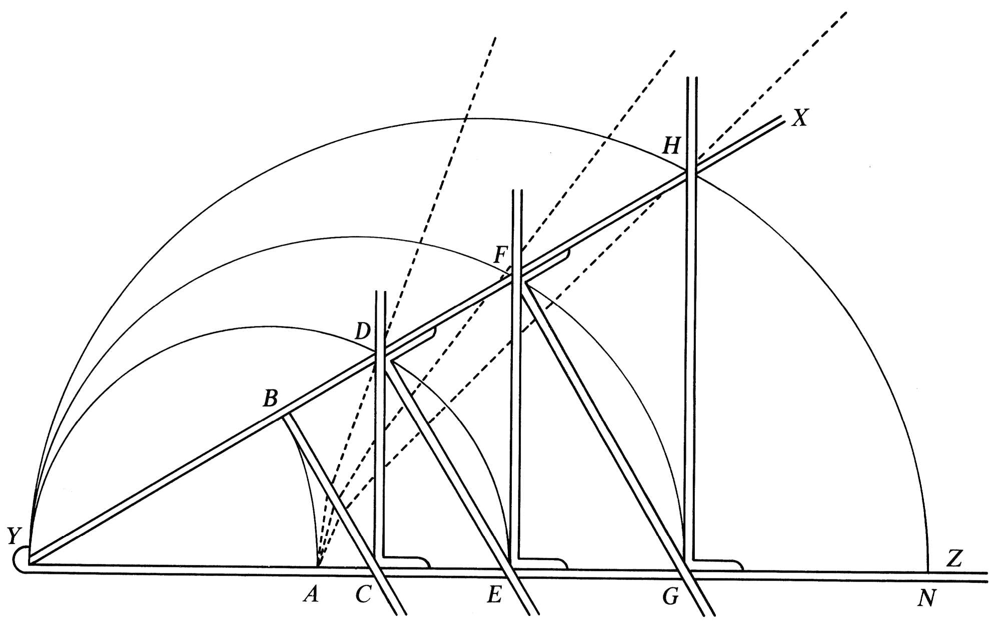

Figure 8: *Descartes's curve-generating instrument. Fauvel & Gray, p. 345.*

(Figure 8) that could generate a whole family of curves from the intersections of moving lines. The key to his conception of these curves as geometrically constructable is that a single continuous motion, the opening of the angle XYZ, causes the parts of the instrument to move in a unique manner, such that each motion is "completely determined by those which precede." In modern terms, we might say that the instrument has only a single degree of freedom, because any given angle XYZ determines a unique point on each curve.[39] Note that the piece BY functions as a simple compass that describes a circular arc AB. This stresses Descartes's point that there is no reasonable criterion for geometric acceptability that could accept the circular arc AB while rejecting the other curves AD, AF, AH, etc., and by extension any curve that could be generated by an instrument whose motions are completely determined by a single continuous motion. Of course, Descartes was still a long way from proving that any algebraic curve could be constructed by such an instrument, but his demonstration served partially to strip away the Greek philosophical seal of approval from the two "simplest" motions, the circle and straight line.

Those who accepted Descartes's views on geometry added an impressive new array of tools to their geometrical toolbox: an infinte variety of curves, and the algebraic tools to manipulate them easily. This came at a cost, of course, for Descartes had glossed over several important points of rigor, confident that his successors would fill in the gaps. As with the conics of Apollonius, the pragmatic

[39]Or, even more succinctly and anachronistically, that points on the curve are functions only of the angle XYZ.

value of this new toolset often outweighed doubts of its rigor, and Descartes's view gradually came to be the accepted one.

The geometry of Descartes may appear radically different from that seen in Euclid's Elements and other classical Greek works, but this change reflects a broader shift in the underlying intuitions that supported geometry through the ages. Just as Euclid took the properties of circles and lines to be self-evident, Descartes perceived certain properties of algebraic equations and curves clearly and distinctly, and he used his understanding of these curves to develop an improved theory of geometry. All throughout, however, the emphasis on simplicity and basic intuition remained a constant source of rigor in Euclidean geometry.

References

[1] Archimedes. The works of Archimedes including the Method. In Robert Maynard Hutchins, editor, *Great Books of the Western World*, volume 11. Chicago: Encyclopedia Britannica, Inc., 1952.

[2] Aristotle. The works of Aristotle Vol. 1. In Robert Maynard Hutchins, editor, *Great Books of the Western World*, volume 8. Chicago: Encyclopedia Britannica, Inc., 1952.

[3] René Descartes. *The Geometry of René Descartes*. Chicago: The Open Court Publishing Company, 1925.

[4] E. J. Dijksterhuis. *Archmedes*. Princeton, NJ: Princeton University Press, 1987.

[5] Euclid. The thirteen books of Euclid's Elements. In Robert Maynard Hutchins, editor, *Great Books of the Western World*, volume 11. Chicago: Encyclopedia Britannica, Inc, 1952.

[6] John Fauvel and Jeremy Gray, editors. *The History of Mathematics: A Reader*. New York: Palgrave Macmillan, 1987.

[7] Michael N. Fried and Sabetai Unguru. *Apollonius of Perga's Conica: Text, Context, Subtext*. Leiden, Netherlands: Koninklijke Brill NV, 2001.

[8] Galileo Galilei. Dialogues concerning the two new sciences. In Robert Maynard Hutchins, editor, *Great Books of the Western World*, volume 28. Chicago: Encyclopedia Britannica, Inc, 1952.

[9] David C. Lindberg. *The Beginnings of Western Science*. Chicago: The University of Chicago Press, 1992.

[10] Reviel. Netz. *The Shaping of Deduction in Greek Mathematics*. New York: Cambridge University Press, 1999.

[11] Reviel. Netz. *The Transformation of Mathematics in the Early Mediterranean World*. New York: Cambridge University Press, 2004.

[12] Apollonius of Perga. Conica. In Robert Maynard Hutchins, editor, *Great Books of the Western World*, volume 11. Chicago: Encyclopedia Britannica, Inc., 1952.

[13] Plato. Timaeus. In Daniel C. Stevenson, editor, *The Internet Classics Archive, http://classics.mit.edu/Plato/timaeus.html*. Cambridge, MA: Massachusetts Institute of Technology, 1994-2000.

[14] Proclus. *A Commentary on the First Book of Euclid's Elements*. New Jersey: Princeton University Press, 1970.