Teaching Geometry According to Euclid

Robin Hartshorne

In the fall semester of 1988, I taught an undergraduate course on Euclidean and non-Euclidean geometry. I had previously taught courses in projective geometry and algebraic geometry, but this was my first time teaching Euclidean geometry and my first exposure to non-Euclidean geometry. I used the delightful book by Greenberg [8], which I believe my students enjoyed as much as I did.

As I taught similar courses in subsequent years, I began to be curious about the origins of geometry and started reading Euclid's *Elements* [12]. Now I require my students to read at least Books I–IV of the *Elements*. This essay contains some reflections and questions arising from my encounters with the text of Euclid.

Euclid's *Elements*

A treatise called the *Elements* was written approximately 2,300 years ago by a man named Euclid, of whose life we know nothing. The *Elements* is divided into thirteen books: Books I–VI deal with plane geometry and correspond roughly to the material taught in high school geometry courses in the United States today. Books VII–X deal with number theory and include the Euclidean algorithm, the infinitude of primes, and the irrationality of $\sqrt{2}$. Books XI–XIII deal with solid geometry, culminating in the construction of the five regular, or platonic, solids.

Throughout most of its history, Euclid's *Elements* has been the principal manual of geometry and indeed the required introduction to any of the sciences. Riccardi [15] records more than one thousand editions, from the first printed edition of 1482 up to about 1900. Billingsley, in his preface to the first English translation of the *Elements* (1570) [1], writes, "Without the diligent studie of Euclides Elements, it is impossible to attaine unto the perfecte knowledge of Geometrie, and consequently of any of the other Mathematical Sciences." Bonnycastle, in the preface to his edition of the *Elements* [4], says, "Of all the works of antiquity which have been transmitted to the present time, none are more universally and deservedly esteemed than the Elements of Geometry which go under the name of Euclid. In many other branches of science the moderns have far surpassed their masters; but, after a lapse of more than two thousand years, this performance still retains its original preeminence, and has even acquired additional celebrity for the fruitless attempts which have been made to establish a different system." Todhunter, in the preface to his edition [18], says simply, "In England the text-book of Geometry consists of the Elements of Euclid." And Heath, in the preface to his definitive English translation [12], says, "Euclid's work will live long after all the text-books of the present day are superseded and forgotten. It is one of the noblest monuments of antiquity; no mathematician worthy of the name can afford not to know Euclid, the real Euclid as distinct from any revised or rewritten versions which will serve for schoolboys or engineers."

These opinions may seem out-of-date today, when most modern mathematical theories have a history of less than one hundred years and the latest logical restructuring of a subject is often the most prized, but they should at least engender some curiosity about what Euclid did to have such a lasting impact.

Robin Hartshorne is professor of mathematics at the University of California at Berkeley. His e-mail address is robin@math.berkeley.edu.

How Geometry Is Taught Today

A typical high school geometry course contains results about congruent triangles, angles, parallel lines, the Pythagorean theorem, similar triangles, and areas of rectilinear plane figures that are familiar to most of us from our own school days. The material is taught mainly as a collection of truths about geometry, with little attention to axioms and proofs. However, one does find in most texts the "ruler axiom", which says that the points of a line can be put into one-to-one correspondence with the real numbers in such a way that the distance between two points is the difference of the corresponding real numbers. This is presumably due to the influence of Birkhoff's article [3], which advocated the teaching of geometry based on measurement of distances and angles by the real numbers. It seems to me that this use of the real numbers in the foundations of geometry is analysis, not geometry. Is there a way to base the study of geometry on purely geometrical concepts?

A college course in geometry, as far as I can tell from the textbooks currently available, provides a potluck of different topics. There may be some "modern Euclidean geometry" containing fancy theorems about triangles, circles, and their special points, not found in Euclid and mostly discovered during a period of intense activity in the mid-nineteenth century. Then there may be an introduction to the problem of parallels, with the discovery of non-Euclidean geometry; perhaps some projective geometry; and something about the role of transformation groups. All of this is valuable material, but I am disappointed to find that most textbooks have somewhere a hypothesis about the real numbers equivalent to Birkhoff's ruler axiom.

OF EUCLID. 51

Book I.

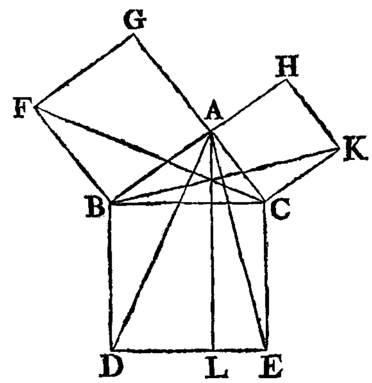
PROP. XLVII. THEOR.

IN any right angled triangle, the ſquare which is deſcribed upon the ſide ſubtending the right angle, is equal to the ſquares deſcribed upon the ſides which contain the right angle.

Let ABC be a right angled triangle having the right angle BAC; the ſquare deſcribed upon the ſide BC, ſhall be equal to the ſquares deſcribed upon BA, AC.

On BC deſcribe [a] the ſquare BDEC, and on BA, AC the ſquares GB, HC; and thro' A draw [b] AL parallel to BD or CE, and join AD, FC. therefor becauſe each of the angles BAC, BAG is a right angle[c], the two ſtraight lines AC, AG upon the oppoſite ſides of AB, make with it at the point A the adjacent angles equal to two right angles; therefor CA is in the ſame ſtraight line [d] with AG. by the ſame reaſon, AB and AH are in the ſame ſtraight line. and becauſe the angle DBC is equal to the angle FBA, for each of them is a right angle, add to each of them the angle ABC, and the whole angle DBA ſhall be equal [e] to the whole FBC. and becauſe the two ſides AB, BD are equal to the two FB, BC, each to each, and the angle DBA equal to the angle FBC; the baſe AD ſhall be equal [f] to the baſe FC, and the triangle ABC to the triangle FBC. now the parallelogram BL is double [g] of the

a. 46. 1.
b. 31. 1.
c. 30. Def.
d. 14. 1.
e. 2. Ax.
f. 4. 1.
g. 41. 1.



G 2 the

Figure 1. The Pythagorean theorem, in Simson's translation [17]. The proof shows that the triangle *ABD* is congruent to the triangle *FBC*. Then the rectangle *BL*, being twice the first triangle, is equal to the square *GB*, which is twice the second triangle. Similarly, the rectangle *CL* equals the square *HC*. Thus the squares on the sides of the triangle *ABC*, taken together, are equal to the square on the base.

This use of the real numbers obscures one of the most interesting aspects of the development of geometry: namely, how the concept of continuity, which belonged originally to geometry only, came gradually by analogy to be applied to numbers, leading eventually to Dedekind's construction of the field of real numbers.

Number versus Magnitude in Greek Geometry

In classical Greek geometry the *numbers* were 2,3,4,... and the unity 1. What we call negative numbers and zero were not yet accepted. Geometrical quantities such as line segments, angles, areas, and volumes were called *magnitudes.* Magnitudes of the same kind could be compared as to size: less, equal, or greater, and they could be added or subtracted (the lesser from the greater). They could not be multiplied, except that the operation of forming a rectangle from two line segments, or a volume from a line segment and an area, could be considered a form of multiplication of magnitudes, whose result was a magnitude of a different kind.

In Euclid's *Elements* there is an undefined concept of equality (what we call *congruence*) for line segments, which could be tested by placing one segment on the other to see whether they coincide exactly. In this way the equality or inequality of line segments is perceived directly from the geometry without the assistance of real numbers to measure their lengths. Similarly, angles form a kind of magnitude that can be compared directly as to equality or inequality without any numerical measure of size.

Two magnitudes of the same kind are *commensurable* if there exists a third magnitude of the same kind such that the first two are (whole number) multiples of the third. Otherwise they are *incommensurable*. So Euclid does not say the square root of two (a number) is irrational (i.e., not a rational number). Instead he says (and proves) that the diagonal of a square is incommensurable with its side.

The difference between classical and modern language is especially striking in the case of area. In the *Elements* there is no real number measure of the area of a plane figure. Instead, equality of plane figures (which I will call equal *content*) is verified by cutting in pieces and adding and subtracting congruent triangles. Thus the Pythagorean theorem (Book I, Proposition 47, "I.47" for short) says that the squares on the sides of a right triangle, taken together, have the same content as the square on the hypotenuse. This is proved as the culmination of a series of propositions demonstrating equal content for various figures (for example, triangles with congruent bases and congruent altitudes have the same content).

For the theory of similar triangles, a modern text will say two triangles are *similar* if their sides are *proportional*, meaning the ratios of the lengths of corresponding sides are equal to a fixed real number. Euclid instead uses the theory of proportion, due to Eudoxus, that is developed in Book V of the *Elements*. Two magnitudes a, b of the same kind are said to have a *ratio* $a : b$. This ratio is not a number, nor is it a magnitude. Its main role is explained by the following fifth definition of Book V: Two ratios $a : b$ and $c : d$ are *equal* (in which case we say that there is a *proportion* a is to b as c is to d, and write $a : b :: c : d$) if for every choice of whole numbers m, n, the multiple ma is less than, equal to, or greater than the multiple nb if and only if mc is less than, equal to, or greater than nd, respectively.

No arithmetic operations (addition, multiplication) are defined for these ratios, but they can be ordered by size. In Book V a number of rules of operation on proportions are proved, using the above definition—for example, one called *alternando* (V.16), which says if $a : b :: c : d$, then also $a : c :: b : d$.

The whole theory of similar triangles is developed in Book VI based on the definition that two triangles are similar if their corresponding sides are proportional in pairs.

Thus Euclid develops his geometry without using numbers to measure line segments, angles, or areas. His theorems have the same appearance as the ones we learn in high school, yet their meaning is different when we look closely. So the questions arise: How and when did this change of perception occur? How and when were the real numbers introduced into geometry? Was Euclid already using something equivalent to the real numbers in disguised form?

Development of the Real Numbers

In Greek mathematics, as we saw, the only numbers were (positive) integers. What we call a rational number was represented by a ratio of integers. Any other quantity was represented as a geometrical magnitude. This point of view persisted even to the time of Descartes. In Book III of *La Géométrie* [6], when discussing the roots of cubic and quartic equations, Descartes considers polynomials with integer coefficients. If there is an integer root, that gives a numerical solution to the problem. But if there are no integral roots, the solutions must be constructed geometrically. A quadratic equation gives rise to a *plane* problem whose solution can be constructed with ruler and compass. Cubic and quartic equations are *solid* problems that require the intersections of conics for their solution. The root of the equation is a certain line segment constructed geometrically, not a number.

Halley [10] improves the method of Descartes to find roots of equations of degree up to six, using intersections of cubic curves in the plane. But he also shows an interest (following Newton) in finding approximate decimal numerical solutions to an equation. He comments that the geometrical method gives an exact theoretical solution but that for practical purposes one can get a more accurate solution—"as near the truth as you please"—by an arithmetical calculus.

One hundred years later the acceptance of approximate numerical solutions had progressed so far that Legendre [14, p. 61], in discussing the theory of proportion, says

> If A, B, C, D are lines [line segments], one can imagine that one of these lines, or a fifth, if one likes, serves as a common measure and is taken as unity. Then A, B, C, D represent each a certain number of unities, whole or fractional, commensurable or incommensurable, and the proportion among the lines A, B, C, D becomes a proportion of numbers.

Legendre's uncritical acceptance of numbers representing geometrical magnitudes makes his proofs easier but at the expense of rigor, for he has not said what kind of numbers these are, nor has he proved that they obey the usual rules of arithmetic.

It was Dedekind [5] who provided a rigorous definition of the real numbers. He was dissatisfied with the appeal to geometric intuition for matters of limits in the infinitesimal calculus and wanted to give a solid theory of continuity based on numbers. He saw the property of continuity expressed in the property of a line: that if one divides its points into two nonempty subsets A, B, with every point of A lying to the left of every point of B, then there exists exactly one point of the line that marks this division. This prompted him to *define* a real number as a partition of the set of rational numbers into two nonempty subsets A, B, with $a < b$ for all $a \in A$ and $b \in B$.[1] He then defined operations of addition, subtraction, multiplication, and division for these new real numbers and proved that they obey the usual laws of arithmetic (hence form a *field* in modern language). He also proved the key property of existence of a least upper bound of a nonempty bounded set of real numbers, needed for the theory of limits in the infinitesimal calculus.

Dedekind's awareness of the abstract nature of his construction is shown by this telling remark: "If space has any real existence at all, it does not necessarily need to be continuous. And if we knew for certain that space was discontinuous, still nothing could hinder us, if we so desired, from making it continuous in our thought by filling up its gaps." The German expression is *die Stetigkeit in die Linie hineindenken* or to think the continuity into the line.

While the gradual acceptance of numbers to measure geometrical quantities was a useful development in thinking about geometrical problems, this approach did not have a rigorous basis until Dedekind's construction of the field of real numbers, and until then there was no adequate substitute for Eudoxus's theory of proportion.

Dedekind's definition implies a criterion for the equality of two real numbers α, β: namely, that a rational number m/n is less than, equal to, or greater than α if and only if it is less than, equal to, or greater than β. This is almost identical to Euclid's definition of equality of ratios in the theory of proportion. So one may ask: Did Euclid already have the concept of real numbers in the back of his mind? Tempting as it may be to impute such a discovery to Euclid, I say no, because Euclid dealt only with those magnitudes that arose in his geometry, magnitudes constructible by ruler and compass, while Dedekind made the amazing mental leap of considering the set of all Dedekind cuts, which for Euclid would have been inconceivable.

The Rise of Analytic Geometry

Analytic geometry, as we understand it today, is based on the principle that by drawing two perpendicular axes in the plane and choosing an interval to serve as unit, one can establish a one-to-one correspondence between the points of the plane and ordered pairs of real numbers. This correspondence creates a dictionary between geometry and algebra, so that geometrical problems can be translated into algebraical properties of polynomial or more general functions of a real variable.

Going one step further, one may *define* the plane to be the set of ordered pairs of real numbers and a straight line to be the subset of all pairs (x, y) satisfying a linear equation $ax + by + c = 0$, with a and b not both 0. Then geometry as an independent discipline disappears; it becomes a branch of algebra or real analysis.

A common misconception is that analytic geometry was invented by Descartes. In the form just described, certainly not. The real numbers had not yet been invented, and even the idea of representing a line segment by any sort of number was not yet clearly developed. If we read the geometry of Descartes carefully, we see that he is applying algebra to geometry. In any problem, he represents known line segments by letters $a, b, c, \ldots$, and unknown line segments by letters $x, y, z, \ldots$. Then from the data of the problem he seeks relations among them that can be expressed as equations with letters $a, b, c, x, y, z, \ldots$. These equations are solved by the usual rules of algebra. The solution gives a recipe for the geometrical construction of the unknown line segments.

Throughout this process, the letters represent line segments, not numbers. What Descartes has really done is to create an *arithmetic of line segments*. Two line segments can be added by placing them end to end. Two line segments a, b can be multiplied, once one has fixed a segment 1 to act as unit, by making a triangle with sides $1, a$, and another similar triangle with sides b, x, so that $x = ab$. His equations are always equations among line segments.

The same approach is followed by Guisnée [9]. It is only by the time of Legendre that his contemporary Biot, in one of the first books on "analytic geometry" [2], allows the letters to be interpreted either as line segments or as the numerical values representing them. But Biot, like Legendre, does not say just what kind of numbers he is using.

Nowhere in the geometry of Descartes does he explain by what right he may assume that the operations he defines on line segments obey the usual laws of arithmetic. For example, it is a nontrivial matter to show that his multiplication of line segments, defined using similar triangles as above, is commutative. This difficulty is answered in the most satisfactory way by Hilbert.

Hilbert, in his *Grundlagen der Geometrie* [13], gives a set of axioms for geometry based on those of Euclid but including others to make explicit some notions, such as betweenness, that were only

[1] *Such a partition (A, B) Dedekind called a* cut. *Strictly speaking, each rational number r defines two such cuts, one in which r is the largest element of A and the other in which r is the smallest element of B. Dedekind said he would regard these two cuts as defining the same real number. Readers of Rudin [16] will recognize that Rudin's cuts are the left-hand halves of Dedekind's cuts, with the ambiguity of rational cuts resolved by requiring that A have no largest element.*

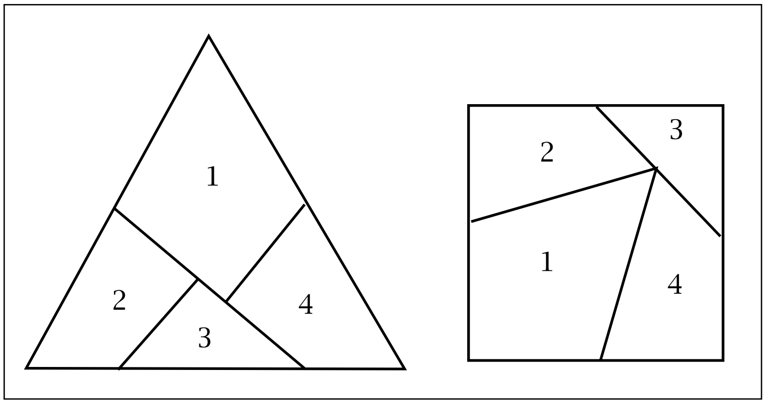


Figure 2. The Haberdasher's puzzle, to dissect an equilateral triangle into a square, from Dudeney [7]. It is an exercise to construct, with ruler and compass, the exact cuts necessary to make this work.

intuitive in the *Elements*. Then he defines arithmetic operations of addition and multiplication on the set of equivalence classes of congruent line segments and proves that there exists an ordered field whose positive elements are the equivalence classes of line segments. This is a wonderful result. First, the field evolves intrinsically from the geometry instead of being imposed from without. Second, we discover that this field is not necessarily the field of real numbers. If we take only the axioms of Euclid's *Elements*, where all constructions must be effected by ruler and compass, we obtain the *constructible* field of all real numbers contained in successive quadratic extensions of $\mathbb{Q}$. If we take Hilbert's slightly weaker axioms, which do not include the circle-circle intersection property, we obtain the smaller field of all totally real elements in the constructible field. Only if we add the extra axiom that every Dedekind cut of a line is determined by a point of that line, is the field generated by our geometry necessarily isomorphic to the field of real numbers.[2]

So why not, in the pure spirit of geometry, consider other geometries in which Dedekind's axiom does not hold? What if we lived in a non-Archimedean universe, with other infinite universes beyond the farthest stars, and infinitesimal subworlds inside every electron?

Conclusion

These reflections suggest another way for a course in geometry to grow, with its roots in the purely geometric tradition and branches making use of modern algebra. We start with the first four books of Euclid's *Elements*, culminating in the elegant construction of the regular pentagon, but not including the theory of proportion and similar triangles. We retrofit the foundations with Hilbert's axioms to bring the treatment up to modern standards of rigor. We use Hilbert's segment arithmetic to obtain a coordinate field intrinsically determined by the geometry. With this field we can develop the usual theory of similar triangles.

The transition to non-Euclidean geometry is natural. If we drop the parallel axiom, we can explore the results of *neutral geometry*, in which the parallel axiom is neither affirmed nor denied. If we add the axiom of existence of limiting parallel rays, we can develop all of *hyperbolic geometry*, including the construction of an intrinsic field of coordinates, and an associated hyperbolic analytic geometry and hyperbolic trigonometry.

With this approach there is no need for the real numbers, no appeal to continuity. In this way the true essence of geometry can develop most naturally and economically.

Dividends

Whenever one approaches a subject from two different directions, there is bound to be an interesting theorem expressing their relation.

In the theory of area, for example, Euclid's notion of equal content is based on cutting up and rearranging a plane figure. The modern approach is to consider a measure of area function that to each figure associates a real number, its area. The equivalence between these two approaches is expressed in the following theorem of Bolyai and Gerwien.

***Theorem**. Two rectilinear plane figures A and B have the same measure of area if and only if it is possible to cut the figure A into triangles $A_1, \ldots, A_n$ and B into triangles $B_1, \ldots, B_n$ in such a way that A_i is congruent to B_i for each i.*

The proof (see, for example, [11, §24]) uses little more than the results on application of areas in the end of Book I of Euclid's *Elements*.

Gauss noted with curiosity that in the treatment of volumes in Book XII of the *Elements*, Euclid does not use the analogous method of dissection to define equal volume content, but instead employs a limiting process called the method of exhaustion. Hilbert asked in his third problem at the International Congress of 1900 whether this limiting process was really necessary. Dehn provided the answer by showing that a cube and a

---

[2] *Hilbert's axioms for plane geometry postulate a set of points and a set of subsets called lines, a notion of betweenness, and undefined relations of congruence for line segments and for angles. The axioms of* incidence *require that two distinct points lie on a unique line, plus conditions of nontriviality. The axioms of* betweenness *govern the relation that a point B lies between points A and C. The axioms of* congruence *include, among others, that it is possible to lay off a segment congruent to a given segment on a given line; that it is possible to lay off an angle congruent to a given angle at a given point on a line; and the side-angle-side (SAS) criterion for congruence of triangles. These are the basic axioms of a* Hilbert plane*. For Euclidean geometry one needs also the* parallel axiom*, that there is at most one line parallel to a given line through a given point; and the* circle-circle intersection axiom*, that if one circle has a point inside and a point outside another circle, then the two circles meet in two points.*

tetrahedron of equal volume cannot be dissected into a finite number of congruent subsolids. The proof (see, for example, [11, §27]) is a nice example of the application of algebra to geometry.

Moral

The moral of my story is: Read Euclid and ask questions. Then teach a course on Euclid and later developments arising out of these questions.

References

[1] H. Billingsley, *The Elements of Geometrie, of the Most Auncient Philosopher Euclide of Megara*, John Daye, London, 1570.

[2] J.-B. Biot, *Essai de Géométrie Analytique*, 6th ed., Bachelier, Paris, 1823 (first published in 1802).

[3] G. D. Birkhoff, A set of postulates for plane geometry (based on scale and protractor), *Ann. of Math.* **33** (1932), 329–345.

[4] J. Bonnycastle, *Elements of Geometry, Containing the Principal Propositions in the First Six, and the Eleventh and Twelfth Books of Euclid*, J. Johnson, London, 1798.

[5] R. Dedekind, *Stetigkeit und Irrationale Zahlen*, Braunschweig, 1872.

[6] R. Descartes, *La Géométrie*, C. Angot, Paris, 1664 (first published in 1637).

[7] H. E. Dudeney, *The Canterbury Puzzles, and Other Curious Problems*, 3rd ed., Thomas Nelson and Sons, London, 1929.

[8] M. J. Greenberg, *Euclidean and Non-Euclidean Geometries, Development and History*, 3rd ed., Freeman, New York, 1993.

[9] N. Guisnée, *Application de l'Algèbre à la Géométrie*, 2nd ed., Quillau, Paris, 1733.

[10] E. Halley, Lectures read in the School of Geometry in Oxford, appendix to: *The Elements of that Mathematical Art Commonly Called Algebra*, by J. Kersey, London, R. and W. Mount, 1717.

[11] R. Hartshorne, *Geometry, Euclid and Beyond*, Springer-Verlag, New York, expected date of publication, May 2000.

[12] T. L. Heath, *The Thirteen Books of Euclid's Elements*, translated from the text of Heiberg, with introduction and commentary, 2nd ed., 3 vols., University Press, Cambridge, 1926 (available in Dover reprint).

[13] D. Hilbert, *Foundations of Geometry*, 2nd English ed., Open Court, La Salle, 1971 (first published as *Grundlagen der Geometrie*, 1899).

[14] A. M. Legendre, *Eléments de Géométrie, avec des notes*, 12th ed., Firmin Didot, Paris, 1823.

[15] P. Riccardi, Saggio di una bibliografia Euclidea, *Mem. Real. Accad. Sci. Inst. Bologna* **8** (1887), 401–523.

[16] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill, New York, 1976.

[17] R. Simson, *The Elements of Euclid, viz. the first six books, together with the eleventh and twelfth. In this Edition, the Errors, by which Theon, or others, have long ago Vitiated these Books, are Corrected, and some of Euclid's Demonstrations are Restored*, Robert and Andrew Foulis, Glasgow, 1756.

[18] I. Todhunter, *The Elements of Euclid, for the Use of Schools and Colleges*, MacMillan, London, 1882 (first published in 1862).