Robin Hartshorne

# Geometry: Euclid and Beyond

With 550 Illustrations

Springer

Robin Hartshorne
Department of Mathematics
University of California
Berkeley, CA 94720
USA
robin@math.berkeley.edu



*Cover illustration*: The diagram illustrates the theorem that the radical axes of three circles meet in a point (Exercise 20.5, p.182).

Mathematics Subject Classification (2000): 51-01

Library of Congress Cataloging-in-Publication Data
Hartshorne, Robin.
Geometry: Euclid and beyond/Robin Hartshorne.
p. cm. – (Undergraduate texts in mathematics)
Includes bibliographical references and index.
ISBN 0-387-98650-2 (hc)
1. Geometry I. Title. II. Series.
QA451.H37 2000
516 21—dc21 99-044789



ISBN 0-387-98650-2 Springer-Verlag New York Berlin Heidelberg SPIN 10696748

# 3 Euclid's Axiomatic Method

One of the remarkable features of Euclid's *Elements* is its orderly logical structure. Euclid took the great mass of geometrical material that had grown in the previous two or three centuries, and organized it into one coherent logical sequence. This is what we now call the *axiomatic method*: Starting from a small number of definitions and assumptions at the beginning, all the succeeding results are proved by logical deduction from what has gone before. Euclid's text has been a model of mathematical exposition, unchallenged for two thousand years, and only recently (in the last hundred years or so) replaced by newer mathematical systems that we consider more rigorous. As we read Euclid, let us observe how he organizes his material, let us be curious about why he does things the way he does, and let us explore the questions that come to mind when we as modern mathematicians read this ancient text.

### Definitions

Euclid begins with definitions. Some of these definitions are akin to the modern notion of definition in mathematics, in that they give a precise meaning to the term being defined. For example, the tenth definition tells us that if a line segment meets a line so that the angles on either side are equal, then these are called right angles. This tells us the meaning of the term *right angle*, assuming that we already know what is meant by a line, a line segment, an angle, and equality of angles. Similarly, the fifteenth definition, rephrased, defines a circle to be a set of points $C$, such that the line segments $OA$ from a fixed point $O$ to any point $A$ of the circle $C$, are all equal to each other, and the point $O$ is called the center of the circle. This tells us what a circle is, assuming that we already know what a line segment is, and what is meant by equality of line segments.

On the other hand, some of Euclid's other definitions, such as the first, "a point is that which has no part," or the second, "a line is breadthless length," or the third, "a straight line is a line which lies evenly with the points on itself," give us no better understanding of these notions than we had before. It seems that Euclid, instead of giving a precise meaning to these terms, is appealing to our intuition, and alluding to some concept we may already have in our own minds of what a point or a line is. Rather than defining the term, he is appealing to our common understanding of the concept, without saying what that is. This may have been very well in a society where there was just one truth and one geometry and everyone agreed on that. But the modern consciousness sees this as a rather uncertain way to set up the foundations of a rigorous discipline. What if we say now, oh yes, we agree on what points and lines are, and then later it turns out we had something quite different in mind? So the modern approach is to say these notions are *undefined*, that is, they can be anything at all, *provided* that they satisfy whatever postulates or axioms may be imposed on them later. In the algebraic definition of an abstract group, for example, you never say what the elements of the group are, nor what the group operation is. Those are undefined. However, they must satisfy the group axioms that the operation is associative, there exists an identity, and that there exist inverses. The elements of the group can then be anything as long as they satisfy these axioms. They could be integers, or they could be cosets of a subgroup of the integers, or they could be rotations of a geometrical object such as a cube, or anything else. So in our reading of Euclid, perhaps we should regard "point" and "line" as undefined terms.

It may be worth noting some differences of language between Euclid's text and modern usage. By a *line* he means something that may be curved, which we would call a curve. He says *straight line* for what we call line. And then he says a *finite straight line* (as in the statement of (I.1)) for what we would call a line segment. For Euclid, a *plane angle* results where two curves meet, and a *rectilineal plane angle* is formed when two line segments meet. Note that Euclid requires the two sides of an angle not to lie in a straight line. So for Euclid there is no zero angle, and there is no straight angle (180°). So we should think of Euclid's concept of angle as meaning an angle of $\alpha$ degrees, with $0 < \alpha < 180°$ (though Euclid makes no mention of the degree measure of an angle).

Euclid's notion of *equality* requires special attention. He never defines equality, so we must read between the lines to see what he means. In Euclid's geometry there are various different kinds of magnitudes, such as line segments, angles, and later areas. Magnitudes of the same kind can be compared: They can be equal, or they can be greater or lesser than one another. Also, they can be added and subtracted (provided that one is greater than the other) as is suggested by the common notions.

Euclid's notion of equality corresponds to what we commonly call *congruence* of geometrical figures. In high-school geometry one has the length of a line segment, as a real number, so one can say that two segments are congruent if they have the same length. However, there are no lengths in Euclid's geometry, so we must regard his equality as an undefined notion. Because of the first common notion, "things which are equal to the same thing are also equal to one another," we may regard equality (which we will call *congruence* to avoid overuse of the word equal) to be an equivalence relation on line segments. Similarly, we will regard congruence of angles as an equivalence relation on angles.

**Postulates and Common Notions**

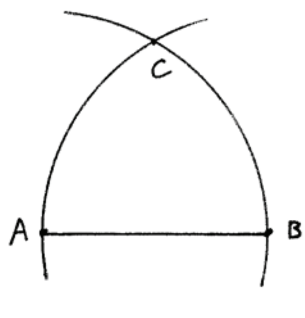
The postulates and common notions are those facts that will be taken for granted and used as the starting point for the logical deduction of theorems. If you think of Euclid's geometry in the classical way as being the one true geometry that describes the real world in its ideal form, then you may regard the postulates and common notions as being self-evident truths for which no proof is required. If you think of Euclid's geometry in the modern way as an abstract mathematical theory, then the postulates and common notions are merely those statements that are arbitrarily selected as the starting point of the theory, and from which other results will be deduced. There is no question of their "truth," because one can begin a mathematical theory from any hypotheses one likes. Later on, however, there may arise a question of relevance, or importance of the mathematical theory constructed. The importance of a mathematical theory is judged by its usefulness in proving theorems that relate to other branches of mathematics or to applications. If you begin a mathematical theory with weird hypotheses as your starting point, you may get a valid logical structure that is of no use. From that point of view the choice of postulates is not so arbitrary. In any case, we can regard Euclid's postulates and common notions collectively as the set of *axioms* on which his geometry is based.

Some commentators say that the postulates (as in Heath's edition) are those statements that have a geometrical content, while the common notions are those statements of a more universal nature, which apply to all the sciences. Other commentators divide them differently, calling "postulates" those statements that allow you to construct something, and calling "axioms" those statements that assert that something is always true. One should also note that some editors give extra axioms not listed in Heath's edition, such as "halves of equals are equal," which is used by Euclid in the proof of (I.37), or "two straight lines cannot contain a space."

We have already noted the constructive nature of Euclid's approach to geometry as expressed in Postulates 1–3. By the way, Euclid makes no explicit statement about the uniqueness of the line mentioned in Postulate 1, though he apparently meant it to be unique, because in the proof of (I.4) he says "otherwise two straight lines will enclose a space: which is impossible."

In the list of Postulates and Common Notions, Postulate 5 stands out as being much more sophisticated than the others. It sounds more like a theorem than an axiom. We will have more to say about this later. For the moment let us just observe that two thousand years of unsuccessful efforts to prove this statement as a consequence of the other axioms have vindicated Euclid's genius in realizing that it was necessary to include Postulate 5 as an axiom.

### Intersections of Circles and Lines

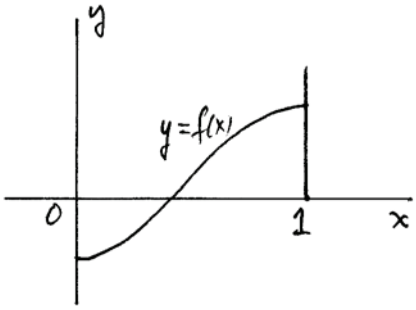



As we read Euclid's *Elements* let us note how well he succeeds in his goal of proving all his propositions by pure logical reasoning from first principles. We will find at times that he relies on "intuition," or something that is obvious from looking at a diagram, but which is not explicitly stated in the axioms. For example, in the construction of the equilateral triangle on a given line segment $AB$ (I.1) how does he know that the two circles actually meet at some point $C$? While the fifth Postulate guarantees that two lines will meet under certain conditions, there is nothing in the definitions, postulates, or common notions that says that two circles will meet. Nor does Euclid offer any reason in his proof that the two circles will meet.

If you carry out the construction with ruler and compass on a piece of paper, you will find that they do meet. Or if you look at the diagram, it seems obvious that they will meet. However, that is not a proof, and we must acknowledge that Euclid is using something that is not explicitly guaranteed by his axioms and yet is essential to the success of his construction.

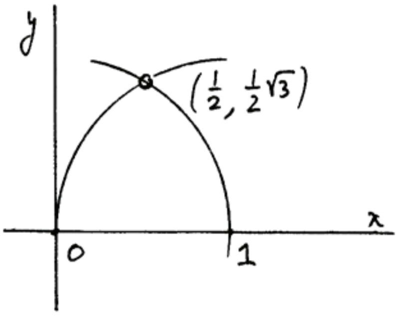
There are two separate issues here. One is the relative position of the two circles. Two circles need not always meet. If they are far apart from each other, or if one is entirely contained in the other, they will not meet. In the present case, part of one circle is inside the other circle, and part outside, so it appears from the diagram that they must cross each other.

The second issue is, assuming that they are in a position so that they appear to meet, does the intersection point actually exist? Today we will immediately think of continuity and the intermediate value theorem: If $y = f(x)$ is a real-valued continuous function defined on the unit interval $[0, 1]$ of the real numbers, and if $f(0) < 0$ and $f(1) > 0$, then there is some point $a \in [0, 1]$ with $f(a) = 0$. In other words, the graph of the function must intersect the $x$-axis at some point in the interval.




However, we must bear in mind that the concepts of real numbers and continuous functions were not made rigorous until the late nineteenth century, and that this kind of mathematical thinking is foreign to the spirit of Euclid's *Elements*.

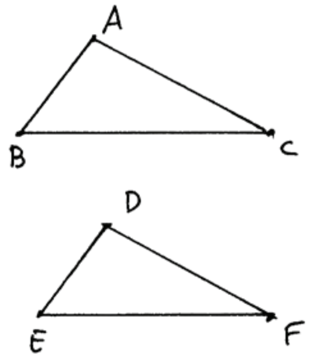
To make the same point in a different way, suppose we consider the *Cartesian plane* over the field of rational numbers $\mathbb{Q}$, where points are ordered pairs of rational numbers, and let $AB$ be the unit interval on the $x$-axis. Then the vertex $C$ of the equilateral triangle, which would have to be the point $(\frac{1}{2}, \frac{1}{2}\sqrt{3})$, actually does not exist in this geometry.




So later on, when we set up a new system of axioms for Euclidean geometry, we will have to include some axiom that guarantees the existence of the intersection points of circles with other circles, or with lines, at least those that arise in the ruler and compass constructions of Euclid's *Elements*. Some modern axiom systems (such as Birkhoff (1932) or the School Mathematics Study Group geometry) build the real numbers into the axioms with a postulate of line measure, or include Dedekind's axiom that essentially guarantees that we are working over the real numbers. In this book, however, we will reject such axioms as not being in the spirit of classical geometry, and we will introduce only those purely geometric axioms that are needed to lay a rigorous foundation for Euclid's *Elements*.

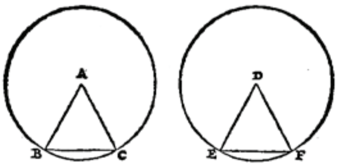
The issue of intersecting circles arises again in (I.22), where Euclid wishes to construct a triangle whose sides should be equal to three given line segments $a$, $b$, $c$. This requires that a circle with radius $a$ at one endpoint of the segment $b$ should meet a circle of radius $c$ at the other end of the segment $b$. Euclid correctly puts the necessary and sufficient condition that this intersection should exist in the statement of the proposition, namely that any two of the line segments should be greater than the third. However, he never alludes to this hypothesis in his proof, so that we do not see in what way this hypothesis implies the existence of the intersection point. While some commentators have criticized Euclid for this, Simson ridicules them, saying "For who is so dull, though only beginning to learn the Elements, as not to perceive ... that these circles must meet one another because *FD* and *GH* are together greater than *FG*." Still, Simson has only discussed the position of the circles and has not addressed the second issue of why the intersection point exists. (See Plate V, p. 109)

### The Method of Superposition




Let us look at the proof of (I.4), the side–angle–side criterion for congruence of two triangles (SAS for short). Suppose that $AB = DE$, and $AC = DF$, and the included angle $\angle BAC$ equals $\angle EDF$. We wish to conclude that the triangles are congruent, that is to say, the remaining sides and pairs of angles are congruent to each other, respectively. Euclid's method is to "apply the triangle" $ABC$ to the triangle $DEF$. That is, he imagines moving the triangle $ABC$ onto the triangle $DEF$, so that the point $A$ lands on the point $D$, and the side $AB$ lands on the side $DE$. Then he goes on to argue that the ray $AC$ must land on the ray $DF$, because the angles are equal, and hence $C$ must land on $F$ because the sides are equal. From here he concludes that the triangles coincide entirely, hence are congruent.

eſſe. Nam quum de re aliqua ſermonem inſtituimus : ea nobis tacitè per definitionem ſubit in animum : Non enim duos angulos æquales eſſe cogitabo , niſi quid ſit æquales eſſe angulos concipiam. Quod reſpiciens Euclides,angulorum æqualitatem proponere , atque eadem opera definire voluit : vt hoc Theorema pro Definitione haberemus. Nemo enim ſignificantius explicabit angulorum æqualitatem, quàm ſi dixerit duos angulos æquales fieri , quum duo latera vnum angulum continentia, duobus alterum angulum continentibus fiunt æqualia , & baſes quæ latera connectunt , æquales. Conſtat enim angulum tantum eſſe , quanta eſt duarum linearum ipſum continentium apertio, ſeu diductio , hanc verò tantam eſſe , quanta eſt baſis, hoc eſt , linea ipſas connectens. Atque vt clarè dicam , tantus eſt angulus B A C, quanta eſt remotio lineæ A C ab ipſa A B : tanta verò efficitur remotio, quantam exhibet linea B C. Hoc autem in Iſoſcelibus eſt euidentius. Sint enim duo Iſoſcelia A B C & D E F : quorum vnius duo latera A B & A C duobus D E & D F alterius ſint ęqualia:anguluſq; A angulo D. Ac poſitis centris in A & D punctis , ducantur duo Circuli:prior ſecundum A B , alter ſecundum D E ſpatium. Horum prior manifeſtò tranſibit per B & C:alter verò per E & F puncta:quum A B & A C, itemq; D E & E F ſint æqualia, & à centro vtrinq; exeuntia. Atque, ex definitione æqualium angulorum , erunt arcus B C & E F æquales. Angulorum enim magnitudo deſignatur ex arcubus Circulorum qui per extremas lineas quæ angulos continēt,tranſeunt. Ac conuerſo modo,æquales anguli atque æqualibus lineis comprehenſi, æquales ſubtendunt peripherias. Quum enim æqualia ſint ſpatia B C & E F , ea æqualibus rectis lineis claudi oportet : propterea quòd recta linea,eſt à puncto ad punctum via breuiſſima. Atque haud diſſimili iudicio , ex laterum ratione & baſium , quanta ſit angulorum magnitudo æſtimabimus. Quur ergo Euclides hoc inter Theoremata repoſuit,non inter Principia præmiſit ? Nimirùm,quum ſpeciem quodammodo mixtam Principij & Theorematis præ ſe ferret : Principij , quòd in communi animi iudicio conſiſteret:Theorematis,quòd ſpeciatim Triangula Triangulis comparanda proponeret : maluit Euclides inter Theoremata reſerre : præſertim quum multa haberet capita, Principium verò ſimplex ac velut nudum eſſe debeat. Ex hoc prætereà Axiomate tanquam ex locupletiſsimo Demonſtrationum themate , multæ Propoſitiones conſequi debebant , eiuſdem propè facilitatis & iudicij : quas, quia erant notiſſimæ,inter Principia annumerari non conueniebat. Paucis enim Principijs Geometriam contentam eſſe oportebat : immò multa Principia conſultò ſupprimuntur,ne ſit oneroſa multitudo:vt etiam quæ exprimuntur,tantùm ad exemplum exprimi videantur. Hûc accedit,quòd primum Theorema facile,perſpicuum, ac ſenſui obuium eſſe debebat , pro Geometriæ lege, quæ ex paruis humilibuſq; initijs, in progreſſus mirabiles ſeſe extollit.




Huius itaque Propoſitionis veritatem non aliunde quàm à communi iudicio petemus : cogitabimuſq; Figuras Figuris ſuperponere , Mechanicum quippiam eſſe: intelligere verò , id demùm eſſe Mathematicum. Iam verò quum fuerit confeſſum duo Triangula inuicem eſſe æquilatera , ipſa quoque inter ſe æqualia fateri erit neceſſarium. Etenim nulla euidentiori ſpecie æqualitas Figurarum dignoſcitur , quàm ex laterum æqualitate : quanquam Circulorum æqualitas ex diametris definitur: ſed non aliam ob cauſam,quàm quòd linea obliqua ſui copiam adeò apertè non facit vt recta : Cuius menſuram facilè capimus, ac per eam, obliquarum inter ſe comparationem facimus.

At ſi hæc ſuperpoſitio aliqua ratione admittenda ſit : tolerabilior ſane fuerit hoc qui ſequitur modo.

Manente duorum Triangulorum A B C & D E F conditione,continuabo E D

Plate II. The commentary on (I.4) from Peletier's Euclid of (1557). He says the truth of this proposition belongs among the common notions, because to superimpose one figure on another is mechanics, not mathematics.

This is another situation where Euclid is using a method that is not explicitly allowed by his axioms. Nothing in the Postulates or Common Notions says that we may pick up a figure and move it to another position. We call this the *method of superposition*.

Euclid uses this method again in the proof of (I.8), but it appears that he was reluctant to use it more widely, because it does not appear elsewhere. If it were a generally accepted method, for example, then Postulate 4, that all right angles are equal to each other, would be unnecessary, because that would follow easily from superposition.

If we think about the implications of this method, it has far-reaching consequences. It implies that one can move figures from one part of the plane to another without changing their sides or angles. Thus it implies a certain homogeneity of the geometry: The local behavior of figures in one part of the plane is the same as in another part of the plane. If you think of modern theories of cosmology, where the curvature of space changes depending on the presence of large gravitational masses, this is a nontrivial assumption about our geometry.

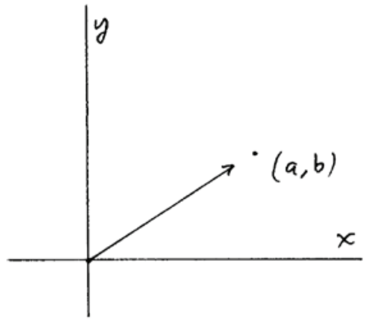
To state more precisely what assumptions the method of superposition is based on, let us define a *rigid motion* of the plane to be a one-to-one transformation of the points of the plane to itself that preserves straight lines and such that segments and angles are carried into congruent segments and angles. To carry out the method of superposition, we need to assume that there exist sufficiently many rigid motions of our plane that

(a) we can take any point to any other point,
(b) we can rotate around any given point, so that one ray at that point is taken to any other ray at that point, and
(c) we can reflect in any line so as to interchange points on opposite sides of the line.

If we were working in the real Cartesian plane $\mathbb{R}^2$ with coordinates $x$, $y$, we could easily show the existence of sufficient rigid motions by using *translations*, *rotations*, and *reflections* defined by suitable formulas in the coordinates.

For example, a translation taking the point $(0, 0)$ to $(a, b)$ is given by

$$\begin{cases} x' = x + a, \\ y' = y + b, \end{cases}$$




and a rotation of angle $\alpha$ around the origin is given by

$$\begin{cases} x' = x \cos\alpha - y \sin\alpha, \\ y' = x \sin\alpha + y \cos\alpha. \end{cases}$$

Thus we can easily justify the use of the method of superposition in the real Cartesian plane. However, since there are no coordinates and no real numbers in Euclid's geometry, we must regard his use of the method of superposition as an additional unstated postulate or axiom.

To formalize this, we could postulate the existence of a group of rigid motions acting on the plane and satisfying the conditions (a), (b), (c) mentioned above. Indeed, there is an extensive modern school of thought, exemplified by Felix Klein's *Erlanger Programm* in the late nineteenth century, which bases the study of geometry on the groups of transformations that are allowed to act on the geometry. This point of view has had wide-ranging applications in differential geometry and in the theory of relativity, for example.

We will discuss the rigid motions in Euclidean geometry in greater detail later (Section 17). For the moment let us just note that the proof of the (SAS) criterion for congruence in (I.4) requires something more than what is in Euclid's axiom system. Hilbert's axioms for geometry actually take (SAS) as an axiom in itself. This seems more in keeping with the elementary nature of Euclid's geometry than postulating the existence of a large group of rigid motions.

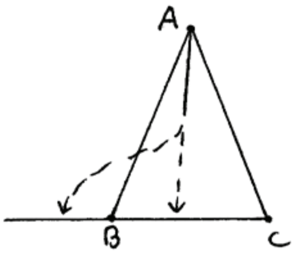
Finally let us note that Euclid's use of the method of superposition in the proof of (I.4) gives us some more insight into his concepts of "equality" for line segments and angles. In Common Notion 4 he says that things that coincide with one another are equal (congruent) to one another. In the proof of (I.4) he also uses the converse, namely, if things (line segments or angles) are equal to one another (congruent), then they will coincide when one is moved so as to be superimposed on the other. So it appears that Euclid thought of line segments or angles being congruent if and only if they could be moved in position so as to coincide with each other.

### Betweenness

Questions of betweenness, when one point is between two others on a line, or when a line through a point lies inside an angle at that point, play an important, if unarticulated, role in Euclid's *Elements*. To explain the notion of points on a line lying between each other, one could simply postulate the existence of a linear ordering of the points. Similarly, for angles at a point one could talk of a circular ordering.

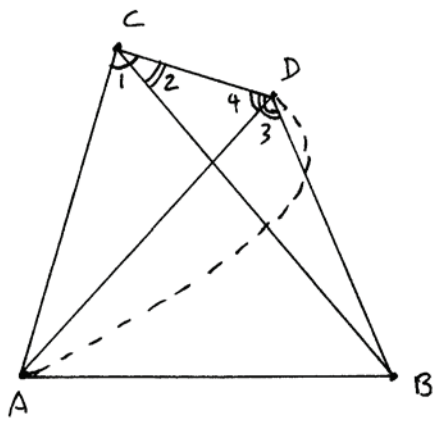
But when a hypothesis of relative position of points and lines in one part of a diagram implies a relationship for other parts of the figure far away, it seems clear that something important is happening, and it may be dangerous to rely on intuition.

For example, how do you know that the angle bisector at a vertex $A$ of a triangle $ABC$ meets the opposite side $BC$ between the points $BC$ and not outside? Of course, it is obvious from the picture, but what if you had to explain why without drawing a picture?




We have already seen that the relative position of two circles may affect whether they meet or not. Let us look at some other instances where betweenness plays an important role in a proof.

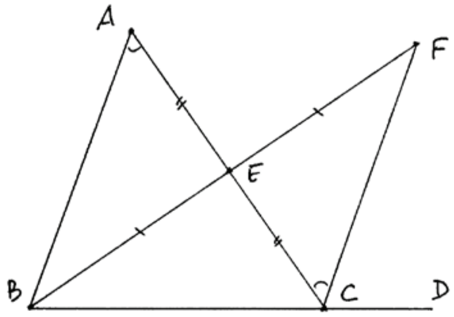
Consider (I.7), which is used in the proof of the side–side–side (SSS) criterion for congruence of triangles (I.8). In (I.7) Euclid shows that it is not possible to have two distinct triangles $ABC$ and $ABD$ on the same side of a segment $AB$ and having equal sides $AC = AD$ and $BC = BD$.




The proof goes like this. Since $AC = AD$, the triangle $ACD$ is isosceles, and so the base angles are equal (I.5). In the diagram $\angle 1 = \angle 4$. On the other hand, since $BC = BD$, the triangle $BCD$ is isosceles, so its base angles are equal (I.5)—in our diagram $\angle 2 = \angle 3$. But now $\angle 2$ is less than $\angle 1$, which is equal to $\angle 4$, which is less than $\angle 3$. So $\angle 2$ is much less than $\angle 3$. But they are also equal, and this is impossible.

Note that this proof depends in an essential way on the relative position of the lines meeting at $C$ and $D$, which determines the inequalities between the angles. If the line $AD$ should reach the point $D$ outside of the triangle $BCD$, as in our second (impossible) picture, then $\angle 2 < \angle 1$ and $\angle 3 < \angle 4$, and there is no contradiction. Thus the original proof depends on a certain configuration of lines being inside certain angles, which in turn depends on some global properties of the entire two-dimensional figure, and these relationships would be hard to explain convincingly without using a diagram. So as soon as we realize that we are depending on a diagram for part of our proof, a mental red flag should pop up to alert us to the question, What exactly is going on here, and what unstated assumptions are we using?

For another example where similar questions arise, look at the proof of (I.16) to show that an exterior angle of a triangle is greater than the opposite interior angle.

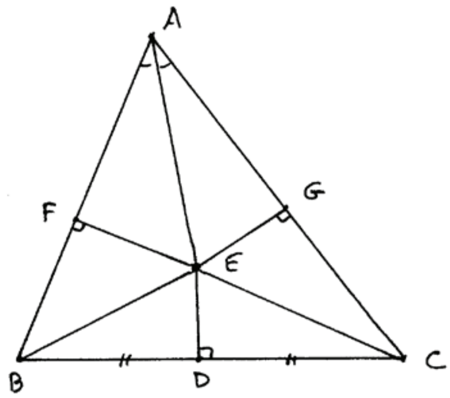



Let $ABC$ be the given triangle. Bisect $AC$ at $E$, draw $BE$, and extend that line to $F$ so that $BE = EF$. Draw $CF$. Then by SAS (I.4), Euclid shows that the triangle $BEA$ is congruent to the triangle $FEC$, and so the angle at $A$ is equal to the angle $\angle ACF$. He then says that the angle $\angle ACF$ is less than the exterior angle $\angle ACD$, which proves the result.

How do we know this relation among the angles? Because the line $CF$ lies inside the angle $ACD$. But why is it inside? Since the line $CF$ was constructed using the point $F$, which in turn was constructed using the point $E$, this is a global property of the whole figure, which is clear from the diagram, but would be hard to explain without a diagram.

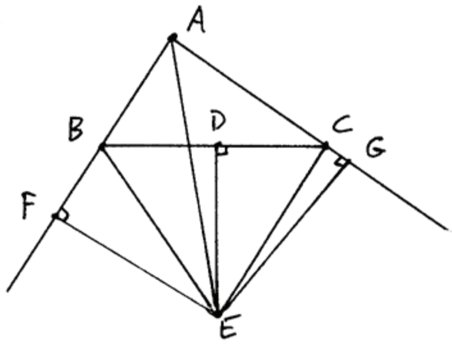
To illustrate the danger of relying on diagrams in geometrical proofs, we will present a well-known fallacy due to W.W. Rouse Ball (1940). The following purports to be a proof that every triangle is isosceles. See if you can find the flaw in the argument.

**Example 3.1**
Let $ABC$ be any triangle. Let $D$ be the midpoint of $BC$. Let the perpendicular to $BC$ at $D$ meet the angle bisector at $A$ at the point $E$. Drop perpendiculars $EF$ and $EG$ to the sides of the triangle, and draw $BE$, $CE$. The triangles $AEF$ and $AEG$ have the side common and two angles equal, so they are congruent by AAS (I.26). Hence $AF = AG$ and $EF = EG$. The triangles $BDE$ and $CDE$ have $DE$ common, two other sides equal, and the included right angles equal. Hence they are congruent by SAS (I.4). In particular $BE = CE$.




Now, the triangles $BEF$ and $CEG$ are right triangles with two sides equal, so they are congruent (see lemma below), and hence $BF = CG$. Adding equals to equals, we find $AB = AF + FB$ is equal to $AC = AG + GC$. So the triangle $ABC$ is isosceles.

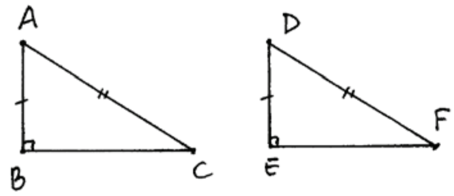
There are several other cases to consider. If the point $E$ lies outside the triangle, one can use this second figure and exactly the same proof to conclude that $AB$ and $AC$ are the differences of equal segments $AF = AG$ and $BF = CG$, hence equal.




If $E$ lands at the point $D$, or if the angle bisector at $A$ is parallel to the perpendicular to $AB$ at $D$, the proof becomes even easier, and we leave it to the reader.

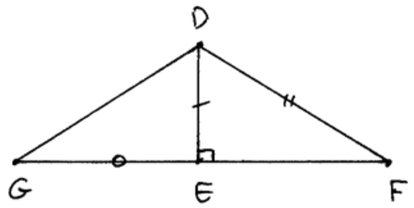
We still need to prove the following lemma.

**Lemma 3.2** (Right-Angle–Side–Side) (RASS)
*If two right triangles have two sides equal, not containing the right angle, they are still congruent.*




*Proof* This result, though not stated by Euclid, is often useful. We give two proofs. The first method is to use (I.47) to conclude that the square on $BC$ is equal to the square on $EF$. Then $BC = EF$, and we can apply (SSS) (I.8).

The second proof does not make use of (I.47) and the theory of area. Extend $FE$ to $G$ and make $EG = BC$. Then the triangles $ABC$ and $DEG$ are congruent by SAS (I.4). Therefore, $AC = DG$. It follows that $DF = DG$, so the triangle $DFG$ is isosceles. Therefore, the angles at $F$ and $G$ are equal. Then the triangles $DEG$ and $DEF$ are congruent by AAS (I.26). But $DEG$ is congruent to $ABC$, so the two original triangles are congruent.

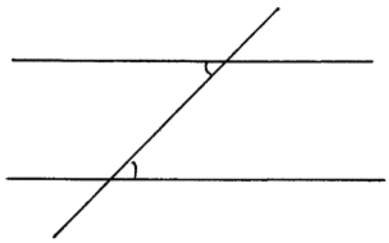



**The Theory of Parallels**

Book I of Euclid's *Elements* can be divided naturally into three parts. The first part, (I.1)–(I.26), deals with triangles and congruence. The second part, (I.27)–

(I.34), deals with parallel lines and their applications, including the well-known (I.32) that the sum of the angles of a triangle is two right angles. The third part, (I.35)–(I.48), deals with the theory of area.

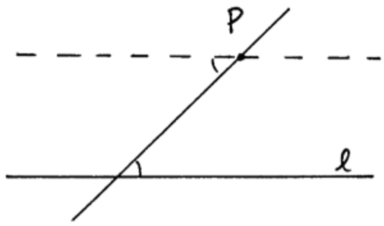
Two lines are *parallel* if they never meet, even if extended indefinitely in both directions (Definition 23). The fifth postulate gives a criterion for two lines to meet under certain conditions, hence to be not parallel, so we often refer to the fifth postulate as the *parallel postulate*. Euclid postponed using this postulate as long as possible so that in fact, the first part of Book I about triangles and congruence does not use the parallel postulate at all. It is first used in (I.29). Let us examine closely Euclid's theory of parallels and his use of the parallel postulate.



The first result about parallel lines, (I.27), says that if a line falling on two other lines makes the alternate interior angles equal, then the lines are parallel. This is proved using (I.16): If not, the lines would meet on one side or the other, and would form a triangle having an exterior angle equal to one of its opposite interior angles, which is impossible.

The next result (I.28) is similar, and follows directly from this one using vertical angles (I.15) or supplementary angles (I.13).

The fifth postulate is used to prove the converse of (I.27), which is (I.29): If the lines are parallel, then the alternate interior angles will be equal. For if not, then one would be greater than the other, and so the sum of the interior angles on one side of the transversal would be less than two right angles. In this situation, the fifth postulate applies and forces the lines to meet, which is a contradiction.




As for the existence of parallel lines, Euclid gives a construction in (I.31) for a line through a point $P$, parallel to a given line $l$. Draw any line through $P$, meeting $l$, and then reproduce the angle it makes with $l$ at the point $P$ (I.23). It follows from (I.27) that this line is parallel to $l$.

Why does Euclid place this construction after (I.29), even though it does not depend on (I.29) and does not make use of the parallel postulate? Presumably, the answer, although Euclid does not say so, is that using (I.29) one can show that this parallel just constructed is *unique*. If there were any other line parallel to $l$ through $P$, it would make the same angle with the transversal (by (I.29)) and hence would be equal to this one. Thus using the parallel postulate we can prove the following statement:

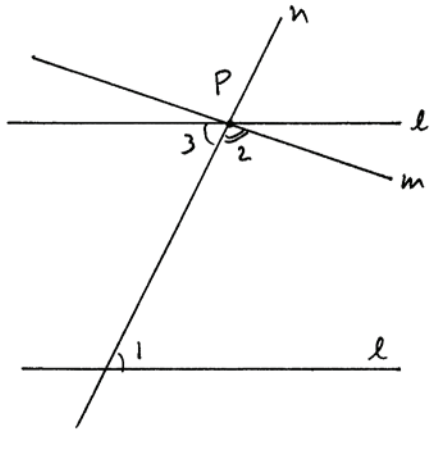
**P.** For each point $P$ and each line $l$, there exists at most one line through $P$ parallel to $l$.

This statement (P) is often called "Playfair's axiom," after John Playfair (1748–1819), even though it already appears in the commentary of Proclus. Of course, in Euclid's development of geometry, this is not an axiom, but a theorem that can be proved from the axioms. Some authors, however, like to take the statement (P) as an axiom instead of using Euclid's fifth postulate. So I would like to explain in what sense we can say that Euclid's fifth postulate is equivalent to Playfair's axiom (P).

Since the parallel postulate plays such a special role in Euclid's geometry, let us make a special point of being aware when we use this postulate, and which theorems are dependent on its use. Let us call *neutral geometry* the collection of all the postulates and common notions *except* the fifth postulate together with all theorems that can be proved without using the fifth postulate. Thus (I.1)–(I.28) and (I.31) all belong to neutral geometry, while for example, (I.32) and (I.47) do not belong to neutral geometry.

If we take neutral geometry and add back the fifth postulate, then we recover ordinary Euclidean geometry, and we can prove (P) as a theorem as we did above.

But now suppose we take neutral geometry and add (P) as an extra axiom. We will show that in this geometry we can prove Euclid's fifth postulate as a theorem.



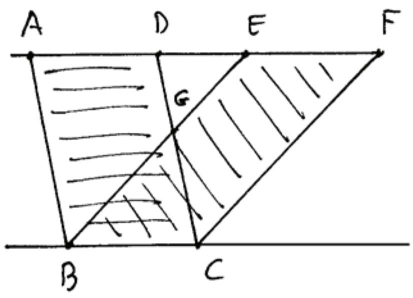

Indeed, suppose we are given two lines $l$, $m$ and a transversal $n$ such that the two interior angles 1, 2 on the same side are less than two right angles. Let $P$ be the intersection of the lines $m$ and $n$, and draw a line $l'$ through $P$, making the alternate angle 3 equal to 1. This is possible by (I.23), which belongs to neutral geometry. Then by (I.27), which also belongs to neutral geometry, $l'$ is parallel to $l$.

Now, since $1+2$ is less than two right angles, it follows that $2+3$ is less than two right angles, and hence the line $l'$ is different from $m$ (I.13). Now we can apply (P). Since $l'$ passes through $P$

and is parallel to $l$, it must be the only line through $P$ that is parallel to $l$. In particular, the line $m$, which is different from $l'$, cannot be parallel to $l$, and so by definition it must meet $l$. This proves the fifth postulate.

Thus in the presence of all the results of neutral geometry, we can use Euclid's fifth postulate to prove Playfair's axiom, or we can use Playfair's axiom to prove Euclid's fifth postulate. In this sense we can say that in neutral geometry, Euclid's fifth postulate is equivalent to Playfair's axiom. This means that adding either one of them as an additional assumption to neutral geometry will give the same body of theorems as consequences.

### The Theory of Area




In (I.35), Euclid says that two parallelograms on the same base and in the same parallels (this means their top sides lie on the same line parallel to the base) are equal to each other. In the figure, the parallelogram *ABCD* is equal to the parallelogram *BCEF*. Clearly, the parallelograms are not congruent.

Looking at the proof, which is accomplished by adding and subtracting congruent figures, we conclude that Euclid must be referring to the area of the parallelograms when he says they are equal. But he has not said what the area of a figure is, so we must reflect a bit to see what he means.

Our intuitive understanding of area comes from high-school geometry, where we learn that the area of a rectangle is the product of the lengths of two perpendicular sides, the area of a triangle is one half the product of the lengths of the base and the altitude, etc. The "area" of high-school geometry is a function that attaches to each plane figure a real number; the area of a nonoverlapping union of figures is the sum of the areas, and so forth. Most likely no one ever told you the definition of area, nor did they prove that such an area function exists. Using calculus, you can define the area of a figure in the real Cartesian plane using definite integrals, and in that way it is possible to prove that a suitable area function exists. But in Euclid's geometry there are no real numbers, and we certainly do not want to use calculus to define the concept of area in elementary geometry.

So what did Euclid have in mind? Since he does not define it, we will consider this new equality as an undefined notion, just as the notions of congruence for line segments and angles were undefined. We will call this new notion *equal content*, to avoid confusion with other notions of equality or congruence. We do not want to use the word area, because this notion is quite different from our common understanding of area as a function associating a real number to each figure.

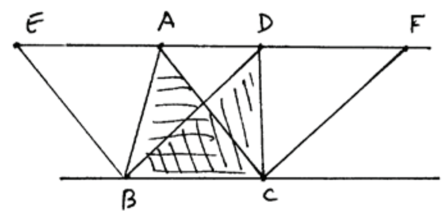
From the way Euclid treats this notion, it is clear that he regards it as an equivalence relation, satisfying the common notions. In particular:

(a) Congruent figures have equal content.
(b) If two figures each have equal content with a third, they have equal content.
(c) If pairs of figures with equal content are *added* in the sense of being joined without overlap to make bigger figures, then these added figures have equal content.
(d) Ditto for subtraction, noting that equality of content of the difference does not depend on *where* the equal pieces were removed.
(e) Halves of figures of equal content have equal content (used in the proof of (I.37)). (Also, doubles of equals are equal, as a consequence of (c) above.)
(f) The whole is greater than the part, which in this case means that if one figure is properly contained in another, then the two figures cannot have equal content (used in the proof of (I.39)).

In terms of the axiomatic development of the subject, at this point Euclid is introducing a new undefined relation, and taking all the properties just listed as new axioms governing this new relation. Later in this book (Section 22), we will discuss Hilbert's reinterpretation of the theory of area where the relationship of having equal content is defined, and all its properties proved, so that it does not require the introduction of new axioms.

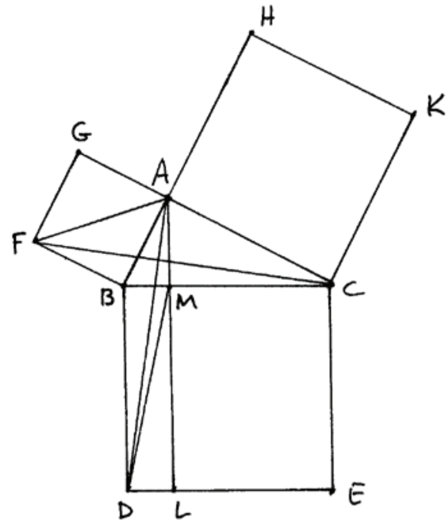
Now let us see what Euclid does with this purely geometric notion of equal content of plane figures. In (I.35) he proves that the two parallelograms have equal content (see diagram above) by first showing that the triangle *ABE* is congruent to the triangle *DCF*, so they have equal content. Then by subtracting the triangle *DGE* from each (in different positions!) and adding the triangle *BGC* to each, he obtains the two parallelograms, which therefore have equal content.

In (I.37) he shows that two triangles *ABC* and *DBC* on the same base and in the same parallels have equal content. The method is to double *ABC* to get a parallelogram *EABC*, and to double *DBC* to get a parallelogram *DFBC*.




By (I.35) the two parallelograms have equal content, and then he applies the axiom that halves of equals are equal to conclude the triangles have equal content.

This is all that is needed to explain Euclid's beautiful proof of (I.47), the theorem of Pythagoras. The statement of the theorem is that if $ABC$ is a right triangle, then the squares on the two legs together have equal content to the square on the hypotenuse. The proof goes like this. The triangle $ABF$ is one half of the square $ABFG$. This triangle $ABF$ has equal content with the triangle $BFC$ by (I.37). The triangle $BFC$ is congruent to the triangle $BAD$. And $BAD$ has equal content to the triangle $BMD$ by (I.37). This latter triangle is equal to one-half of the rectangle $BDLM$. Hence the square $ABFG$ has equal content to the rectangle $BDLM$. Doing the same construction on the other side and adding, one has the result.




Euclid's statement of (I.47) in terms of equal content of the squares constructed on the sides of the triangle may come as a surprise to the modern student who remembers the formula $a^2 + b^2 = c^2$ (which I suppose in the minds of the general public is rivaled in fame only by Einstein's famous formula $E = mc^2$). We are used to thinking of $a$, $b$, $c$ as the lengths of the sides of the triangle, in which case the theorem becomes an equation among real numbers. How can we reconcile these two points of view?

The modern answer to this question, which we will discuss in more detail later (Section 23), is that after introducing coordinates in our geometry we can prove the existence of an area function. The area of a square of side $a$ will be $a^2$. Furthermore, we will show that having equal content in the sense of Euclid is equivalent to having equal area in the sense of the area function. Then the two formulations of the theorem of Pythagoras become equivalent.

This answer makes sense only when we are able to assign numerical lengths to arbitrary line segments, which the Greeks could not do. Yet there is ample evidence that the Greeks did know special cases of this formula when $a$, $b$, $c$ are integers. The equation $3^2 + 4^2 = 5^2$ was known to the Egyptians, and Proclus in his note on (I.47) mentions two general formulas for generating such "Pythagorean triples" of integers, which he ascribes to Plato and to Pythagoras. So we can presume that the Greeks knew some particular right triangles with integer sides, in which case (I.47) can be represented by the equation among integers $a^2 + b^2 = c^2$. But the geometrical proof given by Euclid is then more general, because it applies to all triangles, and not just those for which one can find integers to fit the sides.