1.3 A Critique of Euclid–Modern Axiomatics

At the end of the nineteenth century, Euclid's work received scrutiny far surpassing the preceding 2000 years' efforts. This examination revealed how much more precise and explicit mathematics has become, without calling into question any of Euclid's results. For example, consider Euclid's proof of his first proposition, given in shortened form (from Heath [13, volume I, 241]. Can you find the logical gap in his proof, given his postulates and common notions? (Figure 1.26.)

Proposition I-1 On a given finite straight line to construct an equilateral triangle.

Proof. Let $\overline{AB}$ be the given finite straight line. . . . With center A and radius $\overline{AB}$ let the circle BCD be described; again, with center B and radius $\overline{BA}$ let the circle ACE be described (postulate 3, in Appendix A). From point C, in which the circles cut each other, to points A and B, join the straight lines $\overline{CA}$ and $\overline{CB}$ (postulate 1, in Appendix A). Now, point A is the center of the circle CDB, so AC is equal to AB (definition 15, in Appendix A). Again, . . . BA is equal to BC. . . . And things equal to the same thing are also equal to one another . . . (common notion 1, in Appendix A). Therefore . . . AC, AB, BC are equal to one another. Therefore $\triangle ABC$ is equilateral. ■

Euclid's construction is straightforward, and he followed it with a proof that the three sides are congruent. However, Euclid never showed, nor could he show from his assumptions, that the circles must intersect. This logical gap indicates one of the many implicit assumptions Euclid made that were very hard to detect because they were "obvious." (We made the same assumption in Example 3 of Section 1.2.) In modern terms Euclid assumed that lines, circles, and other figures are continuous. In the fifth postulate and numerous other places Euclid assumes an order to points on lines with no justification or discussion. (In Example 3 of Section 1.2 we assumed such ordering to

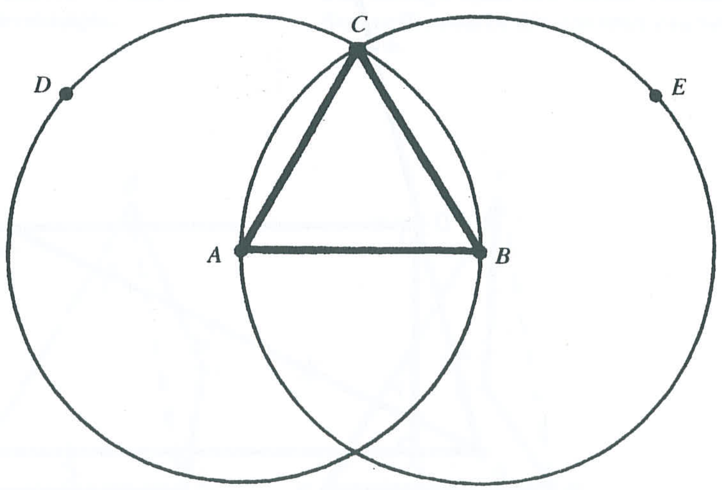

Figure 1.26

talk about a point "outside" a circle.) David Hilbert and others recognized that rigorous proofs require explicit axioms about continuity, order, and other assumptions.

1.3.1 Hilbert's axioms

Hilbert's axiomatic system, first presented in 1898 [15], corrects the logical deficiencies found in Euclid's work. A careful comparison of Euclid's definitions and assumptions (Appendix A) with Hilbert's system (Appendix B) shows some of the differences between a modern axiomatic system and the system Euclid developed. In particular, Hilbert's axioms provide enough explicit properties to prove all the theorems of Euclidean geometry without any logical gaps. Also, Hilbert uses undefined terms, whereas Euclid starts his text with a list of definitions, some having little mathematical value.

The first three groups of Hilbert's axioms codify elementary properties, many previously overlooked. Axioms I-1 and I-2 modernize Euclid's first two postulates. Axiom I-3 guarantees that there are some points in the geometry. The first three axioms of order give the most elementary properties of betweenness on a line. Pasch's axiom (II-4), which says that a line that *enters* a triangle must *exit* as well, deserves more explanation. To help explain Pasch's axiom we derive it from the separation axiom, an alternative (II-4′) that Hilbert gives.

Theorem 1.3.1 The separation axiom (II-4′) implies Pasch's axiom (II-4).

Exercise 1 Illustrate the proof of Theorem 1.3.1.

Proof. Suppose that the separation axiom holds: A line m separates the points that are not on m into two sets such that if X and Y are in the same set, $\overline{XY}$ does not intersect m, and if X and Y are in different sets, $\overline{XY}$ does intersect m. To show Pasch's axiom, suppose that a line l enters $\triangle ABC$ at D on side $\overline{AB}$ but that A, B, and C are not on l. By the separation axiom, A and B are on opposite sides of l; that is, they are in different sets of points. Now C is in just one of those two sets, which means only one of $\overline{AC}$ and $\overline{BC}$ have a point on l, where l exits from $\triangle ABC$, showing Pasch's axiom. ■

The congruence axioms III-2 and III-3 make precise Euclid's first two common notions. Axioms III-1 and III-4 guarantee the existence and uniqueness of congruent segments and angles. These axioms replaced Euclid's use of circles and constructions. Axiom III-5 is Euclid's Proposition I-4, the familiar SAS property of high school geometry. Euclid's proof has a logical gap because he used motion without providing any axioms about movement.

Hilbert's fourth group contains only Playfair's axiom, which by Theorem 1.2.1 is logically equivalent to Euclid's fifth postulate. This axiom distinguishes Euclidean geometry from hyperbolic geometry, which we introduce in Chapter 3.

Although the earlier axioms ensure lines have infinitely many points, group V implies that the points on a line correspond to the real numbers. Axiom V-1, the Archimedean axiom, eliminates infinitely large and infinitely small line segments. [Archimedes (circa 287–212 B.C.) used this idea in some of his proofs.] Hilbert's complicated axiom of linear completeness, V-2, ensures that lines have no gaps while avoiding concepts from analysis.

DAVID HILBERT

But in the present century, thanks in good part to the influence of Hilbert, we have come to see that the unproved postulates with which we start are purely arbitrary. They MUST be consistent; they HAD BETTER lead to something interesting.
—*Coolidge*

The work of David Hilbert (1862–1943) symbolizes the modern abstract, axiomatic approach to mathematics. He contributed significant results in many fields, including algebraic invariants, number theory, partial and ordinary differential equations, integral equations, geometry, and the foundations of mathematics. He won international renown at age 26 when he published a result in algebraic invariants, a precursor of modern abstract algebra, that had seemed beyond possibility. His proof neatly avoided the laborious and limited constructive methods of previous mathematicians. One of these mathematicians, Paul Gordan, at first disdained Hilbert's radical nonconstructive proof saying, "This is not mathematics; it is theology." He later added, "I have convinced myself that theology also has its advantages." Hilbert's work in integral equations led to Hilbert space, an infinite dimensional analog to Euclidean space important in the study of quantum mechanics in physics.

Hilbert was very influential in the foundations of mathematics. Developments in nineteenth century analysis and geometry made clear the need for a careful scrutiny of axiomatics. To make hidden assumptions explicit, Hilbert realized that mathematicians must isolate formal axioms from their meaning. He devised ground-breaking axiom systems for both the real numbers and Euclidean geometry. He proved the relative consistency of geometry, assuming the consistency of the real numbers. Hilbert's program, which he hoped would prove the absolute consistency of all of mathematics, led others to a penetrating analysis of logic and the foundations of mathematics. Gödel's famous incompleteness theorems came out of that program and proved, among other things, the impossibility of Hilbert's goal. However, Gödel's theorems did show how deeply Hilbert's axiomatic approach enabled mathematicians to probe the foundations of mathematics.

Hilbert's formal axiomatics devoid of meaning was never intended to replace mathematics. Throughout his career, Hilbert stressed the dynamic interplay of concrete problems and general, abstract theories. His famous list of 23 problems in 1900 illustrates well this link. Each problem has led mathematicians to a more profound understanding of an area of mathematics. His own research revealed the power of particular problems to inspire deep mathematics and of abstract mathematics to elucidate particular problems.

1.3.2 Axiomatic systems

An axiomatic system provides an explicit foundation for a mathematical subject. Axiomatic systems include seven parts: the logical language, rules of proof, undefined terms, axioms, definitions, theorems, and proofs of theorems. We discuss the last five parts—the ones with geometric content.

Consider Euclid's definition of a point as "that which has no part." This definition is more a philosophical statement about the nature of a point than a way to prove statements. Euclid's definition of a straight line, "a line which lies evenly with the points on itself," is unclear as well as not useful. In essence, points and lines were so basic to Euclid's work that there is no good way to define them. Mathematicians realized centuries ago the need for undefined terms in order to establish an unambiguous beginning. Otherwise, each term would have to be defined with other terms, leading either to a cycle of terms or an infinite sequence of terms. Neither of these options is acceptable for carefully reasoned mathematics. Of course, we then define all other terms from these initial, undefined terms. However, undefined terms are, by their nature, unrestricted. How can we be sure that two people *mean* the same thing when they use undefined terms? In short, we can't. The axioms of a mathematical system become the key: They tell us how the undefined terms *behave.* Axioms are sometimes called *operational definitions* because they describe how to use terms and how they relate to one another, rather than telling us what terms "really mean." Indeed, mathematicians permit any interpretation of undefined terms, as long as all the axioms hold in that interpretation. In Section 1.4 we explore the interplay between axioms and their interpretations in models.

Unlike the Greek understanding of axioms as self-evident truths, we do not claim the truth of axioms. However, this does not mean that we consider axioms to be false. Rather, we are free to choose axioms to formulate the fundamental relationships we want to investigate. From a logical point of view, the choice of axioms is arbitrary; in actuality, though, mathematicians carefully pick axioms to focus on particular features. For example, in perspective drawing parallel lines intersect at a point on the horizon. Projective geometry is an axiomatic system in which any two lines intersect in a point. This system, discussed in Chapter 6, enables us to understand many consequences stemming from perspective. However, we don't need to decide whether "in truth" there are parallel lines or if all lines intersect. Indeed, in the concrete world of atoms and energy, there are no mathematical lines at all. Nevertheless, these axiomatic systems and many others have given us a profound understanding of the world. Axiomatic systems allow us to formulate and logically explore abstract relationships, freed from the specificity and imprecision of real situations.

Mathematicians build two basic types of axiomatic systems. One completely characterizes a particular mathematical system. For example, Hilbert's axioms characterize Euclidean geometry completely.[1] The second focuses on the common features of a family of structures, such as vector spaces. Although infinitely many different vector spaces exist, all satisfy certain essential properties. The general study of vector spaces greatly aids the development of theories in economics, physics, mathematics, and other fields. Such axiomatic systems unite a wide variety of examples within one powerful theoretical framework.

Mathematical definitions are built from undefined terms and previously defined terms. For example, Hilbert defines the angle $\angle ABC$ as a point B and two rays $\overrightarrow{BA}$ and $\overrightarrow{BC}$. In turn, the ray $\overrightarrow{BA}$ is the set of points X on the line $\overleftrightarrow{AB}$ such that X is between A and B, X is A, or A is between B and X. Thus Hilbert reduces the notion of an angle to the undefined terms *point*, *line*, *on*, and *between*. These same undefined terms are sufficient to define *convex*, an important concept in modern mathematics. The Greeks

[1] This statement may appear to contradict Gödel's incompleteness theorem. However, axiom V-2 is a "second-order axiom," a concept beyond the level of this text. (See Delong [6] for information on second-order logic.)

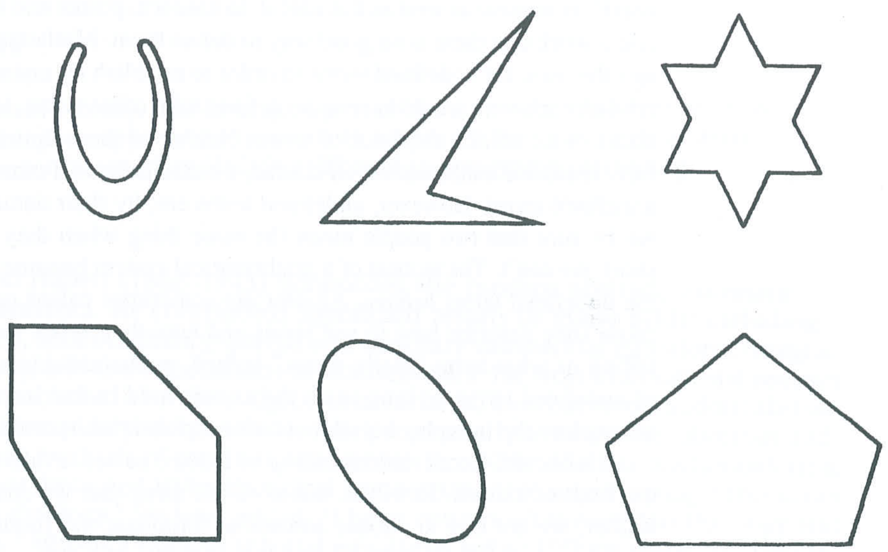

Figure 1.27 Nonconvex and convex sets.

never distinguished between nonconvex sets, which "bend," and convex sets (Fig. 1.27), perhaps because they never considered betweenness. Intuitively, in a convex set, every point can "see" every other point.

Definition 1.3.1 A set S is *convex* iff for distinct points P and Q in S, $\overline{PQ}$ is entirely in S.

Exercise 2 Rewrite the definition of convex, using only the undefined terms *between* and *point*.

Exercise 3 Which of the terms in Euclid's definitions (Appendix A) do you think should be undefined terms? Which of the remaining definitions fit our modern understanding of definitions?

Theorems and their proofs are the most distinctive parts of mathematics, whether in an axiomatic or some other system. In an axiomatic system, a theorem is a statement whose proof depends only on previously proven theorems, the axioms, the definitions, and the rules of logic. This condition ensures that the entire edifice of theorems rests securely on the explicit axioms of the system.

Proofs of theorems in an axiomatic system cannot depend on diagrams, even though diagrams have been part of geometry since the ancient Greeks drew figures in the sand. We need the powerful insight and understanding that such diagrams provide. However, a corresponding risk comes with the use of pictures: We are liable to accept as intuitive a step that does not follow from the given conditions. Euclid's first proof, discussed previously, shows how easy it is to include implicit assumptions. Euclid's minor "sins of omission" never led him to an erroneous result, but the potential remains for a diagram to mislead us, as in Example 1, used with permission from Dubnov [7, 15]. Even though diagrams are not permissible in a proof in an axiomatic system, they certainly can and

should be included to help us understand the ideas. They must be studied critically to ensure that the illustrated relationships are proved or legitimately assumed.

Example 1 *Claim.* A rectangle inscribed in a square is a square.

Verify that the claim is incorrect. Then try to find the error in the "proof."

Proof. Let rectangle $MNPQ$ be inscribed in square $ABCD$ (Fig. 1.28). Drop perpendiculars from P to $\overline{AB}$ and from Q to $\overline{BC}$ at R and S, respectively. Clearly, $\overline{PR} \cong \overline{QS}$ because these segments match the sides of the square $ABCD$. Furthermore, the rectangle's diagonals are congruent: $\overline{PM} \cong \overline{QN}$. So $\triangle PMR \cong \triangle QNS$, and hence $\angle PMR \cong \angle QNS$. Consider the quadrilateral $MBNO$, where O is the intersection of $\overline{QN}$ and $\overline{PM}$. Its exterior angle at vertex N is congruent to the interior angle at vertex M, so the two interior angles at vertices N and M are supplementary. Thus the interior angles at vertices B and O must be supplementary. But $\angle ABC$ is a right angle and hence $\angle NOM$ must also be a right angle. Therefore the diagonals of rectangle $MNPQ$ are perpendicular. Hence $MNPQ$ is a square.

The preceding argument is correct up to the conclusion $\angle PMR \cong \angle QNS$. Then the diagram shown in Fig. 1.28 misleads us to think that $\angle ONB$ is supplementary to $\angle OMR$. However, these angles can be congruent if we switch N and S in the diagram and correspondingly move Q down. Illustrate this second case. ●

How can we possibly make all assumptions explicit and eliminate all risk of incorrect proofs? Mathematical logic, developed by Hilbert and others, involves the use of a formal language so austere that a proof can be checked in a purely mechanical manner, free from human intuition. In principle, a computer could check such a proof to decide its validity. Consider, for example, the statement, "Two distinct points have a unique

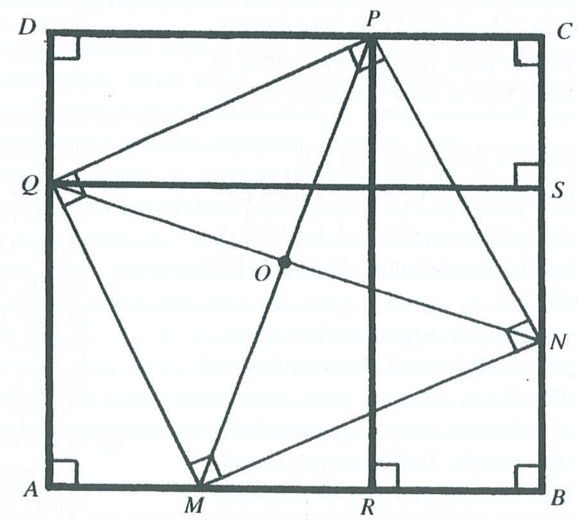

Figure 1.28

line on them." We can make the statement more explicit as follows: "For all points A_1 and A_2 if $A_1 \neq A_2$, then there is a unique line B_1 such that A_1 is on B_1 and A_2 is on B_1." Finally we could turn this more explicit statement into pure symbols:

$$\forall A_1 \forall A_2 (\neg(A_1 = A_2) \Rightarrow (\exists! B_1 : (R_1(A_1, B_1) \wedge R_1(A_2, B_1)))).$$

Clearly, we need to decode these symbols so that people not familiar with them can read the statement. (We don't use these logic symbols, but for your information, $\forall$ $\neg$ $\Rightarrow$ $\exists$! $\wedge$ represent *for all*, *not*, *implies*, *there exists*, *a unique*, *and*, respectively. The variable R_1 represents the relation *is on*.) If the entire axiomatic system, including the rules of proof, is encoded into such a formal language, we can mechanically determine whether a given string of symbols encodes a logical proof of the string of symbols encoding the theorem. There will be no risk of an inappropriate inference, but there will be an incredible barrier to human understanding. Indeed, finding a proof in such a language is a daunting task.

Axiomatic systems are a workable compromise between the austere formal languages of mathematical logic and Euclid's work, with its many implicit assumptions. Mathematicians need both the careful reasoning of proofs and the intuitive understanding of content. Axiomatic systems provide more than a way to give careful proofs. They enable us to understand the relationship of particular concepts, to explore the consequences of assumptions, to contrast different systems, and to unify seemingly disparate situations under one framework. In short, axiomatic systems are one important way in which mathematicians obtain insight. (Heath's edition of Euclid's *Elements* [13] provides detailed commentary on its logical shortcomings. Wilder [29] explores axiomatic systems in more detail.)

PROBLEMS FOR SECTION 1.3

1. a) From Group I of Hilbert's axioms, how many points and lines must exist? Prove your answer.

b) Given two distinct lines prove from Group I of Hilbert's axioms that at most one point lies on both lines.

c) Include axiom IV-1 and repeat part (a).

2. a) Compare Hilbert's axiom III-4 with Euclid I-7.

b) Read in Euclid's *Elements* [13, vol. I, 247ff] the proofs of I-4 and I-7 and discuss the logical gaps in these proofs.

3. Discuss how well Euclid's definition of an angle (number 8) fits your intuition and how easy it would be to apply in a proof.

4. a) Use Hilbert's definition of an angle to define the interior of an angle. Define a triangle and its interior. How does the interior of a triangle relate to the interiors of its angles? Illustrate your definitions.

b) Define a convex set from Hilbert's undefined terms. Are the interior of an angle and of a triangle, as you defined them, convex? Justify your answer.

5. a) List which of Hilbert's axioms can refer to points on just one line.

b) Use the axioms in part (a) that are in groups I and II to find infinitely many points on any line. [*Hint:* Use induction.]

c) Fix P_0 and P_1 on a line m. Use the axioms in part (a) to prove by induction that for all positive integers n there are points P_n on m such that P_n is between P_0 and P_{n+1} and such that $\overline{P_nP_{n+1}} \cong \overline{P_0P_1}$.

d) In the notation of part (c), what properties should P_{-n} satisfy? Prove that there are points that satisfy these properties.

e) Which axiom(s) guarantee(s) that between two points on a line there is a third point?

6. Consider the axiomatic system with the undefined terms *point* and *adjacent* and the following axioms. Use $\bowtie$ to denote "is adjacent to."

i) There is at least one point.

ii) If $P \bowtie Q$, then $Q \bowtie P$ and $P \neq Q$.

iii) Every point has exactly three distinct points adjacent to it.

iv) If $P \bowtie Q$ and $P \bowtie R$, then not $Q \bowtie R$.

a) Prove that there are at least six points.

b) Suppose that you omit axiom (iv). Can you still prove that there are at least six points? If so, prove it; otherwise find the largest number of points that must exist and prove your answer.

c) Suppose that we change axiom (iii) to require exactly four distinct points adjacent to any point but leave the other axioms unchanged. How many points must exist? Prove your answer. Generalize.

7. Consider the axiomatic system with the undefined terms *point*, *line*, and *on* and the following axioms.

i) There are a line and a point not on that line.

ii) Every two distinct points have a unique line on them both.

iii) Every two distinct lines have at least one point on them both.

iv) Every line has at least three points on it.

a) Given two distinct lines prove that they have exactly one point on them.

b) Prove that there are at least seven points.

c) Given any point prove that it has at least three lines on it. [*Hint:* First consider a point not on the line of axiom (i).]

d) Prove that there are at least seven lines.

8. In the axiomatic system with the undefined terms *point* and *between*, use $P(Q)R$ to denote "Q is between P and R." Define a set S to be *convex* iff whenever P and R are in S and $P(Q)R$, then Q is in S. The axioms are:

i) If $P(Q)R$, then $R(Q)P$.

ii) If $P(Q)R$, then not $P(R)Q$ and $P \neq R$.

a) Prove that, if $P(Q)R$, then not $Q(R)P$, not $R(P)Q$, and not $Q(P)R$.

b) Prove that, if $P(Q)R$ then P, Q, and R are three distinct points.

c) Compare the axioms and parts (a) and (b) of this problem with Hilbert's axioms of order II-1, II-2, and II-3.

d) Prove that, if S and T are convex, then $S \cap T$ is also convex.

e) If each S_i is convex, for i in a finite or infinite index set I, prove that $\bigcap_{i \in I} S_i$ is convex, where $\bigcap_{i \in I} S_i = \{P : \text{for all } i \in I,\ P \in S_i\}$.