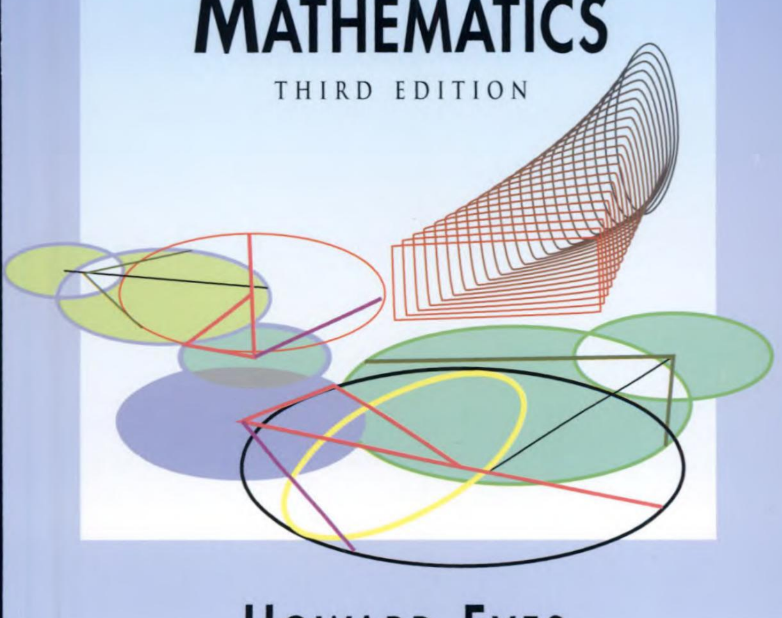
FOUNDATIONS AND FUNDAMENTAL CONCEPTS OF MATHEMATICS

THIRD EDITION



HOWARD EVES

FOUNDATIONS AND FUNDAMENTAL CONCEPTS OF MATHEMATICS

THIRD EDITION

Howard Eves

University of Maine
University of Central Florida

DOVER PUBLICATIONS, INC.
Mineola, New York

Copyright

Bibliographical Note

This Dover edition, first published in 1997, is an unabridged and unaltered republication of the third edition originally published by PWS-Kent Publishing Company, Boston, in 1990.

Library of Congress Cataloging-in-Publication Data

Eves, Howard Whitley, 1911–
Foundations and fundamental concepts of mathematics / Howard Eves. — 3rd ed.
p. cm.
Originally published: 3rd ed. Boston : PWS-Kent, c1990.
Includes bibliographical references and index.
ISBN 0-486-69609-X (pbk.)
1. Mathematics—Philosophy. 2. Mathematics—History. I. Title.
[QA9.E89 1997]
510'.1—dc21 97-8057
CIP

Manufactured in the United States of America
Dover Publications, Inc., 31 East 2nd Street, Mineola, N.Y. 11501

1

MATHEMATICS BEFORE EUCLID

1.1 The Empirical Nature of Pre-Hellenic Mathematics

The thesis can be advanced that mathematics arose from necessity. The annual inundation of the Nile Valley, for example, forced the Egyptians to develop some system for redetermining land markings; in fact, the word *geometry* means "measurement of the earth." The need for mensuration formulas was especially imperative if, as Herodotus remarked, taxes in Egypt were paid on the basis of land area. The Babylonians likewise encountered an urgent need for mathematics in the construction of the great engineering structures for which they were famous. Marsh drainage, irrigation, and flood control made it possible to convert the land along the Tigris and Euphrates rivers into a rich agricultural region. Similar undertakings undoubtedly occurred in early times in south-central Asia along the Indus and Ganges rivers, and in eastern Asia along the Hwang Ho and the Yangtze. The engineering, financing, and administration of such projects required the development of considerable technical knowledge and its attendant mathematics. A useable calendar had to be computed to serve agricultural needs, and this required some basic astronomy with its concomitant mathematics. Again, the demand for some system of uniformity in barter was present in even the earliest civilizations; this fact also furnished a pronounced stimulus to mathematical development. Finally, early religious ritual found need for some basic mathematics.[1]

Thus there is a basis for saying that mathematics, beyond that implied by primitive counting, originated during the period of the fifth, fourth, and third millennia B.C. in certain areas of the ancient orient as a practical science to assist in engineering, agricultural, and business pursuits and in religious ritual. Although the initial emphasis was on mensuration and practical arithmetic, it

[1] See A. Seidenberg [1] and [2]. (References by author's name only are to the Bibliography at the end of the book.)

was natural that a special craft should come into being for the application, instruction, and development of the science and that, in turn, tendencies toward abstraction should then assert themselves and the subject be studied, to some extent, for its own sake. In this way a basis for the beginnings of theoretical geometry grew out of mensuration, and the first traces of elementary algebra evolved from practical arithmetic.[2]

In our study of early mathematics we are restricted essentially to that of Egypt and Babylonia. The ancient Egyptians recorded their work on stone and papyrus, the latter fortunately enduring because of Egypt's unusually dry climate; the Babylonians used imperishable baked clay tablets. In contrast to the use of these media, the early Indians and Chinese used very perishable writing materials like bark and bamboo. Thus it has come to pass that we have a fair quantity of definite information, obtained from primary sources, about the science and the mathematics of ancient Egypt and Babylonia, while we know very little indeed, with any degree of certainty, about these fields of study in ancient India and China.

It is the nature, rather than the content, of this pre-Hellenic mathematics that concerns us here, and in this regard it is important to note that, outside of very simple considerations, the mathematical relations employed by the Egyptians and by the Babylonians resulted essentially from "trial and error" methods. In other words, to a great extent the earliest mathematics was little more than a practically workable empiricism—a collection of rule-of-thumb procedures that gave results of sufficient acceptability for the simple needs of those early civilizations. Thus the Egyptian and Babylonian formulas for volumes of granaries and areas of land were arrived at by trial and error, with the result that many of these formulas are definitely faulty. For example, an Egyptian formula for finding the area of a circle was to take the square of eight ninths of the circle's diameter. This is not correct, as it is equivalent to taking $\pi = (4/3)^4 = 3.1604 \cdots$. The even less accurate value of $\pi = 3$ is implied by some Babylonian formulas.[3] Another incorrect formula found in ancient Babylonian mathematics is one that says that the volume of a frustum of a cone or of a square pyramid is given by the product of the altitude and half the sum of the bases. It seems that the Babylonians also used, for the area of a quadrilateral having a, b, c, d for its consecutive sides, the incorrect formula $K = (a + c)(b + d)/4$. This formula gives the correct result only if the quadrilateral is a rectangle; in every other instance the formula gives too large an answer. It is curious that this same incorrect formula was reproduced 2000 years later in an Egyptian inscription found in the tomb of Ptolemy XI, who died in 51 B.C.

In general, simple empirical reasoning may be described as the formulation of conclusions based upon experience or observation; no real understanding is involved, and the logical element does not appear. Empirical reasoning often entails stodgy fiddling with special cases, observation of coincidences, experience

[2] For comments on a possible prehuman origin of mathematics see D. E. Smith [1], vol. 1, chap. 1, and H. Eves [3], Items 1°, 2°, 3°, 4°.

[3] This value for π is also found in the Bible; see I Kings 7:23, and II Chron. 4:2.

at good guessing, considerable experimentation, and flashes of intuition. Perhaps a very simple hypothetical illustration of empirical reasoning might clarify what is meant by this type of procedure.

Suppose a farmer wishes to enclose with 200 feet of fencing a rectangular field of greatest possible area along a straight river bank, no fencing being required along the river side of the field. If we designate as the *depth* of the field the dimension of the field perpendicular to the river bank and as the *length* of the field the dimension parallel to the river bank (see Figure 1.1), the farmer could soon form the following table:

Depth in feet	Length in feet	Area in square feet
10	180	1800
20	160	3200
30	140	4200
40	120	4800
50	100	5000
60	80	4800
70	60	4200
80	40	3200
90	20	1800

Examination of the table shows that the maximum area recorded occurs when the depth is 50 feet and the length is 100 feet. The interested farmer might now try various depths close to, but on each side of, 50 and would perhaps make the following additional table:

Depth	Length	Area
48	104	4992
49	102	4998
50	100	5000
51	98	4998
52	96	4992

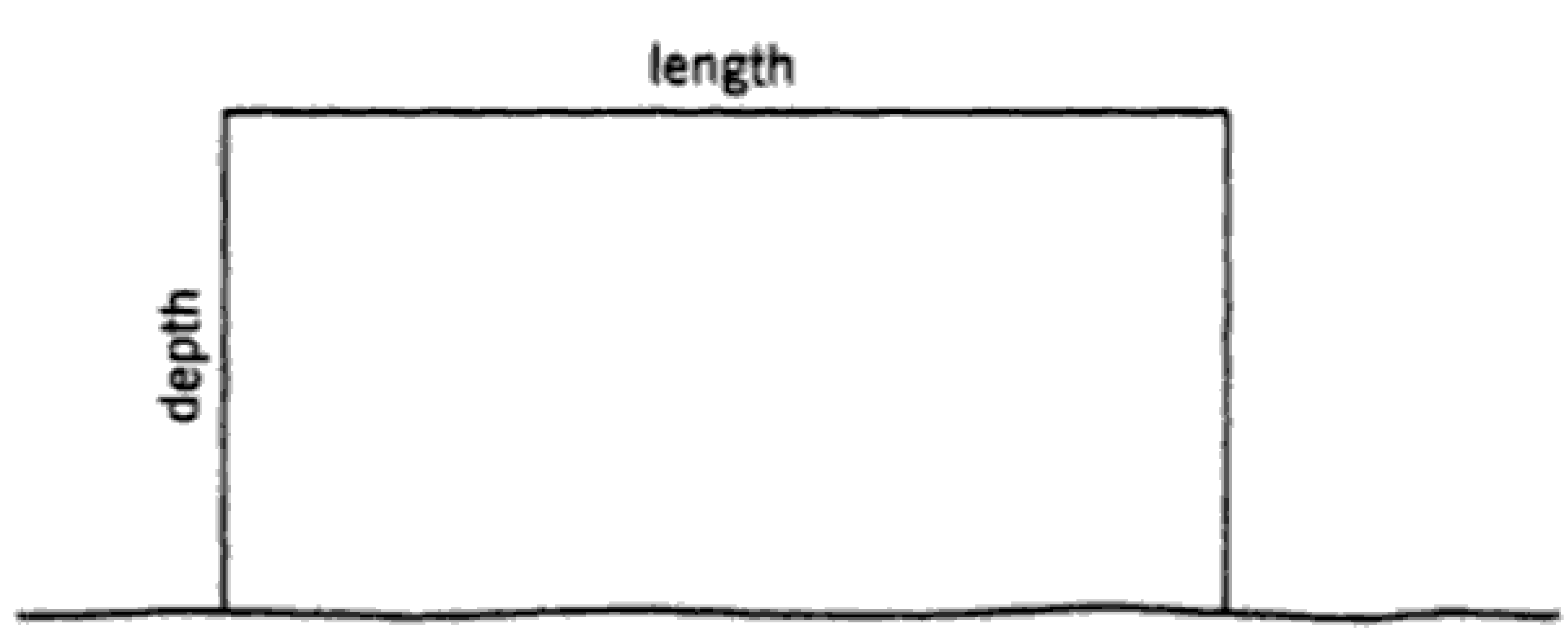


FIGURE 1.1

By now the farmer would feel quite certain that the maximum area is obtained when the depth is 50 feet and the length is 100 feet; that is, he would accept the proposition that *the maximum area occurs when the length of the field is twice the depth of the field*. A further strengthening of this belief would result from his examination of the symmetry observed in his table, and he would no doubt use his conjecture and pass it along to others as a reliable mathematical fact. Of course, the farmer's conclusion is by no means established, and no present-day student of mathematics would be permitted to "prove" the conjecture in this fashion. Shrewd guessing has taken the place of deductive logic; patience has replaced brilliance.

In spite of the empirical nature of ancient oriental mathematics, with its complete neglect of proof and the seemingly little attention paid to the difference between exact and approximate truth, one is nevertheless struck by the extent and the diversity of the problems successfully attacked. Particularly has this become evident in recent years with the scholarly deciphering of many Babylonian mathematical tablets. Apparently a great deal of elementary mathematical truth can be discovered by empirical methods when supplemented by extensive experimentation carried on patiently over a long period of time.

How were the mathematical findings of the ancient orient stated? Here we must rely on such primary sources as the Rhind, the Moscow, and other Egyptian mathematical papyri and on the approximately three hundred Babylonian mathematical tablets that have so far been deciphered.

The Rhind, or Ahmes, papyrus is a mathematical text dating from about 1650 B.C. Partaking of the nature of a practical handbook, it contains 85 problems copied by the scribe Ahmes from a still earlier work. Now possessed by the British Museum, it was originally purchased in Egypt by the Scottish antiquarian, A. Henry Rhind. This papyrus and the somewhat older Moscow papyrus, a similar mathematical text containing 25 problems, constitute our chief sources of information concerning ancient Egyptian mathematics. All of the 110 problems found in these papyri are numerical, and many of them are very easy. In general, each problem is first formulated and then followed by a step-by-step solution using the special numbers given at the beginning. Although special numbers are employed in this fashion, one feels that they are incidental and are being used merely to illustrate a general procedure. Many of the problems require nothing more than a simple linear equation, and are generally solved by the method known later in Europe as the *rule of false position*. This rule clearly reflects the empirical nature of the mathematical procedures of the time. As an example, suppose we are to solve the simple equation $x + (x/5) = 24$. Assume any convenient value for x, say, $x = 5$. Then $x + (x/5) = 6$, instead of 24. Since 6 must be multiplied by 4 to give the required 24, the correct value of x must be 4(5), or 20.

The Babylonian mathematical tablets are of two types, *table texts* and *problem texts*. There must be at least 500,000 Babylonian tablets now scattered among various museums of the world; of these only about 100 problem texts, and somewhat more than twice this number of table texts, are known to us. The table texts exhibit a wide variety of mathematical tables, such as multiplication tables, tables of reciprocals (for reducing division to multiplication), tables of squares and square roots and cubes and cube roots, tables of sums of squares and cubes (for solving certain types of cubic equations), exponential tables (for computing compound interest), and many others. The ancient Babylonians were indefatigable table makers, as one might have expected, for the construction of tables is indispensable to empirical procedure.

The problem texts also show considerable variety and are all more or less concerned with the formulation and solution of algebraic and geometric problems. A large group of the problem texts, like the Egyptian papyri considered above, formulate a problem in terms of specific numbers and then proceed with a step-by-step solution using the specific numbers. Such texts often terminate with the phrase, "such is the procedure." Again it is apparent that it is the general procedure, and not the numerical result, that is considered important. If, in a multiplication, a factor has the value 1, multiplication by this 1 will be explicitly performed, for this step is necessary in the general case. The remaining problem texts contain on a single tablet, often not as large as a page of this book, a large number of related numerical problems carefully arranged from the simplest cases up through the more complicated ones. The apparent purpose of such a text was to teach, by repetition and gradual introduction of complexities, a certain method or procedure, and the accompanying numbers serve merely as a guide to illustrate the underlying general procedure. The solution of quadratic equations, for example, both by general formula and by the method of completing the square, is explained in this way on ancient Babylonian tablets.

In summary, then, we find that pre-Hellenic mathematics was empirical. Nowhere do we find in ancient oriental mathematics a single instance of what we today call a logical demonstration. Instead of an argument we find a description of a process explained by means of specific numerical cases. In short, we are instructed to "Do thus and so." It is very interesting to note that although today confirmed students of the scientific method find this "Do thus and so" procedure highly unsatisfactory it is the procedure employed in much of our elementary teaching.

1.2 Induction Versus Deduction

Empirical conclusions, we have seen, are generalizations based on a limited number of observations or experiments. For example, the farmer of the previous section obtained a general rule by observing a limited number of computed areas. Another farmer may observe that unusually good crops have followed a number of winters of heavy snow, and empirically conclude that snowy winters are beneficial to crops. As a further example, a scientist may observe that particularly fine displays of the aurora borealis always occurred in his experience during periods of pronounced sun-spot activity and conclude that there must be a connection between the two phenomena. This type of reasoning, which concludes on the basis of a limited number of instances that something is always true, is known as *induction*.[4] Modern probability considerations have served to introduce refinements into inductive procedures. It is important to note,

[4] *Induction* should not be confused with so-called *mathematical induction*, which is considered in Section 7.3.

however, that no matter how fully the conclusions of inductive reasoning may seem warranted by the facts, these conclusions are not established beyond all possible doubt; conclusions obtained by induction are only more or less probable.

Empirical conclusions are sometimes reached by using a primitive form of induction known as reasoning by analogy. For example, if we cut off the top of a triangle by a line parallel to the base of the triangle, a trapezoid will remain, and the area of a trapezoid is given by the product of its altitude and the arithmetic average of its two bases. Now, if we cut off the top of a pyramid by a plane parallel to the base of the pyramid, a frustum will remain. By analogy one might expect the volume of a frustum of a pyramid to be given, as before, by the product of the altitude and the arithmetic average of the two bases. This is the incorrect Babylonian formula noted in the previous section. Reasoning by analogy certainly is useful, but obviously its conclusions cannot be regarded as established.

In sharp contrast to reasoning by analogy or by induction is reasoning by deduction, because the conclusions reached by deduction, provided one accepts the premises that are adopted and the system of logic that is employed, are incontestable. To illustrate deductive procedure, consider the following two statements: (1) All Canadians are North Americans; (2) Two particular men under consideration are Canadians. If we accept these two statements, we are logically compelled, following accepted principles of Aristotelian logic, to accept a third statement—namely, (3) The two men under consideration are North Americans. This is an example of deductive reasoning, which at this point may be described as those ways of deriving new statements from accepted ones that *compel* us to accept the derived statements. In the example, the first two statements are called *premises*, and the third statement the *conclusion*.

It is very important to realize that in deductive reasoning we are not concerned with the *truth* of the conclusion but rather whether the conclusion does or does not follow from the premises. If the conclusion follows from the premises, we say that our reasoning is *valid*; if it does not, we say that our reasoning is *invalid*. For example, from the two

Premises: (1) All college students are clever,
(2) All freshmen are college students,

follows the

Conclusion: All freshmen are clever.

Now the last statement certainly is not regarded generally as true, but the reasoning leading to it is valid. *If both of the premises had been true, the conclusion also would have been true*; it is essential that one understand early in the treatment of this book this meaning of the deductive process.

A useful and easy way to test the validity of a piece of deductive reasoning, like either of the examples given above, is by a diagrammatic procedure ascribed to the eminent Swiss mathematician Leonhard Euler (1707–1783). Consider our last example. We may represent the class of all clever people by a planar region within a closed boundary, and we may do likewise for the class of all college students and for the class of all freshmen. But statement (1) insists that the class of all college students is contained in the class of all clever people, and statement (2) insists that the class of all freshmen is contained in the class of all college students. Thus our various classes must be represented by their corresponding regions as shown in Figure 1.2. Clearly, the requirements of our premises *forced* us to place the class of all freshmen entirely within the class of all clever people, which is exactly what our conclusion asserts. Hence, although the conclusion is undoubtedly false, the reasoning leading to it is valid. It cannot be overemphasized at this point that the expert in the use of deduction is not fundamentally concerned with *truth* but with *validity*; he merely wants to be able to assert that his conclusions are implied by the premises. It would then follow that *if* the premises should happen to be true, the conclusion *must*, of necessity, also be true.

Consider, as a second example, the following:

Premises: (1) All parallelograms are polygons.
(2) All quadrilaterals are polygons.
Conclusion: All parallelograms are quadrilaterals.

Here all three statements are true, but the reasoning is invalid, for the premises do not *force* us to place the region representing the class of all parallelograms entirely within the region representing the class of all quadrilaterals; we are able to satisfy the requirements of our premises by a diagram like that shown in Figure 1.3.

As a third example, consider the following:

Premises: (1) All parallelograms are circles.
(2) All circles are polygons.
Conclusion: All parallelograms are polygons.

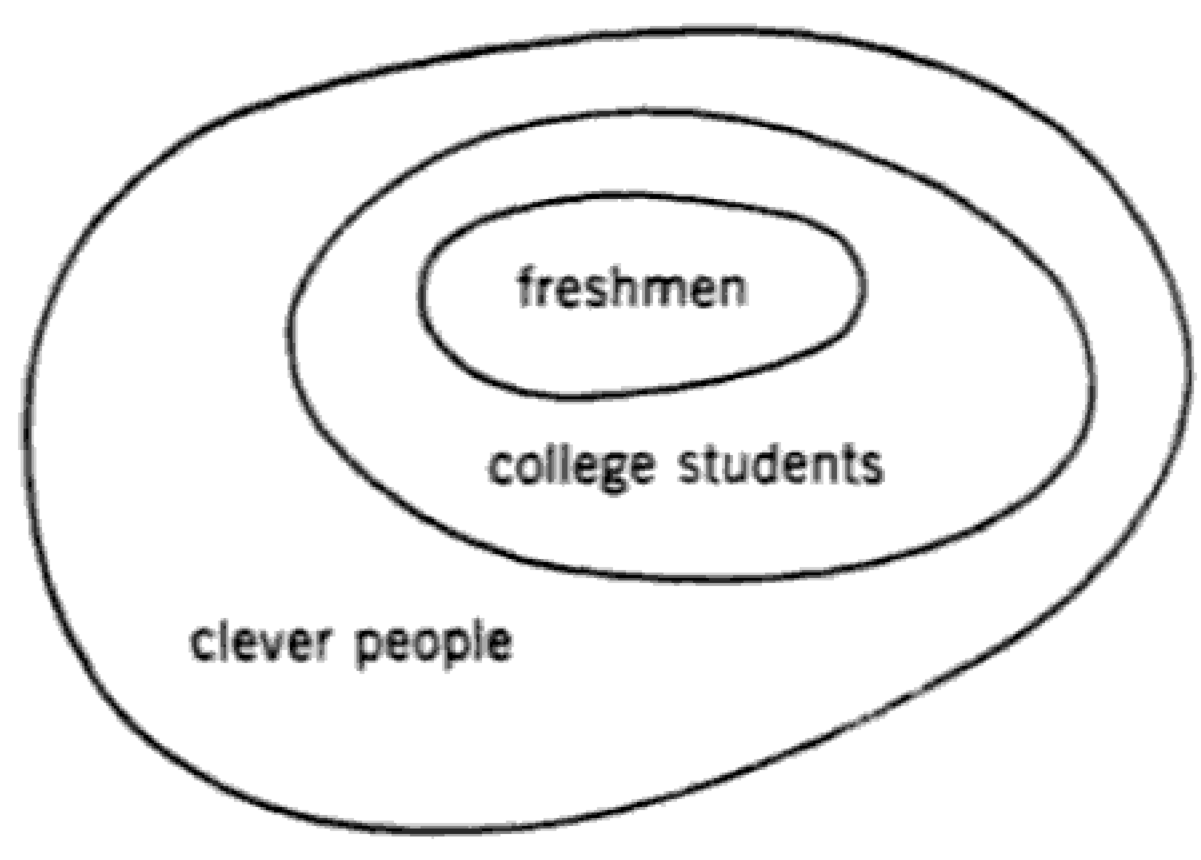


FIGURE 1.2

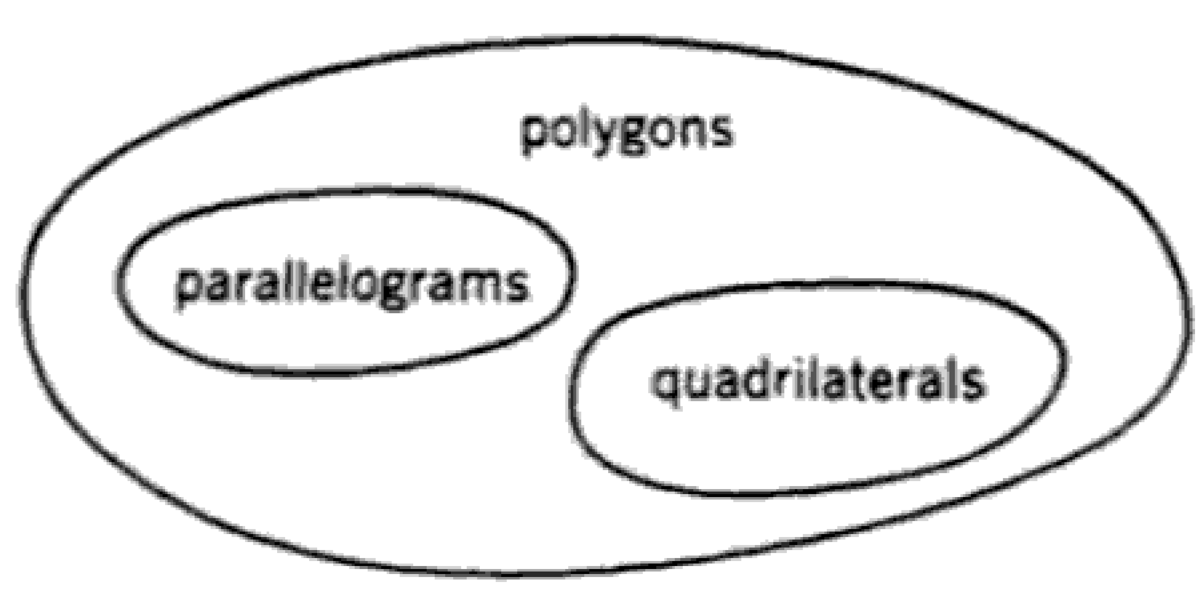


FIGURE 1.3

Here the premises are both false, the conclusion is true, and, as tested by the diagram of Figure 1.4, the reasoning is valid. Thus false assumptions may actually yield a true conclusion. True premises can yield only true conclusions when deductive logic is applied, but false premises may or may not yield true conclusions.

Finally, we shall examine the following:

Premises: (1) No quadrilaterals are triangles.
(2) Some quadrilaterals are parallelograms.
Conclusion: Some parallelograms are not triangles.

Since, by (1), the region representing the class of all quadrilaterals and that representing the class of all triangles *cannot* overlap, and, by (2), the region representing the class of all quadrilaterals and that representing the class of all parallelograms *must* overlap, the conclusion (see Figure 1.5) certainly follows, and the reasoning is valid. Note, however, that we cannot conclude, from our premises, that *no* parallelogram is a triangle, for there is nothing that *forces* us to keep the region representing the class of all parallelograms from cutting into the region representing the class of all triangles.

Euler's diagrammatic device can be used in a great variety of situations, and it is recommended to the person unfamiliar with logical procedure.

We shall not, for the present, go beyond the above superficial study of inductive and deductive reasoning. As already indicated, deductive reasoning has the advantage that its conclusions are unquestionable if the premises are accepted, and it has the additional advantage of considerable economy: before a

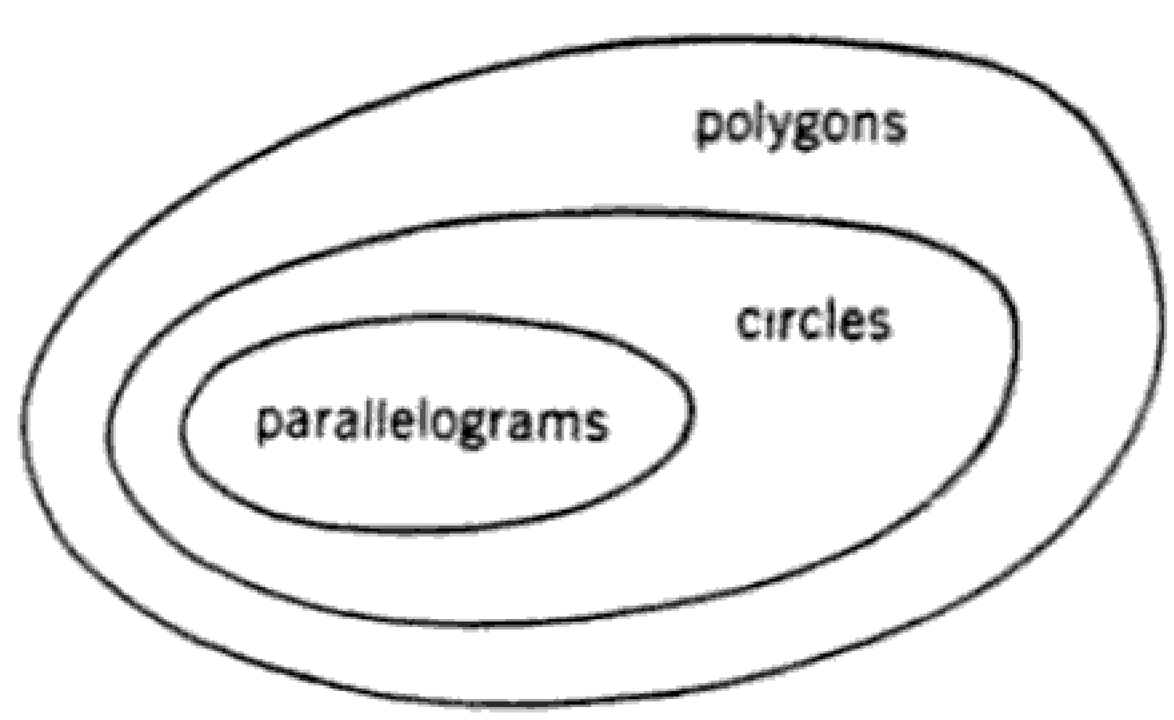


FIGURE 1.4

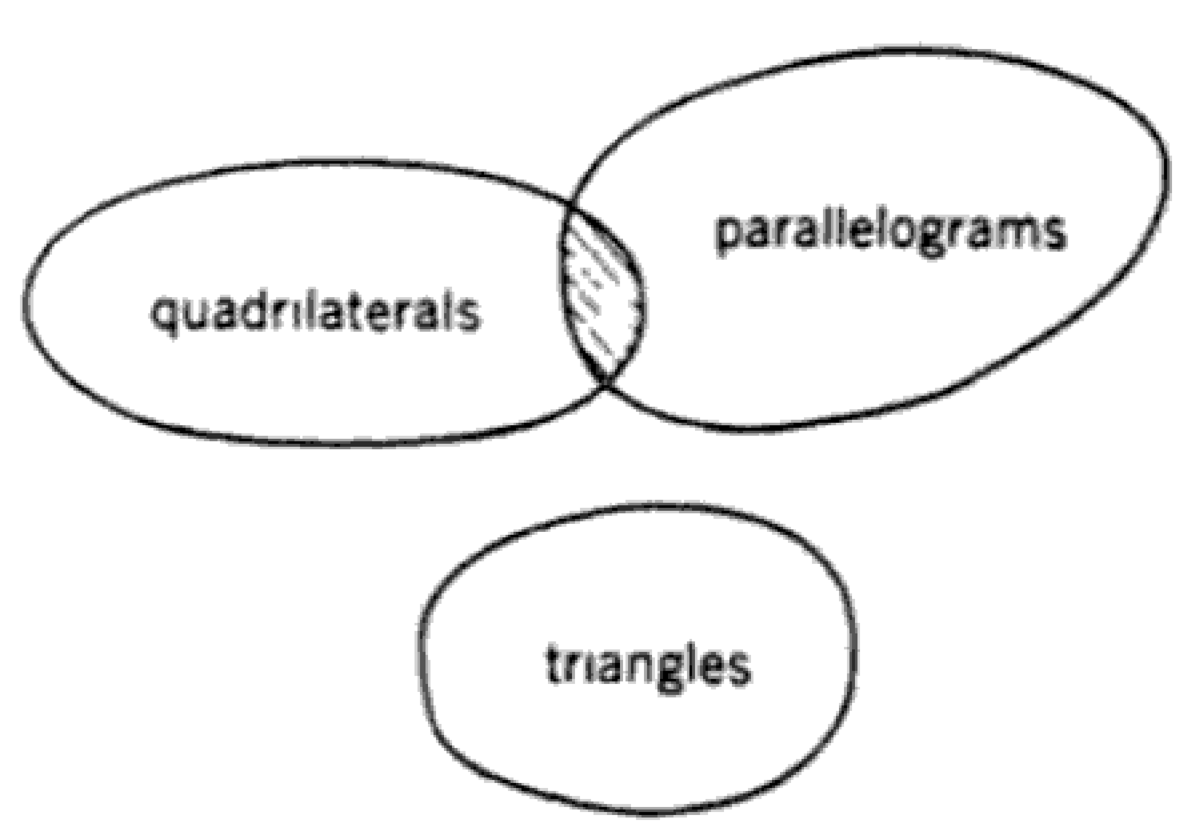


FIGURE 1.5

bridge is built and put into use, deductive reasoning can determine the outcome. But in spite of these particular advantages, deductive reasoning does not supplant the inductive approach; actually, each way of obtaining knowledge has its advantages and disadvantages. The significant thing, from the point of view of our present study, is that the ancient Greeks found in deductive reasoning the vital element of the modern mathematical method.

1.3 Early Greek Mathematics and the Introduction of Deductive Procedures

The origin of early Greek mathematics is clouded by the greatness of Euclid's *Elements*, written about 300 B.C., because this work so clearly excelled many preceding Greek writings on mathematics that the earlier works were thenceforth discarded. As the great mathematician David Hilbert (1862–1943) once remarked, one can measure the importance of a scientific work by the number of earlier publications rendered superfluous by it.

The debt of Greek mathematics to ancient oriental mathematics is difficult to evaluate, nor has the path of transmission from the one to the other yet been satisfactorily uncovered. That the debt is considerably greater than formerly believed became evident with twentieth-century researches of Babylonian and Egyptian records. Greek writers themselves expressed respect for the wisdom of the East, and this wisdom was available to anyone who could travel to Egypt and Babylonia. There are also internal evidences of a connection with the East. Early Greek mysticism in mathematics smacks strongly of oriental influence, and some Greek writings, like those of Heron and Diophantus, exhibit a Hellenic perpetuation of the more arithmetic tradition of the orient. Also, there are strong links connecting Greek and Mesopotamian astronomy.

But whatever the strength of the historical connection between Greek and ancient oriental mathematics, the Greeks transformed the subject into something vastly different from the set of empirical conclusions worked out by their predecessors. The Greeks insisted that mathematical facts must be established, not by empirical procedures, but by deductive reasoning; mathematical conclusions must be assured by logical demonstration rather than by laboratory experimentation.

This is not to say that the Greeks shunned preliminary empirical and experimental methods in mathematics, for it is probably quite true that few, if any, significant mathematical facts have ever been found without some preliminary empirical work of one form or another. Before a mathematical statement can be proved or disproved by deduction, it must first be thought of, or conjectured, and a conjecture is nothing but a guess made more or less plausible by intuition, observation, analogy, experimentation, or some other form of empirical procedure. Deduction is a convincing formal mode of exposition, but it is hardly a means of discovery. It is a set of complicated machinery that needs material to work upon, and the material is frequently furnished by empirical considerations. Even the steps of a deductive proof or disproof are not dictated to us by the deductive apparatus itself but must be arrived at by trial and error, experience, and shrewd guessing. Indeed, skill in the art of good guessing is one of the prime ingredients in the make-up of a worthy mathematician. What is important here is that the Greeks insisted that a conjectured or laboratory-obtained mathematical statement must be followed up with a rigorous proof or disproof by deduction and that no amount of verification by experiment is sufficient to *establish* the statement.

It is difficult to give a wholly adequate explanation of just why the Greeks of 600 to 400 B.C. decided to abandon empirical methods of establishing mathematical knowledge and to insist that all mathematical conclusions be established only by deductive reasoning. This completely new viewpoint on mathematical method is usually explained by the peculiar mental bias of the Greeks of classical times toward philosophical inquiries. In philosophical speculations, reasoning centers about abstract concepts and broad generalizations and is concerned with inevitable conclusions following from assumed premises. Now the empirical method affords no way of discriminating between a valid and an invalid argument and so is hardly applicable to philosophic considerations. It is deductive reasoning that philosophers find to be their indispensable tool, and so the Greeks naturally gave preference to this method when they began to consider mathematics.

Another explanation of the Greek preference for deduction stems from the Hellenic love for beauty. Appreciation of beauty is an intellectual as well as an emotional experience, and from this point of view the orderliness, consistency, completeness, and conviction found in deductive argument are very satisfying.

A still further explanation for the Greek preference for deductive procedures has been found in the nature of Greek society in classical times. Philosophers, mathematicians, and artists belonged to a social class that in general disdained manual work and practical pursuits, which were carried on by a large slave class. In Greek society the slave class ran the businesses and managed the industries, took care of households, and did both the technical and the unskilled work of the time. This slave basis naturally fostered a separation of theory from practice and led the members of the privileged class to a preference for deduction and abstraction and a disdain for experimentation and practical application.

It is disappointing that, unlike the situation with ancient Egyptian and Babylonian mathematics, there exist virtually no source materials for contemporary study that throw much light on early Greek mathematics. We are forced to rely on manuscripts and accounts that are dated several hundred years after the original treatments were written. In spite of this difficulty, however, scholars of classicism have been able to build up a rather consistent, though somewhat hypothetical, account of the history of early Greek mathematics and have even plausibly restored many of the original Greek texts. This work required amazing ingenuity and patience; it was carried through by painstaking comparisons of derived texts and by the examination of countless literary fragments and scattered remarks made by later authors, philosophers, and commentators.

Our principal source of information concerning very early Greek mathematics is the so-called *Eudemian Summary* of Proclus. This summary constitutes a few pages of Proclus's *Commentary on Euclid, Book I*, and is a very brief outline of the development of Greek geometry from the earliest times to Euclid. Although Proclus lived in the fifth century A.D., a good thousand years after the inception of Greek mathematics, he still had access to a number of historical and critical works that are now lost to us except for the fragments and allusions preserved by him and others. Among these lost works was apparently a full history of Greek geometry, covering the period before 335 B.C., written by Eudemus, a pupil of Aristotle. The *Eudemian Summary* is so named because it is based on this earlier work.

According to the *Eudemian Summary*, Greek mathematics appears to have started in an essential way with the work of Thales of Miletus in the first half of the sixth century B.C. This versatile genius, declared to be one of the "seven wise men" of antiquity, was a worthy founder of systematic mathematics and is the first known individual with whom the use of deductive methods in mathematics is associated. Thales, the summary tells us, sojourned for a time in Egypt and brought back geometry with him to Greece, where he began to apply to the subject the deductive procedures of philosophy. In particular, he is credited with the following elementary geometrical results:

1. A circle is bisected by any diameter.

2. The base angles of an isosceles triangle are equal.

3. Vertical angles formed by two intersecting straight lines are equal.

4. Two triangles are congruent if two angles and a side in one are equal respectively to two angles and the corresponding side of the other. (It is thought that Thales used this result to determine the distance of a ship from shore.)

5. An angle inscribed in a semicircle is a right angle. (The Babylonians of some 1400 years earlier were acquainted with this geometrical fact.)

We are not to measure the value of these results by their content but rather by the belief that Thales supported them with a certain amount of logical reasoning instead of intuition and experiment. For the first time a student of mathematics was committed to a form of deductive reasoning, crude and incomplete though it may have been. Moreover, the fact that the first deductive thinking was done in the field of geometry instead of algebra, for instance, inaugurated a tradition in mathematics that was maintained, as we shall see, until very recent times.

The next outstanding Greek mathematician mentioned in the *Eudemian Summary* is Pythagoras, who is claimed to have continued the purification of geometry that was begun some fifty years earlier by Thales. Pythagoras was born about 572 B.C. on the island of Samos, one of the Aegean islands near Thales's home city of Miletus, and it may be that he studied under the older man. It seems that he visited Egypt and perhaps traveled even more extensively about the orient. When, on returning home, he found Samos under the tyranny of Polycrates and Ionia under Persian dominion, he decided to migrate to the Greek seaport of Crotona in southern Italy. Here he founded the celebrated

Pythagorean school, a brotherhood knit together with secret and cabalistic rites and observances and committed to the study of philosophy, mathematics, and natural science.

The philosophy of the Pythagorean school was built on the mystical assumption that whole number is the cause of the various qualities of man and matter. This oriental outlook, perhaps acquired by Pythagoras in his eastern travels, led to the exaltation and study of number relations and to a perpetuation of numerological nonsense that has lasted even into modern times. However, in spite of the unscientific nature of much of Pythagorean study, members of the society contributed, during the two hundred or so years following the founding of their organization, a good deal of sound mathematics. They developed the properties of parallel lines and used them to prove that the sum of the angles of any triangle is equal to two right angles. They contributed in a noteworthy manner to Greek geometrical algebra; they effected the geometrical equivalent of addition, subtraction, multiplication, division, extraction of roots, and even the complete solution of the general quadratic equation insofar as it has real roots. They developed a fairly complete theory of proportion, though it was limited only to commensurable magnitudes, and used it to deduce properties of similar figures. They were aware of the existence of at least three of the regular polyhedral solids, and they discovered the incommensurability of a side and a diagonal of a square. Although much of this information was already known to the Babylonians of earlier times, the deductive aspect of mathematics is thought to have been considerably exploited and advanced in this work of the Pythagoreans. Chains of propositions in which successive propositions were derived from earlier ones in the chain began to emerge. As the chains lengthend, and some were tied to others, the bold idea of developing all of geometry in one long chain suggested itself. It is claimed in the *Eudemian Summary* that the Pythagorean, Hippocrates of Chios,[5] was the first to attempt, with at least partial success, a logical presentation of geometry in the form of a single chain of propositions based upon a few initial definitions and assumptions.

The famous Greek philosopher, Plato, was strongly influenced by the Pythagoreans, and Plato, in turn, exerted a considerable influence on the development of mathematics in Greece. Plato's influence was not due to any mathematical discoveries he made but rather to his enthusiastic conviction that the study of mathematics furnished the finest training field for the mind, and hence was essential in the cultivation of philosophers and those who should govern his ideal state. This belief explains the renowned motto over the door of his Academy, "Let no one unversed in geometry enter here." Thus, because of its logical element and the pure attitude of mind that he felt its study creates, mathematics seemed of utmost importance to Plato, and for this reason it occupied a valued place in the curriculum of the Academy. Some see in certain of Plato's dialogues what may perhaps be considered the first serious attempt at a philosophy of mathematics. Certainly mathematics in Greece at the time of Plato had advanced a long way from the empirical mathematics of ancient Egypt and Babylonia.

[5]He is not to be confused with Hippocrates of Cos, the eminent Greek physician of antiquity.

1.4 Material Axiomatics

Much was accomplished by the Greeks during the three hundred years between Thales in 600 B.C. and Euclid in 300 B.C. Not only did the Pythagoreans and others develop the material that ultimately was organized into the *Elements* of Euclid, but there were developed notions concerning infinitesimals and summation processes (notions that did not attain final clarification until the rigorization of the calculus in modern times) and also considerable higher geometry (the geometry of curves other than the circle and the straight line and of surfaces other than the sphere and plane). Curiously enough, much of this higher geometry originated in continued attempts to solve the three famous construction problems of antiquity—the duplication of a cube, the trisection of an arbitrary angle, and the quadrature of a circle—illustrating the principle that the growth of mathematics is stimulated by the presence of outstanding unsolved problems.

Also, some time during the first three hundred years of Greek mathematics, there developed the Greek notion of a logical discourse as a sequence of statements obtained by deductive reasoning from an accepted set of initial statements. Certainly, if one is going to present an argument by deductive procedure, any statement of the argument will have to be derived from some previous statement or statements of the argument, and such a previous statement must itself be derived from some still more previous statement or statements. Clearly this cannot be continued backward indefinitely, nor should one resort to illogical circularity by deriving statement q from statement p and then later deriving statement p from statement q. The only way out of the difficulty is to set down, toward the start of the discourse, a collection of fundamental statements whose truths are to be accepted and then to proceed by purely deductive reasoning to derive all the other statements of the discourse. Now both the initial and the derived statements of the discourse are statements about the technical matter of the discourse and hence involve special or technical terms. The meanings of these terms must be made clear to the reader, and so, the Greeks felt, the discourse should start with a list of explanations and definitions of these technical terms. After these explanations and definitions have been given, the initial statements, called *axioms* and/or *postulates* of the discourse, are to be listed. These initial statements, according to the viewpoint held by some of the Greeks, should be so carefully chosen that their truths are quite acceptable to the reader in view of the explanations and definitions already cited.

A discourse that is conducted according to the above plan is described today as a development by *material axiomatics*. Certainly the most outstanding contribution of the early Greeks to mathematics was the formulation of axiomatic procedure and the insistence that mathematics be systematized by such a procedure. Euclid's *Elements* is the earliest extensively developed example of axiomatic procedure that has come down to us; it largely follows the pattern of material axiomatics, and we shall certainly want to examine it in some detail. In more recent years, the pattern of material axiomatics has been significantly refined to yield a more abstract form of discourse known as *formal axiomatics* (see Chapter 6). For the time being we will content ourselves by summarizing the pattern of material axiomatics.

Pattern of Material Axiomatics

1. Initial explanations of certain basic technical terms of the discourse are given, the intention being to suggest to the reader what is to be meant by these basic terms.
2. Certain primary statements that concern the basic terms and that are felt to be acceptable as true on the basis of the properties suggested by the initial explanations are listed. These primary statements are called the *axioms* or *postulates* of the discourse.
3. All other technical terms of the discourse are defined by means of previously introduced terms.
4. All other statements of the discourse are logically deduced from previously accepted or established statements. These derived statements are called the *theorems* of the discourse.

To gain a feeling for the pattern of material axiomatics, let us consider an example. Suppose one is faced with the task of developing a logical discourse on carpentry. The subject of carpentry contains many special or technical terms, such as nail, spike, brad, screw, wood, hard wood, soft wood, board, strut, beam, hammer, saw, screw driver, plane, or chisel. Some of these technical terms can be defined in terms of others. For example, a spike and a brad can each be defined as a special kind of nail; hard wood and soft wood can be defined as certain special kinds of wood; board, strut, and beam can be defined as pieces of wood of certain shapes used for certain purposes; various kinds of hammers and saws can be defined in terms of the basic hammer and saw. It is certainly logical, then, to commence the discourse with some sort of explanation or description of the basic technical terms—say, nail, wood, hammer, saw, and others—and then to define further technical terms, either at the start or as needed, in terms of the basic ones. After giving these initial explanations and possible definitions, the next thing to do is to list some fundamental statements about the explained and defined terms that will be assumed so that the discourse may get under way. Now these assumed statements, from the point of view of material axiomatics, should be such that the reader is perfectly willing to accept them on account of the initial explanations of the basic terms involved. For example, one may wish to assume that *it is always possible to drive a nail with a hammer into a piece of wood*, that *it is always possible with a saw to cut a piece of wood in two by a planar cut*, etc. That two boards of desired lengths can be fastened together with nails now follows as a consequence of these assumptions, and is thus a theorem of the discourse. Probably enough has been said to illustrate the Greek notion of material axiomatics.

The theory of some simple games can be rather easily developed by material axiomatics. Consider, for example, the familiar game of tic-tac-toe. Among the technical terms of this game are the *playing board*, *nought*, *cross*, a *win*, a *draw*, and so on. These technical terms are to be explained or defined. The rules of the game are then stated as the postulates of the discourse, these rules being perfectly acceptable once one understands the basic terms of the discourse. From these rules one can then proceed to deduce the theory of the game, proving as a theorem, for example, that *with sufficiently good playing, the player who starts a game need not lose the game.*

1.5 The Origin of the Axiomatic Method

We do not know with whom the axiomatic method originated. By the account given in the *Eudemian Summary*, the method seems to have evolved with the Pythagoreans as a natural outgrowth and refinement of the early application of deductive procedures to mathematics. This is the traditional and customary account and is based principally on Proclus's summary, which, in turn, is based on the lost history of geometry written by Eudemus about 335 B.C. The account may be the true one, and, if so, we must concede to the Pythagoreans a very high place in the history of the development of mathematics.

There are some historians of ancient mathematics who find the account of the early history of Greek mathematics, as reconstructed from the *Eudemian Summary*, somewhat difficult to believe and who feel that the traditional stories about Thales and Pythagoras must be discredited as purely legendary and unhistorical in content. For example, the *Eudemian Summary* says that Thales proved that a circle is bisected by any one of its diameters. The realization that so obvious a matter as this should need demonstration seems to reflect a mathematical sophistication of a much more advanced period, when the importance and delicacy of initial assumptions had become much clearer. Eudemus may have hypothetically restored the sequence of events so that they accorded with the state of the theory of his time, as many historians do when source material is not available. Actually, we can have very little idea of the roles played in the history of mathematics by Thales and Pythagoras, and it may be much closer to reality to assume that early Greek mathematics cannot have differed greatly from the oriental type. An essential turn in the development of a subject is usually brought about by some crucial circumstance, and in mathematics such a circumstance arose some time in the fifth century B.C. with the devastating discovery of the irrationality of $\sqrt{2}$.

Let us pause a moment to consider the significance of the last statement. Since the rational numbers consist of all numbers of the form p/q, where p and q are integers with $q \neq 0$, the discovery alluded to states that there are no integers p and q such that $p/q = \sqrt{2}$; that is, $\sqrt{2}$ is *not* a rational number and hence, by definition, is an *irrational* (nonrational) number. The traditional proof of this fact, apparently known to Aristotle (384–322 B.C.), is simple and runs as follows: Suppose, on the contrary, that there are two integers p and q such that $p/q = \sqrt{2}$, where, without any loss of generality, we may assume that p and q have no common positive integral factor other than unity. Then $p^2 = 2q^2$. Since p^2 is twice an integer, we see that p^2, and hence p, must be even. So we may put $p = 2r$. Then we find $4r^2 = 2q^2$, or $2r^2 = q^2$, from which we conclude that q^2, and hence q, must be even. But this is impossible, since we assumed that p and q have no common integral factor different from unity. The supposition that $\sqrt{2}$ is rational has led to a contradictory situation, whence it follows that $\sqrt{2}$ must be irrational. This result was surprising and disturbing on several grounds. First of all, it seemed to deal a mortal blow to the Pythagorean philosophy that all depends on the integers. Next, it seemed contrary to common sense, for it was felt intuitively that any magnitude could be expressed by *some* rational number. The geometrical counterpart was equally startling, for who could doubt that for any two given line segments one is able to find some third line segment, perhaps very very small, that can be marked off a whole number of times into each of the given segments. But take as the two given segments a side s and a diagonal d of a square. Now if there exists a third segment t which can be marked off a whole number of times into s and d we would have $s = qt$ and $d = pt$, where p and q are integers. But $d = s\sqrt{2}$, whence $pt = qt\sqrt{2}$, or $\sqrt{2} = p/q$, a rational number. Contrary to intuition, then, there exist line segments having no common unit of measure. But the whole Pythagorean theory of proportion was built on the seemingly obvious assumption that any two line segments are commensurable, that is, do have some common unit of measure.

No wonder the discovery of the irrationality of $\sqrt{2}$ led to some consternation in the Pythagorean ranks. The situation must have caused a profound reaction in mathematical thinking, and must have very considerably emphasized the extreme importance of careful agreement on what can be taken for basic assumptions. A crisis, like this one of the discovery of irrational numbers, could well account for the origin of the axiomatic method, and, if so, the credit for the invention might largely go to Eudoxus, the genius of the time who finally resolved the crisis that had arisen.[6]

This second explanation of the possible origin of the axiomatic method has other points in its favor. For example, it places less stress on any peculiar mentality possessed by the Greeks of very early times, and it accounts for the relatively large number of Greek papyrus fragments containing texts after the pattern of oriental mathematics. These texts, like the similar ones from Babylonian times, probably formed the backbone of instruction in elementary mathematics. At this elementary level the highly sophisticated axiomatic method had as little influence as it has today in much of our elementary teaching. Writings of this sort, then, do not reflect any degeneration of the so-called Greek spirit in mathematics but simply exhibit the continuance, on an elementary level, of older traditions. Heron's geometry, for example, can be properly considered a Hellenic form of oriental tradition; it should not be regarded as a sign of decline in Greek mathematics just because it does not employ the refined procedures of the axiomatic method.

Perhaps it is needless to hypothesize about the origin of the axiomatic method. Certainly, by the middle of the fourth century B.C., the method had been fairly well developed, for in Aristotle's *Analytica posteriora*, we find a good deal of light thrown on some of its features. Aristotle was not a mathematician, but as the systematizer of classical logic, he found in elementary mathematics excellent models of logical reasoning, and his mathematical illustrations tell us a great deal about the principles of the axiomatic method as accepted in his time. By the turn of the century the stage was set for Euclid's magnificent and epoch-making application of the axiomatic method.

---

[6] See Appendix, Section A.6.

PROBLEMS[7]

1.1.1 In the Rhind papyrus the area of a circle is taken as equal to that of a square on 8/9 of the circle's diameter. Show that this is equivalent to taking $\pi = 3.1604 \cdots$.

1.1.2 The *Śulvasūtras*, ancient Hindu religious writings dating from about 500 B.C., are of interest in the history of mathematics because they embody certain geometrical rules for the construction of altars and show an acquaintance with the Pythagorean theorem. Among the rules furnished there appear empirical solutions of the circle-squaring problem that are equivalent to taking $d = (2 + \sqrt{2})s/3$ and $s = 13d/15$, where d is the diameter of the circle and s is the side of the equal square. These formulas are equivalent to taking what values for π?

1.1.3 Show that the ancient Babylonian formula $K = (a + c)(b + d)/4$, for the area of a quadrilateral having a, b, c, d for consecutive sides, gives too large an answer for all nonrectangular quadrilaterals.

1.1.4 (*for students who have studied calculus*) Prove, by elementary differential calculus, the farmer's conjecture that the rectangular field of maximum area lying along a straight river bank and utilizing a given amount of fencing has a length that is twice the depth of the field.

1.1.5 A disc of radius R spins vertically on a horizontal axis held above the surface of a liquid. As the disc spins it cuts into the liquid. Estimate, by empirical methods, how high the axis must be above the liquid's surface so that the wetted area of the spinning disc above the surface of the liquid shall be a maximum. (This problem arose in the manufacture of fruit syrups from fruit juices, and was solved empirically by the manufacturer, who found the required height r of the axis above the surface of the liquid to be about $(3/10)R$. It is not very difficult to show by differential calculus that $r = R/(1 + \pi^2)^{1/2}$. It is interesting that the General Electric Company began studies of this evaporation method in the early 1960s in connection with the design of a diffusion still.)

1.1.6 Two ladders, 60 ft long and 40 ft long, lean from opposite sides across an alley lying between two buildings, the feet of the ladders resting against the bases of the buildings. If the ladders cross each other at a height of 10 ft above the alley, how wide is the alley? Solve this problem empirically from drawings. [An algebraic treatment of this problem requires the solution of a quartic equation. If a and b represent the lengths of the ladders, c the height at which they cross, and x the width of the alley, it can be shown that $(a^2 - x^2)^{-1/2} + (b^2 - x^2)^{-1/2} = c^{-1}$.]

1.1.7 How good is the following empirical straightedge and compass trisection of an angle of 30°? Let AOB be the given angle, with $OA = OB$. On AB as diameter draw a semicircle lying on the same side of AB as is the point O. Take D and E on the semicircle such that $AD = DE = EB$. Take F on DE such that $DF = DE/4$. Then OF is a sought trisector.

1.1.8 Solve, by the rule of false position, the following problem found in the Rhind papyrus: "A quantity, its 2/3, its 1/2, and its 1/7, added together, become 33. What is the quantity?"

1.1.9 Find the length of side BC in the quadrilateral pictured in Figure 1.6.

1.1.10 In the study of geometrical constructions there is a counterpart of the rule of false position, generally known as the *method of similitude*. The method lies in constructing a figure similar to the one desired, and then, by the use of proportion, "blowing it up" to proper size. Suppose, for example, we wish to inscribe a square

[7]Note that triple numbering is used in the problems. The first number is the chapter number, the second is the section number, and the third is the sequence number.

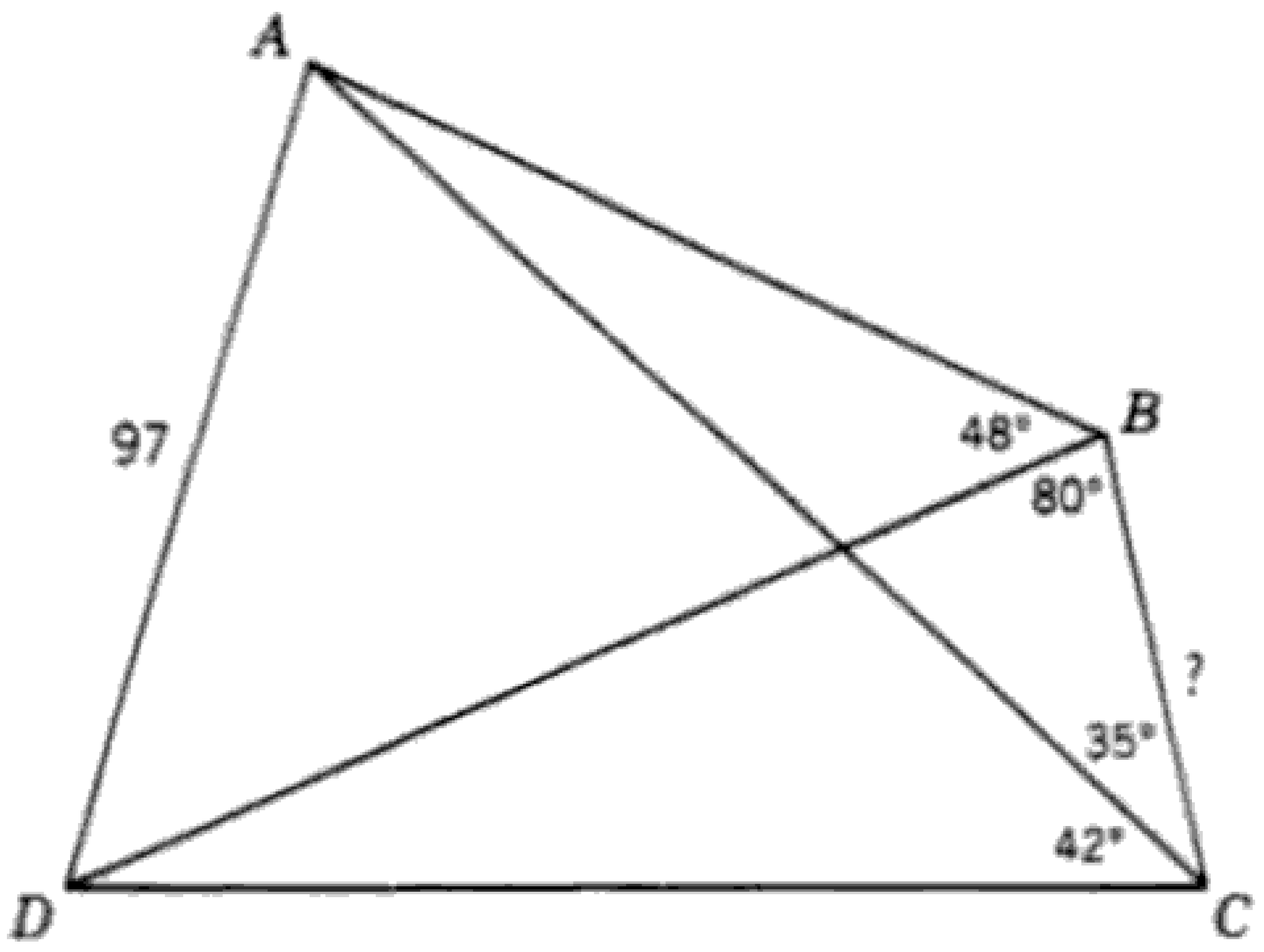


FIGURE 1.6

in a given triangle ABC so that one side of the square lies along the base BC of the triangle (see Figure 1.7). First draw a square $D'E'F'G'$ of any convenient size, as indicated. If F' falls on AC, the problem is solved. Otherwise we have solved the problem for a triangle $A'BC'$ similar to triangle ABC and having B as a center of similitude. It follows that line BF' cuts AC in the vertex F of the sought square inscribed in triangle ABC.

Construct, by the method of similitude, a line segment DE, where D is on side AB and E on side AC of a given triangle ABC, so that $BD = DE = EC$.

1.1.11 In the Rhind papyrus we find, "If you are asked, what is 2/3 of 1/5, take the double and the sixfold; that is 2/3 of it. One must proceed likewise for any other fraction." Interpret this and prove the general statement.

1.1.12 In the Moscow papyrus we find the following numerical example: "If you are told: A truncated pyramid of 6 for the vertical height by 4 on the base by 2 on the top. You are to square this 4, result 16. You are to double 4, result 8. You are to square 2, result 4. You are to add the 16, the 8, and the 4, result 28. You are to take one third of 6, result 2. You are to take 28 twice, result 56. See, it is 56. You will find it right." Show that this illustrates the general formula

$$V = \frac{h(a^2 + ab + b^2)}{3},$$

giving the volume of a frustum of a square pyramid in terms of the height h and the sides a and b of the bases.

1.1.13 Interpret the following, found on a Babylonian tablet dating from about 2600 B.C.:

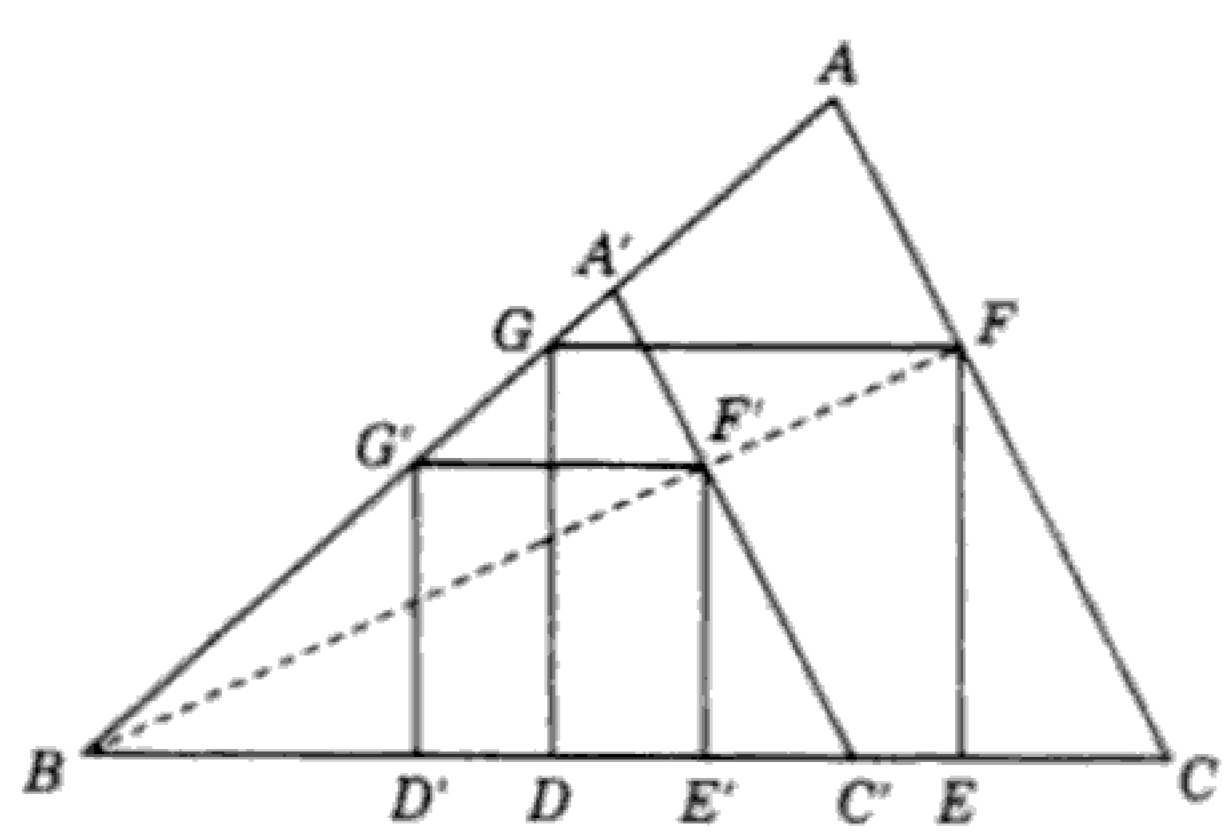


FIGURE 1.7

"60 is the circumference, 2 is the perpendicular, find the chord." "Thou, double 2 and get 4, dost thou not see? Take 4 from 20, thou gettest 16. Square 20, thou gettest 400. Square 16, thou gettest 256. Take 256 from 400, thou gettest 144. Find the square root of 144. 12, the square root, is the chord. Such is the procedure."

1.1.14 In 1936 a group of Old Babylonian tablets was lifted at Susa, about 200 miles east of Babylon. On one of the tablets the ratio of the perimeter of a regular hexagon to the circumference of the circumscribed circle is given as $57/60 + 36/3600$. Show that this leads to $3\frac{1}{8}$ as an approximation of π.

1.1.15 **(a)** A Babylonian tablet has been discovered that gives the values of $n^3 + n^2$ for $n = 1$ to 30. Make such a table for $n = 1$ to $n = 10$, and use it to find a root of the cubic equation $x^3 + 2x^2 - 3136 = 0$.

(b) A Babylonian problem of about 1800 B.C. seems to call for the solution of the simultaneous system $xyz + xy = 7/6$, $y = 2x/3$, $z = 12x$. Solve this system using the table of part (a).

1.1.16 It is known that the infinite series obtained by expanding $(a^2 + h)^{1/2}$ by the process of the binomial theorem converges to $(a^2 + h)^{1/2}$ if $-a^2 < h < a^2$.

(a) Establish the approximation formula

$$(a^2 + h)^{1/2} \approx a + \frac{h}{2a}, \qquad 0 < h < a^2.$$

(b) Take $a = 4/3$ and $h = 2/9$ in the approximation formula of part (a), and thus find a Babylonian rational approximation for $\sqrt{2}$. Find a rational approximation for $\sqrt{5}$ by taking $a = 2$, $h = 1$.

1.1.17 The Hindu mathematician, Āryabhata, wrote early in the sixth century A.D. His work is a poem of 33 couplets called the *Ganita*. Following are translations of two of the couplets: (1) The area of a triangle is the product of the altitude and half the base; half of the product of this area and the height is the volume of the solid of six edges. (2) Half the circumference multiplied by half the diameter gives the area of the circle; this area multiplied by its own square root gives the volume of the sphere. Show that, in each of these couplets, Āryabhata is correct in two dimensions but wrong in three.

1.1.18 An early Chinese work that dates probably from the second century B.C. and that had considerable influence on the development of mathematics in China was the *K'ui-ch'ang Suan-shu*, or *Arithmetic in Nine Sections*. In this work we find the empirical formula $s(c + s)/2$ for the area of a circular segment of chord c and depth s.

(a) Show how this formula might have been obtained.

(b) Obtain a correct formula.

1.1.19 **(a)** Devise an empirical procedure, using templates and a balance, for showing that the area under one arch of the cycloid curve is equal to three times the area of the generating circle. (An experiment of this nature was performed by Galileo in 1599. The first published mathematical demonstration that the area of a cycloidal arch is exactly three times that of the generating circle was furnished in 1644, by Galileo's pupil, Evangelista Torricelli.)

(b) Devise an empirical procedure, using a right circular cone, a right circular cylinder of the same radius and altitude, and some sand, for showing that the volume of a right circular cone is one third the product of its altitude and the area of its base.

(c) Devise an empirical procedure, using a circular disc, a hemisphere of the same radius, and a long piece of thick cord, for showing that the area of a sphere is equal to four times that of a great circle.

(d) Show empirically, by folding paper, that the sum of the angles of a triangle is equal to a straight angle.

1.2.1 Criticize the following inductions:

(a) Mr. Smith and Mr. Brown were both born in January, and both suffer from colds. They resigned themselves to their fate on the ground that all people born in January must suffer from colds.

(b) John had never eaten yeast, and at the beginning of the year weighed 120 pounds. For the next six months he ate three yeast cakes a day, and at the end of that time weighed 150 pounds. Therefore eating yeast makes people gain weight.

1.2.2 Criticize the following inductions:

(a) During a certain summer someone noted that the number of pounds of butter sold in New York City each month varied more or less directly with the number of inches of rainfall in New York City each month and conjectured that there must be some connection between the two.

(b) During another summer a high degree of correlation was observed between the number of people each day at a beach resort and the corresponding number of people each day taking a boat ride on a river leading from a large city to the beach resort. From this correlation it was induced that many people travel to the beach by boat.

(c) From an observation that students who made high grades in English also generally made high grades in mathematics it was induced that English helps mathematics.

(d) Statistics show that over the years our principal roads have been made wider and wider, and at the same time accidents have increased. Apparently wide roads (perhaps because they cause people to drive faster) are a cause of accidents.

1.2.3 The three altitudes of a triangle are concurrent. Would you expect the four altitudes of a tetrahedron to be concurrent? (Many theorems concerning tetrahedra were first suggested by the corresponding theorems about triangles. In this case, however, the analogy leads to an incorrect result. Only for the so-called *orthocentric tetrahedra* are the four altitudes concurrent. An orthocentric tetrahedron is a tetrahedron each edge of which is perpendicular to its opposite edge.)

1.2.4 Two lines through the vertex of an angle and symmetrical with respect to the bisector of the angle are called a pair of *isogonal conjugate lines* of the angle. There is an attractive theorem about triangles that states that if three lines through the vertices of a triangle are concurrent, then the three isogonal conjugate lines through the vertices of the triangle are also concurrent. Try to construct an analogous definition and theorem for the tetrahedron.

1.2.5 List from the following statements those that are equivalent to the statement, "All parallelograms are quadrilaterals":

(a) Every parallelogram is a quadrilateral.

(b) If a figure is a quadrilateral, then it must be a parallelogram.

(c) If a figure is not a quadrilateral, then it is not a parallelogram.

(d) If a figure is a parallelogram, then it surely is not a quadrilateral.

1.2.6 List the following statements that are equivalent to the statement, "When the sunset is red, it is sure to rain the next day":

(a) If it is raining today, then the sunset last evening must have been red.

(b) If it does not rain today, then the sunset last evening must have been red.

(c) If it does not rain today, then the sunset last evening must not have been red.

(d) Whenever it rains during the day, the sunset of the previous evening was red.

1.2.7 List the following statements that are equivalent to the statement, "It never rains in June":

(a) If it is June, it is not raining.
(b) If it is not raining, it is not June.
(c) In June it never rains.
(d) Never in June does it rain.
(e) If it is raining, it is not June.
(f) Sometimes in June it does not rain.

1.2.8 Draw diagrams illustrating each of the following types of categorical propositions:

(a) **Universal Affirmative:** All *a* are *b*.
(b) **Universal Negative:** No *a* are *b*.
(c) **Particular Affirmative:** Some *a* are *b*.
(d) **Particular Negative:** Some *a* are not *b*.

1.2.9 Test the following arguments for validity:

(a) **Premise:** All *x* are *y*.
Conclusion: All non-*x* are non-*y*.

(b) **Premises:** (1) All games played in the street are dangerous.
(2) No bull fighting is played in the street.
Conclusion: Bull fighting is not a dangerous game.

(c) **Premises:** (1) No *x* are *y*.
(2) Some *x* are *z*.
Conclusion: Some *z* are not *y*.

(d) **Premises:** (1) All trapezoids are quadrilaterals.
(2) All parallelograms are quadrilaterals.
Conclusion: All parallelograms are trapezoids.

(e) **Premises:** (1) All useful books are amusing.
(2) All books of tables are useful books.
Conclusion: All books of tables are amusing.

(f) **Premise:** All knowledge is useful.
Conclusion: No knowledge is useless.

(g) **Premises:** (1) Some doctors are not paid enough.
(2) Some doctors are college professors.
Conclusion: Some college professors are not paid enough.

(h) A student must study to deserve good grades. John studied. Therefore he deserves good grades.

(i) In a certain triangle the sum of the squares on two sides equals the square on the third. Hence the triangle is a right triangle by the Pythagorean theorem.

(j) "He that is of God heareth God's words; ye therefore hear them not, because ye are not of God." (John 8:47.)

(k) "I have tasted eggs, certainly," said Alice, . . . "but little girls eat eggs quite as much as serpents do, you know."

"I don't believe it," said the Pigeon; "but if they do, why, then they're a kind of serpent: that's all I can say." (Lewis Carroll, *Alice in Wonderland.*)

1.2.10 Let *T* stand for *true*, *F* for *false*, *V* for *valid*, and *I* for *invalid*. Try to construct simple arguments satisfying each of the following possibilities:

Premises:	*T*	*T*	*T*	*T*	*F*	*F*	*F*	*F*
Argument:	*V*	*V*	*I*	*I*	*V*	*V*	*I*	*I*
Conclusion:	*T*	*F*	*T*	*F*	*T*	*F*	*T*	*F*

1.2.11 Consider the following four statements, called, respectively, the *direct* statement, the *converse* statement, the *inverse* statement, and the *contrapositive* statement:

1. All a are b.
2. All b are a.
3. All non-a are non-b.
4. All non-b are non-a.

(a) Show that the direct and contrapositive statements are equivalent.
(b) Show that the converse and inverse statements are equivalent.
(c) Taking "All parallelograms are quadrilaterals" as the direct statement, state the converse, inverse, and contrapositive statements.

1.2.12 **(a)** The categorical statement, "All a are b," may be stated in the equivalent hypothetical form, "If w is an a, then w is a b." State the corresponding converse, inverse, and contrapositive statements in hypothetical form.
(b) State the converse, inverse, and contrapositive of "If a triangle is isosceles, then the bisectors of its base angles are equal." (The direct proposition is very easily established. The converse proposition is known as the *Steiner-Lehmus theorem* and is troublesome to establish. If one can manage to establish the inverse proposition, then, of course, by Problem 1.2.11(b), the Steiner-Lehmus theorem will follow.)

1.3.1 We are told that Thales measured the distance of a ship from shore by using the fact that two triangles are congruent if two angles and the included side of one are equal to two angles and the included side of the other. Thomas L. Heath, the historian, has conjectured that this computation was probably made by an instrument consisting of two rods AC and AD, hinged together at A. The rod AD was held vertically over a point B on shore, while rod AC was pointed toward the ship P. Then, without changing the angle DAC, the instrument was revolved about AD, and point Q noted on the ground at which arm AC was directed. What distance must be measured in order to find the distance from B to the inaccessible point P?

1.3.2 The *Eudemian Summary* says that in Pythagoras's time there were three means, the *arithmetic*, the *geometric*, and the *subcontrary*, the last name being later changed to *harmonic* by Archytas and Hippasus. We may define these three means of two positive numbers a and b as

$$A = (a + b)/2, \qquad G = \sqrt{ab}, \qquad H = 2ab/(a + b),$$

respectively.
(a) Show that $A \geqq G \geqq H$, equality holding if and only if $a = b$.
(b) Show that H is the harmonic mean between a and b if there exists a number n such that $a = H + a/n$ and $H = b + b/n$. This was the Pythagorean definition of the harmonic mean of a and b.
(c) Since 8 is the harmonic mean of 12 and 6, Philolaus, a Pythagorean of about 425 B.C., called the cube a "geometrical harmony." Explain this.

1.3.3 Tradition is unanimous in ascribing to Pythagoras the independent discovery of the theorem on the right triangle that now universally bears his name (the square on the hypotenuse of a right triangle is equal to the sum of the squares on the two legs). This theorem was known to the Babylonians of Hammurabi's time, more than a thousand years earlier, but the first general proof of the theorem may well have been given by Pythagoras. There has been considerable conjecture regarding the proof Pythagoras might have offered; the common belief is that it probably was a dissection type of proof such as is suggested by Figure 1.8. Supply the proof. (To prove that the central piece of the second dissection is actually a square

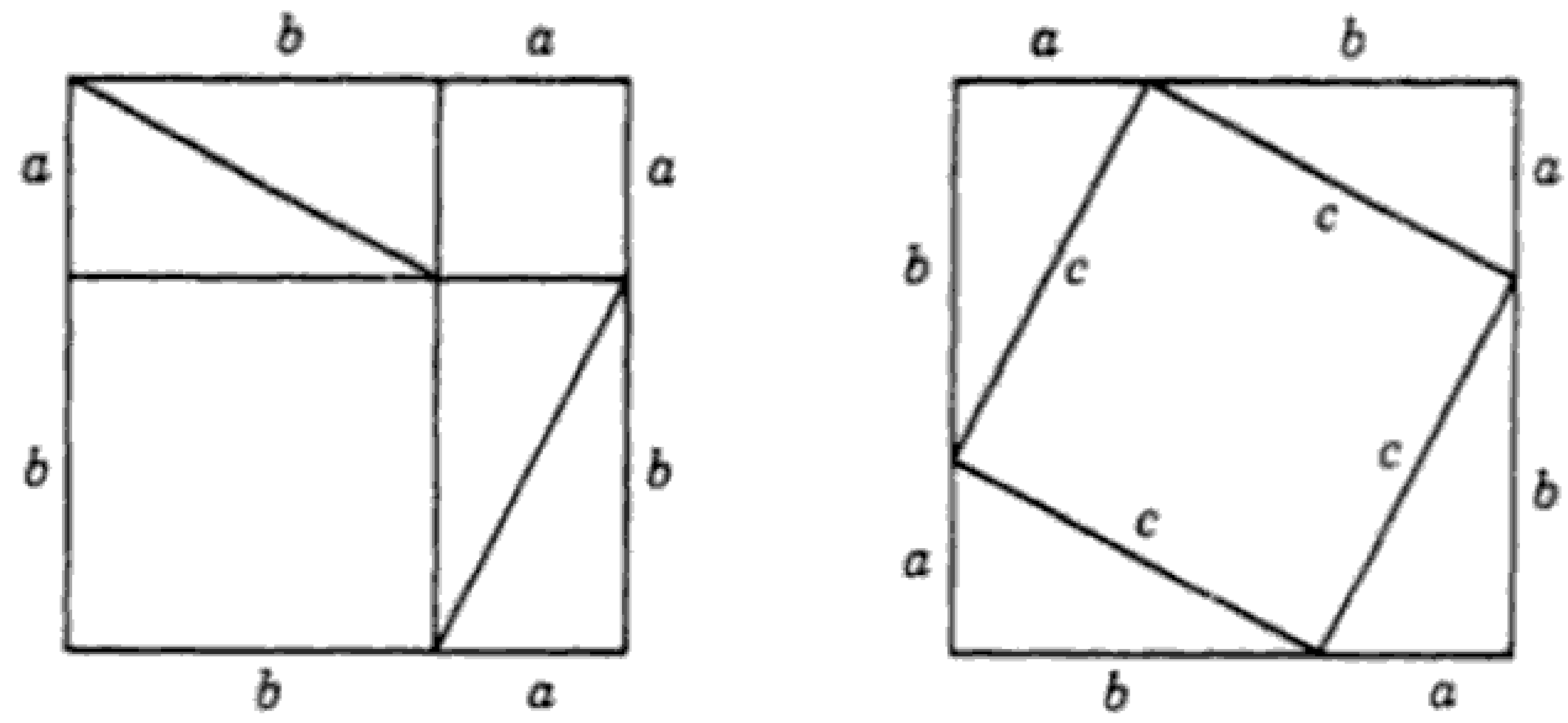


FIGURE 1.8

of side c we need to employ the fact that the sum of the angles of a right triangle is equal to two right angles. But the *Eudemian Summary* attributes this theorem for the general triangle to the Pythagoreans. Since a proof of this theorem requires, in turn, a knowledge of some properties of parallels, the early Pythagoreans are also credited with the development of that theory.)

1.3.4 State and prove the converse of the Pythagorean theorem.

1.3.5 Closely allied to the Pythagorean theorem is the problem of finding integers a, b, c to represent the legs and hypotenuse of a right triangle. Such a triple of numbers is known as a *Pythagorean triple*, and there is fairly convincing evidence that the ancient Babylonians knew how to calculate such triples.

(a) Show that, for any *odd* integer m, the three numbers m, $(m^2 - 1)/2$, and $(m^2 + 1)/2$ yield a Pythagorean triple. (The Pythagoreans have been credited with this discovery.)

(b) Show that for *any* integer m, the three numbers $2m$, $m^2 - 1$, and $m^2 + 1$ yield a Pythagorean triple. (This slight generalization of the above result is attributed to Plato. Neither formula yields all Pythagorean triples.)

1.3.6 Draw three unequal line segments. Label the longest one a, the medium one b, and take the shortest one as 1 unit. With straightedge and compasses construct line segments of lengths

(a) $a + b$ and $a - b$,

(b) ab,

(c) a/b,

(d) $\sqrt{a}$,

(e) a/n, n a positive integer.

1.3.7 Show that there can be no more than five regular polyhedra.

1.4.1 Explain how the commonly heard statement, "An axiom is a self-evident truth," reflects part of the pattern of material axiomatics.

1.4.2 As a simple example of a discourse conducted by material axiomatics, consider a certain (finite and nonempty) collection S of people and certain clubs formed among these people, a club being a (nonempty) set of people organized for some common purpose. Our basic terms are thus the *collection* S *of people* and the *clubs* to which these people belong. About these people and their clubs we assume:

Postulate 1: *Every person of* S *is a member of at least one club.*
Postulate 2: *For every pair of people of* S *there is one and only one club to which both belong.*

Definition Two clubs having no members in common are called *conjugate clubs.*

Postulate 3: *For every club there is one and only one conjugate club.*

From these postulates deduce the following theorems:

Theorem 1: *Every person of* S *is a member of at least two clubs.*
Theorem 2: *Every club contains at least two members.*
Theorem 3: S *contains at least four people.*
Theorem 4: *There exist at least six clubs.*

1.4.3 Using the same basic terms as in Problem 1.4.2, let us assume:

Postulate 1: *Any two distinct clubs have one and only one member in common.*
Postulate 2: *Every person of* S *belongs to two and only two clubs.*
Postulate 3: *There are exactly four clubs.*

From these postulates deduce the following theorems:

Theorem 1: *There are exactly six people in* S.
Theorem 2: *There are exactly three people in each club.*
Theorem 3: *For each person in* S *there is exactly one other person in* S *not in the same club.*

1.4.4 Establish the theorem about the game of tic-tac-toe cited in the text.

1.4.5 The *cigar game* is played on a rectangular table top by two players with a large stock of cigars. The two players, taking turns, lay (at each turn) a cigar on the table top so that it does not overlap any other cigar nor protrude over the edge of the table top. The last player able to place a cigar on the table top wins the game. Prove the following theorem about this game: *With proper strategy, the player who starts the game can win the game.*

1.5.1 **(a)** Prove that the straight line through the points $(0, 0)$ and $(1, \sqrt{2})$ passes through no point, other than $(0, 0)$, of the coordinate lattice.
(b) Show how the coordinate lattice may be used for finding rational approximations of $\sqrt{2}$.

1.5.2 If p is a prime number and n is an integer greater than 1, show that $\sqrt[n]{p}$ is irrational.

1.5.3 Give a purely geometric proof of the irrationality of $\sqrt{2}$.

1.5.4 The most important of Heron's geometrical works is his *Metrica*, discovered in Constantinople by R. Schöne as recently as 1896. In this work is found Heron's method of approximating the square root of a nonsquare integer, a process frequently used by computers today. If $n = ab$, then $\sqrt{n}$ is approximated by $(a + b)/2$, the approximation improving with the closeness of a to b. The method permits of successive approximations. Thus, if a_1 is a first approximation to $\sqrt{n}$, then $a_2 = (a_1 + n/a_1)/2$ is a better approximation, and $a_3 = (a_2 + n/a_2)/2$ is still better, and so on. Approximate successively, by Heron's method, $\sqrt{3}$ and $\sqrt{720}$.

1.5.5 In some problems in the Heronian collection appear the formulas

$$a, b = \frac{(r + s) \pm \{(r + s)^2 - 8rs\}^{1/2}}{2},$$

for the legs a and b of a right triangle of perimeter $2s$ and inradius r. Obtain these formulas.

1.5.6 **(a)** In his work *Catoptrica*, Heron proves, on the assumption that light travels by the shortest path, that the angles of incidence and reflection in a mirror are equal. Prove this.
(b) A man wishes to go from his house to the bank of a straight river for a pail of water, which he will then carry to his barn, on the same side of the river as his

house. Find the point on the riverbank that will minimize the distance the man must travel.

1.5.7 A regular heptagon (seven-sided polygon) cannot be constructed with straightedge and compasses. In his work *Metrica*, Heron takes, for an approximate construction, the side of the heptagon equal to the apothem of a regular hexagon having the same circumcircle. How good an approximation is this?

1.5.8 Assuming the equality of alternate interior angles formed by a transversal cutting a pair of parallel lines, prove the following:

(a) The sum of the angles of a triangle is equal to a straight angle.

(b) The sum of the interior angles of a convex polygon of n sides is equal to $n - 2$ straight angles.

1.5.9 Assuming (1) A central angle of a circle is measured by its intercepted arc, (2) The sum of the angles of a triangle is equal to a straight angle, (3) The base angles of an isosceles triangle are equal, (4) A tangent to a circle is perpendicular to the radius drawn to the point of contact, establish the following chain of theorems:

(a) An exterior angle of a triangle is equal to the sum of the two remote interior angles.

(b) An inscribed angle in a circle is measured by one half its intercepted arc.

(c) An angle formed by two intersecting chords in a circle is measured by one half the sum of the two intercepted arcs.

(d) An angle formed by two intersecting secants of a circle is measured by one half the difference of the two intercepted arcs.

(e) An angle formed by a tangent to a circle and a chord through the point of contact is measured by one half the intercepted arc.

(f) An angle formed by a tangent and an intersecting secant of a circle is measured by one half the difference of the two intercepted arcs.

(g) An angle formed by two intersecting tangents of a circle is measured by one half the difference of the two intercepted arcs.

1.5.10 Assuming the area of a rectangle is given by the product of its two dimensions, establish the following chain of theorems:

(a) The area of a parallelogram is equal to the product of its base and altitude.

(b) The area of a triangle is equal to half the product of any side and the altitude on that side.

(c) The area of a right triangle is equal to half the product of its two legs.

(d) The area of a triangle is equal to half the product of its perimeter and the radius of its inscribed circle.

(e) The area of a trapezoid is equal to the product of its altitude and half the sum of its bases.

(f) The area of a regular polygon is equal to half the product of its perimeter and the radius of its inscribed circle.

(g) The area of a circle is equal to half the product of its circumference and its radius.

2

EUCLID'S *ELEMENTS*

2.1 The Importance and Formal Nature of Euclid's *Elements*

The earliest extensively developed example of the use of the axiomatic method that has come down to us is the very remarkable and historically important *Elements* of Euclid. The production of this treatise is generally regarded as the first great landmark in the history of mathematical thought and organization, and its subsequent influence on scientific thinking can hardly be overstated.

Of Euclid himself, however, disappointingly little is known. It is from Proclus's *Commentary on Euclid, Book I*, that we obtain our most satisfying information about Euclid. He writes,

> Euclid, who put together the *Elements*, collected many of the theorems of Eudoxus. He perfected many of the theorems of Theaetetus, and also brought to irrefragable demonstration the things which were only somewhat loosely proved by his predecessors. This man lived in the time of the first Ptolemy, for Archimedes, who came immediately after the first Ptolemy, makes mention of Euclid, and furthermore, it is said that Ptolemy once asked him if there was in geometry any shorter way than that of the *Elements*, and Euclid answered that there was no royal road to geometry. It is evident, then, that Euclid came after the time of Plato, but preceded Eratosthenes and Archimedes.[1]

This statement would imply that Euclid lived about 300 B.C. Also, from other evidence, it seems quite certain that Euclid was the first professor of mathematics at the famous University of Alexandria,[2] and that he was the founder of the

[1] The quotations from Proclus and Aristotle that appear in this and the next chapter are adapted, by permission, from T. L. Heath, pp. 1, 115, 116, 117–118, 119, 121–122, 153–155, 202–203, 241–242.

[2] For an interesting exposition on Alexandria, see R. E. Langer.

distinguished and long-lived Alexandrian School of Mathematics. Even his birthplace is not known, but there is some reason to believe that he received his mathematical training in the Platonic School at Athens.

Although Euclid wrote at least ten treatises on mathematics, posterity has come to know him chiefly through his *Elements*, a monumental work written in thirteen books, or parts. This extraordinary work so quickly and so completely superseded all previous works of the same nature that now no copies remain of the earlier efforts. Apparently from its very first appearance it was accorded the highest respect, and the mere citation of Euclid's book and proposition numbers has been regarded ever since as sufficient to identify a particular theorem or construction. With the single exception of the Bible, no work has been more widely studied or edited. For more than two millennia it has dominated all teaching of geometry, and over a thousand editions of it have appeared since the first one printed in 1482. And, as the prototype of the axiomatic or postulational method, its impact on the development of mathematics has been enormous.

Proclus has clarified for us the meaning of the term *elements*. It seems that the elements of any demonstrative study are to be regarded as the leading, or key, theorems that are of wide and general use in the subject. Their function has been compared to that of the letters of the alphabet in relation to language; as a matter of fact, letters are called by the same name in Greek. The selection of the theorems to be taken as the elements of the subject requires the exercise of considerable judgment. As Proclus says,

> Now it is difficult, in each science, both to select and arrange in due order the elements from which all the rest is resolved. And of those who have made the attempt some were able to put together more and some less; some used shorter proofs; some extended their investigations to an indefinite length; some avoided the method of *reductio ad absurdum*; some avoided proportion; some contrived preliminary steps directed against those who reject the principles; and, in a word, many different methods have been invented by various writers of elements.
>
> It is essential that such a treatise should be rid of everything superfluous (for this is an obstacle to the acquisition of knowledge); it should select everything that embraces the subject and brings it to a point (for this is of supreme service to science); it must have great regard at once to clearness and conciseness (for their opposites trouble our understanding); it must aim at the embracing of theorems in general terms (for the piecemeal division of instruction into the more partial makes knowledge difficult to grasp). In all these ways Euclid's system of elements will be found to be superior to the rest.

And elsewhere, Proclus says,

> Starting from these elements, we shall be able to acquire knowledge of the other parts of this science as well, while without them it is impossible for us to get a grasp of so complex a subject, and knowledge of the rest is unattainable. As it is, the theorems which are most of the nature of principles, most simple, and most akin to the first hypotheses are here collected, in their appropriate order; and the proofs of all other propositions use these theorems

> as thoroughly known, and start from them. Thus Archimedes in the books on the sphere and cylinder, Apollonius, and all other geometers, clearly use the theorems proved in this very treatise as constituting admitted principles.

Aristotle, in his *Metaphysics*, speaks of "elements" in the same sense when he says, "Among geometrical propositions we call those 'elements' the proofs of which are contained in the proofs of all or most of such propositions."

It is no reflection on the brilliance of Euclid's work that there had been other *Elements* anterior to his own. According to the *Eudemian Summary*, Hippocrates of Chios made the first effort along this line, and the next attempt was that of Leon, who in age fell somewhere between Plato and Eudoxus. It is said that Leon's work contained a more careful selection of propositions than did that of Hippocrates, and that these propositions were more numerous and more serviceable. The textbook of Plato's Academy was written by Theudius of Magnesia and was praised as an admirable collection of elements. The geometry of Theudius seems to have been the immediate precursor of Euclid's work and was undoubtedly available to Euclid, especially if he studied in the Platonic School. Euclid was acquainted also with the important work of Theaetetus and Eudoxus. Thus it is probable that Euclid's *Elements* is, for the most part, a highly successful compilation and systematic arrangement of works of earlier writers. No doubt Euclid had to supply a number of the proofs and to perfect many others, but the chief merit of his work lies in the skillful selection of the propositions and in their arrangement into a logical sequence presumably following from a small handful of initial assumptions.

In the thirteen books that comprise Euclid's *Elements* there is a total of 465 propositions. Contrary to popular impression, many of these propositions are concerned, not with geometry, but with number theory and with elementary (geometric) algebra. Book I contains the necessary preliminary material, together with theorems on congruence, parallel lines, and rectilinear figures. Book II is devoted to geometric algebra, Book III to circles, and Book IV to the construction of regular polygons. Books V and VI contain the Eudoxian theory of proportion and its application to geometry. Books VII, VIII, and IX, containing a total of 102 propositions, deal with elementary number theory. Book X is devoted to the study of irrationals, much of the material probably from Theaetetus. The remaining three books are concerned with solid geometry. The material of Books I, II, and IV was, in all likelihood, developed by the early Pythagoreans. The material found in current American high school plane and solid geometry texts is largely that found in Euclid's Books I, III, IV, VI, XI, and XII.

Certainly there is a good deal in the contents of Euclid's *Elements* that is of considerable interest, but in the present study our concern is with the formal nature of the *Elements* rather than with its mathematical contents. In fact, the various consequences of the formal character of this great work will constitute some of our chief avenues of investigation. At the moment, we are especially interested in Euclid's conception of the axiomatic method and in the precise manner in which he applied the method to the development of his *Elements*. We consider these matters in the two following sections.

2.2 Aristotle and Proclus on the Axiomatic Method

It is a misfortune that no copy of Euclid's *Elements* has been found that actually dates from the author's own time. Modern editions of the work are based on a revision that was prepared by the Greek commentator Theon of Alexandria, who lived almost 700 years after the time of Euclid. Theon's revision was, until the early nineteenth century, the oldest edition of the *Elements* known to us. In 1808, however, when Napoleon ordered valuable manuscripts to be taken from Italian libraries and to be sent to Paris, F. Peyrard found, in the Vatican library, a tenth-century copy of an edition of Euclid's *Elements* that predates Theon's recension. A study of this older edition and a careful sifting of citations and remarks made by early commentators indicate that the introductory material of Euclid's original treatise undoubtedly underwent some editing in the subsequent revisions, but that the propositions and their proofs, except for minor additions and deletions, have remained essentially as Euclid wrote them.

Because of our lack of a copy of Euclid's original treatise, and because of the changes and additions made by later editors, it is not certain precisely what statements Euclid assumed at the start of his work, nor even how many such statements he had. Also, unfortunately, there is no known commentary by Euclid himself on the nature of the deductive organization used so successfully in his mathematical studies. It would be valuable to have Euclid's own point of view on the meaning of proof or on the significance that he attached to such terms as *definition*, *axiom*, and *postulate*. Even partially to understand Euclid, therefore, we must study the ideas held by Euclid's contemporaries. Aristotle, in particular, is an important source of information. Since Aristotle studied at Plato's Academy, his scholastic background may have been quite similar to that of Euclid.

A student of mathematics would do well to study Aristotle's *Analytica posteriora*. The following passage from that work is particularly full and enlightening:

> By the first principles of a subject I mean those the truth of which it is not possible to prove. What is *denoted* by the first terms and those derived from them is assumed; but, as regards their *existence*, this must be assumed for the principles but proved for the rest. Thus what a unit is, what a straight line is, or what a triangle is, must be assumed; and the existence of the unit and of magnitude must also be assumed, but the existence of the rest must be proved. Now of the premises used in demonstrative sciences some are peculiar to each science and others common to all, the latter being common by analogy, for of course they are actually useful insofar as they are applied to the subject-matter included under the particular science. Instances of first principles peculiar to a science are the assumptions that a line is of such and such a character, and similarly for a straight line; whereas it is a common principle, for instance, that if equals be subtracted from equals, the remainders are equal. But it is enough that each of the common principles is true as regards the particular subject-matter; in geometry, for instance, the effect will be the same even if the common principles be assumed to be true, not of everything, but only of magnitudes, and, in arithmetic, of numbers.

> Now the things peculiar to the science, the existence of which must be assumed, are the things with reference to which the science investigates the essential attributes, for example arithmetic with reference to units, and geometry with reference to points and lines. With these things it is assumed that they exist and that they are of such and such a nature. But, with regard to their essential properties, what is assumed is only the meaning of each term employed; thus arithmetic assumes the answer to the question what is meant by "odd" or "even," "a square" or "a cube," and geometry to the question what is meant by "the irrational," or "deflection," or the so-called "verging" to a point; but that there are such things is proved by means of the common principles and of what has already been demonstrated. It is similar with astronomy. For every demonstrative science has to do with three things, (1) the things which are assumed to exist, namely the subject-matter in each case, the essential properties of which the science investigates, (2) the so-called common axioms, which are the primary source of demonstration, and (3) the properties, with regard to which all that is assumed is the meaning of the respective terms used.

This remarkable passage is almost modern in its point of view. It says that a demonstrative science must start from a set of assumptions, known as the *first principles* of the subject. These first principles constitute a sort of platform of initial agreement from which the rest of the discourse can be launched by purely deductive procedures. Of these principles, according to Aristotle, some are common to all sciences and others are peculiar to the particular science being studied. The first principles common to all sciences are called *axioms* (illustrated by, "if equals be subtracted from equals, the remainders are equal"). Among the first principles, or initial assumptions, peculiar to the science being studied, we have, first of all, statements of the *existence* of the subject matter and of the fundamental things whose properties the science intends to investigate (for example, in geometry, we must assume the *existence* of "magnitude," of "points," and of "lines"). Also among the first principles peculiar to the science being studied we have the *connotation* of the technical terms employed in the discourse. That is, we must accept certain definitions concerning manifestations or attributes of our subject matter (for example, in geometry, we must assume what is *meant* by *triangle* and by *irrational*). These definitions, however, say nothing of the existence of the things defined but must be merely understood. The existence of only the subject matter and the fundamental things is assumed; the existence of all other things defined must be proved.

In addition to the definitions, one might expect to find among the first principles that are peculiar to the particular science being studied some statements concerning properties or relationships of the technical terms of the discourse. Certainly, since we cannot prove all the statements of our discourse, we anticipate the need for some such assumed statements for the purpose of getting started. About such assumptions Aristotle, again in his *Analytica posteriora*, has the following to say:

> Now anything that the teacher assumes, though it is matter of proof, without proving it himself, is a hypothesis if the thing assumed is believed by the learner, and it is moreover a hypothesis, not absolutely, but relatively to the particular pupil; but if the same thing is assumed when the learner either has no opinion on the subject or is of contrary opinion, it is a postulate. This is

> the difference between a hypothesis and a postulate; for a postulate is that which is rather contrary than otherwise to the opinion of the learner, or whatever is assumed and used without being proved, although matter for demonstration. Now definitions are not hypotheses, for they do not assert the existence or non-existence of anything, while hypotheses are among propositions. Definitions only require to be understood; a definition is therefore not a hypothesis, unless indeed it be asserted that any audible speech is a hypothesis. A hypothesis is that from the truth of which, if assumed, a conclusion can be established.

It must be admitted that Aristotle's notion of a postulate and of the role that a postulate plays in a demonstrative science is not too clear. His remarks imply that a postulate represents the assumption of a thing which is properly a subject of demonstration, and that the assumption is made without, perhaps, the assent of the student. In other words, a postulate may not appeal to a person's sense of what is right, but it has been adopted as basic in order that the work may proceed. From this point of view, then, a postulate is a first principle. In contradistinction to this, a hypothesis is an assumption believed in by the learner, and thus is introduced apparently in order to continue an argument. For example, once a theorem has been established, and hence is acceptable to the learner, that theorem may be taken as a hypothesis from which to deduce some later theorem. If we read further in the works of Aristotle we find other passages that are of special significance in comprehending the organization of Euclid's *Elements*. In several places we find that Aristotle regards an axiom as a universal assumption that is so self-evident that no sane person would question it; also he considers an axiom to be too fundamental ever to be regarded as matter for demonstration. We thus seem to have, according to Aristotle, the following four distinctions between an axiom and a postulate. An axiom is common to all sciences, whereas a postulate is related to a particular science; an axiom is self-evident, whereas a postulate is not; an axiom cannot be regarded as a subject for demonstration, whereas a postulate is properly such a subject; an axiom is assumed with the ready assent of the learner, whereas a postulate is assumed without, perhaps, the assent of the learner. Some of Aristotle's statements appear somewhat contradictory, but the interpretations just given seem especially appropriate in any attempt to understand Euclid's work.

Aristotle's characterizations of definitions, axioms, and postulates are further clarified by the following account given by Proclus in his *Commentary on Euclid, Book I.*[3]

> The compiler of elements in geometry must give separately the principles of the science, and, after that, the conclusions from those principles, not giving any account of the principles but only of their consequences. No science proves its own principles, or even discourses about them; they are treated as self-evident. . . . Thus, the first essential was to distinguish the principles from their consequences. Euclid carries out this plan practically in every book and, as a preliminary to the whole enquiry, sets out the common principles of this science. Then he divides the common principles themselves into *definitions*, *postulates*, and *axioms*. For all these are different from one another; an axiom,

[3]We have everywhere corrected a confusion that exists in the original statement caused by Proclus's consistent misuse of the term *hypothesis* for the term *definition*.

> a postulate, and a definition are not the same thing, as the inspired Aristotle has somewhere pointed out. Whenever that which is assumed and ranked as a principle is both known to the learner and convincing in itself, such a thing is an *axiom*, for example the statement that things which are equal to the same thing are also equal to one another. When, on the other hand, the pupil has not the notion of what is told him which carries conviction in itself, but nevertheless lays it down and assents to its being assumed, such an assumption is a *definition*. Thus we do not preconceive by virtue of a common notion, and without being taught, that the circle is such and such a figure, but, when we are told so, we assent without demonstration. When, again, what is asserted is both unknown and assumed even without the assent of the learner, then, he says, we call this a *postulate*, for example that all right angles are equal. This view of a postulate is clearly implied by those who have made a special and systematic attempt to show, with regard to one of the postulates, that it cannot be assented to by any one straight off. According then to the teaching of Aristotle, an axiom, a postulate, and a definition are thus distinguished.

That there was no unanimity of opinion, even among the early Greek mathematicians themselves, concerning the precise nature of, and the difference between, an axiom and a postulate is borne out by remarks made by Proclus. Proclus points out the following three distinctions advocated by various parties: (1) An axiom is a self-evident assumed statement about something, and a postulate is a self-evident assumed construction of something; thus axioms and postulates bear a relation to one another much like that between theorems and construction problems. (2) An axiom is an assumption common to all sciences, whereas a postulate is an assumption peculiar to the particular science being studied. (3) An axiom is an assumption of something that is both obvious and acceptable to the learner; a postulate is an assumption of something that is neither necessarily obvious nor necessarily acceptable to the learner. (This last is essentially the Aristotelian distinction.) Further confusion is indicated by Proclus when he points out that some preferred to call them all postulates.

In summary, then, according to the Greek conception of the axiomatic method, every demonstrable science must start from assumed first principles. These first principles consist of definitions, axioms (or common notions), and postulates. The definitions describe the technical terms used in the discourse and, except in the case of a few fundamental terms, are not meant to imply the existence of the entities described. The axioms and the postulates are initial statements that must be assumed so that the discourse may proceed. Just which of these statements should be called axioms and which postulates was a matter of varying opinion.

2.3 Euclid's Definitions, Axioms, and Postulates

Adhering to the Greek conception of the axiomatic method, we find, at the very start of Book I of Euclid's *Elements*, a list of the definitions, postulates, and common notions that are to serve as the first principles of the work. Some of the succeeding books of the work commence with additional lists of definitions. It is presumed by the author that all of the 465 propositions included in the treatise are logically deduced from these principles. For reference, we now give here the complete set of first principles for Book I essentially as furnished by T. L. Heath[4] in his translation of the distinguished Heiberg text of Euclid's *Elements*.

Definitions

1. A *point* is that which has no part.

2. A *line* is length without breadth.

3. The extremities of a line are points.

4. A *straight line* is a line which lies evenly with the points on itself.

5. A *surface* is that which has only length and breadth.

6. The extremities of a surface are lines.

7. A *plane surface* is a surface which lies evenly with the straight lines on itself.

8. A *plane angle* is the inclination to one another of two lines in a plane if the lines meet and do not lie in a straight line.

9. When the lines containing the angle are straight lines, the angle is called a *rectilinear angle*.

10. When a straight line erected on a straight line makes the adjacent angles equal to one another, each of the equal angles is called a *right angle*, and the straight line standing on the other is called a *perpendicular* to that on which it stands.

11. An *obtuse angle* is an angle greater than a right angle.

12. An *acute angle* is an angle less than a right angle.

13. A *boundary* is that which is an extremity of anything.

14. A *figure* is that which is contained by any boundary or boundaries.

15. A *circle* is a plane figure contained by one line such that all the straight lines falling upon it from one particular point among those lying within the figure are equal.

16. The particular point (of Definition 15) is called the *center* of the circle.

17. A *diameter* of a circle is any straight line drawn through the center and terminated in both directions by the circumference of the circle. Such a straight line also bisects the circle.

18. A *semicircle* is the figure contained by a diameter and the circumference cut off by it. The center of the semicircle is the same as that of the circle.

19. *Rectilinear figures* are those which are contained by straight lines, *trilateral* figures being those contained by three, *quadrilateral* those contained by four, and *multilateral* those contained by more than four straight lines.

---

[4]T. L. Heath, 1, 153–155.

20. Of the trilateral figures, an *equilateral triangle* is one which has its three sides equal, an *isosceles triangle* has two of its sides equal, and a *scalene triangle* has its three sides unequal.
21. Furthermore, of the trilateral figures, a *right-angled triangle* is one which has a right angle, an *obtuse-angled triangle* has an obtuse angle, and an *acute-angled triangle* has its three angles acute.
22. Of the quadrilateral figures, a *square* is one which is both equilateral and right-angled; an *oblong* is right-angled but not equilateral; a *rhombus* is equilateral but not right-angled; and a *rhomboid* has its opposite sides and angles equal to one another but is neither equilateral nor right-angled. Quadrilaterals other than these are called *trapezia*.
23. *Parallel* straight lines are straight lines which, being in the same plane and being produced indefinitely in both directions, do not meet one another in either direction.

Postulates

Let the following be postulated:

1. A straight line can be drawn from any point to any point.
2. A finite straight line can be produced continuously in a straight line.
3. A circle may be described with any center and distance.
4. All right angles are equal to one another.
5. If a straight line falling on two straight lines makes the interior angles on the same side together less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which the angles are together less than two right angles.

Common notions

1. Things which are equal to the same thing are also equal to one another.
2. If equals be added to equals, the wholes are equal.
3. If equals be subtracted from equals, the remainders are equal.
4. Things which coincide with one another are equal to one another.
5. The whole is greater than the part.

We observe that the first principles of Euclid's *Elements* fit quite well the Aristotelian account of definitions, postulates, and axioms as given in Section 2.2. It would also seem that Euclid strove to keep his list of postulates and axioms to an irreducible minimum. This economy, too, is in keeping with Aristotle's views, for in his *Analytica posteriora* he says, "other things being equal, that proof is the better which proceeds from the fewer postulates, or hypotheses, or propositions."

We shall pass over Euclid's definitions without much comment. Most of them probably were taken from earlier works, which would account for the fact that some terms, like *oblong*, *rhombus*, and *rhomboid*, are included but are never used anywhere in the work. It is curious that after having defined parallel lines Euclid does not give a formal definition of *parallelogram*. The existence of a parallelogram is established in I 33,[5] and in I 34 it is referred to as a *parallelogramic area*; then in I 35 this latter expression is shortened to *parallelogram*. We note that to the definition of a diameter of a circle (Definition 17) is appended the statement, "Such a straight line also bisects the circle." This addition is, of course, really a theorem (one of those attributed by Proclus in the *Eudemian Summary* to Thales), but its statement in Definition 17 is necessary in order to justify the definition of a semicircle that immediately follows. There are indications for believing that the definitions of a straight line and of a plane (Definitions 4 and 7) were original with Euclid. These definitions are not easy to understand but can be comprehended, at least partially, if we appeal to sight by considering an eye placed at an extremity of the line or the plane and looking, respectively, along the line or the plane. Other interpretations of these definitions have been given. A number of Euclid's definitions are vague and virtually meaningless; we shall return to this in the next section. The work of Heath previously referred to contains a full and valuable commentary on Euclid's definitions.

Some aspects of Euclid's postulates are of especial interest. The first three are postulates of construction, for they assert what we are permitted to draw. Since these postulates restrict constructions to only those that can be made in a permissible manner with straightedge and compasses, these instruments, so limited, have become known as *Euclidean tools*, although their use under these restrictions certainly predates Euclid. The construction of figures with only straightedge and compasses, viewed as a game played according to the rules set down in Euclid's first three postulates, has proved to be one of the most fascinating and absorbing games ever devised. One is surprised at the really intricate constructions that can be accomplished in the allowed manner, and accordingly it is hard to believe that certain seemingly simple construction problems, like that of trisecting a given arbitrary angle, for example, cannot also be so accomplished. The energetic efforts of early Greek geometers to solve legitimately some of the construction problems that are now known to be beyond the use of Euclidean tools profoundly influenced the development of much of the content of early geometry. For example, the invention of the conic sections, of many cubic and quartic curves, and of several transcendental curves resulted from this work. A later outgrowth was the development, in modern times, of portions of the theory of equations, of the theory of algebraic numbers, and of group theory. This whole line of mathematical development, so intimately tied to Euclid's first three postulates, has little connection with our present line of investigation and so will not be further considered here. We shall return to the subject, however, in Section A.2 of the Appendix.

Postulates 1 and 3 refer to existence. In other words, the existence of a straight line joining any two given points is assumed, as is the existence of a circle having any given center and radius. From applications that Euclid makes of

[5] I 33 means Proposition 33 of Book I.

Postulates 1 and 2, it appears that these postulates are meant also to imply that the straight line segment joining two points in the one case, and the produced portion in the other case, are *unique*, although it must be admitted that the postulates do not explicitly say as much. Postulate 3 may be construed as implying something in regard to the continuity and extent of the space under consideration, since the radius of the circle may be as small or as large as one desires.

Postulates 4 and 5 are quite different from the first three postulates. The meaning of Postulate 4 is certainly evident, but there has been much debate on whether it is properly classified when placed among the postulates. If it should be classified as a theorem its proof would have to be accomplished by applying one pair of adjacent right angles to another such pair, but Euclid preferred to shun, as much as possible, such *proofs by superposition*. In any event, Euclid had to place Postulate 4 before his Postulate 5, since the condition in Postulate 5 that a certain pair of interior angles be together less than two right angles would be useless unless it were first made clear that all right angles are equal.

Postulate 5, known as Euclid's *parallel postulate*, has become, as we shall see, one of the most famous statements in mathematical history. There is more evidence for the origin of this postulate with Euclid than for the origin of any of the other four. Aristotle alludes to a *petitio principii*, or a circularity in reasoning, that was involved in the theory of parallels current in his time. It is a mark of Euclid's mathematical acumen that he perceived that the only way out of the difficulty was to lay down *some* postulate as a basis for the theory of parallels that is so essential to the development of his geometry. The postulate that he formulated serves this purpose admirably and also, at the same time, furnishes a criterion for determining whether two straight lines in a figure will or will not meet if extended. This fact is an advantage of Euclid's postulate over the substitutes that were later suggested to take its place, and this advantage is actually employed in the *Elements* as early as I 44. The consequences of investigations carried on in connection with Euclid's fifth postulate proved to be very far-reaching. Not only did these investigations supply the stimulus for the development of much of the mathematics that we characterize as modern, but they led to a far deeper examination, and consequent refinement, of the axiomatic method. These investigations are therefore vital to our present study and will constitute the dramatic story of the next chapter.

Of the common notions, or axioms, there is reason to believe that the first three were given by Euclid but that the last two may have been added at a later time. Axiom 4 has been criticized on the ground that its subject matter is special rather than general and that it ought therefore to be listed as a postulate instead of as an axiom. Objections that can be raised to the method of superposition, used by Euclid with apparent reluctance to establish some of his early congruence theorems, can be at least partially met by Axiom 4. Again the student is referred to the excellent commentary given by Heath.

In conclusion, we may summarize Euclid's conception and use of the axiomatic method as follows: Every deductive system requires assumptions from which the deduction may proceed. Therefore, as initial premises, Euclid puts down five postulates, or assumed statements about his subject matter. In addition to the five postulates, Euclid lists five axioms, or common notions, that he also needs for his proofs. These axioms are not peculiar to his subject matter but are general principles valid in any field of study. Now in the postulates a number of terms occur, such as *point*, *straight line*, *right angle*, and *circle*, of which it is not certain that the reader has a precise notion. Hence some definitions are also given. These definitions are not, like the postulates, assumptions about the nature of the subject matter but are merely explanations of the meanings of the terms. Definition 10, for example, tells what a right angle is and how an angle may be identified as a right angle, but it says nothing about the existence of right angles, nor does it state what is assumed about such angles. These latter functions are left to the postulates and to deduced propositions. Thus Postulate 4 informs us that all right angles are equal, and Proposition I 11 proves that right angles exist. On the other hand, Postulate 4 gives no clue regarding the nature of a right angle, nor does it tell how the term is to be employed; it merely states a fundamental assumption about such angles. Finally, the natural order for presenting the postulates, axioms, and definitions to the student is, first, the definitions explaining the meanings of the technical terms of the discourse, next, the postulates that are so closely related to the definitions, and last, the axioms or common notions.

2.4 Some Logical Shortcomings of Euclid's Elements

It would be very surprising indeed if Euclid's *Elements*, because it is such an early and extensive application of the axiomatic methods, should be free of logical blemishes. Therefore it is no great discredit to the work that critical investigations have revealed a number of defects in its logical structure. Probably the gravest of these defects are certain tacit assumptions that are employed later in the deductions and are not granted by the first principles of the work. This danger exists in any deductive study when the subject matter is overly familiar to the author. Usually a thorough grasp of the subject matter in a field of human endeavor is regarded as an indispensable prerequisite to serious work, but in developing a deductive system such knowledge can be a definite disadvantage unless proper precautions are taken.

A deductive system differs from a mere collection of statements in that it is organized in a very special way. The key to the organization lies in the fact that all statements of the system other than the original assumptions must be deducible from these initial hypotheses, and that if any additional assumptions should creep into the work the desired organization is not realized. Now anyone formulating a deductive system knows more about his subject matter than just the initial assumptions he wishes to employ. He has before him a set of statements belonging to his subject matter, some of which he selects for postulates and the rest of which he presumably deduces from his postulates as theorems. But with a large body of information before one, it is very easy to employ in the proofs some piece of this information that is not embodied in the postulates. Any piece of information used in this way may be so apparently obvious or so seemingly elementary that it is assumed unconsciously. Such a tacit assumption, of course, spoils the rigidity of the organization of the deductive system. Moreover, should that piece of information involve some misconception, its introduction may lead to results that not only do not strictly follow from the postulates but that may actually contradict some previously established theorem. Herein, then, lies the pitfall of too great a familiarity with the subject matter of the discourse; at all times in building up a deductive system one must proceed with the appearance of being completely ignorant of the developing material. This does not mean that in building up a deductive system one refrains from making any use of one's intuitive appreciation of the significance of the axioms and of possible interpretations of the primitive terms. On the contrary one makes *full* use of these things, but only to conjecture possible theorems and possible avenues of investigation. In the actual establishment of these theorems and in the actual development of these avenues of investigation, one must be careful to proceed only in terms of the accepted assumptions.

The tacit assumption by Euclid of something that is not contained in his first principles is exemplified in the very first deduced proposition of the *Elements*. In order to examine the difficulty we shall quote Proposition I 1 verbatim from Heath's translation.[6]

> *On a given finite straight line to construct an equilateral triangle* [see Figure 2.1].
> Let *AB* be the given finite straight line.
> Thus it is required to construct an equilateral triangle on the straight line *AB*.
> With center *A* and distance *AB*, let the circle *BCD* be described. [Postulate 3]
> Again, with center *B* and distance *BA*, let the circle *ACE* be described. [Postulate 3]
> And from the point *C*, in which the circles cut one another, to the points *A*, *B*, let the straight lines *CA*, *CB* be joined. [Postulate 1]
> Now, since the point *A* is the center of the circle *CDB*, *AC* is equal to *AB*. [Definition 15]
> Again, since the point *B* is the center of the circle *CAE*, *BC* is equal to *BA*. [Definition 15]

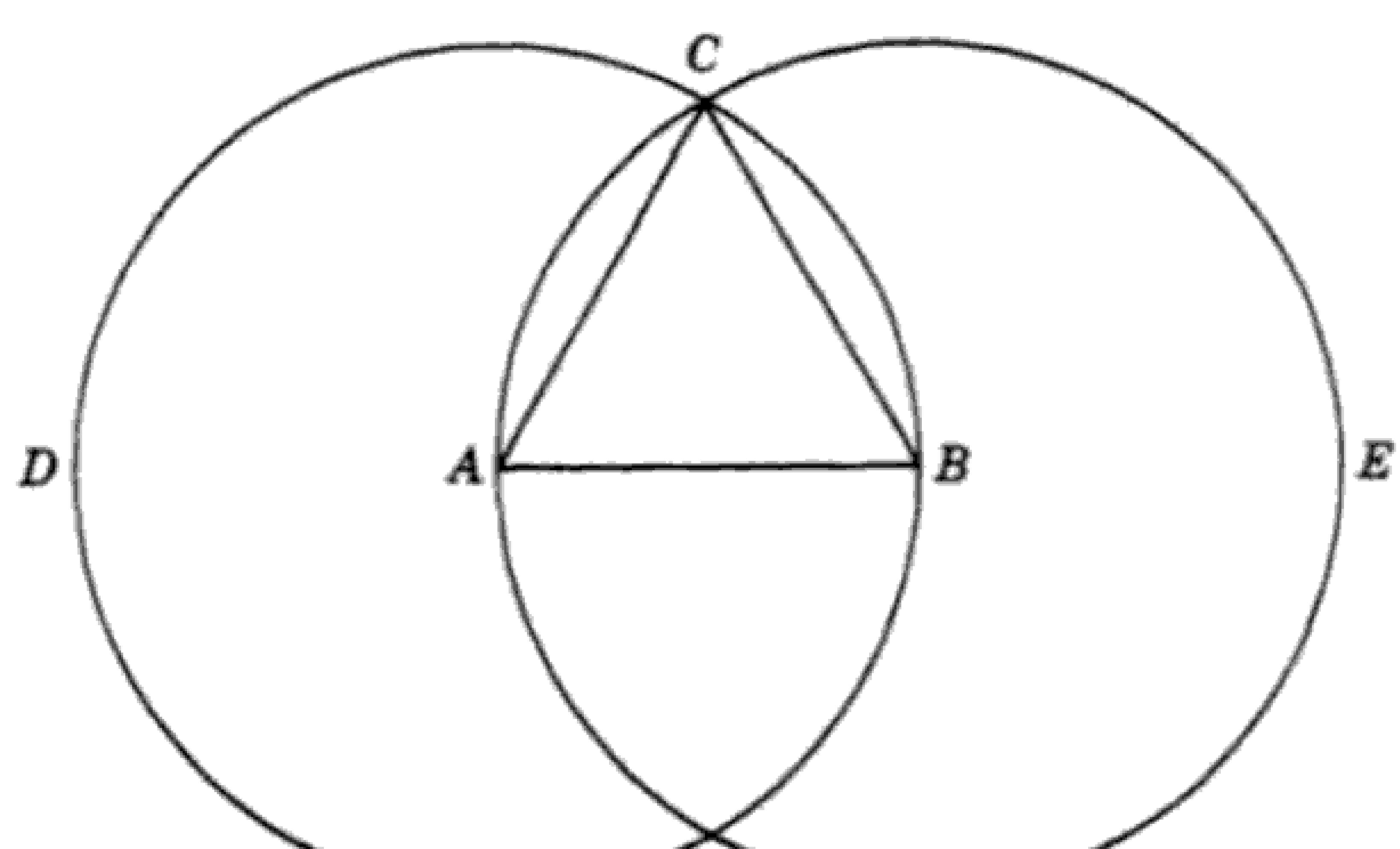


FIGURE 2.1

[6] T. L. Heath, 241, 242.

> But CA was also proved equal to AB; therefore each of the straight lines CA, CB is equal to AB. And things which are equal to the same thing are also equal to one another; therefore CA is also equal to CB. [Axiom 1]
>
> Therefore the three straight lines CA, AB, BC are equal to one another.
>
> Therefore the triangle ABC is equilateral; and it has been constructed on the given finite straight line AB.
>
> (Being) what it was required to do.

Now the construction of the two circles in this demonstration is certainly justified by Postulate 3, but there is nothing in Euclid's first principles which explicitly guarantees that the two circles shall intersect in a point C, and that they will not, somehow or other, slip through each other with no common point. The existence of this point, then, must be either postulated or proved, and it can be shown that Euclid's postulates are insufficient to permit the latter (see Problem 2.4.3). Only by the introduction of some additional assumption can the existence of the point C be established. Therefore the proposition does not follow from Euclid's first principles, and the proof of the proposition is invalid.

The fallacy here lies not in assuming something contrary to our concept of circles but in assuming something that is not implied by our accepted first principles. This is an example where the tacit assumption is so evident and elementary that there does not appear to be any assumption. The fallacy is a subtle one, but had Euclid known nothing more about circles than what his first principles say of them, he certainly could not have fallen into this error.

What is needed here is some additional postulate that will guarantee that the two circles concerned will intersect. Postulate 5 gives a condition under which two straight lines will intersect. We need similar postulates telling when two circles will intersect and when a circle and a straight line will intersect. What is essentially involved here is the continuity of circles and straight lines, and in modern treatments of geometry the existence of the desired points of intersection is taken care of by some sort of continuity postulate.

Another tacit assumption made by Euclid is that the straight line is of infinite extent. Although Postulate 2 asserts that a straight line may be produced indefinitely, it does not necessarily imply that a straight line is infinite in extent but merely that it is endless, or boundless. The arc of a great circle joining two points on a sphere may be produced indefinitely along the great circle, making the prolonged arc endless, but certainly it is not infinite in extent. Now it is conceivable that a straight line may behave similarly, and that after a finite prolongation it, too, may return on itself. It was the great German mathematician Bernhard Riemann (1826–1866) who, in his famous probationary lecture, *Über die Hypothesen welche der Geometrie zu Grunde liegen*, of 1854, distinguished between the boundlessness and the infinitude of straight lines. There are numerous occasions where Euclid unconsciously assumes the infinitude of a straight line. Let us briefly consider, for example, Proposition I 16:

> *In any triangle, if one of the sides be produced, the exterior angle is greater than either of the interior and opposite angles.*

A précis of Euclid's proof runs as follows. Let ABC (Figure 2.2) be the given triangle, with BC produced to D. Let E be the midpoint of AC. Draw BE and extend it its own length to F. Draw CF. Then triangles BEA and FEC can easily

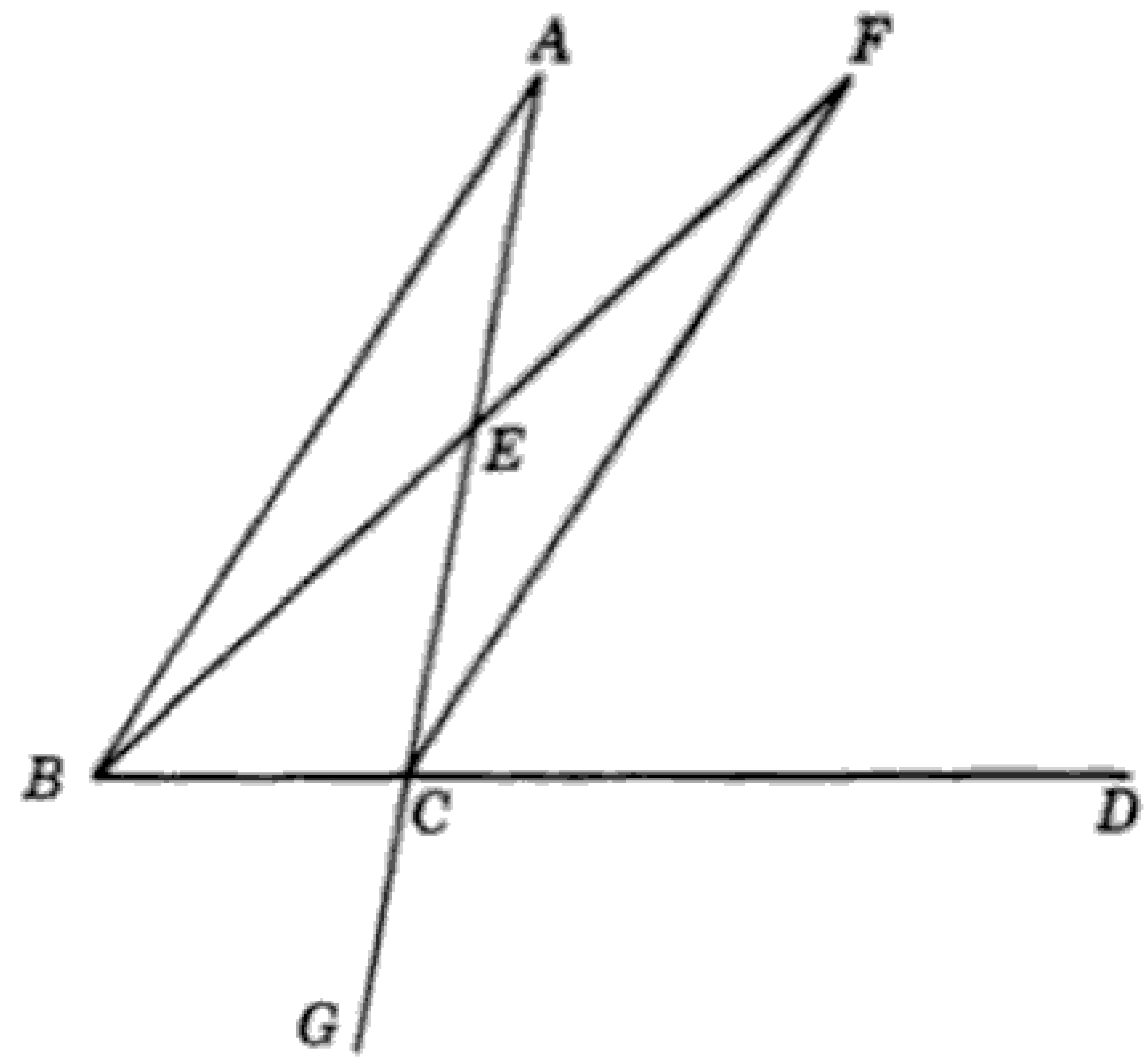


FIGURE 2.2

be shown to be congruent, whence $\measuredangle FCE = \measuredangle BAC$. But $\measuredangle ACD > \measuredangle FCE$, whence $\measuredangle ACD > \measuredangle BAC$. By producing AC to G, we may similarly show that $\measuredangle BCG$, which is equal to $\measuredangle ACD$, is also greater than $\measuredangle ABC$.

Now if a straight line should return on itself, like the great circle arc considered here, BF may be so long that F will coincide with B or lie on the segment BE. Should this be the case, the proof would certainly fail. The author has been misled by his visual reference to the figure rather than to the principles that should be the basis of his argument. Clearly, then, to make the proof universally valid we must either prove or postulate the infinitude of straight lines.

One can point out many other tacit assumptions that, like the preceding one, were unconsciously made by Euclid and that vitiate the true deductive character of his work. For example, in Proposition I 21, Euclid unconsciously assumes that if a straight line enters a triangle at a vertex it must, if sufficiently produced, intersect the opposite side. It was Moritz Pasch (1843–1930) who recognized the necessity of a postulate to take care of this situation. Again, Euclid makes no provision for *linear order*, and his concept of "betweenness" is without any postulational foundation, with the result that paradoxes are possible. We have already pointed out that Postulate 1, which guarantees the existence of at least one straight line joining two points A and B, probably was meant to imply uniqueness of this line, but the postulate fails to assert so much. Also, the objections that can be raised against the principle of superposition, employed in many early popular textbooks, are only partially met by Euclid's Axiom 4.

In short, the truth of the matter is that Euclid's first principles are simply not sufficient for the derivation of all of the 465 propositions of the *Elements*. In particular, the set of postulates needs to be considerably amplified. The work of perfecting Euclid's initial assumptions, so that all of his geometry can rigorously follow, occupied mathematicians for more than two thousand years. Not until the end of the nineteenth century and the early part of the twentieth century, after the foundations of geometry had been subjected to an intensive study, were satisfactory sets of postulates supplied for Euclidean plane and solid geometry. The history of this struggle is of major concern in our present study, and in following it we shall encounter the device that mathematicians contrived for avoiding the pitfall, into which Euclid so often fell, of overfamiliarity with the subject matter.

Not only is Euclid's work marred by numerous tacit assumptions, but some of his preliminary definitions are also open to criticism. Euclid, following the Greek pattern of material axiomatics, makes some sort of attempt to define, or at least to explain, all the terms of his discourse. Actually, it is as impossible to define *explicitly* all of the terms of a discourse as it is to prove all of the statements of the discourse, for a term must be defined by means of other terms, and these other terms by means of still others, and so on. In order to get started, and to avoid circularity of definition where term x is defined by means of term y, and then later term y by means of term x, one is forced to set down at the very start of the discourse a collection of primitive, or basic, terms whose meanings are not to be questioned. All subsequent terms of the discourse must be defined, ultimately, by means of these initial primitive ones. The postulates of the discourse are, then, in final analysis, assumed statements about the primitive terms. From this point of view, the primitive terms may be regarded as defined *implicitly*, in the sense that they are any things or concepts that satisfy the postulates, and this implicit definition is the only kind of definition that the primitive terms can receive.

In Euclid's development of geometry the terms *point* and *line*, for example, could well have been included in a set of primitive terms for the discourse. At any rate, Euclid's definition of a point as "that which has no part" and of a line as "length without breadth" are easily seen to be circular and therefore, from a logical viewpoint, woefully inadequate. One distinction between the Greek conception and the modern conception of the axiomatic method lies in this matter of primitive terms; in the Greek conception there is no listing of the primitive terms. The excuse for the Greeks is that to them geometry was not just an abstract study but was an attempted logical analysis of idealized physical space. Points and lines were, to the Greeks, idealizations of very small particles and of very thin threads. It is this idealization that Euclid attempts to express in his two initial definitions.

Other differences between the Greek and the modern views of the axiomatic method will be discussed in a later chapter.

2.5 The End of the Greek Period and the Transition to Modern Times[7]

Very little in the further development of the axiomatic method took place after Euclid until relatively modern times. We must mention, however, the brilliant exploitation of the method by Archimedes (*ca.* 287–212 B.C.), one of the greatest mathematicians of all time, and certainly *the* greatest of antiquity. Although Archimedes lived most of his long life in the Greek city of Syracuse, on the island of Sicily, it seems that he studied for a time at the University of Alexandria. He was thoroughly schooled in the Euclidean tradition, and he left deep imprints on

[7]This section is largely skimmed from the appropriate places in H. Eves [1].

both geometry and mechanics. Archimedes' works are masterpieces of mathematical exposition and resemble to a remarkable extent, because of their high finish, economy of presentation, and rigor in demonstration, the articles found in present-day research journals. It is interesting that Archimedes employed the axiomatic method in his writings on theoretical mechanics, as well as in his purely geometrical studies, always laying down the first principles of the work and then deducing a sequence of propositions. Thus, in his treatise *On Plane Equilibriums*, Archimedes establishes twenty-five theorems of mechanics on the basis of three simple postulates suggested by common experience. The postulates are as follows:

1. *Equal weights at equal distances balance; equal weights at unequal distances do not balance but incline toward the weight that is at the greater distance.*
2. *If, when weights at certain distances balance, something is added to one of the weights, equilibrium will not be maintained, but there will be inclination on the side of the weight to which the addition was made; similarly, if anything is taken away from one of the weights, there will be inclination on the side of that weight from which nothing was taken.*
3. *When equal and similar plane figures coincide if placed on one another, their centroids similarly coincide; and in figures that are unequal but similar, the centroids will be similarly situated.*

From these simple postulates Archimedes locates, for example, the centroid of any parabolic segment and of any portion of a parabola lying between two parallel chords. Problems of this sort would today be worked out by means of the integral calculus.

Again, in his work *On Floating Bodies*, Archimedes rests the establishment of the nineteen propositions of the work on two fundamental postulates. This treatise is the first recorded application of mathematics to hydrostatics, and it begins by developing those familiar laws of hydrostatics that nowadays are encountered in an elementary physics course. The treatise then goes on to consider several rather difficult problems, culminating with a remarkable investigation of the positions of rest and of stability of a right segment of a paraboloid of revolution floating in a fluid. Not until the sixteenth-century work of Simon Stevin did the science of statics and the theory of hydrodynamics appreciably advance beyond the points reached by Archimedes. It is worthy of note that these early researches in theoretical physics were developed by the use of the axiomatic method.

A geometrical assumption explicitly stated by Archimedes in his work *On the Sphere and Cylinder* deserves special mention; it is one of the five postulates assumed at the start of Book I of the work and it has become known as the *postulate of Archimedes*. A simple statement of the postulate is as follows: *Given two unequal linear segments, there is always some finite multiple of the shorter one which is longer than the other.* In some modern treatments of geometry this postulate serves as part of the postulational basis for introducing the concept of continuity. It is a matter of interest that in the nineteenth and twentieth centuries geometric systems were constructed that denied the Archimedean postulate, thus giving rise to so-called non-Archimedean geometries. Although named after Archimedes, this postulate had been considered earlier by Eudoxus.

There were other able Greek mathematicians in ancient times after Euclid besides Archimedes—for example, Apollonius, Eratosthenes, Menelaus, Claudius Ptolemy, Heron, Diophantus, and Pappus—but these men did little to advance the development of the axiomatic method and so have slight connection with our present study. After Pappus, who flourished toward the end of the third century A.D., Greek mathematics practically ceased as a living study, and thenceforth merely its memory was perpetuated by minor writers and commentators, such as Theon and Proclus. This closing period of ancient times was dominated by Rome. One Greek center after another had fallen before the power of the Roman armies; in 146 B.C. Greece had become a province of the Roman Empire, although Mesopotamia was not conquered until 65 B.C., and Egypt held out until 30 B.C. The economic structure of the empire was based essentially on agriculture and an increasing use of slave labor. Conditions proved more and more stifling to original scientific work, and a gradual decline in creative thinking set in. The eventual collapse of the slave market, with its disastrous effect on Roman economy, found science reduced to a mediocre level. The famous Alexandrian school gradually faded with the breakup of ancient society, and finally, in A.D. 641, Alexandria was taken by the Arabs, who put the torch to what the Christians had left. The long and glorious era of Greek mathematics was over.

The period starting with the fall of the Roman Empire in the middle of the fifth century and extending into the eleventh century is known as Europe's Dark Ages, for during this period civilization in western Europe reached a very low ebb. Schooling became almost nonexistent, Greek learning all but disappeared, and many of the arts and crafts bequeathed by the ancient world were forgotten. Only the monks of the Christian monasteries, and a few cultured laymen, preserved a slender thread of Greek and Latin learning. The period was marked by great physical violence and intense religious faith. The old social order gave way, and society became feudal and ecclesiastical.

The Romans had never taken to abstract mathematics but had contented themselves with merely a few practical aspects of the subject that were associated with commerce and civil engineering. With the fall of the Roman Empire and the subsequent closing of much of east–west trade and the abandonment of state engineering projects, even these interests waned, and it is no exaggeration to say that very little in mathematics, beyond the development of the Christian calendar, was accomplished in the West during the whole of the half millennium covered by the Dark Ages.

During this bleak period of learning the people of the east, especially the Hindus and the Arabs, became the major custodians of mathematics. However, the Greek concept of rigorous thinking—in fact, the very idea of proof—seemed distasteful to the Hindu way of doing things. Although the Hindus excelled in computation, contributed to the devices of algebra, and played an important role in developing our present positional numeral system, they produced nothing of importance as far as basic methodology is concerned. Hindu mathematics of this period is largely empirical and lacks those outstanding Greek characteristics of clarity and logicality in presentation and of insistence on rigorous demonstration.

The spectacular episode of the rise and decline of the Arabian empire occurred during the period of Europe's Dark Ages. Within a decade following Mohammed's flight from Mecca to Medina in A.D. 622, the scattered and disunited tribes of the Arabian peninsula were consolidated by a strong religious fervor into a powerful nation. Within a century, force of arms had extended the Moslem rule and influence over a territory reaching from India, through Persia, Mesopotamia, northern Africa, and into Spain. Of considerable importance for the preservation of much of world culture was the manner in which the Arabs seized on Greek and Hindu erudition. The Baghdad caliphs not only governed wisely and well but many became patrons of learning and invited distinguished scholars to their courts. Numerous Hindu and Greek works in astronomy, medicine, and mathematics were industriously translated into the Arabic tongue and thus were saved until later European scholars were able to retranslate them into Latin and other languages. But for the work of the Arabian scholars a great part of Greek and Hindu science would have been irretrievably lost over the long period of the Dark Ages.

Not until the latter part of the eleventh century did Greek classics in science and mathematics begin once again to filter into Europe. There followed a period of transmission during which the ancient learning preserved by Moslem culture was passed on to the western Europeans through Latin translations made by Christian scholars traveling to Moslem centers of learning, and through the opening of western European commercial relations with the Levant and the Arabian world. The loss of Toledo by the Moors to the Christians in 1085 was followed by an influx of Christian scholars to that city to acquire Moslem learning. Other Moorish centers in Spain were infiltrated, and the twelfth century became, in the history of mathematics, a century of translators. One of the most industrious translators of the period was Gherardo of Cremona, who translated into Latin more than ninety Arabian works, among which were Ptolemy's *Almagest* and Euclid's *Elements*. At the same time Italian merchants came in close contact with eastern civilization, thereby picking up useful arithmetical and algebraical information. These merchants played an important part in the European dissemination of the Hindu-Arabic system of numeration.

The thirteenth century saw the rise of the universities at Paris, Oxford, Cambridge, Padua, and Naples. Universities were to become potent factors in the development of mathematics, since many mathematicians associated themselves with one or more such institutions. During this century Campanus made a Latin translation of Euclid's *Elements*, which later, in 1482, became the first printed version of Euclid's great work.

The fourteenth century was a mathematically barren one. It was the century of the Black Death, which swept away more than a third of the population of Europe; and during this century the Hundred Years' War, with its political and economic upheavals in northern Europe, got well under way.

The fifteenth century witnessed the beginning of the European Renaissance in art and learning. With the collapse of the Byzantine Empire, culminating in the fall of Constantinople to the Turks in 1453, refugees flowed into Italy, bringing with them treasures of Greek civilization. Many Greek classics, up to that time known only through the often inadequate Arabic translations, could now be studied from original sources. Also, the middle of the century witnessed the invention of printing, which revolutionized the book trade and enabled knowledge to be disseminated at an unprecedented rate. Mathematical activity in this century was largely centered in the Italian cities and in the central European cities of Nuremberg, Vienna, and Prague, and it concentrated on arithmetic, algebra, and trigonometry, under the practical influence of trade, navigation, astronomy, and surveying.

In the sixteenth century the development of arithmetic and algebra continued, the most spectacular mathematical achievement of the century being the discovery, by Italian mathematicians, of the algebraic solution of cubic and quartic equations. In 1572 Commandino made a very important Latin translation of Euclid's *Elements* from the Greek. This translation served as a basis for many subsequent translations, including a very influential work by Robert Simson, from which, in turn, so many English editions were derived.

The seventeenth century proved to be particularly outstanding in the history of mathematics. Early in the century Napier revealed his invention of logarithms, Harriot and Oughtred contributed to the notation and codification of algebra, Galileo founded the science of dynamics, and Kepler announced his laws of planetary motion. Later in the century Desargues and Pascal opened a new field of pure geometry, Descartes launched modern analytic geometry, Fermat laid the foundations of modern number theory, and Huygens made distinguished contributions to the theory of probability and other fields. Then, toward the end of the century, after many mathematicians had prepared the way, the epoch-making creation of the calculus was made by Newton and Leibniz. Thus, during the seventeenth century, many new and vast fields were opened for mathematical investigation. The dawn of modern mathematics was at hand, and it was perhaps inevitable that sooner or later some aspect of the axiomatic method itself should once again claim the attention of researchers.

PROBLEMS

2.1.1 Which of the following two theorems should more likely appear among the "elements" of a course in plane geometry, and why? (1) The three altitudes of a triangle, produced if necessary, meet in a point. (2) The sum of the three angles of a triangle is equal to two right angles.

2.1.2 A mathematics instructor is going to present the subject of geometric progressions to his college algebra class. After defining this type of progression, what theorems about geometric progressions should the instructor offer as the "elements" of the subject?

2.1.3 Imagine yourself building up an elementary treatment of trigonometric identities. Which identities would you select for the "elements" of your treatment, and in what order would you arrange them?

2.1.4 As an illustration of nongeometrical material found in Euclid's *Elements*, let us consider the *Euclidean algorithm*, or process, for finding the greatest common integral divisor (g.c.d.) of two positive integers. The process is found at the start of Euclid's Book VII, although perhaps it was known before Euclid's time. This algorithm is basic to several developments in modern mathematics. Stated in the form of a rule, the process is this: *Divide the larger of the two positive integers by the smaller one. Then divide the divisor by the remainder. Continue this process, of* *dividing the last divisor by the last remainder, until the division is exact. The final divisor is the sought g.c.d. of the two original positive integers.*

(a) Find, by the Euclidean algorithm, the g.c.d. of 5913 and 7592.

(b) Find, by the Euclidean algorithm, the g.c.d. of 1827, 2523, and 3248.

(c) Prove that the Euclidean algorithm does lead to the g.c.d.

(d) Let h be the g.c.d. of the positive integers a and b. Show that there exist integers p and q (not necessarily positive) such that $pa + qb = h$.

(e) Find p and q for the integers of part (a).

(f) Prove that a and b are relatively prime if and only if there exist integers p and q such that $pa + qb = 1$.

2.1.5 **(a)** Prove, using Problem 2.1.4 (f), that if p is a prime and divides the product uv then either p divides u or p divides v.

(b) Prove, from part (a), the "fundamental theorem of arithmetic": *Every integer greater than* 1 *can be uniquely factored into a product of primes.* This is essentially Proposition IX 14 of Euclid's *Elements.*

2.1.6 The fundamental theorem of arithmetic says that, for any given positive integer a, there are unique non-negative integers $a_1, a_2, a_3, \ldots$, only a finite number of which are different from zero, such that

$$a = 2^{a_1}3^{a_2}5^{a_3}\ldots,$$

where 2, 3, 5, . . . are the consecutive primes. This suggests a useful notation. We shall write

$$a = (a_1, a_2, a_3 \ldots, a_n),$$

where a_n is the last nonzero exponent. Thus we have $12 = (2, 1)$, $14 = (1, 0, 0, 1)$, $27 = (0, 3)$, and $360 = (3, 2, 1)$.

Prove the following theorems:

(a) $ab = (a_1 + b_1, a_2 + b_2, \ldots)$.

(b) b is a divisor of a if and only if $b_i \leqq a_i$ for each i.

(c) The number of divisors of a is $(a_1 + 1)(a_2 + 1)\cdots(a_n + 1)$.

(d) A necessary and sufficient condition for a number n to be a perfect square is that the number of divisors of n be odd.

(e) Set g_i equal to the smaller of a_i and b_i if $a_i \neq b_i$ and equal to either a_i or b_i if $a_i = b_i$. Then $g = (g_1, g_2, \ldots)$ is the g.c.d. of a and b.

(f) If a and b are relatively prime and b divides ac, then b divides c.

(g) If a and b are relatively prime and if a divides c and b divides c, then ab divides c.

(h) Show that $\sqrt{2}$ and $\sqrt{3}$ are irrational.

2.1.7 Prove the famous Proposition IX 20 of Euclid's *Elements: The number of prime numbers is infinite.*

2.1.8 A number is said to be perfect if it is the sum of its proper divisors. For example, 6 is a perfect number, since $6 = 1 + 2 + 3$. The last proposition of the ninth book of Euclid's *Elements* proves that *if* $2^n - 1$ *is a prime number, then* $2^{n-1}(2^n - 1)$ *is a perfect number.* The perfect numbers given by Euclid's formula are even numbers, and it has been shown that every even perfect number must be of this form. The existence or nonexistence of odd perfect numbers is one of the celebrated unsolved problems in number theory. There is no number of this type having less than 100 digits.

(a) Show that in Euclid's formula for perfect numbers, n must be prime.

(b) What are the first four perfect numbers given by Euclid's formula?

(c) Prove that the sum of the reciprocals of *all* the divisors of a perfect number is equal to 2.

2.2.1 Discuss Euclid's axioms and postulates, as listed in Section 2.3, in relation to the three distinctions that, according to Proclus, were advocated by various early Greeks.

2.3.1 How does the modern definition of a circle differ from Euclid's definition?

2.3.2 "Prove" Euclid's Postulate 4 by the method of superposition.

2.3.3 One should understand precisely the intention of Euclid's Postulate 3. When Euclid says that "a circle may be described with any center and distance," he means that *a circle may be described with any point as center and having any straight line segment radiating from this center as a radius.* It follows that the Euclidean compasses differ from our modern compasses, for with the modern compasses we are permitted to draw a circle having any point A as center and any segment BC as radius. In other words, we are permitted to transfer the distance BC to the center A, using the compasses as dividers. The Euclidean compasses, on the other hand, may be supposed to collapse if either leg is lifted from the paper.

A student reading Euclid's *Elements* for the first time might experience surprise at the opening propositions of Book I. The first three propositions are the construction problems:

1. *To describe an equilateral triangle upon a given finite straight line.*
2. *From a given point to draw a straight line equal to a given straight line.*
3. *From the greater of two given straight lines to cut off a part equal to the lesser.*

These three constructions are trivial with straightedge and *modern* compasses but require some ingenuity with straightedge and *Euclidean* compasses.

(a) Solve Proposition 1 of Book I with Euclidean tools.

(b) Solve Proposition 2 of Book I with Euclidean tools.

(c) Solve Proposition 3 of Book I with Euclidean tools.

(d) Show that Proposition 2 of Book I proves that the straightedge and *Euclidean* compasses are equivalent to the straightedge and *modern* compasses.

2.4.1 If an assumption tacitly made in a deductive development should involve a misconception, its introduction may lead not only to a result that does not follow from the postulates of the deductive system but to one that may actually contradict some previously established theorem of the system. From this point of view, criticize the following three geometrical paradoxes:

(a) *To prove that any triangle is isosceles.*

Let ABC be any triangle (see Figure 2.3). Draw the bisector of $\measuredangle C$ and the perpendicular bisector of side AB. From their point of intersection E, drop perpendiculars EF and EG on AC and BC, respectively, and draw EA and EB. Now right triangles CFE and CGE are congruent, since each has CE as hypotenuse and since $\measuredangle FCE = \measuredangle GCE$. Therefore $CF = CG$. Again, right triangles EFA and EGB are congruent, since leg EF of one equals leg EG of the other (any point E on the bisector of an angle C is equidistant from the sides of the angle) and since hypotenuse EA of one equals hypotenuse EB of the other (any point E on the perpendicular bisector of a line segment AB is equidistant from the extremities of that line segment). Therefore $FA = GB$. It now follows that $CF + FA = CG + GB$, or $CA = CB$, and the triangle is isosceles.

(b) *To prove that a right angle is equal to an obtuse angle.*

Let $ABCD$ be any rectangle (see Figure 2.4). Draw BE outside the rectangle and equal in length to BC, and hence to AD. Draw the perpendicular bisectors of DE and AB; since they are perpendicular to

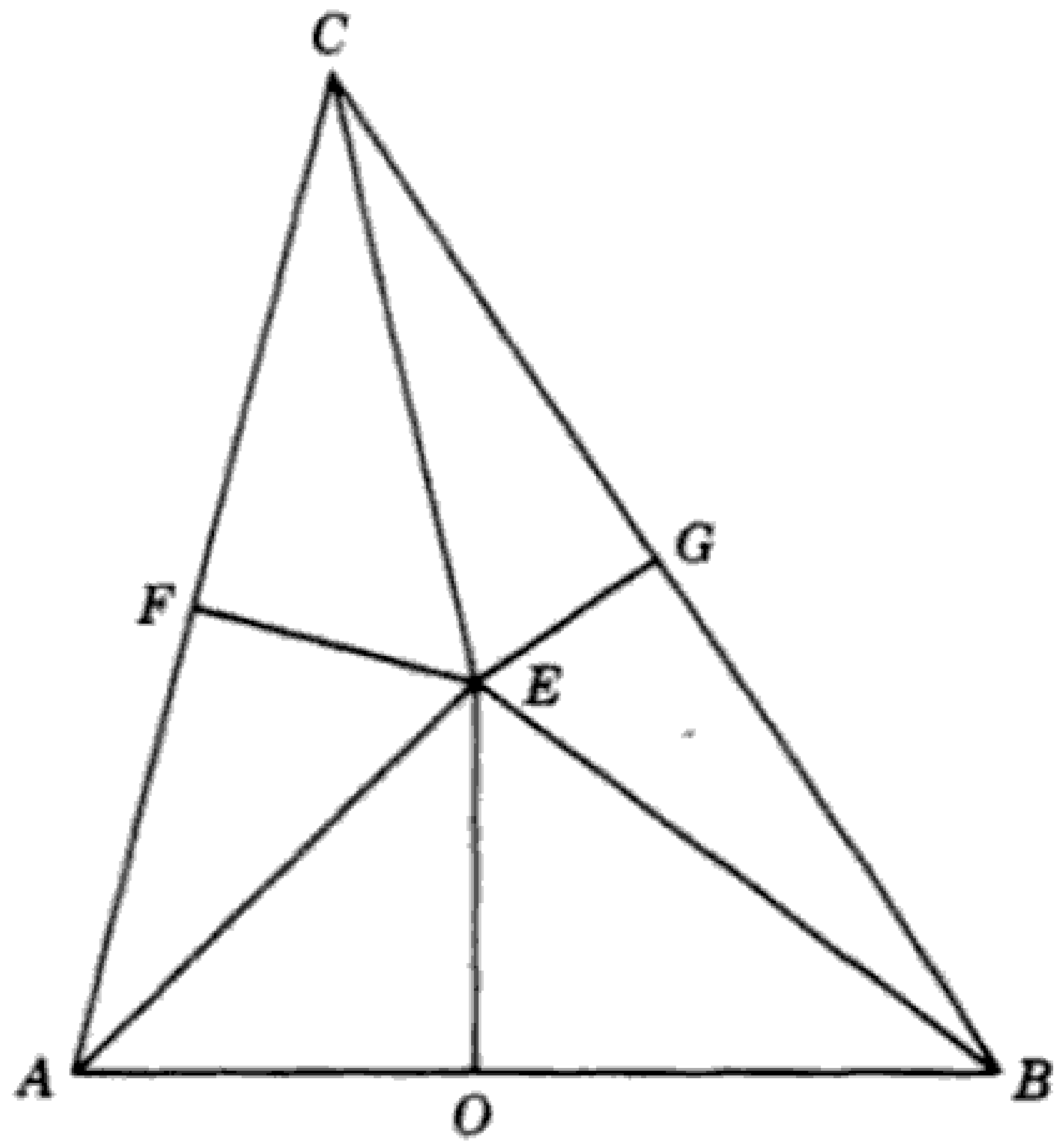


FIGURE 2.3

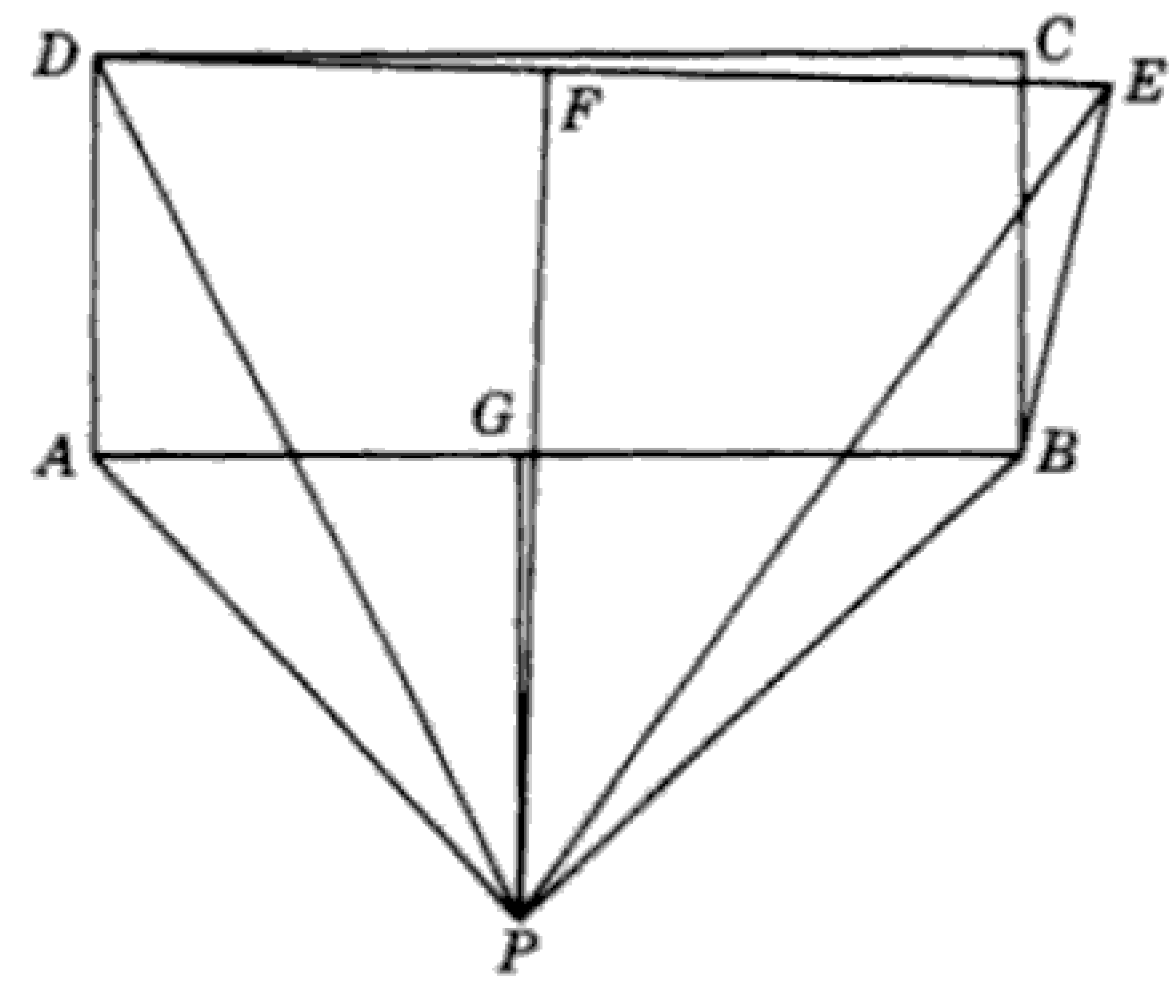


FIGURE 2.4

nonparallel lines, they must intersect in a point P. Draw AP, BP, DP, EP. Then $PA = PB$ and $PD = PE$ (any point on the perpendicular bisector of a line segment is equidistant from the extremities of the line segment). Also, by construction, $AD = BE$. Therefore triangles APD and BPE are congruent, since the three sides of one are equal to the three sides of the other. Hence $\measuredangle DAP = \measuredangle EBP$. But $\measuredangle BAP = \measuredangle ABP$, since these angles are base angles of the isosceles triangle APB. By subtraction it now follows that right angle DAG = obtuse angle EBA.

(c) *To prove that there are two perpendiculars from a point to a line.*

Let two circles intersect in A and B (see Figure 2.5). Draw the diameters AC and AD, and let the join of C and D cut the respective circles in M and N. Then angles AMC and AND are right angles, since each is inscribed in a semicircle. Hence AM and AN are two perpendiculars to CD.

2.4.2 To guarantee the existence of certain points of intersection (of line with circle and circle with circle) Richard Dedekind (1831–1916) introduced into geometry the following continuity postulate:

If all points of a horizontal straight line fall into two classes, such that every point of the first class lies to the left of every point of the second class, then

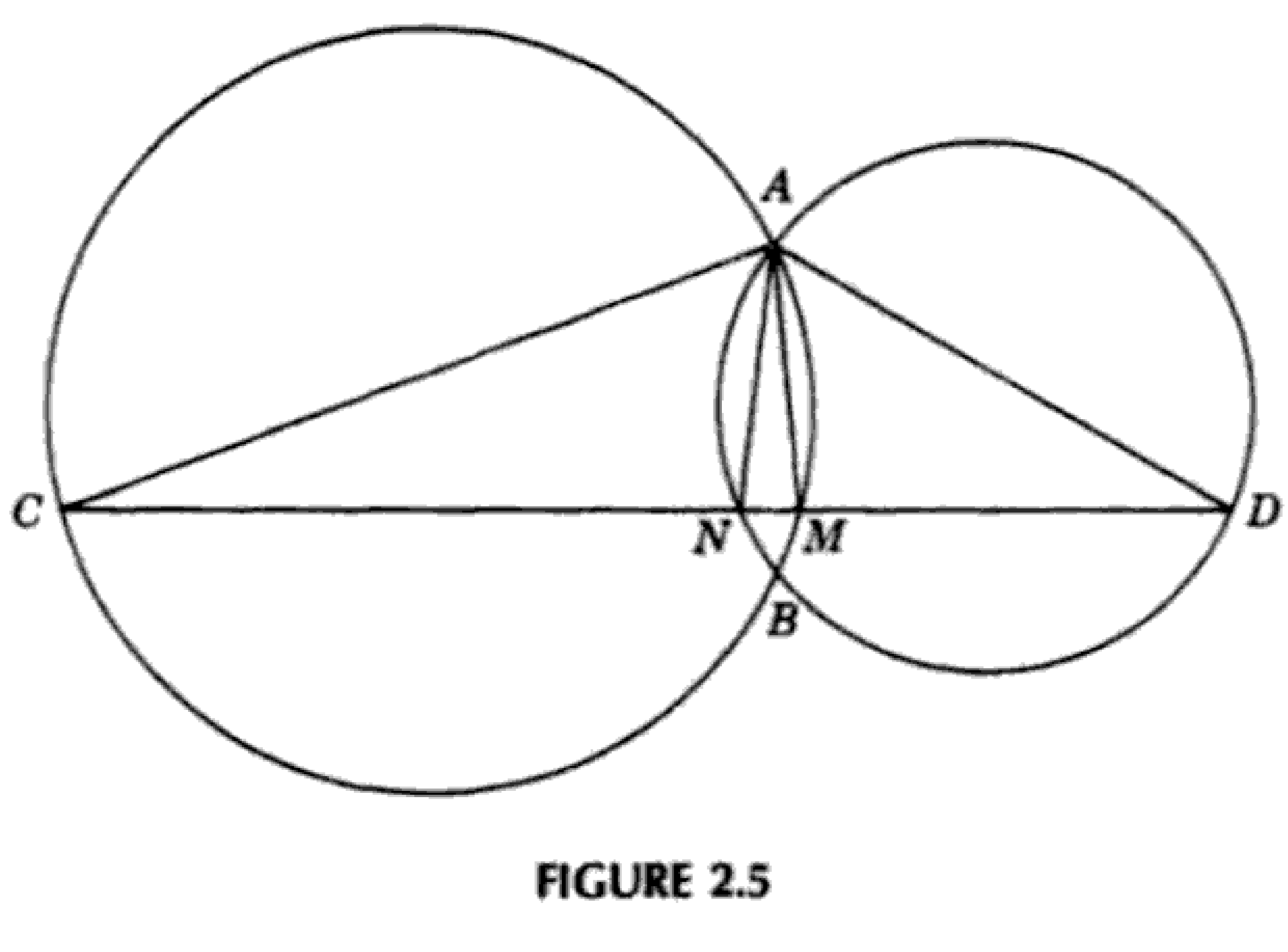


FIGURE 2.5

there exists one and only one point that produces this division of all points into two classes—that is, this severing of the straight line into two portions.

(a) Complete the details of the following indicated proof of the theorem:

> *The straight line segment joining a point* A *inside a circle to a point* B *outside the circle has a point in common with the circle.*

Let O be the center and r the radius of the given circle (see Figure 2.6), and let C be the foot of the perpendicular from O on the line determined by A and B. The points of the segment AB can be divided into two classes: those points P for which $OP < r$ and those points Q for which $OQ \geqq r$. It can be shown that, in every case, $CP < CQ$. Hence, by Dedekind's postulate, there exists a point R of AB such that all points that precede it belong to one class and all that follow it belong to the other class. Now $OR \nless r$, for otherwise we could choose S on AB, between R and B, such that $RS < r - OR$. But since $OS < OR + RS$, this would imply the absurdity that $OS < r$. Similarly, it can be shown that $OR \ngtr r$. Hence we must have $OR = r$, and the theorem is established.

(b) How might Dedekind's postulate be extended to cover angles?

(c) How might Dedekind's postulate be extended to cover circular arcs?

2.4.3 Let us, for convenience, restate Euclid's first three postulates in the following equivalent forms:

1. *Any two distinct points determine a straight line.*
2. *A straight line is boundless.*
3. *There exists a circle having any given point as center and passing through any second given point.*

Show that Euclid's postulates, partially restated above, hold if the points of the plane are restricted to those whose rectangular Cartesian coordinates for some fixed frame of reference are rational numbers. Show, however, that under this restriction a circle and a line through its center need not intersect each other.

2.4.4 Show that Euclid's postulates (as partially restated in Problem 2.4.3) hold if we interpret the plane as the surface of a sphere, straight lines as great circles on the sphere, and points as points on the sphere. Show, however, that in this interpretation the following are true:

(a) Parallel lines do not exist.

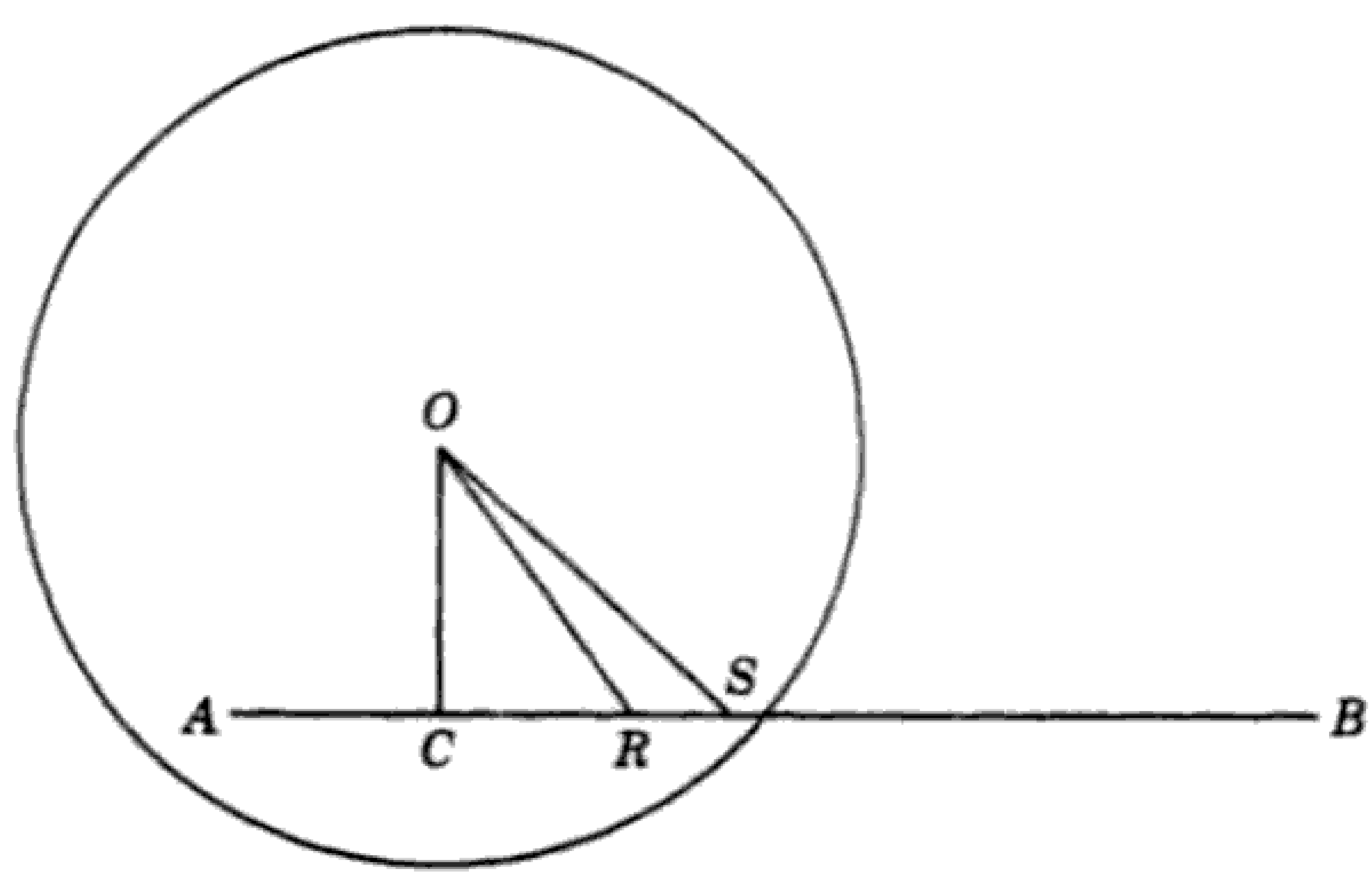


FIGURE 2.6

(b) All perpendiculars to a given line erected on one side of the line intersect in a point.
(c) It is possible to have two distinct lines joining the same two points.
(d) The sum of the angles of a triangle exceeds two right angles.
(e) There exist triangles having all three angles right angles.
(f) An exterior angle of a triangle is not always greater than either of the two remote interior angles.
(g) The sum of two sides of a triangle can be less than the third side.
(h) A triangle with a pair of equal angles may have the sides opposite them unequal.
(i) The greatest side of a triangle does not necessarily lie opposite the greatest angle of the triangle.

2.4.5 In 1882 Moritz Pasch formulated the following postulate:

> *Let* A, B, C *be three points not lying in the same straight line, and let* m *be a straight line lying in the plane of* ABC *and not passing through any of the points* A, B, C. *Then, if the line* m *passes through a point of the segment* AB, *it will also pass through a point of the segment* BC *or a point of the segment* AC.

This postulate is one of those assumptions classified by modern geometers as a *postulate of order*, and it assists in bringing out the idea of "betweenness."
(a) Prove, as a consequence of Pasch's postulate, that *if a line enters a triangle at a vertex, it must cut the opposite side.*
(b) Show that Pasch's postulate does not always hold for a spherical triangle cut by a great circle.

2.4.6 New terms defined by means of more primitive terms are not essential to a deductive system but are convenient in the drawing of inferences, for the new terms serve as shorthand for complex and unmanageable phrases involving the more primitive terms. To illustrate the cumbersomeness that would result if, in Euclid's *Elements*, we should dispense with, say, the terms *point* and *line*, we would have to describe a straight line as "a breadthless length which lies evenly with all the entities on itself which have no parts." State Euclid's first two postulates without using the terms *point* and *line*.

2.5.1 Let W_1 and W_2 be two weights at distances d_1 and d_2, respectively, from a fulcrum. On the basis of the first two postulates of Archimedes' treatise *On Plane Equilibriums*, establish the following theorems:
(a) If we have equilibrium and $W_1 = W_2$, then $d_1 = d_2$.
(b) If we have equilibrium and $d_1 = d_2$, then $W_1 = W_2$.
(c) If we have equilibrium and $W_1 \neq W_2$, then $d_1 \neq d_2$.
(d) If, when weights at certain distances balance, one of the distances should be increased, equilibrium will not be maintained, but there will be inclination on the side of the distance which was increased.
(e) Try to show that we have equilibrium if and only if $W_1 d_1 = W_2 d_2$.

2.5.2 **(a)** Assuming the existence of the usual one-to-one correspondence between real numbers and points on an x-axis, show that an arithmetized form of the postulate of Archimedes is, "If a and b are any two positive real numbers, there exists a positive integer n such that $na > b$."
(b) State the postulate of Archimedes for angles, and indicate how it might be deduced from the arithmetized form of the postulate.

3

NON-EUCLIDEAN GEOMETRY

3.1 Euclid's Fifth Postulate

Postulate 5 of Euclid's *Elements* has been described as "perhaps the most famous single utterance in the history of science."[1] Certainly it has been the source of much controversy, and the dissatisfaction of mathematical scholars with its statement as a postulate is indicated by the fact that many reputable geometers attempted over a period of some twenty centuries either to prove it as a theorem or to replace it by a more acceptable equivalent. As we shall soon see, this concern over Euclid's fifth postulate furnished the stimulus for the development of a great deal of modern mathematics and also led to deep and revealing inquiries into the logical and philosophical foundations of the subject.

A rigorous development of the theory of parallels apparently gave the early Greeks considerable trouble. Euclid met the difficulties by defining parallel lines as coplanar straight lines that do not meet one another however far they may be produced in either direction, and by adopting as an assumption his now famous fifth postulate. Proclus tells us that this postulate was attacked from the very start. Even a cursory reading of Euclid's five postulates discloses a very noticeable difference between the fifth postulate and the other four; the fifth postulate lacks the terseness and the simple comprehensibility possessed by the other four, and it certainly does not have that quality of ready acceptance demanded by material axiomatics. A more studied examination reveals that the fifth postulate is actually the converse of Proposition I 17.[2] It is not surprising that it seemed more like a proposition than a postulate. Moreover, Euclid himself made no use of it until he reached Proposition I 29. It was very natural to wonder whether the postulate was really needed at all and to think that perhaps it could be derived as a theorem from the remaining nine "axioms" and "postulates" or,

[1] C. J. Keyser [1], p. 113.

[2] See Appendix, Section A.1, for the statements of the first twenty-eight propositions of Euclid's Book I.

at least, that it could be replaced by a more acceptable equivalent. Proclus, who was under the illusion that he possessed a proof of the postulate, favored deleting the postulate from the first principles. A quotation from Proclus might be of interest:

> This ought to be struck out of the postulates altogether; for it is a theorem involving many difficulties. Ptolemy, in a certain book, set himself to solve it, and it requires for its demonstration a number of definitions as well as theorems. Moreover, its converse is actually proved by Euclid himself as a theorem. It may be that some persons would be deceived, and would think it proper to place the assumption in question among the postulates as affording ground for an instantaneous belief that the straight lines converge and meet when the two angles are made less than two right angles. To such persons, Geminus correctly replied that we have learned from the very pioneers of this science not to have any regard for mere plausible imaginings when it is really a question of the reasonings to be included in our geometrical doctrine. Aristotle says that it is as justifiable to ask scientific proofs of a rhetorician as to accept mere plausibilities from a geometer, and Simmias is made by Plato to say that he recognizes as quacks those who fashion for themselves proofs from probabilities. So in this case, when the two right angles are lessened, the fact that the straight lines converge is true and necessary; but the statement that they will meet sometime, since they converge more and more as they are produced, is plausible, but it is not necessary in the absence of some argument showing that this is true. It is a known fact that some lines exist which approach each other indefinitely, but yet remain nonintersecting; this seems improbable and paradoxical, but nevertheless it is true and fully ascertained with regard to other species of lines. May not the same thing which happens in the case of the lines referred to be possible in the case of straight lines? Indeed, until the statement in the postulate is clinched by proof, the facts shown in the case of other lines may direct our imagination the opposite way. Though the controversial arguments against the meeting of the straight lines should contain much that is surprising, is that not all the more reason why we should expel from our body of doctrine this merely plausible and unreasoned hypothesis?

There were many attempts to "prove" the fifth, or parallel, postulate and many substitutes devised for its replacement. Of the various substitutes, the one most commonly favored is that made well known in modern times by the Scottish physicist and mathematician, John Playfair (1748–1819), although this particular alternative had been used by others and had even been stated as early as the fifth century by Proclus. This substitute is the one most often encountered in present-day high school geometry texts—namely, *Through a given point not on a given line can be drawn only one line parallel to the given line.*[3] Some other alternatives for the parallel postulate that have been either proposed or tacitly assumed over the years are these:

1. (Posidonius and Geminus) *There exists a pair of coplanar straight lines everywhere equally distant from one another.*

[3]Propositions I 27 and I 28 guarantee, under the tacit assumption of the infinitude of straight lines, the existence of at least *one* parallel.

2. (Wallis, Saccheri, Carnot, and Laplace) *There exists a pair of similar noncongruent triangles.*
3. (Saccheri) *If in a quadrilateral a pair of opposite sides are equal and if the angles adjacent to a third side are right angles, then the other two angles are also right angles.*
4. (Lambert and Clairaut) *If in a quadrilateral three angles are right angles, the fourth angle is also a right angle.*
5. (Legendre) *There exists at least one triangle having the sum of its three angles equal to two right angles.*
6. (Legendre) *Through any point within an angle less than* 60° *there can always be drawn a straight line intersecting both sides of the angle.*
7. (Legendre and W. Bolyai) *A circle can be passed through any three noncollinear points.*
8. (Gauss) *There is no upper limit to the area of a triangle.*

It constitutes an interesting and challenging collection of exercises for the student to try to show the equivalence of these alternatives to the original postulate stated by Euclid. To show the equivalence of Euclid's postulate and a particular one of the alternatives, one must show that the alternative follows as a theorem from Euclid's assumptions and also that Euclid's postulate follows as a theorem from Euclid's system of assumptions with the parallel postulate replaced by the considered alternative.

It would be difficult to estimate the number of attempts that have been made, throughout the centuries, to deduce Euclid's fifth postulate as a consequence of the other Euclidean assumptions, either explicitly stated or tacitly implied. All these attempts ended unsuccessfully, and most of them were sooner or later shown to rest on an assumption equivalent to the postulate itself. The earliest effort, of which we are today aware, to prove the postulate was made by Claudius Ptolemy (*ca.* A.D. 150), alluded to by Proclus in the quotation given above. Claudius Ptolemy was the author of the famous and very influential *Almagest*, the great definitive Greek work on astronomy. Proclus exposed the fallacy in Ptolemy's attempt by showing that Ptolemy had unwittingly assumed that through a point only one parallel can be drawn to a given line; this assumption is the Playfair equivalent of Euclid's postulate. Proclus submitted an attempt of his own, but his "proof" rests on the assumption that parallel lines are always a bounded distance apart, and this assumption can be shown to imply Euclid's fifth postulate. Among the more noteworthy attempts of somewhat later times is one made in the thirteenth century by Nasir-ed-din (1201–1274), a Persian astronomer and mathematician who compiled, from an earlier Arabic translation, an improved edition of the *Elements* and who wrote a treatise on Euclid's postulates, but his attempt, too, involves a tacit assumption equivalent to the postulate being "proved."

An important stimulus to the development of geometry in western Europe after the Renaissance was a renewal of the criticism of Euclid's fifth postulate. Hardly any critical comments are to be found in the early printed editions of the *Elements* made at the end of the fifteenth century and at the beginning of the sixteenth century. However, after the translation, in 1533, of Proclus's *Commentary on Euclid, Book I*, many men once again embarked upon a critical analysis of the fifth postulate. For example, John Wallis (1616–1703), while lecturing at Oxford University, became interested in the work of Nasir-ed-din and in 1663 offered his own "proof" of the parallel postulate, but this attempt involves the equivalent assumption that similar noncongruent triangles exist. So it was with all the many attempts to derive Euclid's postulate as a theorem; each attempt involved the vitiating circularity of assuming something equivalent to the thing being established or else committed some other form of fallacious reasoning. Most of this vast amount of work is of little real importance in the actual evolution of mathematical thought until we come to the remarkable investigation of the parallel postulate made by Girolamo Saccheri in 1733.

3.2 Saccheri and the *Reductio ad Absurdum* Method

Every student of elementary geometry has encountered the so-called *indirect*, or *reductio ad absurdum*,[4] method of proof. It is a powerful, and at times seemingly indispensable, method that is employed frequently by Euclid in his *Elements*. The method, it will be recalled, consists of assuming, by way of hypothesis, that a proposition that is to be established is false; if an absurdity follows, one concludes that the hypothesis is untenable and that the original proposition must then be true. It was this method of proof that we employed in Section 1.5 to show that $\sqrt{2}$ is irrational.

To illustrate further the *reductio ad absurdum* method, let us briefly consider Euclid's Proposition I 6, the first proposition in the *Elements* established by this type of proof. We wish to prove the theorem: *If in a triangle two angles are equal to each other, then the two sides opposite these angles are also equal to each other*. Let ABC (Figure 3.1) be the triangle, and suppose $\measuredangle ABC = \measuredangle ACB$. We wish to show that side AB = side AC. Suppose the sides AB and AC are not equal to each other. Then one of them—say, AB—is greater than the other, and we may mark off on BA a segment BD equal to the lesser side AC. Now in triangles ABC and DCB we have $CB = BC$, $CA = BD$, $\measuredangle BCA = \measuredangle CBD$. It follows that the triangles are congruent. But this conclusion is impossible, since triangle DCB is only a part of triangle ABC. Our hypothesis that $AB \neq AC$ has led to an absurd situation and hence is untenable. We must conclude, therefore, that $AB = AC$, and our theorem is established.

The *reductio ad absurdum* method rests on two cardinal principles of classical logic—namely, the *law of contradiction* and the *law of the excluded middle*. Somewhat loosely described, the law of contradiction says that *if* S *is any statement, then* S *and a contradiction* (that is, the denial) *of* S *cannot both hold*, and the law of the excluded middle says that *either* S *or the denial of* S *must hold* (that

[4] In a more refined treatment one distinguishes several slight variations in the indirect method of proof, and then it is customary to assign the technical terminology, *reductio ad absurdum*, to a particular one of these variations. We do not make this refinement here.

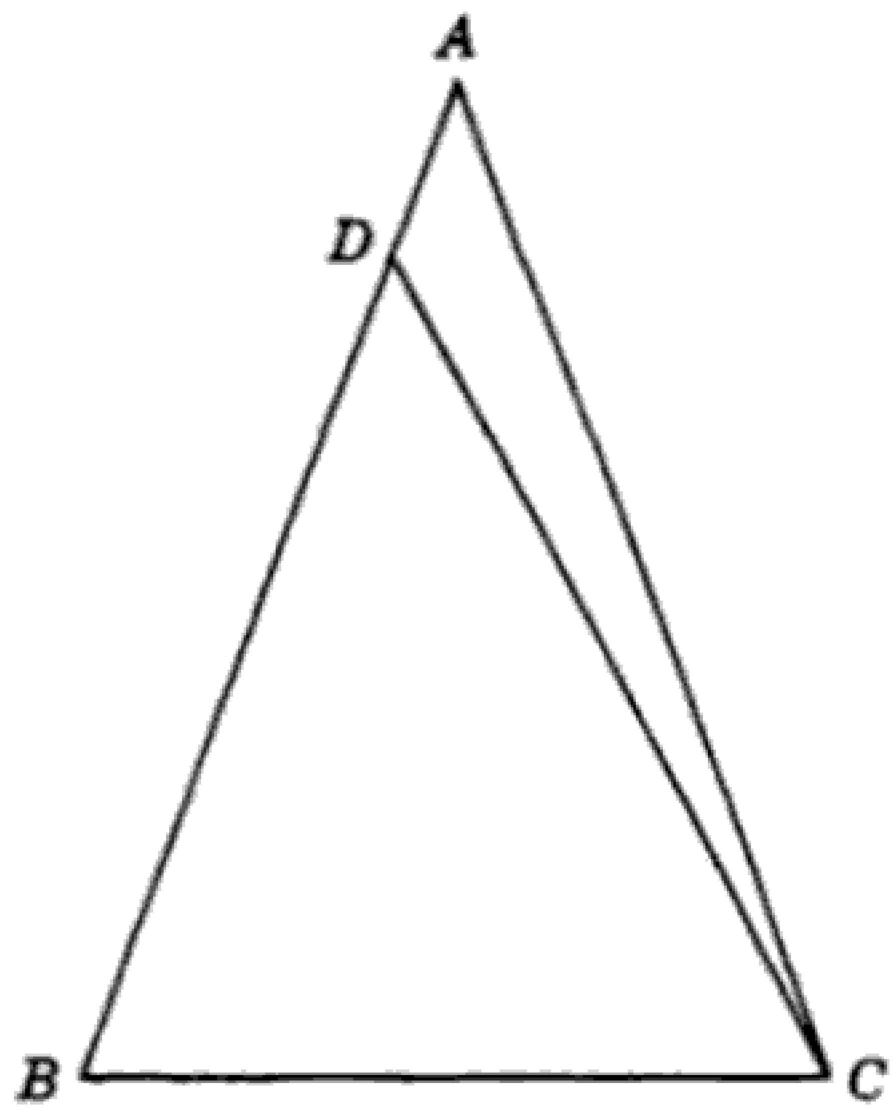


FIGURE 3.1

is, there is no third, or middle, possibility). As an illustration, suppose S is the statement, "Archimedes was born in 287 B.C." The law of contradiction then asserts that Archimedes cannot have been born in both 287 B.C. and A.D. 400, for example, and the law of the excluded middle asserts that Archimedes either was born in 287 B.C. or he was not born in 287 B.C. Now let S be the statement of any proposition to be established by the *reductio ad absurdum* method—for example, the statement of Proposition I 6 above. By the method, we set about and show that the denial of S implies the denial of some previously assumed or established statement T. By the law of contradiction, T and the denial of T cannot both be true (that is, cannot both follow from the postulates). Since T is true, the denial of T is then false, from which, since a true statement can never imply a false one, it follows that the denial of S must also be false. By the law of the excluded middle, however, either S is true or the denial of S is true. Since the denial of S is false, it follows that S is true, and our proposition is established.

The law of contradiction and the law of the excluded middle have settled so deeply into the warp and woof of human thinking that it is difficult to conceive of questioning their validity. We shall see later, however, that although these laws are usually pronounced as universally true, some sort of limitation must be made concerning their applicability. Indeed, since 1912 some mathematicians have felt that we must drastically restrict the free use of the law of the excluded middle as part of the logical machinery used in deducing theorems from postulates. But more of this in its proper place. For the time being we shall accept these laws, particularly insofar as they apply to the *reductio ad absurdum* method of proof.

One concluding remark about the *reductio ad absurdum* method seems appropriate here. In the game of chess a *gambit* is one of various possible openings in which a pawn or a piece is risked in order to obtain an advantageous attack. The eminent English mathematician, G. H. Hardy (1877–1947), delightfully pointed out that *reductio ad absurdum* "is a far finer gambit than any chess gambit: a chess player may offer the sacrifice of a pawn or even a piece, but a mathematician offers *the game*."[5] *Reductio ad absurdum* emerges as the most stupendous gambit conceivable.

[5] G. H. Hardy [2], p. 34.

Reductio ad absurdum certainly constitutes one of the finest weapons in a mathematician's armory of attack; with this weapon Girolamo Saccheri in 1733 made the first really scientific assault on the problem of Euclid's parallel postulate.

Little is known of Saccheri's life. He was born in San Remo in 1667, showed marked precocity as a youngster, completed his novitiate for the Jesuit Order at the age of twenty-three, and then spent the rest of his life filling a succession of university teaching posts. While instructing rhetoric, philosophy, and theology at a Jesuit College in Milan, Saccheri read Euclid's *Elements* and became enamored with the powerful method of *reductio ad absurdum.* Later, while teaching philosophy at Turin, Saccheri published his *Logica demonstrativa*, in which the chief innovation is the application of the method of *reductio ad absurdum* to the treatment of formal logic. Some years after, while a professor of mathematics at the University of Pavia, it occurred to Saccheri to apply his favorite method of *reductio ad absurdum* to a study of Euclid's parallel postulate. He was well prepared for the task, having dealt ably in his earlier work on logic with such matters as definitions and postulates. Also, he was acquainted with the work of others regarding the parallel postulate and had succeeded in pointing out the fallacies in the attempts of Nasir-ed-din and Wallis.

Saccheri's effort to establish Euclid's parallel postulate by attempting to institute a *reductio ad absurdum* was apparently the first time anyone had conceived the idea of denying the postulate and of studying the consequences of a contradiction of the famous assumption. The result of these researches was a little book entitled *Euclides ab omni naevo vindicatus* (Euclid Freed of Every Flaw), which was printed in Milan in 1733, only a few months before the author's death. In this work Saccheri accepts the first twenty-eight propositions of Euclid's *Elements*, which, as we have previously stated, do not require the fifth postulate for their proof. With the aid of these theorems he then proceeds to study the *isosceles birectangle*—that is, a quadrilateral $ABDC$ in which (see Figure 3.2) $AC = BD$ and the angles at A and B are right angles. By drawing the diagonals AD and BC and then using simple congruence theorems (which are found among Euclid's first 28 propositions), Saccheri easily shows that the angles at C and D are equal to each other. But nothing can be ascertained in regard to the magnitude of these angles. Of course, as a consequence of Euclid's fifth postulate, it follows that these angles are both right angles, but the assumption of this postulate is not to be employed. As a result, the two angles might both be right angles, obtuse angles, or acute angles. Here Saccheri

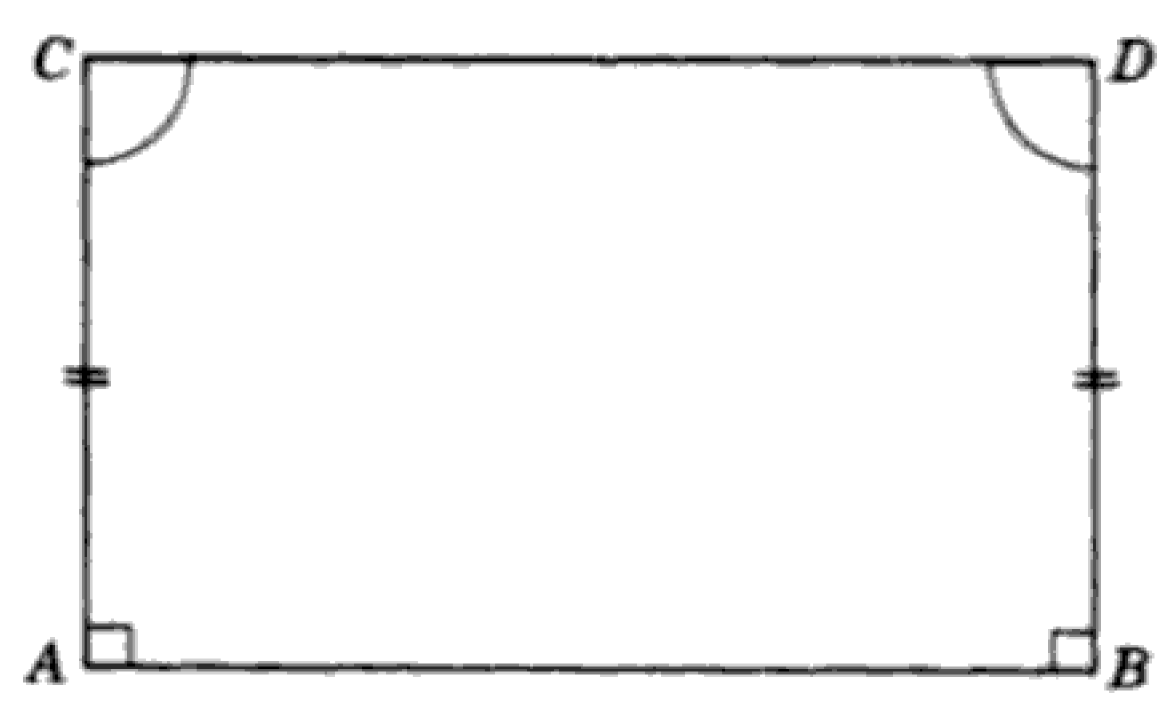


FIGURE 3.2

maintains an open mind and names the three possibilities: the *hypothesis of the right angle*, the *hypothesis of the obtuse angle*, and the *hypothesis of the acute angle*. The plan of the work is to rule out the last two possibilities by showing that their respective assumptions lead to absurdities, thus leaving, by the *reductio ad absurdum* method, the first hypothesis. But this hypothesis can be shown to be equivalent to Euclid's fifth postulate. In this way the parallel postulate is to be established and the blemish of its assumption by Euclid removed.

The task of eliminating the hypothesis of the obtuse angle and the hypothesis of the acute angle turns out to be rather arduous. With real geometrical skill and fine logical penetration, Saccheri establishes a number of theorems, of which the following are among the more important:

1. *If one of the hypotheses is true for a single isosceles birectangular quadrilateral, it is true for every such quadrilateral.*
2. *On the hypothesis of the right angle, the obtuse angle, or the acute angle, the sum of the angles of a triangle is respectively equal to, greater than, or less than two right angles.*
3. *If there exists a single triangle for which the sum of the angles is equal to, greater than, or less than two right angles, there follows the truth of the hypothesis of the right angle, the obtuse angle, or the acute angle.*
4. *On the hypothesis of the right angle two distinct straight lines intersect, except in the one case in which a transversal cuts them under equal corresponding angles. On the hypothesis of the obtuse angle two straight lines always intersect. On the hypothesis of the acute angle there is an infinitude of straight lines through a given point not on a given straight line and which do not meet the given straight line.*
5. *The locus of the extremity of a perpendicular of constant length that moves with its other end on a fixed straight line is a straight line on the hypothesis of the right angle, a curve convex to the fixed line on the hypothesis of the obtuse angle, and a curve concave to the fixed line on the hypothesis of the acute angle.*

After establishing a chain of thirteen propositions, Saccheri manages to dispose of the hypothesis of the obtuse angle, but in so doing he makes the same tacit assumption that Euclid made concerning the infinitude of the straight line. With this tacit assumption (introduced by using Euclid's Proposition I 18, which depends on I 16) Saccheri shows that the hypothesis of the obtuse angle implies Euclid's fifth postulate, which, in turn, implies that the sum of the angles of a triangle is equal to two right angles. But this second implication contradicts the theorem that, on the hypothesis of the obtuse angle, the sum of the angles of a triangle is greater than two right angles.

The case of the hypothesis of the acute angle proves to be even more stubborn, and Saccheri requires nearly twenty more propositions before he feels he can dispose of it. After obtaining many of the theorems that were later to become classical in non-Euclidean geometry, Saccheri weakly forces into his development an unconvincing contradiction involving vague concepts about elements at infinity. The contradiction that he reaches is that there exist two straight lines that when produced to infinity merge into one another and there have a common perpendicular. Coming after the careful work that has been presented up to this point, it is difficult to believe that Saccheri himself was really convinced by his lame ending. Indeed, in a second part to his work, he attempted, with no greater success, a second attack on the hypothesis of the acute angle. Had Saccheri not been so eager to exhibit a contradiction but rather had boldly admitted his inability to find one, he would today unquestionably be credited with the discovery of non-Euclidean geometry.

It is difficult to evaluate the influence that Saccheri's work may have had on later researches connected with the parallel postulate, for subsequently his little publication was lost for a long time. It was dramatically resurrected in 1889 by Saccheri's compatriot, Eugenio Beltrami (1835–1900), a mathematician who, as we shall soon see, made notable contributions of his own to the subject of non-Euclidean geometry. The first part of Saccheri's work has been translated into English[6] and can be easily read by any student of elementary plane geometry.

3.3 The Work of Lambert and Legendre

In 1766, thirty-three years after Saccheri's publication, the German mathematician Johann Heinrich Lambert (1728–1777) wrote an investigation of the parallel postulate entitled *Die Theorie der Parallellinien*, which, however, was not published until eleven years after the author's death. Lambert's treatise is in three parts. The first part considers whether Euclid's fifth postulate can be proved from Euclid's other assumptions or only if additional assumptions are made. The second part is concerned with the reduction of the parallel postulate to various equivalent propositions. It is the last part of the study that closely resembles the earlier work by Saccheri. Here Lambert chooses as a fundamental figure the *trirectangle*, or quadrilateral containing three right angles, which can be regarded as the half of a Saccheri isosceles birectangle formed by joining the midpoints of the latter's bases. As with Saccheri, three hypotheses arise, according to whether the fourth angle of the trirectangle is right, obtuse, or acute.

Lambert went considerably beyond Saccheri in deducing propositions under the hypotheses of the obtuse and acute angles. Thus, not only did he show that for the three hypotheses the sum of the angles of a triangle is equal to, greater than, or less than two right angles, respectively, but in addition he showed that the excess above two right angles in the hypothesis of the obtuse angle, or the deficiency below two right angles in the hypothesis of the acute angle, is proportional to the area of the triangle. This result led him to observe the resemblance to spherical geometry of the geometry following from the hypothesis of the obtuse angle (in spherical geometry the area of a triangle is proportional to its spherical excess), and he conjectured that the geometry following from the hypothesis of the acute angle could perhaps be verified on a sphere of imaginary radius.

[6] See G. B. Halsted [2] or D. E. Smith [2], pp. 351–359.

Another notable discovery made by Lambert concerns the measurement of lengths in the two geometries that follow from the obtuse-angle and acute-angle hypotheses. In Euclidean geometry, because similar noncongruent figures exist, lengths can be measured only in terms of some arbitrary unit that has no structural connection with the geometry. Angles, on the other hand, possess a natural unit of measure, such as the right angle or the radian, which is capable of geometrical definition. This is what is meant when mathematicians say that in Euclidean geometry lengths are *relative* but angles are *absolute.* Lambert discovered that under the hypotheses of the obtuse and acute angles, angles are still absolute, but lengths are absolute also! In fact, it can be shown for these geometries that for every angle there is a corresponding line segment, so that to a natural unit of measure for angles there corresponds a natural unit of measure for lengths.

Lambert eliminated the hypothesis of the obtuse angle by making the same tacit assumption that Saccheri made, but his conclusions with regard to the hypothesis of the acute angle were indefinite and unsatisfactory. Indeed, it was this incomplete and unsettled state of affairs with regard to the acute hypothesis that held Lambert from publishing his work, with the result that it did not appear until friends finally put it through the press after his death.

Lambert was a mathematician of high quality. As the son of a poor tailor he was largely self-taught. He possessed a fine imagination and established his results with great attention to rigor. In fact, Lambert was the first to prove rigorously that the number π is irrational. He showed that if x is rational but not zero, then $\tan x$ cannot be rational; since $\tan \pi/4 = 1$, it follows that $\pi/4$, or π, cannot be rational. We also owe to Lambert the first systematic development of the theory of hyperbolic functions and, indeed, our present notation for these functions. Lambert was a many-sided scholar who contributed to the mathematics of numerous other topics, such as descriptive geometry, the determination of comet orbits, and the theory of projections employed in the making of maps.

A third distinguished effort to establish Euclid's parallel postulate by the *reductio ad absurdum* method was essayed, over a long period of years, by the eminent French analyst Adrien-Marie Legendre (1752–1833). He began anew and considered three hypotheses according to whether the sum of the angles of a triangle is equal to, greater than, or less than two right angles. Tacitly assuming the infinitude of a straight line, he was able to eliminate the second hypothesis, but although he made several attempts, he could not dispose of the third hypothesis. These various endeavors appear in the successive editions of his very popular *Éléments de géométrie*,[7] which ran from a first edition in 1794 to a twelfth in 1823. Legendre's first effort is vitiated by the assumption that the choice of a unit of length will not affect the correctness of his propositions, but this, of

[7] This work is an attempted pedagogical improvement of Euclid's *Elements* made by considerably rearranging and simplifying the propositions. The work won high regard in continental Europe and was so favorably received in the United States that it became the prototype of the elementary geometry textbooks in this country. The first English translation was made in the United States in 1819 by John Farrar of Harvard University. The next English translation was made in 1822 by the famous Scottish litterateur, Thomas Carlyle, who early in life was a teacher of mathematics. Carlyle's translation ran through thirty-three American editions.

course, is equivalent to assuming the existence of similar noncongruent figures. The next attempt is vitiated by assuming the existence of a circle through any three noncollinear points. Later Legendre independently observed the fact already discovered by Lambert that, under the third hypothesis, the deficiency of the sum of the angles of a triangle below two right angles is proportional to the area of the triangle. Hence, Legendre reasoned, if by starting with any given triangle one could obtain another triangle containing the given triangle at least twice, then the deficiency for this new triangle would be at least twice the deficiency for the given triangle. By repeating the operation a sufficient number of times, one could finally end with a triangle whose angle sum has become negative, a situation that is absurd. But in order to solve the problem of obtaining a triangle containing a given triangle at least twice, Legendre found he had to assume that through any point within a given angle less than 60° there can always be drawn a straight line intersecting both sides of the angle, and this, as we have pointed out, is equivalent to Euclid's fifth postulate. Legendre gave an elegant proof of the theorem: *If there exists a single triangle having the sum of its angles equal to two right angles, then the sum of the angles of every triangle is equal to two right angles*. Although this theorem is contained in the results given by Saccheri, it is generally referred to as *Legendre's second theorem. Legendre's first theorem* is: *The sum of the three angles of a triangle cannot be greater than two right angles*. Of course, in proving this theorem, Legendre tacitly assumed the infinitude of straight lines. In fact, in proving both his first and second theorems, Legendre assumed the postulate of Archimedes. Max Dehn (1878–1952) has shown that this assumption is unavoidable in proving the first theorem but not necessary in proving the second.

Legendre's last paper on parallels, essentially a collection of his earlier efforts, was published in 1833, the year of his death. He perhaps holds the record for persistence in attempting to prove the famous postulate. The simple and straightforward style of his proofs, widely circulated because of their appearance in his *Éléments,* and his high eminence in the world of mathematics, created marked popular interest in the parallel postulate. Actually, however, Legendre had scarcely progressed as far as had Saccheri a hundred years earlier. Moreover, even before the appearance of his last paper, a Russian mathematician, separated from the rest of the scientific world by barriers of distance and language, had taken a most significant step, the boldness and importance of which were far to transcend anything Legendre had done on the subject.

3.4 The Discovery[8] of Non-Euclidean Geometry

We have seen that, in spite of considerable effort exerted over a long period of time, no one was able to find a contradiction under the hypothesis of the acute angle. It is no wonder that no contradiction was found under this hypothesis, for it is now known that the geometry developed from a certain basic set of

[8] We are not here concerned with any philosophical distinction between *discovery* and *invention.*

assumptions plus the acute angle hypothesis is as consistent as the Euclidean geometry developed from the same basic set of assumptions plus the hypothesis of the right angle. In other words, it is now known that the parallel postulate *cannot* be deduced as a theorem from the other assumptions of Euclidean geometry but is independent of those other assumptions. It took unusual imagination to entertain such a possibility, for the human mind had for two millennia been bound by the prejudice of tradition to the firm belief that Euclid's geometry was most certainly the only possible one and that any contrary geometric system simply could not be consistent.

The first to suspect the independence of the parallel postulate were Carl Friedrich Gauss (1777–1855) of Germany, Johann Bolyai (1802–1860) of Hungary, and Nicolai Ivanovitch Lobachevsky (1793–1856) of Russia. These men independently approached the subject through the Playfair form of the parallel postulate by considering the three possibilities: Through a given point not on a given straight line can be drawn *just one* line, *no* line, or *more than one* line parallel (in Euclid's sense) to the given line. These three situations are equivalent, respectively, to the hypotheses of the right, the obtuse, and the acute angle. Assuming, as did their predecessors, the infinitude of a straight line, the second case was easily eliminated. Inability to find a contradiction in the third case, however, led each of the three mathematicians to suspect, in time, a consistent geometry under that hypothesis, and each, unaware of the work of the other two, carried out, for its own intrinsic interest, an extensive development of the resulting new geometry.

Gauss was perhaps the first person really to anticipate a non-Euclidean geometry. Although he meditated a good deal on the matter from very early youth on, probably not until his late twenties did he begin to suspect the parallel postulate to be independent of Euclid's other assumptions. Unfortunately, Gauss failed, throughout his life, to publish anything on the subject, and his advanced conclusions are known to us only through copies of letters to interested friends, a couple of published reviews of works of others, and some notes found among his papers after his death. Although he refrained from publishing his own findings, he strove to encourage others to persist in similar investigations, and he called the new geometry *non-Euclidean*.

Apparently the next person to anticipate a non-Euclidean geometry was Johann Bolyai, who was a Hungarian officer in the Austrian army and the son of the mathematician Wolfgang Bolyai, a long-time personal friend of Gauss. The younger Bolyai undoubtedly received considerable stimulus for his study from his father, who had earlier shown an interest in the problem of the parallel postulate. As early as 1823 Johann Bolyai began to understand the real nature of the problem that faced him, and a letter written during that year to his father shows the enthusiasm he held for his work. In this letter he discloses a resolution to publish a tract on the theory of parallels as soon as he can find the time and opportunity to put the material in order, and exclaims, "Out of nothing I have created a strange new universe." The father urged that the proposed tract be published as an appendix to his own large two-volume semiphilosophical work on elementary mathematics. The expansion and arrangement of ideas proceeded more slowly than Johann had anticipated, but finally, in 1829, he submitted the finished manuscript to his father, and three years later, in 1832, the tract appeared as a twenty-six–page appendix to the first volume of his father's work.[9] Johann Bolyai never published anything further, although he did leave behind a great pile of manuscript pages. His chief interest was in what he called "the absolute science of space," by which he meant the collection of those propositions which are independent of the parallel postulate and which consequently hold in both the Euclidean geometry and the new geometry. For example, the familiar law of sines for triangle ABC,

$$a:b:c = \sin A : \sin B : \sin C,$$

holds only in the Euclidean geometry, but if modified to read

$$O(a):O(b):O(c) = \sin A : \sin B : \sin C,$$

where $O(r)$ denotes the circumference of a circle of radius r, then the modified law holds in both of the geometries. This, then, is the form that the law of sines takes in Bolyai's work. It is not difficult to show that this same form of the sine law also holds in the geometry of triangles on a sphere.

Although Gauss and Johann Bolyai are acknowledged to be the first to conceive a non-Euclidean geometry, actually the Russian mathematician Lobachevsky published the first really systematic development of the subject. Lobachevsky spent the greater part of his life at the University of Kasan, first as a student, later as a professor of mathematics, and finally as rector, and his earliest paper on non-Euclidean geometry was published in 1829–1830 in the *Kasan Bulletin*, two to three years before Bolyai's work appeared in print. This memoir attracted only slight attention in Russia, and, because it was written in Russian, practically no attention elsewhere. Lobachevsky followed this initial effort with other presentations. For example, in the hope of reaching a wider group of readers, he published in 1840 a little book written in German entitled *Geometrische Untersuchungen zur Theorie der Parallellinien* (Geometrical Researches on the Theory of Parallels),[10] and then still later, in 1855, a year before his death and after he had become blind, he published in French a final and more condensed treatment entitled *Pangéométrie* (Pangeometry).[11] So slowly did information of new discoveries spread in those days that Gauss probably did not hear of Lobachevsky's work until the appearance of the German publication in 1840, and Johann Bolyai was unaware of it until 1848. Lobachevsky himself did not live to see his work accorded any wide recognition, but the non-Euclidean geometry which he developed is nowadays frequently referred to as *Lobachevskian geometry*.

The characterizing postulate of Lobachevskian geometry, which replaces Euclid's parallel postulate, is that *through a given point* P, *not on a given line* m, *more than one line can be drawn lying in the plane of* P *and* m *and not intersecting* m. On the basis of this postulate, together with the other assumptions of Euclidean geometry, it is not difficult to show (see Figure 3.3) that there are always two lines through P that do not intersect m, that make equal acute angles α with the

[9] For a translation of this appendix, see R. Bonola, or D. E. Smith [2], pp. 375–388.

[10] N. Lobachevsky.

[11] D. E. Smith [2], pp. 360–374.

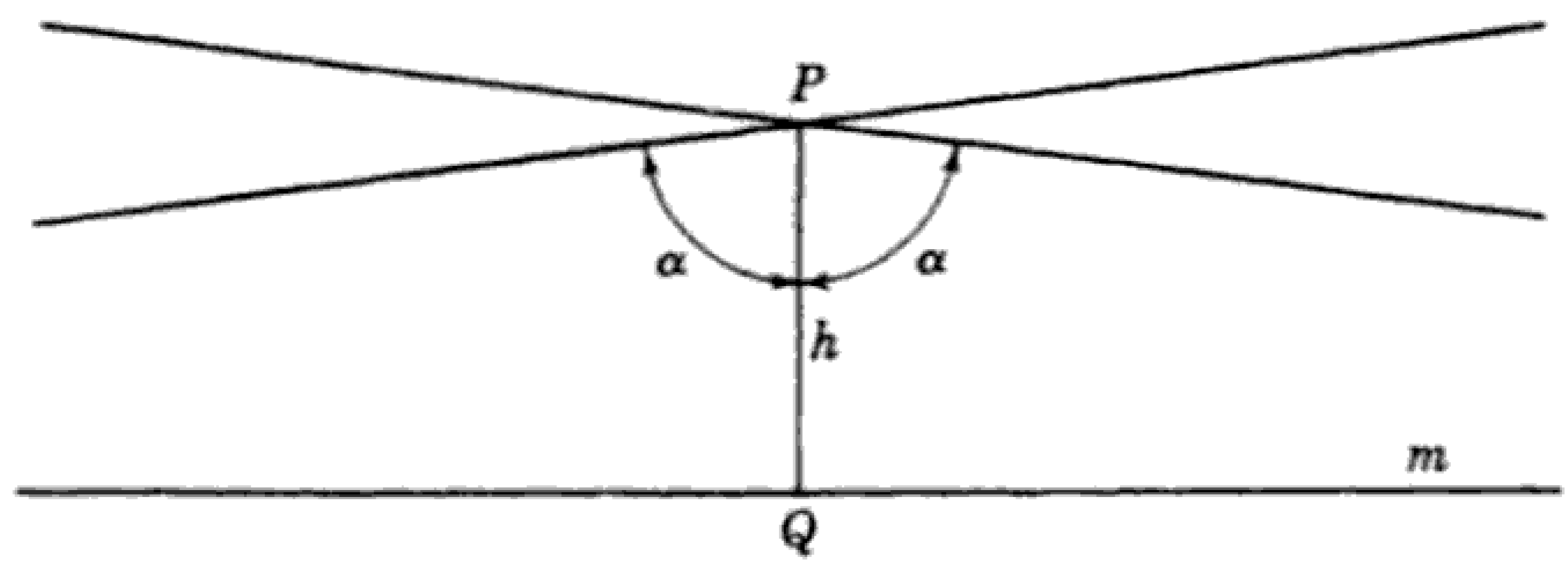


FIGURE 3.3

perpendicular PQ from P to m, and that are such that any line through P lying within the angle formed by the two lines and containing the perpendicular PQ intersects m, while any other line through P does not intersect m. The two lines through P are, then, boundary lines separating all lines through P into two classes, those that cut m and those that do not. From the viewpoint of Euclid, all the lines through P that do not cut m should be called parallels to m. Lobachevsky, however, uses this term more reservedly and refers to only the two boundary lines as being *parallel* to m. Other lines passing through P but not cutting m may be said to be *hyperparallel* to m. The acute angle α is called the *angle of parallelism*, and it plays an important role in Lobachevsky's development. He shows that the size of α depends on the length h of the perpendicular PQ, and emphasizes this by denoting α by the functional symbol $\Pi(h)$. In fact, he shows that if the unit of length is chosen as the distance that corresponds to the particular angle of parallelism

$$\alpha = 2 \text{ Arc tan } e^{-1},$$

where e is the base for natural logarithms, then

$$\Pi(h) = 2 \text{ Arc tan } e^{-h} \quad \text{and} \quad h = \ln \cot \frac{\Pi(h)}{2}.$$

We note that the angle of parallelism, $\Pi(h)$, increases from 0 to $\pi/2$ as h decreases from ∞ to 0, so that, "in the small," Lobachevskian geometry approximates Euclidean geometry. Also, since to each angle $\Pi(h)$ is associated a definite distance h, we see why distances, as well as angles, are absolute in Lobachevskian geometry. It further turns out that the trigonometrical formulas in Lobachevskian geometry are nothing but the familiar formulas of spherical trigonometry when the sides a, b, c of the triangle are replaced by a/i, b/i, c/i, and we are reminded of Lambert's suggestion about an imaginary sphere, mentioned in Section 3.3.

It is not the purpose of our study to go deeply into the Lobachevskian non-Euclidean geometry resulting from the hypothesis of the acute angle, and perhaps we have already indicated a sufficient number of propositions in the geometry to give the reader some idea of its content. We have seen that the hypothesis of the obtuse angle was discarded by all who did research in this subject because it contradicted the assumption that a straight line is infinite in length. Recognition of a second non-Euclidean geometry, based on the hypothesis of the obtuse angle, was not fully achieved until some years later, when Bernhard Riemann, in his probationary lecture of 1854, discussed the concepts of boundlessness and infiniteness. With these concepts clarified, one can realize an equally consistent geometry satisfying the hypothesis of the obtuse angle if Euclid's Postulates 1, 2, and 5 are modified to read:

(1′) *Two distinct points determine at least one straight line.*

(2′) *A straight line is boundless.*

(5′) *Any two straight lines in a plane intersect.*

Much in this new non-Euclidean geometry is interesting. For example (see Figure 3.4), it can be shown, without great difficulty, that all the perpendiculars erected on the same side of a given straight line m are concurrent in a point O, and that the lengths along these perpendiculars from O to the line m are all equal to one another. Moreover, this common length, which we shall denote by q, is independent of which straight line in the plane is chosen for m. It can also be shown that, if A, B, P are any three points on line m, then

$$AP:AB = \measuredangle AOP : \measuredangle AOB,$$

and that if AB is taken equal in length to q, then $\measuredangle AOB = \pi/2$. It now follows that all straight lines are finite and of the same constant length $4q$, for we observe that OP coincides with OA when $\measuredangle AOP = 2\pi$, so that, under such circumstances, AP becomes the total length of the line m. But now

$$AP:AB = 2\pi : \measuredangle AOB,$$

from which, by taking $AB = q$ and therefore $\measuredangle AOB = \pi/2$, we find that

$$AP = 4q.$$

Thus straight lines, though boundless, are finite in length. Also, as in the case of Lobachevskian geometry, lengths, as well as angles, are absolute.

Riemann's celebrated lecture[12] of 1854 is not detailed or specific in its development but is extraordinarily rich in the depth and generality of its concepts and in the originality of its powerful new points of view. It would be difficult to point out another paper that has so greatly influenced modern geometrical research. This paper inaugurated a second period in the development of non-Euclidean geometry, a period characterized by the employment of

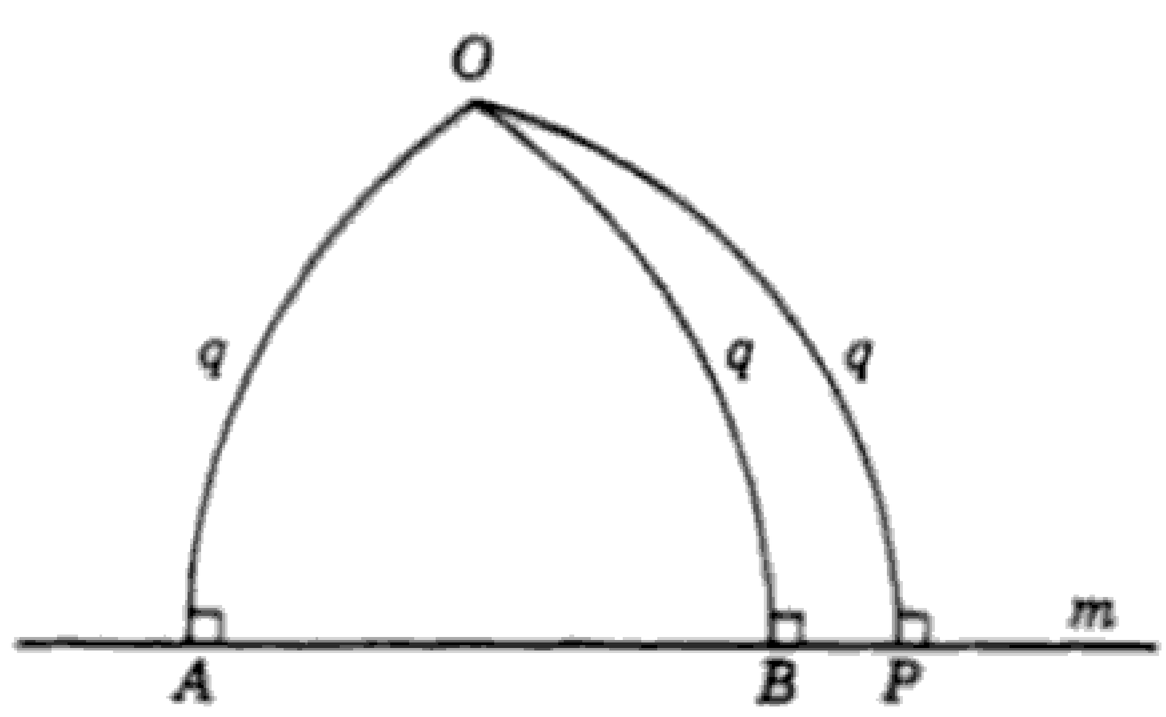


FIGURE 3.4

[12] The lecture was published in 1866, shortly after Riemann's death. For an English translation, see D. E. Smith [2], pp. 411–425.

the methods of differential geometry rather than the previously used methods of elementary synthetic geometry. To this paper we owe a considerable generalization of the concept of space that has led, in more recent times, to the extensive and important theory of abstract spaces; some of this theory has found application in the physical theory of relativity. Literally volumes of modern mathematical research can be traced to ideas advanced in this remarkable paper.

3.5 The Consistency and Significance of Non-Euclidean Geometry

It was some years after the appearance of the work of Lobachevsky and Bolyai that the mathematical world in general paid much attention to the subject of non-Euclidean geometry, and several decades elapsed before the full implication of the invention was appreciated. Most of the development of the subject beyond the historical point to which we have carried it is of too advanced a nature to be adequately considered here. One important matter, however, in this later development must be at least briefly touched on. Although Lobachevsky and Bolyai encountered no contradiction in their extensive investigations of the non-Euclidean geometry based on the hypothesis of the acute angle, and although they even felt confident that no contradiction would arise, the possibility still remained that such a contradiction or inconsistency might appear if the investigations should be sufficiently continued. To Beltrami goes the credit for the first proof of the consistency of this non-Euclidean geometry. In a brilliant paper,[13] published in 1868, Beltrami showed that the plane non-Euclidean geometry of Lobachevsky and Bolyai can be represented, with certain restrictions, on a surface of so-called constant negative curvature. It can be similarly shown that the plane non-Euclidean geometry of Riemann can be represented on a surface of constant positive curvature. Although Beltrami's methods are those of differential geometry and cannot be fully appreciated without an understanding of that field of mathematics, we can rather simply explain the gist of his idea.

Of the surfaces of constant positive curvature, the simplest is the sphere. Now the *geodesics* on the sphere—that is, the curves of shortest length lying on the sphere and joining pairs of points on the sphere—are the great circles of the sphere. If we should interpret the plane of the non-Euclidean geometry of Riemann as the surface of a sphere, and the straight lines of that non-Euclidean geometry as the great circles on the sphere, then it is a very simple matter to show that the postulates of the non-Euclidean geometry hold in our interpretation. For example:

(1′) *Two distinct points on the sphere determine at least one great circle on the sphere. (In fact, the great circle is unique, unless the points on the sphere happen to be diametrically opposite to each other, in which case any number of great circles may be passed through the two points.)*

[13] E. Beltrami.

(2′) *A great circle on the sphere is boundless. (A great circle is not infinite in length, however; in fact, all great circles on the sphere have the same finite length.)*

(3′) *With any point on the sphere as center and any great circle arc as polar distance a circle can be drawn on the sphere.*

(4′) *All right angles on the sphere are equal to one another.*

(5′) *Any two great circles on the sphere intersect.*

In view of our success in finding on the surface of a sphere a representation of Riemann's non-Euclidean geometry, it now follows that the plane non-Euclidean geometry of Riemann is consistent if Euclidean geometry is consistent, for if a deduced inconsistency were in this plane non-Euclidean geometry, there would be a corresponding deduced inconsistency in the ordinary geometry of great circles on a sphere, and this geometry is a part of the Euclidean geometry of space. What we have shown to be true of the great circles on a sphere can, by the methods of differential geometry, be shown to be true of the geodesics on any surface of constant positive curvature.

Just as the plane non-Euclidean geometry of Riemann can be realized on a surface of constant positive curvature, so also can the plane non-Euclidean geometry of Lobachevsky and Bolyai be similarly realized on a surface of constant negative curvature. Perhaps the simplest surface of constant negative curvature is the *pseudosphere*, or *tractoid*. To define this surface we first define a plane curve known as the *tractrix*. The tractrix may be generated as follows: Imagine a piece of inextensible cord lying along the positive y-axis (see Figure 3.5), one end of the cord lying at the origin, and the other end having attached to it a small heavy pellet. If the end lying at the origin is now pulled along the x-axis, the pellet will trace a kind of curve of pursuit; this curve is the tractrix, as shown in Figure 3.5. The curve is symmetrical in the y-axis and has the x-axis for an asymptote. Now the pseudosphere is the surface of revolution obtained by revolving the tractrix about its asymptote as an axis of rotation (see Figure 3.6). It can be shown that the geometry of the geodesics on this surface satisfies the postulates of the non-Euclidean geometry of Lobachevsky and Bolyai, after there has been suitable particularization of terms, but the proof here is not as simple as in the previous case of the sphere and can perhaps best be

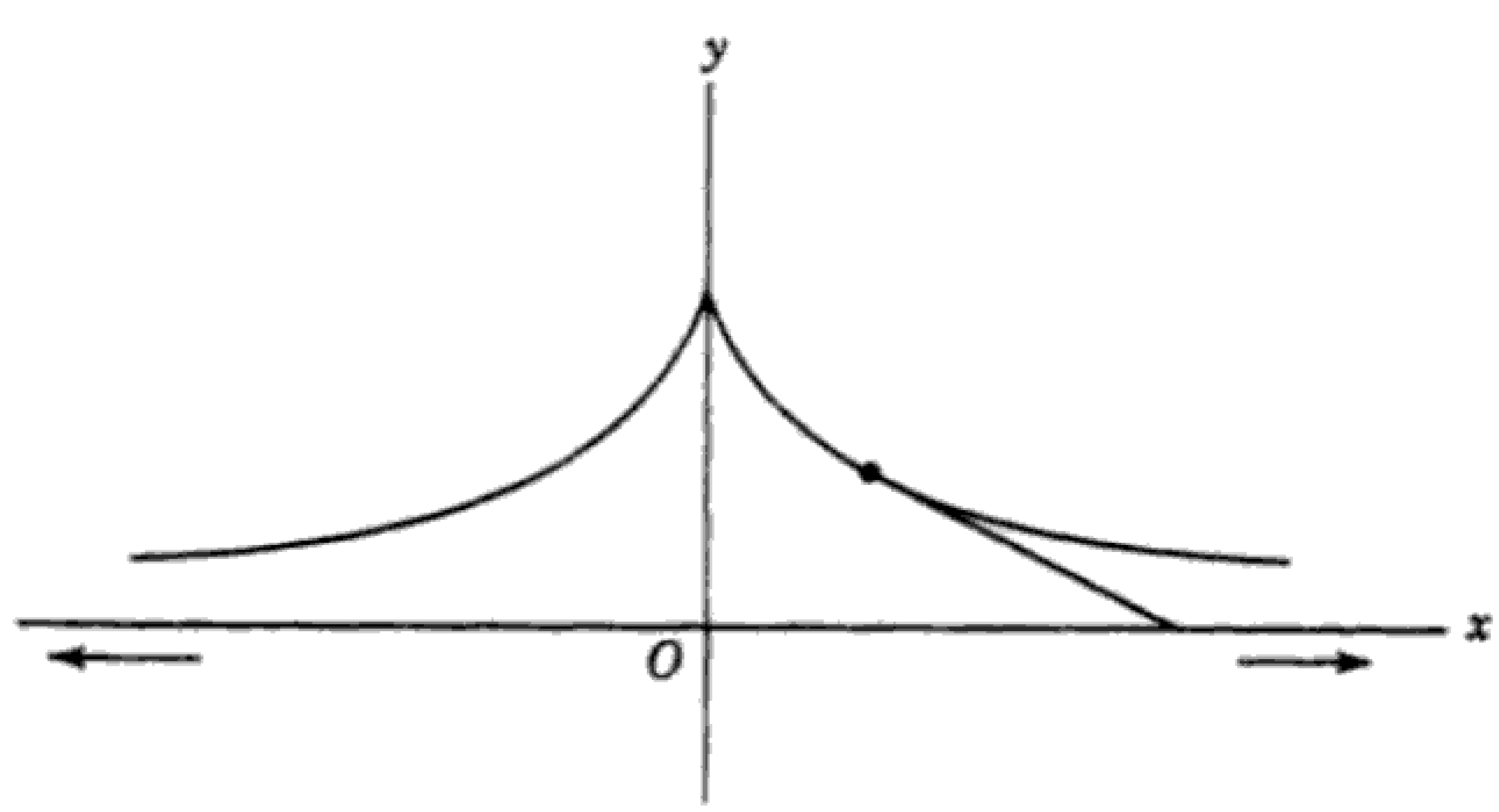


FIGURE 3.5

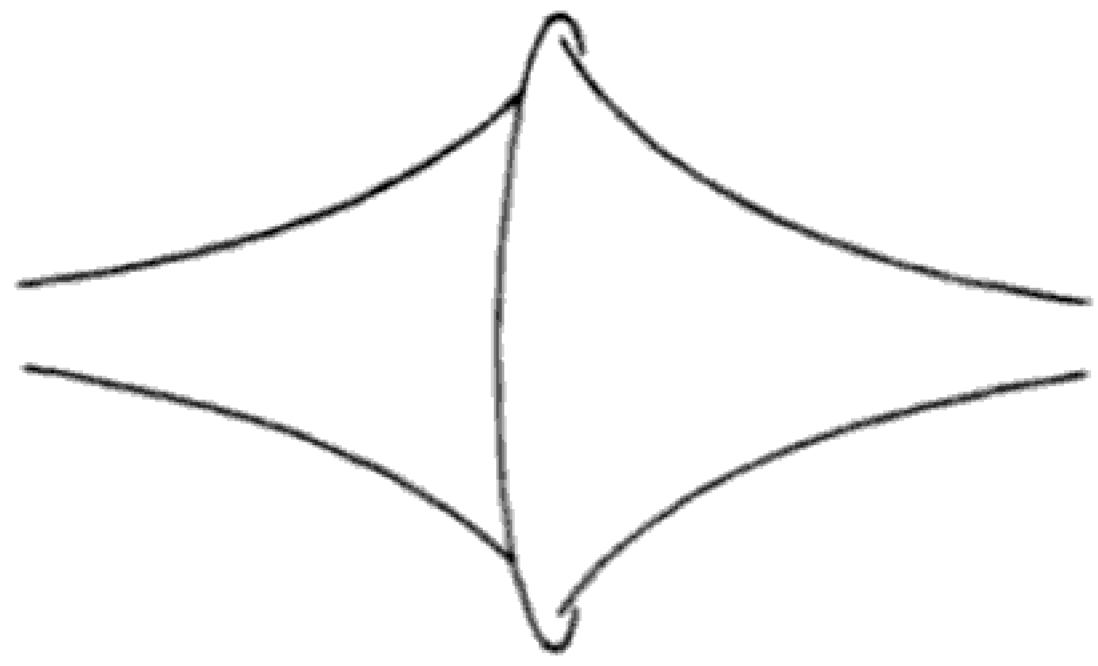


FIGURE 3.6

accomplished by the differential geometry methods employed by Beltrami for the general surface of constant negative curvature. This model, or representation, in Euclidean space, of the plane non-Euclidean geometry of Lobachevsky and Bolyai shows that this plane non-Euclidean geometry, too, is consistent if Euclidean geometry is consistent; for, once more, any inconsistency in the plane non-Euclidean geometry would imply a corresponding inconsistency in the Euclidean geometry of geodesics on the pseudosphere.

The pseudosphere and its geodesics, considered as a representation of the plane non-Euclidean geometry of Lobachevsky and Bolyai, is not as satisfactory as the sphere and its geodesics, considered as a representation of the plane non-Euclidean geometry of Riemann, for the pseudosphere represents only a limited part of the one non-Euclidean plane, whereas the sphere represents the whole of the other non-Euclidean plane. Beltrami conjectured, and it has since been proved, that no surface of constant negative curvature can represent the entire plane in the first case. Also, it is to be noted that neither of our representations takes into account any solid non-Euclidean geometry.

After the discovery of the above representations of the two classical non-Euclidean geometries on surfaces of constant curvature, many other, and in some ways more satisfying, representations of a different nature were devised. In Chapter 4 we shall examine an elementary representation of Lobachevskian non-Euclidean geometry that was devised by the great French mathematician Henri Poincaré (1854–1912). This method of models, or representations, however, does not establish absolute consistency but merely relative consistency. All we can assert from such models is that the two classical non-Euclidean geometries are consistent if Euclidean geometry is consistent. The possibility of an absolute test of consistency for a postulate set will be considered in a later chapter.

One consequence of the consistency of the non-Euclidean geometries is, of course, the final settlement of the ages-old problem of the parallel postulate. The consistency established the fact that the parallel postulate is independent of the other assumptions of Euclidean geometry and proved the impossibility of deducing the postulate as a theorem from those other assumptions, for if the parallel postulate could be so deduced there would have to be an inconsistency in the non-Euclidean systems.

But some consequences of the consistency of the non-Euclidean geometries are much more far-reaching than the settlement of the parallel postulate problem. One of the chief of these is the liberation of geometry from its traditional mold. The postulates of geometry become, for the mathematician, mere hypotheses whose physical truth or falsity need not concern him; the mathematician may take his postulates to suit his pleasure, as long as they are consistent with one another. A postulate, as the word is employed by the mathematician, has nothing to do with "self-evidence" or "truth." With the possibility of inventing such purely "artificial" geometries it became apparent that physical space must be viewed as an empirical concept derived from our external experiences, and that the postulates of a geometry designed to describe physical space are simply expressions of this experience, like the laws of a physical science. Euclid's parallel postulate, for example, insofar as it tries to interpret actual space, appears to have the same type of validity as Galileo's law of falling bodies; that is, they are both laws of observation that are capable of verification within the limits of experimental error. This point of view, that geometry when applied to actual space is an experimental science, or a branch of applied mathematics, is in striking contrast to the Kantian theory of space that dominated philosophical thinking at the time of the discovery of the non-Euclidean geometries. The Kantian theory claimed that space is a framework already existing intuitively in the human mind, that the axioms and postulates of Euclidean geometry are *a priori* judgments imposed on the mind, and that without these axioms and postulates no consistent reasoning about space can be possible. That this viewpoint is untenable was incontestably demonstrated by the invention of the non-Euclidean geometries.

Indeed, the consistency of the non-Euclidean geometries not only liberated geometry but had a similar effect on mathematics as a whole. Mathematics emerged as an arbitrary creation of the human mind, and not as something essentially dictated to us of necessity by the world in which we live. The matter is very neatly put in the following words of E. T. Bell:

> In precisely the same way that a novelist invents characters, dialogues, and situations of which he is both author and master, the mathematician devises at will the postulates upon which he bases his mathematical systems. Both the novelist and the mathematician may be conditioned by their environments in the choice and treatment of their material; but neither is compelled by any extrahuman, eternal necessity to create certain characters or invent certain systems.[14]

The invention of the non-Euclidean geometries, by puncturing a traditional belief and breaking a centuries-long habit of thought, dealt a severe blow to the *absolute truth* viewpoint of mathematics. In the words of Georg Cantor, "The essence of mathematics lies in its freedom."

Since we have a number of geometries of space—the Euclidean and the two classical non-Euclidean geometries—the question is often asked, "Which is the true geometry?" This question is, of course, quite meaningless when geometry is considered a branch of mathematics, because all we can say about truth with respect to a branch of mathematics is that if the postulates are true then the theorems are true. If, on the other hand, geometry is considered a branch of physics, then the question becomes more meaningful. But even here we cannot give a simple and definite answer. When it comes to the applications of several

[14] Quoted by permission from E. T. Bell [3], p. 330.

mathematical theories to a given physical situation, we are interested in that mathematical theory that best explains, or most closely agrees with, the observed facts of the physical situation and that will stand the kinds of tests customarily placed on hypotheses in any field of scientific inquiry. In the present case, then, we are interested in which of the Euclidean and non-Euclidean systems of geometry most closely agrees with the observed facts of physical space. It is not difficult to show that all three geometries under consideration fit our very limited portion of physical space equally well, and so it would seem we must be content with an indeterminate answer until some crucial experimental test on a great scale can be devised to settle the matter. Such a crucial test would appear to be the measurement of the sum of the three angles of a large physical triangle. To date no deviation, exceeding expected errors in measurement, from 180° has been found in the sum of the angles of any physical triangle. But, we recall, the discrepancy of this sum from 180° in the two non-Euclidean geometries is proportional to the area of the triangle, and the area of any triangle so far measured may be so small that any existing discrepancy is swallowed by the allowed errors in measurement. There are even some reasons for believing that physical experiments will never be able to resolve the matter anyway. In this event, then, we would do better to ask not which is the *true* geometry but which is the *most convenient* geometry, and this convenience may depend on the application at hand. Certainly, for drafting, for terrestrial surveying, and for the construction of ordinary buildings and bridges, Euclidean geometry is probably the most convenient simply because it is the easiest with which to work.

There are physical studies where geometries other than the Euclidean have been found to be more acceptable. For example, Einstein found in his study of the general theory of relativity that none of the three geometries that we have been considering is, in itself, adequate, and he adopted a suitable generalization of the Riemannian non-Euclidean geometry wherein the curvature of space may vary from point to point of the space. Again, a recent study[15] of *visual space* (the space psychologically observed by persons of normal vision) came to the conclusion that such a space can best be described by Lobachevskian non-Euclidean geometry. Other examples can be given.

Though it may be logical to call any geometry whose postulate system is not equivalent to a postulate system of Euclidean geometry a non-Euclidean geometry, custom has reserved this term only for the two geometries that result from the hypotheses of the acute and obtuse angle. Many other geometries other than these two, and that differ from Euclidean geometry, have been devised. Riemann was the originator of a whole class of these other geometries, usually referred to as *Riemannian geometries*, of which the Riemannian non-Euclidean geometry is a particular example. One of the accomplishments of the twentieth century was the development of general *non-Riemannian geometries*. Another geometry different than that of Euclid, invented through a deliberate application of the postulational method, is one by Max Dehn (1878–1952) in which the postulate of Archimedes is denied; such a geometry is referred to as a *non-Archimedean geometry*. The creation of these new geometries considerably

[15] R. K. Luneburg.

FOUNDATIONS AND FUNDAMENTAL CONCEPTS OF MATHEMATICS

Howard Eves

This third edition of a popular, well-received text offers undergraduates an opportunity to obtain an overview of the historical roots and the evolution of several areas of mathematics. The selection of topics conveys not only their role in this historical development of mathematics but also their value as bases for understanding the changing nature of mathematics.

Among the topics covered in this wide-ranging text are: mathematics before Euclid, Euclid's *Elements*, non-Euclidean geometry, algebraic structure, formal axiomatics, the real numbers system, sets, logic and philosophy and more. The emphasis on axiomatic procedures provides important background for studying and applying more advanced topics, while the inclusion of the historical roots of both algebra and geometry provides essential information for prospective teachers of school mathematics.

The readable style and sets of challenging exercises from the popular earlier editions have been continued and extended in the present edition, making this a very welcome and useful version of a classic treatment of the foundations of mathematics. "...a truly satisfying book."—Dr. Bruce E. Meserve, Professor Emeritus, University of Vermont.

Unabridged Dover (1997) republication of the third edition, published by PWS-KENT Publishing Company, Boston, 1990. (1st ed., 1958; 2nd ed., 1965). Foreword. Preface. Problems at ends of chapters. Bibliography. Index. Solution Suggestions for Selected Problems. 362pp. 6⅛ x 9¼. Paperbound.

cover design by Manuela Paul

ISBN 0-486-69609-X

90000

9 780486 696096

$16.95 IN USA
$25.50 IN CANADA