CALIFORNIA STATE UNIVERSITY, NORTHRIDGE

THE HISTORY OF MATHEMATICS

A Course of Study

A thesis submitted in partial satisfaction of the requirements for the degree of Master of Science in Mathematics,

Option II

by

Lawrence Douglas Boone, Jr.

June, 1973
The thesis of Lawrence Douglas Boone, Jr. is approved:

Committee Chairman

California State University, Northridge

June, 1973
DEDICATION: To my wife, Judy, without whose infinite patience and untiring assistance this work would not have been possible.
ACKNOWLEDGMENT: My sincere thanks to two fine scholars—Dr. Barnabas Hughes, O.F.M., teacher and mathematical historian extraordinaire whose inspiration, constructive criticisms and manuscript suggestions were extremely helpful, and to Dr. William Karush who spent many hours reading the manuscript and made excellent constructive suggestions to correct and strengthen mathematical content, and also to Mrs. Cecilia Tooker who typed the final copy.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>vii</td>
</tr>
<tr>
<td>I. The Development of Geometry</td>
<td></td>
</tr>
<tr>
<td>Ancient Beginnings</td>
<td>1</td>
</tr>
<tr>
<td>Babylonian and Egyptian</td>
<td>3</td>
</tr>
<tr>
<td>Indian and Chinese</td>
<td>6</td>
</tr>
<tr>
<td>The Greek Contribution</td>
<td>6</td>
</tr>
<tr>
<td>Pythagorean Geometry</td>
<td>8</td>
</tr>
<tr>
<td>Three Famous Problems from Greek Antiquity</td>
<td>16</td>
</tr>
<tr>
<td>Euclid's Elements</td>
<td>30</td>
</tr>
<tr>
<td>Logical Shortcomings</td>
<td>33</td>
</tr>
<tr>
<td>Post-Euclidean Geometry</td>
<td>38</td>
</tr>
<tr>
<td>Greek Trigonometry</td>
<td>44</td>
</tr>
<tr>
<td>India, Arabia, and the Orient</td>
<td>47</td>
</tr>
<tr>
<td>Non-Euclidean Geometries</td>
<td>50</td>
</tr>
<tr>
<td>Projective Geometry</td>
<td>61</td>
</tr>
<tr>
<td>Analytic Geometry</td>
<td>71</td>
</tr>
<tr>
<td>Topology</td>
<td>73</td>
</tr>
<tr>
<td>Klein's Erlangen Programme</td>
<td>80</td>
</tr>
<tr>
<td>Footnotes</td>
<td>86</td>
</tr>
<tr>
<td>II. The Development of Algebra</td>
<td></td>
</tr>
<tr>
<td>Origin of Numeration</td>
<td>90</td>
</tr>
<tr>
<td>Babylonian and Egyptian Algebra</td>
<td>96</td>
</tr>
<tr>
<td>Greek Mathematics</td>
<td>102</td>
</tr>
<tr>
<td>Greek Number Theory</td>
<td>112</td>
</tr>
<tr>
<td>Topic</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Hindu and Arabic Algebra</td>
<td>122</td>
</tr>
<tr>
<td>European Algebra to 1600</td>
<td>122</td>
</tr>
<tr>
<td>Probability and Statistics</td>
<td>126</td>
</tr>
<tr>
<td>The Growth of Modern Algebra</td>
<td>128</td>
</tr>
<tr>
<td>Quaternions</td>
<td>133</td>
</tr>
<tr>
<td>Modern Systems</td>
<td>139</td>
</tr>
<tr>
<td>Footnotes</td>
<td>144</td>
</tr>
</tbody>
</table>

III. The Development of Mathematical Logic and Set Theory

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Postulational System</td>
<td>147</td>
</tr>
<tr>
<td>David Hilbert</td>
<td>153</td>
</tr>
<tr>
<td>Cantor's Theory of the Infinite</td>
<td>155</td>
</tr>
<tr>
<td>Boolean Algebra</td>
<td>168</td>
</tr>
<tr>
<td>Footnotes</td>
<td>178</td>
</tr>
</tbody>
</table>

Bibliography                                                | 180  |

Appendices                                                   | 183  |

<table>
<thead>
<tr>
<th>Series</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Pythagoras and the Pythagoreans</td>
<td>184</td>
</tr>
<tr>
<td>B. The Fundamental Theorem of Algebra</td>
<td>191</td>
</tr>
<tr>
<td>C. Gödel's Proof</td>
<td>195</td>
</tr>
<tr>
<td>D. Euclid's Proof of Pythagorean Theorem</td>
<td>199</td>
</tr>
<tr>
<td>E. Babylonian Texts</td>
<td>201</td>
</tr>
</tbody>
</table>
PREFACE

It is my conviction that the history of mathematics as a teaching tool has lead to greater interest and appreciation of its importance on the part of my students. Mathematics is much more meaningful and relevant when students see its incredible humanness, how its ideas originated and developed through history. Today's challenges may provide tomorrow's common notions. The greatest minds of the mathematical past rejected today's common acceptances—Descartes considered negative numbers "false"; Gauss had a "horror of the infinite"; and negative square roots were not even considered numbers until the early nineteenth century.

In this paper, I examine the history of mathematics from three vantage points: the development of geometry, the development of arithmetic and algebra, and the history of logic and set theory. My purpose is to provide a course of study in the history of mathematics which could be used by a secondary teacher in the absence of a textbook for a one-semester's course for academically qualified high school students. It could also be used as a reference source by any mathematics teacher for enrichment material to supplement courses in geometry, advanced algebra, and trigonometry.

Viewed through its history, students will see mathematics as a vibrantly living, constantly changing and developing body of knowledge.
ABSTRACT

THE HISTORY OF MATHEMATICS
A Course of Study
by
Lawrence Douglas Boone, Jr.
Master of Science in Mathematics, Option II
June, 1973

This paper is designed to provide a source of material for use in a one semester's course in the history of mathematics at the high school level. It has been written for use either as a supplement for a textbook or as an independent source in the absence of a textbook. It also provides material to enrich any other mathematics course. Emphasis has been placed on mathematical content and aspects of the history. While there is sufficient descriptive material, a considerable amount of material is included designed for the mathematically mature student. For example, there is a rather complete discussion in Chapter I of the three famous construction problems including their historical importance, solutions, and a proof of their inconstructibility using traditional Euclidean tools—compass and unmarked straight edge.

The work has been divided into three chapters—the history of geometry, the development of arithmetic and algebra, and the development of logic and set theory.
Chapter I traces the development of geometry from its beginnings in antiquity through Euclidean fruition during the magnificent period of ancient Greece. The remainder of the chapter examines the nature of each of the more important varieties including non-Euclidean geometry, projective and analytic geometry, topology, and the Erlangen Programme of Felix Klein.

Chapter II examines the history of numeration, the Eudoxian solution to the problem of irrational magnitudes, major theorems from Greek number theory, Cardano's solution of the cubic equation, modern algebra and quaternions.

Chapter III discusses the important properties of a postulational system, Cantor's theory of infinite sets, and Boole's universal language. There are also five appendices—A., The life and times of Pythagoras and the Pythagoreans; B., A proof of the fundamental theorem of algebra; C., A summary of Gödel's proof of the incompleteness theorem; D., Euclid's proof of the Pythagorean theorem; and E., Babylonian texts.
CHAPTER I: THE DEVELOPMENT OF GEOMETRY

Geometry had its beginnings way back in the fog and mists of antiquity. There seems to be some evidence to support the conclusion that the earliest uses of geometry were definitively enough "earth measurements". Early man must have used it as an aid in measuring sections of land. Gradually, using concrete geometrical problems, discovery and intuition led to perception of certain rules and relationships which were to form the basis of early scientific geometry. There is no evidence to support an abstract or deductive nature of geometry prior to the Greeks of about 500 B.C.

Herodotus, a Greek historian of the 5th century B.C., states the case for practical geometry in the Nile Valley of ancient Egypt.

"They said that this king (Sesostris) divided the land among all Egyptians so as to give each one a quadrangle of equal size and to draw from each his revenues, by imposing a tax to be levied yearly. But every one from whose part the river tore away anything, had to go before him and notify what had happened. He then sent the overseers, who had to measure out by how much the land became smaller in order that the owner might pay on what was left, in proportion to the entire tax imposed. In this way, it appears to me, geometry originated, which passed thence to Hellas."

There is no reason to assume that development of practical geometry was restricted to the Nile River Valley. Evidence seems to indicate a source in other flourishing civilizations of the time in the great river basins such as the Tigris and Euphrates in Mesopotamia (modern day Iraq), the Indus and Ganges of south-central Asia, and the Huang-Ho and Yangtze of eastern Asia. We know that these fertile areas gave birth to much building, drainage, irrigation, and flood control projects, all of which demand a knowledge of practical geometry. Unfortunately, there is almost
no written source to indicate just what sort of knowledge there was in any of the areas except for Mesopotamia and Egypt. There our tale begins with the geometry of Babylonia.

The oldest written records exist on baked clay tablets which have been unearthed in Mesopotamia and have been dated approximately 3000 B.C. We have also numerous supplies of cuneiform (wedge shaped) characters scratched on clay tablets dating from about 2000 B.C. to 1600 B.C. Politically this would put us at the time of the First Babylonian Dynasty of King Hammurabi's era. Later tablets date from the New Babylonian Empire of Nebuchadnezzar II. The myriad of concrete examples show that these people were familiar with the general rules for calculating the areas of a rectangle, right and isosceles triangles (possibly the general triangle), the areas of a trapezoid with one leg perpendicular to the bases, the volumes of a rectangular prism, a right prism with aforementioned trapezoidal base. The circumference of a circle was assumed to be three times the diameter, giving the area of a circle as one-twelfth the area of a square whose side is the circumference of the circle (a recently discovered tablet uses 3 1/8 as a value for \( \pi \)). There are a number of incorrect but fairly accurate formulas coming down to us: the volume of a frustum of a cone or square pyramid as the product of the altitude and the average of the areas of its parallel bases, the area of a quadrilateral having \( a, b, c, d \) as consecutive sides, \( K = \frac{(a+c)(b+d)}{4} \). There are also a number of facts which were known: corresponding sides of two similar right triangles are proportional, the altitude through the vertex of an isosceles triangle bisects the base, and that the angle inscribed in a
The semicircle is a right angle. Interestingly enough, the Babylonians used concrete examples of the Pythagorean right triangle principle as far back as 2000 B.C.

We owe much of our present knowledge of ancient Babylonia to Sir Henry Creswicke Rawlinson. In 1846 he deciphered and interpreted a key to cuneiform inscriptions. Our mathematical knowledge is due primarily to two prominent mathematical historians, Otto Neugebauer and F. Thureau-Dangin. There are over a half million tablets which have been excavated, of which about 300 are strictly mathematical problems. Collections of these tablets are in the great museums in Paris, Berlin, and London; and also in archaeological exhibits at Yale, Columbia, and the University of Pennsylvania. Physically these tablets vary in size from a few square inches to ones approximately 50 to 60 square inches and about an inch and one-half thick. Writing appears on both sides and often around the rounded edges.

One of the most interesting features of Babylonian geometry is its algebraic nature. Let us consider one such example here and others will be found in the appendix of student exercises. This particular problem was found on a tablet dated approximately 1800 B.C.

"An area A consisting of the sum of two squares is 1000. The side of one square is 10 less than 2/3 the side of the other square. What are the sides of the squares?"

Solution: Let the side of one of the squares be $x$ (in modern notation), then the side of the other is represented by $\frac{2}{3} x - 10$. Thus we have $x^2 + \left(\frac{2}{3} x - 10\right)^2 = 1000$. This gives a quadratic equation in $x$, whose Babylonian solution can be found in Chapter II.

According to the research of Neugebauer, we owe to these
ancient people the division of our circle into 360 equal parts. The explanation is the following. The Babylonians used a large distance unit (equal to about seven of our miles) which we will refer to as a Babylonian mile. Later during the first millennium B.C., this became a unit of time, namely the time required to travel a Babylonian mile. According to their calendar, a day (one revolution of the sun) was 12 time-miles and since the time-miles had been subdivided into 30 equal parts, we have \(12 \cdot (30) = 360\) equal parts in a complete circuit.

**Egypt**

As with the Babylonians, we are fortunate to have a wealth of information regarding the mathematics of ancient Egypt. These records are found on papyri and inscriptions on tomb walls. Until recently, Egypt was the richest area for archaeological research probably because of the unusually dry climate and the practice of burying all sorts of artifacts and records with their dead in air tight tombs. It is interesting to note that, contrary to public opinion, the mathematics of ancient Egypt never reached the sophistication achieved by the Babylonians. Several reasons are asserted for this. Babylon was the hub of great trade routes and Egypt was relatively isolated. Another is the fact that the relatively peaceful Nile did not require the extensive engineering efforts that the undependable Tigris and Euphrates did.

The great wealth of mathematical information is found primarily on the Moscow and the Rhind (Ames) Papyri dating from 1650 and 1650 B.C., respectively. Twenty-six of the 110 problems are geometric, primarily concrete examples which concern measurement of land areas and granary volumes. These problems seem to indicate that by experi-
mentations the area of a circle was taken to be equal to that of a
square on 8/9 of the diameter.\footnote{We thus have}
\[\pi r^2 = \left[\frac{8}{9} (2r)\right]^2\]
\[\pi r^2 = \left(\frac{16}{9}\right) r^2\]
\[= \left(\frac{16}{9}\right)^2 = \frac{256}{81} = 3.1605\]

which is a respectable approximation of the value of \(\pi\). We must
mention that neither the Babylonians or the Egyptians had any
conception of the constant ratio of the circumference to the dia-
meter of a circle. The above equation is a modern deduction from
ancient records. Recent information has shown that they apparently
recognized the area of any triangle given by one-half the product
of the base and altitude. There also is no documentary evidence
to indicate a knowledge of the Pythagorean theorem contrary to
some popular misconceptions. The same incorrect Babylonian formula
\[K = \frac{(a + c)(b + d)}{h}\]
is used for the area of a quadrilateral of
successive sides \(a, b, c, d\). There is further, one remarkable problem
in the Moscow Papyrus which leads to the correct formula for the
volume of the frustum of a square pyramid.

"If you are told: A truncated pyramid of 6 for
the vertical height by \(\frac{1}{2}\) on the base by 2 on the top.
You are to square this \(\frac{1}{2}\), result 16. You are to double
\(\frac{1}{2}\), 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 \(\frac{1}{3}\) twice,
result 56. See it is 56. You will find it right.*

It can be shown quite easily that this illustrates the general
formula \(V = \frac{1}{3} h (a^2 + ab + b^2)\) where \(h\) is the height, \(a\) and \(b\) are
the sides of the bases. There is one minor flaw in the problem if
the sides of the bases are not in the ratio of 1:2. The instruction
to double \( h \) would give 2a which is not equal to ab unless \( a:b = 1:2 \). Nevertheless this is quite remarkable considering this formula was known some 1800 to 1900 years before Christ.

We know very little of the progress of Egyptian Geometry from 1800 B.C. to the time of Pythagoras who studied in Egypt. Indeed the above same incorrect formula for the area of a quadrilateral given the sides which was known to the Babylonians 3000 years before was inscribed on the walls of the tomb of Ptolemy who died in 51 B.C.

**Indian and Chinese Geometry**

Unlike Babylonian and Egyptian geometry we know singularly very little of Oriental and Asian Geometries. An explanation is apparently that records were kept on perishable materials like bark and bamboo. The earliest surviving works were written between the first and third centuries B.C. and the period of greatest productivity was between 200 B.C. and 500 A.D. Like the other early geometries we have examined, there seems to be no interest in demonstrative geometry, only the practical, experimental results they could apply. Several works, specifically, Chiu Chang Suan Shu deals partially with the areas of plane figures and the volumes of solids with statements of problems and answers written in prose form. Another work deals with the 3-4-5-triangle and the Pythagorean relationship using rope stretching.

**The Greek Contribution**

The world was changing, the power of Egypt and Babylonia declined, and new peoples came to take their places, the Hebrews, Assyrians, Phoenicians, and Greeks. Of these groups, the latter is of primary concern to us. This new civilization was developing
of the age of deductive mathematics, which according to modern scholars is a fundamental feature of mathematics.

During the first half of the sixth century B.C. this new age was ushered in by Thales of Miletus, one of the "seven wise men" of antiquity. Thales spent his early years as a merchant, becoming wealthy enough to devote his genius to study and his time to travel. There is evidence to support the belief that Thales was the first to prove several elementary geometrical facts by logical reasoning instead of experiment and guess. Among these facts are:

1. A circle is bisected by any diameter.
2. The base angles of an isosceles triangle are congruent.
3. Vertical angles formed by two intersecting lines are equal.
4. Two triangles are congruent if they have two angles and one side in each respectively congruent.
5. An angle inscribed in a semicircle is a right angle. (known to the Babylonians, but "proved" by Thales).

Anecdotes about Thales, of doubtful truth but of interesting entertainment, may catch the attention of some students. For instance there is the story of the mule which was transporting salt; it found that by rolling in a stream it could dissolve the contents of his load and travel lighter. Thales cured the mule of this by loading him with sponges one trip instead. Again, asked how we might lead more upright lives, Thales replied, "By refraining from doing what we blame in others." Another story goes, having fallen into a ditch while engrossed in the heavens, an old woman asked him how he expected to see anything in the stars when
It must be understood that sources for knowledge of Greek mathematics, unlike the Babylonian and Egyptian mathematics, are comparatively recent. The most important source of information is the Eudemian Summary of Proclus who lived and wrote in the 5th century A.D., who had access to works which are now lost. The Eudemian Summary is based on a full history of Greek Geometry, covering the period prior to 335 B.C. written by Eudemus, a pupil of Aristotle.12

At this point in a course, I would include a lesson prepared by the author on the life and contributions of Pythagoras and the Pythagoreans (see Appendix A).

The Pythagorean Theorem:

The first discovery of the Pythagorean theorem, as we have seen, certainly cannot be credited to the master or to the brotherhood. The first rigorous proof may have been given by them although this is only conjecture. Some mathematical historians attribute proof by dissection to the Pythagoreans (Figure 1).13

![Fig.1(a)](image)

![Fig.1(b)](image)

Let \( a, b, c \) denote the legs and hypotenuse of a right triangle. In using the fact attributed to the Pythagoreans that the sum of the angles of a right triangle is two right angles, we can show that the quadrilateral of side \( c \) in Fig.1b is a square. Since the
areas of the squares are both equal to \((a + b)^2\),
\[a^2 + 2ab + b^2 = c^2 + 2ab\]

or
\[a^2 + b^2 = c^2\]

One of the most important discoveries attributed to the Pythagoreans is the discovery and proof of the irrationality of the square root of 2. As was seen in the lesson on Pythagoreanism (Appendix A), the whole Pythagorean philosophy of the universe was based upon the integers and their simple ratios. Here suddenly was irrefutable evidence that not all magnitudes were commensurable in terms of their sacred cows. This shook the very foundations of their mathematics. Indeed the unfortunate discoverer was asked (as the legend goes) to drown himself, so great was his ostracism. It was many years before the existence of the square root of two would be admitted by the brotherhood. It was Eudoxus' Theory of Proportions (ca 370 B.C.) which finally settled the whole issue (see Chapter II).

A geometrical proof of the incommensurability of the diagonal of the unit square with its side is interesting. Suppose otherwise, then there exists a segment \(AP\) (Fig. 2) such that both diagonal

![Fig. 2](image-url)
AC and side AB of square ABCD are integral multiples of AP, in other words are commensurable with respect to AP, i.e., that
AC = k_1AP and AB = k_2AP where k_1 and k_2 are integers. On AC, lay
off CB_1 = AB and draw B_1C_1 \perp CA. Draw CC_1 forming two congruent
triangles CB_1C_1 and CBC_1, from which we obtain C_1B = B_1C_1. Now
since \angle B_1AC_1 = 45^\circ, triangle AB_1C_1 is isosceles right, thus
BC_1 = B_1C_1 = AB_1. Then AC_1 = AB - AB_1 and AB_1 are commensurable
with respect to AP. Now lay off on C_1A a length C_1B_2 = B_1A and
construct B_2C_2 \perp AC_1 at B_2. By the above reasoning we have AB_2
and AC_2 both commensurable with respect to AP. Eventually we will
have a diagonal AC_n and side AB_n are commensurable with respect to
AP, i.e., AC_n = k_nAP for some integer k_n; and AC_n < AP. This
absurdity proves the theorem. It must be noted that this
geometric proof of incommensurability is not usually associated
with the brotherhood, but is nevertheless relevant to the present
discussion. The algebraic proof actually attributed to the
Pythagoreans may be found in the next chapter.

Theodorus of Cyrene (ca 425 B.C.) showed that an isosceles
right triangle with unit legs gives rise to the square root spiral
(Fig.3). Repeated application of the Pythagorean theorem gives
rise to the other incommensurables \(\sqrt{3}, \sqrt{5}, \sqrt{6}, \sqrt{7}, \ldots\).
It is relevant and instructive at this point to consider the ingenious, albeit cumbersome, geometric methods devised by the Greeks to carry out algebraic operations. We shall illustrate a geometric interpretation of various algebraic identities and two geometric methods for solving quadratic equations, both of which were probably known to the Pythagoreans.

Proposition 5 of Book II of Euclid's Elements states:

"If a straight line is divided equally and also unequally, the rectangle contained by the unequal parts together with the square on the line between the points of section is equal to the square on half the line."

The proof by dissection follows (see Fig.4):

Let AB be the given straight line with P and Q, the equal and unequal divisions respectively. The statement may be expressed

\[(AQ)(QB) + (PQ)^2 = (PB)^2\]

Proof: (1) \[(AQ)(QB) = \text{Area rect. AGFQ} = A1 + A2\]

\[(PQ)^2 = \text{Area sq. HCEF} = A3\]

\[(PB)^2 = \text{Area sq. PCDB} = (A2 + A4) + A2 + A3\]

(2) Now \[A1 = A2 + A4\]

we have \[A1 + A2 + A3 = (A2 + A4) + A2 + A3\]

or \[\text{AGFQ} + \text{HCEF} = \text{PHLB} + \text{FEDL} + \text{HCEF}\]

and \[(AQ)(QB) + (PQ)^2 = (PB)^2.16\]
Algebraically the proof is easily seen:

Let \( AB = 2a, \ PQ = b, \)

we have \( AGFQ = (a + b)(a - b), \ HCEF = b^2, \ PCDB = a^2 \)

so \( (a + b)(a - b) + b^2 = a^2 \)

or \( (a + b)(a - b) = a^2 - b^2, \)

a familiar identity.

Geometric solutions of certain types of quadratic equations were accomplished by the method of proportions.

Example (1) \( x^2 = ab. \) This is the familiar mean proportional construction (Fig. 5).

![Fig. 5](image)

Pythagorean solutions of equations were accomplished by what was known as "application of areas". Consider (Fig. 6) a segment \( AB \) and a parallelogram \( AQRS \) having side \( AQ \) lying along the ray \( AB. \)

If \( Q \neq B, \) choose \( C \) so that \( QBCR \) is a parallelogram. If \( Q \) is between \( A \) and \( B, \) then parallelogram \( AQRS \) is said to be applied to segment \( AB, \) falling short by parallelogram \( QBCR; \) when \( Q = B, \) parallelogram \( AQRS \) is applied to segment \( AB; \) when \( Q \) lies on \( AB \) extended through \( B, \) parallelogram \( AQRS \) is applied to segment \( AB \) exceeding by parallelogram \( QBCR. \)

Application of areas illustrates the solution of \( x^2 - ax + b^2 = 0. \) The construction is given by Proposition 28 of Book VI of the Elements:
"To apply to a given line segment AB a parallelogram AQRS equal in area to a given rectilinear figure F, and falling short by a parallelogram QBCR similar to a given parallelogram; the area of F is not to exceed that of the parallelogram described on half of AB and similar to the defect QBCR."

Fig. 6

In the problem above we want to apply to a given line segment a rectangle which falls short (is deficient) by a square. We may restate the problem as follows: "To divide a given line segment so that the rectangle contained by its parts will equal a given square, the square not exceeding the square on half the given segment." Algebraically it appears,

\[ x(a - x) = b^2 \]

where \( a \) is the given segment, \( x \) is the applied segment, and \( b \) is the side of the given square, \( (b < \frac{1}{2} a) \). By observation it is seen that this is a variation of the mean proportion construction above. We lay off a given segment \( AB = a \) (Fig. 7) and construct the

semi-circle with diameter \( a \). Construct a line parallel to \( AB \) at
distance b above it. At the intersection of this parallel with
the semi-circle we construct a perpendicular to AB whose inter-
section with AB determines point Q. AQ = x and QB = a - x are the
roots of the equation:

\[ x(a - x) = b^2 \quad \text{or} \]
\[ x^2 - ax + b^2 = 0 \]

We recall from algebra that if r and s are roots of a quadratic,
r + s = a and rs = b^2. Here AQ = r and QB = s. Note that the
roots of the equation

\[ x^2 + ax + b^2 = 0 \]

are represented by the negatives of the lengths of AQ and QB.

Another interesting application of these constructions,
originated by the Pythagoreans, is transformation of areas. This
means constructing one polygon different from but having the same
area as another polygon. The method is exemplified in Figure 8:

![Figure 8](image)

Let ABCDE be a given pentagon. Through B construct a parallel to
diagonal AC intersecting DC extended in R. Triangles ABC and ARC
are equal in area since they have a common base AC and equal alti-
tudes to the common base. Notice that Area EDCA + Area ACR =
Area EDCA + Area ACB. We thus have a quadrilateral ARDE equal in area to pentagon ABCDE. By repeated application of this process we can reduce the quadrilateral to a triangle equal in area to the original pentagon.

If \( b \) is any side of this triangle, \( h \) is the altitude to \( b \), and \( s \) is the side of a square equivalent to the area of the triangle, then we have

\[
\frac{1}{2}bh = s^2
\]

\[
\sqrt{\frac{1}{2}bh} = s
\]

or \( s \) is the mean proportional between \( b \) and \( h/2 \). We have geometrically constructed a square equal in area to a given triangle.

There exist only five different regular polyhedra. These are the tetrahedron with four triangular faces, the hexahedron (cube) with six square faces, the octahedron with eight triangular faces, the dodecahedron with twelve pentagonal faces, and the icosahedron with twenty triangular faces. Plato, a well-known Pythagorean mathematician and a philosopher of no mean repute, showed in his "Timaeus" how to construct models of the solids by putting triangles, squares, and pentagons together to form the solids. In accordance with Pythagorean philosophy each of the first four polyhedrons were associated with the four primal elements of all material bodies—fire, air, water, and earth. The fifth, the dodecahedron, is associated with the "enveloping universe".

There are a number of fascinating aspects of these solids. If time permits, they could be examined in a more advanced class; for instance, (1) finding the volume and surface area of each; (2) if \( v \) = the number of vertices, \( e \) = the number of edges, and
any simply-oriented (convex) polyhedron, \( v - e + f = 2 \).

**Three Famous Problems From Greek Antiquity**

No history of geometry would be complete without an analysis of the three classical problems: duplication of the cube, trisection of a general angle, and squaring the circle. We shall consider the historical significance, solutions, and why, under the specified conditions, their solutions are impossible.

The first is the Delian Problem or the Duplication of the Cube. The story goes that Delians consulted the Oracle at Delphi as to how to be rid of a plague. They were instructed to construct an altar whose volume would be twice the size of the given one. As was the custom of the time only an unmarked straight edge and non-rigid compasses were allowed in geometrical solution of problems. Plato submitted the problem to the geometers at his academy. In the course of time many solutions were offered, but none which met the specified conditions of using no mechanical means. The two Platonic tools were not considered mechanical since they are merely tangible representations of two geometric postulates:

1. to draw a straight line from any point to any point,
2. to describe a circle with any center and any radius.

Menaechmus (ca 350 B.C.) solved the problem by inventing conic sections for the purpose. Let the original volume be one cubic unit and the new volume two cubic units. Calling an edge \( x \), we would have

\[
x^3 = 2
\]

or \( x = 3\sqrt{2} \).

A visiting Babylonian would, of course, consult his bag full of
tables for the cube root of two. Menaechmus, instead, drew a parabola and a hyperbola, which, it must be noted, cannot be constructed using straight edge and compasses. The intersection of the hyperbola \( xy = 2 \) with the parabola \( y = x^2 \) provides the geometric distance for substituting in \( xy = 2 \), thus \( x(x^2) = x^3 = 2 \), (Figure 9) and \( x = 1,2 \), approximately. 21

![Fig. 9](image)

Menaechmus also solved the problem by intersecting two parabolas, \( y^2 = 2x \) and \( x^2 = y \), from which

\[
\begin{align*}
  x^4 &= 2x \\
  \text{or} & \quad x(x^3 - 2) = 0.
\end{align*}
\]

The roots are 0 and \( \sqrt[3]{2} \) (Figure 10).

![Fig. 10](image)

The trisection of a general angle seems to the uninitiated as though it should be easy. Even when told that it is impossible to trisect a given angle with straight edge and compasses alone, students invariably respond that they can do it, "Just watch." Several even turn in their solutions which are, of course, wrong.
The solution of the problem is quite elementary if one is allowed to mark a straight edge. The method is as follows:

Let ADBC be a rectangle and AB a diagonal. Acute \( \angle ABC \) is the angle to be trisected. Draw segment BE in such a manner that its extension will intersect the extension of DA so that \( EF = 2AB \) and G is the midpoint of EF. Now \( AB = EG = GF \). We must now show \( AG = EG \). If right triangle AEF is inscribed in a semicircle with diameter EF, then \( AG = EG = GF \) are radii.

We have
\[ \angle 3 = \angle 4 \]
\[ \angle 1 = \angle 2 \]
\[ \angle 3 = 2 \angle 2 = 2 \angle 5 \]

thus \( \angle 4 = 2 \angle 5 \) and \( \angle ABC \) is trisected.

Note that in order to lay off \( EF = 2AB \), we had to locate the point E on segment AC in such a manner that \( EF = 2AB \). This could only be accomplished by adjusting the position of a straight edge on which is marked a segment equal to in length to \( 2AB \). The Euclidean postulate upon which a straight edge is based allows the drawing of a line between two given points. If either point E or point F were constructible, then the construction would be possible.

Several higher curves were invented by Greek geometers to solve this problem. Nicomedes (ca 240 B.C.) invented the conchoid
(blue line) for this purpose (Fig.12).

It is defined as follows: Let 0 be a fixed point at distance "a" from a fixed line. Now pass a pencil of rays through 0 and passing through the fixed line. At the points of intersection with the given line lay off a given distance b, in both directions. The locus of points thus obtained is a conchoid. If b < a, a conjugate point is produced; if b > a, as shown, there is a small loop; if b = a, the loop becomes a single point.

The conchoid was originally constructed to locate the required point F on the extension of DA. If point B is taken as a pole, AC as the given line, known as directrix, and given distance b equal to twice AB, the conchoid will cut DA extended in the required point F.\textsuperscript{22}

A proof of the impossibility of constructing these solutions with compass and unmarked straight edge depends upon a general examination of constructible numbers.

Given a segment of unit length, it is possible to construct all segments of the form a + b, a - b, ab, and a/b, where a and b are two segments constructible from the unit segment. For the sake of simplicity at this point, we refer to a set of numbers...
closed under the four fundamental operations as a rational number field. (A more accurate definition may be found on page 134.)

It was shown above that we transcend the rational number field by constructing a segment \( x = \sqrt{2} \). Then any number of the form \( a + b \sqrt{2} \) may be constructed where \( a \) and \( b \) are rational and are themselves constructible. Likewise it is easily seen that a number such as \( a + b \frac{\sqrt{2}}{c + d} \) or \( (a + b \sqrt{2})(c + d \sqrt{2}) \) can be written in the form \( a + b \frac{\sqrt{2}}{c + d} \):

\[
\frac{a + b \sqrt{2}}{c + d \sqrt{2}} \cdot \frac{c - d \sqrt{2}}{c - d \sqrt{2}} = \frac{ac + 2bd}{c^2 - 2d^2} + \frac{bc - ad}{c^2 - 2d^2} \sqrt{2}
\]

where \( p \) and \( q \) are rational. The denominator \( c^2 - 2d^2 \) is not zero, for that would imply that \( \frac{c}{d} = \sqrt{2} \) which is impossible. In a similar manner \( (a + b \sqrt{2})(c + d \sqrt{2}) \) can be shown to be of this form.

It follows from the preceding discussion that all numbers of the form \( a + b \sqrt{2} \) constitute a field. More generally, let \( F_0 \) represent the rational number field and \( F_1 \) represent the field of numbers of the form \( a + b \sqrt{k} \) where \( k \) is any member of \( F_0 \). \( F_1 \) may be termed an "extension field" of \( F_0 \) since all elements of \( F_0 \) are also elements of \( F_1 \). If we take \( k \) to be a perfect square, then \( \sqrt{k} \) and thus \( a + b \sqrt{k} \) is a member of \( F_0 \).

The totality of all constructible numbers can now be described precisely. Start with a field \( F_0 \) defined by whatever quantities are given. \( F_0 \) will be rational if a single element chosen as the unit is given. Then form \( F_1 \) by constructing \( \sqrt{k_0} \) where \( k_0 \) is a member of \( F_0 \) but \( \sqrt{k_0} \) is not, and forming all numbers \( a_0 + b_0 \sqrt{k_0} \), where \( a_0 \) and \( b_0 \) are members of \( F_0 \). \( F_2 \) will be a new extension field of \( F_1 \) if elements of the form \( a_1 + b_1 \sqrt{k_1} \) are constructed.
where $a_1$, $b_1$, and $k_1$ are elements of $F_1$, and $k_1$ is not. Continuing in this manner after $n$ adjunctions of square roots a field, $F_n$, is reached.

It must be shown that application of compass and straight edge yield only numbers of the form $a + b\sqrt{k}$. More specifically, that the intersections of two circles; the intersections of a circle and a straight line; or the intersection of two straight lines yields only such numbers.

An equation of a circle is $x^2 + y^2 + ax + by + c = 0$ and of a straight line is $px + qy + r = 0$, where $a$, $b$, $c$, $p$, $q$, $r$ are in $F_0$. Solving these two equations simultaneously for one variable, it is easily seen that we obtain a quadratic equation in one variable of the form $Ax^2 + Bx + C = 0$. The roots of this equation are obviously of the form $m + n\sqrt{k}$, with $m$, $n$, $k$ in $F_0$. Similarly if the equation of a second circle is given by $x^2 + y^2 + a'x + b'y + c' = 0$, then subtracting this equation from that of the original circle, we obtain the linear equation

$$(a - a')x + (b - b')y + (c - c') = 0$$

which may be solved simultaneously with either circle equation as above.

Third, it is trivial that the intersection of two linear equations whose coefficients are in $F_0$ yields coordinates which are in $F_0$.

We have thus shown that the intersection of lines or circles yields only numbers which lie in $F_0$ or in the first extension field $F_1$.

It may be deduced from the foregoing that all "constructible numbers are those and only those which can be reached by such a
sequence of extension fields; that is, lie in some field $F_n$ of the type described.  

Next we show that all constructible numbers are algebraic. An algebraic number is a number which is a root of the equation

$$a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0$$

with integral coefficients.

Assume that $F_0$ is the field of rational numbers, then $F_1$ contains numbers of the form $a + b \sqrt{w}$, where $a$, $b$, and $w$ are in $F_0$ but $\sqrt{w}$ is not. Now $x = a + b \sqrt{c}$ implies that $(x - a)^2 = bc$, which is quadratic. Thus $x$ in $F_1$ satisfies a quadratic equation. Proceeding in this way it can be seen that $x$ in $F_2$ would satisfy an equation of the $2^2$ or 4th degree. Assume inductively that $x$ in $F_k$ implies $x$ of the $2^k$ degree. Now $x_k$ in $F_k$ implies that $x_k = p + q\sqrt{w_{k-1}}$ where $w_{k-1}$ is in $F_{k-1}$. If $x_{k+1}$ is in $F_{k+1}$, then $w_k$ is in $F_k$ and $w_k$ is in $F_{k+1}$ and the degree of $x_{k+1}$ must be twice the degree of $x_k$. Consequently $x_{k+1}$ is $2 \cdot 2^k = 2^{k+1}$ degree, as was to be proven. We have proven that all constructible numbers are members of extension fields, $F_k$ whose elements satisfy equations of $2^k$ degree.

With this background the non-constructibility of the first two classical problems with compass and straight edge follows.

First consider the problem of duplication of the cube. As was shown, a constructible number must satisfy the cubic equation $x^3 - 2 = 0$. The proof is indirect. Assume that such a number is constructible, thus $x$ is in some extension field $F_k$ obtained from the rational field by successive additions of square roots. It can be shown in a proof similar to the one given for $\sqrt{2}$, that $\sqrt[3]{2}$ is irrational, thus $x$ cannot be in the rational field $F_0$. 
Assume then that \( x = p + q\sqrt{w} \) where \( p, q, \) and \( w \) belong to some field \( F_{k-1} \), but \( \sqrt{w} \) does not. It will be shown that if \( x = p + q\sqrt{w} \) is a solution of the cubic equation \( x^3 - 2 = 0 \), that \( p - q\sqrt{w} \) is also a root. Since \( x \) is in \( F_k \), then \( x^3 - 2 \) is also in \( F_k \), and
\[
 x^3 - 2 = a + b\sqrt{w} = 0
\]
where \( a \) and \( b \) are in \( F_{k-1} \). Since
\[
 x = p + q\sqrt{w}
\]
\[
 (p + q\sqrt{w})^3 - 2 = a + b\sqrt{w}
\]
which gives
\[
 a = p^3 + 3pq^2w - 2
\]
and
\[
 b = 3p^2q + q^3w.
\]
Now let \( y = p - q\sqrt{w} \), and by substituting \(-q\) for \( q \) in the equation for \( a \) and multiplying equation for \( b \) by \(-1\),
\[
 (2) \quad y^3 - 2 = a - b\sqrt{w}.
\]
Since \( x \) is a root of \( x^3 - 2 = 0 \), then
\[
 (3) \quad a + b\sqrt{w} = 0.
\]
But this implies that \( a \) and \( b \) are both zero; for if \( b \) were not zero, then from \( (3) \) \( w = -\frac{a}{b} \) and hence is a member of \( F_{k-1} \) contrary to our original assumption. If \( b = 0 \), \( a \) is immediately seen to be zero also. If \( a = b = 0 \), equation \( (2) \) implies that \( y^3 - 2 = 0 \), thus that \( y = p - q\sqrt{w} \) is a root of equation \( (1) \).
Furthermore \( y \neq x \), for \( x = y \) implies
\[
 x - y = 0 = 2q\sqrt{w}
\]
which can only vanish if \( q = 0 \) or \( \sqrt{w} = 0 \). But this is impossible since if either were zero we would have \( x = p \) in \( F_{k-1} \) contrary to assumption. The two roots \( p + q\sqrt{w} \) and \( p - q\sqrt{w} \) are obviously real since \( p, q, \) and \( w \) are real.

It has been proven that if \( x = p + q\sqrt{w} \) is a root that \( y = p - q\sqrt{w} \) is a different root. This leads to an immediate
contradiction since there exists only one real cube root of 2, the other two being complex. This last point is easily seen by writing $x^3 - 2$ in the form

$$(x - 2^{1/3})(x^2 + 2^{1/3}x + 2^{2/3}).$$

Using the discriminant of the quadratic formula

$$B^2 - 4AC = 2^{2/3} - 2^{8/3} < 0$$

implying the roots of the second factor are complex as was to be shown.

The basic assumption that $x$ lies in some field $F_k$ leads to a falsehood; hence it must be incorrect. Thus doubling a cube by straight edge and compass is impossible.

In order to show the impossibility of trisecting the general angle with Euclidean tools we must prove the theorem: "If a cubic equation with rational coefficients has no rational root, then none of its roots is constructible from the rational field $F_0$."

The proof is indirect again. Suppose $x$ were a constructible root of

$$(1) \quad z^3 + az^2 + bz + c = 0.$$  

Then $x$ would lie in some extension field $F_k$. Assume without loss of generality that $k$ is the smallest such integer such that $F_k$ contains $x$. Now $k > 0$ since $x$ is not rational by hypothesis. Let

$$x = p + q\sqrt{w}$$

where $p$, $q$, and $w$ are in $F_{k-1}$, but $\sqrt{w}$ is not. By the preceding proof of impossibility of doubling the cube, the second root is

$$y = p - q\sqrt{w}$$

where as before $x \neq y$.

Now in a cubic equation the sum of the roots is equal to $-a$ in equation (1). Thus the third root is $u = -a - x - y$; whence
\[ u = -a - 2p; \]

u then is a number in the field \( F_{k-1} \). But this contradicts the assumption that \( k \) is the smallest integer such that \( F_k \) contains a root of (1). Thus no root of a cubic without rational roots can lie in \( F_k \).

If a particular angle cannot be trisected, then in general an angle cannot be trisected. From trigonometry we obtain the identity:

\[ \cos \theta = 4 \cos^3 \left( \frac{\theta}{3} \right) - 3 \cos \left( \frac{\theta}{3} \right). \]

Taking \( \theta = 60^\circ \) and setting \( x = \cos \left( \frac{\theta}{3} \right) \):

\[ \frac{1}{2} = 4x^3 - 3x \]
\[ 0 = 8x^3 - 6x - \frac{1}{2} \]

Multiplying by 2, \( 8x^3 - 6x - 1 = 0 \).

Recall from algebra that if a polynomial with integral coefficients, \( a_0 x^n + a_1 x^{n-1} + \ldots + a_n = 0 \), has rational roots, they will be of the form \( \frac{a}{b} \) where \( a \) is a factor of \( a_n \) and \( b \) is a factor of \( a_0 \). Thus if \( 8x^3 - 6x - 1 = 0 \) is to have rational roots they must be \( \pm \frac{1}{4}, \pm \frac{1}{4}, \pm \frac{1}{8} \). Substitution reveals that none of these satisfy. Therefore the equation has no rational roots. This does not mean that no angle can be trisected with Euclidean tools, in point of fact a \( 90^\circ \) angle and a number of other specific angles can be trisected.

The third classical problem, to construct a square equal in area to a given circle, has attracted a great deal of attention throughout history. The ancient Egyptians "solved" the problem by taking the side of the square to be \( 8/9 \) the diameter of the given circle. Archimedes (ca 225 B.C.) invented a curve known as the spiral of Archimedes for this purpose (see Fig. 13). The
polar equation is \( r = a\theta \) for some constant \( a \). We take \( OA \) as the polar axis and \( P \) as a point which moves uniformly along the radius vector as it rotates counterclockwise, also uniformly. The ratio of the length of \( OP \) to the measure of \( \angle AOP \), in radians, is 'a' (see Fig. 13). To see how to construct the side, \( s \), of the desired square using this curve, a circle is drawn at \( O \) with radius \( a \). Then \( OP' = \) the length of \( \overline{AB} = a\theta \). If we take \( OP' \perp OA \), then \( OP' \) will be equal in length to one-fourth the circumference of the circle. Further, since area \( K \) of the circle is half the product of its radius and its circumference, i.e.,

\[
\text{Area} = \pi a^2 = \frac{1}{2} a (2\pi a)
\]

we have

\[
s^2 = K = \frac{3}{2} \cdot \theta \cdot (OP') = (2a)(OP')
\]

and \( s \) is the mean proportional between \( 2a \) and \( OP' \).

Figure 14 illustrates how the spiral of Archimedes can be used to trisect (multisect) an angle \( AOB \). Let \( OB \) cut the spiral in \( P \). Trisect the segment \( OP \) in \( P_1 \) and \( P_2 \). If the circles with \( O \) as center and \( OP_1 \) and \( OP_2 \) as radii cut the spiral in \( T_1 \) and \( T_2 \), then \( OT_1 \) and \( OT_2 \) trisect the angle \( AOB \).
The impossibility of squaring the circle can be understood after we recall the definition of transcendental numbers. Algebraic numbers are numbers which are roots of an equation

\[ a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0 \]

with integral coefficients. Transcendental numbers cannot be roots of such an equation. Indeed they transcend, go beyond, classical algebra. Georg Cantor, in 1873, demonstrated that these new numbers together with algebraic numbers form a continuum. We shall deal more with this concept in Chapter III.

In 1882 the German mathematician, Ferdinand Lindemann, proved that the number \( e \) cannot satisfy an equation of the form

(1) \[ C_0 + C_1e^k + C_2e^m + \ldots = 0, \]

where the non-zero exponents \( k, m, \ldots \) and the coefficients \( C_0, C_1, \ldots \) are algebraic. Another way of stating this is that an equation of the form (1) cannot have non-zero exponents and coefficients, all of which are algebraic. If all coefficients are algebraic numbers, then at least one exponent must be transcendental. 29

In order to relate \( e \) and \( \pi \) we must have a further result. Using the infinite series expansion for \( \sin x \) and \( \cos x \)
(2) \( \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots, \)

and (3) \( \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \)

we obtain

(4) \( \sin x + \cos x = 1 + x - \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} - \frac{x^6}{6!} - \frac{x^7}{7!} + \cdots \)

Now from the series

\( e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots, \)

we obtain

\[
e^{ix} = 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \frac{(ix)^4}{4!} + \frac{(ix)^5}{5!} + \cdots,
\]

\[
= 1 + ix - \frac{x^2}{2!} - i \frac{x^3}{3!} + \frac{x^4}{4!} + i \frac{x^5}{5!} - \cdots + \cdots
\]

Collecting real and imaginary parts,

\[ e^{ix} = (1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots) + i(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots) \]

or

\[ e^{ix} = \cos x + i \sin x, \]

which is Euler's Formula. For \( x = \pi \), this formula becomes

\[ e^{i\pi} = \cos \pi + i \sin \pi = -1. \]

This result relating \( e \) and \( \pi \) becomes \( e^{i\pi} + 1 = 0 \), which is of the form of equation (1). It is evident that the coefficients, both equal to one, are algebraic, thus \( i \pi \) must be non-algebraic. Now since \( i \) is clearly algebraic (a solution of \( x^2 + 1 = 0 \)), \( \pi \) is non-algebraic and thus transcendental. The reason for this last statement can be seen in the following argument. Assume \( \pi \) were algebraic, then, since the product of two algebraic numbers is algebraic, the product \( i\pi \) must be algebraic, a contradiction. It was pointed out above that all constructible numbers are algebraic. Thus since \( \pi \) is non-algebraic, it is not constructible.
A number of other quite interesting results are obtainable from the general theorem. Consider the function \( y = e^x \) and its graph in the xy-plane. Except for the point (0,1), \( x \) and \( y \) cannot be simultaneously algebraic. This means that the graph of \( y = e^x \) will not pass through any points whose coordinates are both algebraic. Likewise the common function \( y = \sin^{-1} x \) which is defined by the equation \( 2ix = e^{iy} - e^{-iy} \) cannot pass through points whose coordinates are both simultaneously algebraic numbers.

As our story progresses, we should pause at this time to consider what events were taking place in the world at this time which were to have an effect on the astounding intellectual pursuits which were being made.

The separate Greek city states were in great disarray following the Peloponnesian War and were easy prey for the strong armies of King Philip of Macedonia who defeated Athens at Choeronia in 338 B.C. and Greece finally became part of the Macedonian Empire. In 336 B.C. Alexander the Great succeeded his father and continued the conquest of the then known world, finally conquering what today is the northern coast of Egypt, and founded the city which bears his name in 332 B.C. Following his death in 323 B.C., the control of the far-flung empire was subdivided among his military generals. Egypt fell under the rule of Pтолomy, who selected Alexandria for his capital. As good fortune would have it, Pтолomy was greatly interested in intellectual pursuits and gathered the finest minds in the world around him and built the great University of Alexandria which was to be the center of intellectual progress for the next one thousand years. Euclid was called to head the department of mathematics. The nerve
center of the institution was its library which soon boasted over 600,000 papyrus rolls.

Euclid

Not surprisingly we know very little about Euclid's birth and early life. We do not know when or where he was born. It seems probable that like most other Pythagoreans of his time, he received his mathematical training at the Platonic School of Athens. The story is told of him that when asked by Ptolomy for a short-cut to geometry he replied, "There is no royal road to geometry."

Another tale has it that a student, studying geometry under Euclid, asked what he would get from learning the subject; whereupon Euclid gave the boy a "penny", since he must gain from what he learns.

The Elements

Euclid's original Elements are lost to us. Our copies are based upon a work compiled by Theon of Alexandria (400 A.D.). Careful comparison by historians has revealed little significant differences in the propositions and their proofs from what Euclid is believed to have actually written. Probably no book with the exception of the Bible has been more widely translated and more carefully studied by more people than has Euclid's Elements. They were the only immutable truths for nearly 2000 years. No mathematicians dared disclose any findings which contradicted the Elements until Bolyai and Lobachevsky published their revolutionary non-Euclidean geometries which violated the famous fifth postulate.

Euclid was probably not a first rate mathematician in his own right. But there can be no doubt that he was an organizational genius. His work is largely a collection and a compilation of the
mathematical works which preceded him. The chief merit of the work lies in the skillful selection of propositions and their arrangement into a logical sequence. He collected the scattered and disorganized work of Eudoxus, Plato, Pythagoras, and the works of the Babylonians and Egyptians. He created a logical structure which is the model of mathematics today. Upon a foundation of basic assumptions, some definitions, and axioms (self-evident truths), he built a succession of "elements"; today we would call them theorems. Each is provable in terms of preceding propositions and/or the assumptions (postulates), definitions, and axioms.

The Elements are divided into 13 Chapters or Books which contain a total of 465 propositions with proofs for all. High school plane and solid geometry is contained in much of the material of Books I, III, IV, VI, XI, and XII.

Book I contains 48 propositions which fall into 3 major categories, the first 26 deal mainly with properties of triangles and include the three congruence theorems. Proposition I,5 is the famous "pons asinorum" or "Bridge of Asses", which, in the Middle Ages the mastery of it and its proof marked the culmination of the mathematical training required for the degree. Propositions 26 to 32 establish the theory of parallels and prove the theorem about the sum of the angles of a triangle. The remainder of Book I concerns the spiral properties, quadrilaterals, triangles, and area relationships. Proposition I,47 is the Pythagorean Theorem, the proof of which is the only one universally credited to Euclid himself.\textsuperscript{31} Proposition I,48 is the converse of the Pythagorean Theorem.

Book II deals with the transformation of areas and geometri-
cal algebra we examined above. The last 2 propositions at the end of the book are generalizations of the Pythagorean Theorem which we refer to as the "law of cosines".

Book III contains theorems about circles, chords and tangents which are common to high school geometry texts.

Book IV discusses Pythagorean constructions with straight edge and compass of regular polygons of 3, 4, 5, 6, and 15 sides. It was not known until 1796 when Gauss proved that only regular polygons having a prime number of sides can be constructed with Euclidean tools if and only if that number is of the form $2^{2^n} + 1$. At the age of 19 he showed that a regular polygon of 17 sides can be constructed with Euclidean tools. This was to him his crowning achievement and the one which convinced him to devote his life to mathematics—how fortunate for us, we are so much the richer for it!

Book V is an ingenius presentation of the Eudoxus' theory of proportions to which we will return in Chapter II.

Book VI applies the theory of proportions to plane geometry giving as the familiar similar triangle, mean proportionals, the geometric solution of quadratic equations we discussed earlier.

Books VII, VIII, and IX, a total of 102 propositions, deal with elementary number theory. Included are the well-known Euclidean proof of the infinity of primes, the Euclidean algorithm for finding the greatest common factor of two or more numbers, the fundamental theorem of arithmetic which says that any composite number greater than one can be expressed as a product of primes in one and only one way. We also find the useful theorem concerning the proposition, if $a:b = b:c = c:d$ forms a continued proportion, then $a$, $b$, $c$, $d$ form a geometric progression. Proposition IX,35
derives geometrically the formula for the first n terms of a geometric progression. Proposition IX,36 is the formula for perfect numbers which we will examine in Chapter II.

Book X develops the theory of irrationals or segments which are incommensurable with respect to some given line segment. The subject matter is credited to Thaetetus, but the organization, completeness, and classification belong to Euclid. The first proposition in Book X is the basis for the method of extraction (or Greek theory of limits) which we will examine in more detail later. Its statement is that

"if from any magnitude there be subtracted a part not less than its half, from the remainder another part not less than its half, and so on, there will at length remain a magnitude less than any preassigned magnitude of the same kind."

The remaining three books XI, XII, XIII are solid geometry. The treatment of volumes in Book XII is heavily dependent upon the method of exhaustion.32

Logical Shortcomings

It would be most remarkable indeed if the Elements were a paragon of perfection with no discrepancies in assumptions or in logic, that the subjection of 2000 years of scrutiny by the world's foremost mathematical minds could not turn up a few shortcomings. Such is not the case. There are numerous tacit assumptions made, and the inevitable cloudiness of language. For example, Euclid in Postulate P2 asserts that a straight line may be produced indefinitely but does not imply that it is infinite, merely endless. The great circle on a sphere joining two points is endless but certainly not infinite in length. Another problem is his reliance on appearances in drawings. For example, he assumes in Proposition
that circles with centers at the ends of a line segment and having the segment as a common radius intersect and do not somehow or other slip through each other with no common point. Modern mathematics recognizes the need for a continuity axiom similar to the one furnished by Dedekind which we will consider later in Chapter III. An example is the following: take the circle $x^2 + y^2 = 1$ and the line $y = x$ defined in the plane whose rectangular coordinates are rational numbers. The two figures will never intersect even though the line joins the origin at the center of the circle to the point $(1, 1)$ outside the circle.

To illustrate the danger in relying on figures in proofs, we offer the classic and well-known "proof" that every triangle is isosceles:

1. construct DA bisecting $\angle A$ and let it intersect the perpendicular bisector of BC in D.
2. from D perpendiculars are drawn to sides AB and AC.
3. $\triangle ADF \cong \triangle ADE$ (AAS)
4. $\triangle BDG \cong \triangle CDG$ (SAS)
5. $\triangle FDB \cong \triangle EDC$ (HL)
6. $BF = EC$ and $AF = AE$
7. $BF + FA = CE + EA$
8. $AB = AC$.

Of course the fallacy is that the angle bisector will not necessarily intersect the perpendicular bisector of the opposite side inside the triangle. It is assumed naturally that in teaching
this course the teacher would allow students to discover for themselves the flaw involved.

Another problem is Euclid's wording of his first postulate "to draw a straight line from any point to any point," in that he does not allow for the uniqueness of a line between any two points. An improvement is Hilbert's first two axioms: first, "through any two points there is always a line m" postulating existence and second, "through any two distinct points A and B there is not more than one line m" postulating uniqueness. Euclid did not consider the necessity for undefined terms. The reader will quickly realize that a definition must use previously defined terms and so on back. Upon what do we base our first definition? Today it is a common practice to list point, line, and plane as "undefined" terms and build definitions upon these concepts which we merely describe. We shall consider in more detail the Hilbert and Birkhoff set of axioms for Euclidean plane geometry.

It might be instructive in this connection to explain the philosophical differences between ancient Greek geometrical thought and that of present western culture. The Greek Platonists were idealistic and transcendental. Their geometrical points, and lives were merely physical manifestations of perceptions they had about how the world appeared to them. The Greeks were not concerned about whether their propositions actually fit the physical world, for they were truths which transcend reality. A Platonic geometer would no more consider actually measuring the angles of a triangle to see whether or not their sum was in reality 180° than would a modern scientist accept on faith the existence of something which he could not see either with his eyes, his microscope, or prove
existed with other laboratory paraphernalia. The laws of geometry appeared to the ancients as "the manifestation of the divine spirit which permeates creation." To us reality means that which we can see, touch, feel, or hear from the external world. We actually shape our conceptions of things or ideas to match that which we observe or can prove scientifically. To the Greeks sensual images were fleeting shadows, distorted images of the true world of the spirit. A Greek artist did not copy exactly what he saw, rather he transformed real images to imagined ideals in an impressionistic sense.

It was only around the turn of the last century that mathematicians began to scrutinize from an entirely different frame of reference the Euclidean postulates and proofs. That frame of reference was that nothing could be left for belief or intuition, but must be put on a solid formalistic and strictly defined basis. Terms must be precisely defined and proofs rigorously argued. An example will suffice: to the Greeks as to the modern day man-on-the-street it was self evident that to go from a point inside a circle to a point outside one must of necessity pass through the circle; it was included in their concept of a circle as a closed continuous curve. To a formalistic mathematician this is not necessarily true for it depends upon the frame of reference, i.e., type of coordinate system for which the circle was defined; rational numbers or all real numbers.

Thus we see that to the Greeks and to numerous excellent mathematicians, namely Pascal, Newton, Gauss, etc., Euclidean definitions and postulates were sufficient to prove propositions without a shadow of a doubt. In a word, the Greek system is
cosmically oriented whereas modern science is man oriented.\(^{34}\)

Let us at this point examine the "improvements" which the modern mind has brought to bear on Euclidean geometry by considering David Hilbert's (1862-1943) axiom system for plane geometry published in his "Foundations of Geometry" (1899). Hilbert takes as undefined the notions of point, line, plane, on, between, and congruent. The set of axioms are divided into five groups: incidence, order, congruence, parallelism, and continuity. Nothing is taken intuitively evident. His axioms of congruence eliminate the troublesome Euclidean habit of proving congruence by tautly superimposing one figure on another without thought to whether movement changed size or shape.

As we noted above Euclid assumed in Proposition I.321 that if a line enters a triangle through a vertex, it must if extended far enough intersect a side of the triangle. To the modern mathematical mind this is not necessarily the case unless provided specifically in an axiom. Thus Hilbert's axiom of order which was actually written earlier by Moritz Pasch states:\(^{35}\)

"given 3 non-collinear points A, B, C and a line not passing through any of these points, if a point of the segment AB lies on a given line, a point of AC or a point of BC also lies on the line." (Fig.16)
Post-Euclidean Geometry

The greatest mathematician of antiquity and one of the greatest of all times was Archimedes (287–212 B.C.). His work in higher mathematics was 2000 years ahead of his time. He planted the seeds of the calculus which was not to come to first fruition until the age of Newton almost 800 years later. He was an engineering genius whose inventions contributed to the defense of Syracuse against the Roman siege. He invented catapults and cranes and used a system of compound pulleys to easily and singlehandedly move a heavily laden ship. This led credence to his boast, "Give me a place to stand and I will move the earth."36 He was capable of super powers of concentration so that he was completely oblivious to events surrounding him. The story is told of his discovery of the first law of hydrostatics, that a floating object will displace its weight in water. Using this principle he solved the problem given to him by King Hieron of Syracuse to determine whether his gold crown was actually filled with silver. When he actually discovered the principle he was taking a bath, from which he arose and forgetting his clothes, he ran through the streets shouting "Eureka, Eureka! (I have found it)".

What follows is an explanation of two of his mathematical accomplishments. First, calculating the area of a circle and the area bounded by numerous other closed curves. His method consisted of dividing a circle into a number of rectangles of equal width laid in parallel strips across the circle (Fig.17). By doubling the number of such rectangles we can make the area outside the rectangles but inside the circle as small as we please. The limit of the sum of the rectangular areas is always bounded by the area
of the circle. This method is based upon the Method of Exhaustion of Eudoxus. Archimedes used it to calculate numerous other areas. Leibnitz and Newton called this method integration. By whatever name, the result is the same. All the more remarkable is that Archimedes worked without the aid of a coordinate system or even algebraic notation.

One further application of this limiting process is the classical method for approximating the value of π devised by Archimedes. His method was this: first he calculated the perimeters of inscribed and circumscribed polygons of 6, 12, 24, 48, 96 sides. The inscribed and circumscribed perimeters always provide lower and upper bounds for the circumference of the circle. By constantly increasing the number of sides the difference between the upper and lower bounds becomes smaller and smaller. In modern notation we have,

\[ P_1 < P_2 < P_3 < \cdots \quad \text{where } P_i = \text{perimeters of inscribed polygons} \]

and

\[ P_1 > P_2 > P_3 > \cdots \quad \text{where } p_i = \text{perimeters of circumscribed polygons}. \]

Letting \( E_i \) represent the differences, we have

\[ p_1 - P_1 = e_1 \quad p_2 - P_2 = e_2 \]
and \( e_1 > e_2 > e_3 > \ldots \)
for any \( e \), no matter how small, we can find an \( N \) such that \( e_n < e \)
for all \( n > N \).

In other words, for any arbitrarily small difference, \( E \), we
can find a minimum number of sides, \( N \), for our polygon to ensure
that the difference between perimeters will always be less than \( E \).

Archimedes used \( n = 96 \) to find
\[ 3 \frac{10}{71} < \pi < 3 \frac{1}{7} \]
or
\[ 3.1408 < \pi < 3.1428. \]

This of course we recognize from the modern theory of limits.

Archimedes also calculated many volumes using integration.

His treatise "On Spheres" contains the problem of cutting a sphere
by a plane such that the volumes of the two segments are in a given
ratio to one another. This problem leads to a cubic equation
solved by an ingenious geometrical method. We can illustrate his
famous Method of Equilibrium by showing how it can be used to
obtain the volume of a sphere. This method was not an acceptable
method of proof to him so he used the method of exhaustion to
establish rigorously his formula as correct once he had discovered
it by equilibrium. With the modern theory of limits equilibrium
can be made quite rigorous.

Let \( r \) be the radius of the sphere. Orient it so that its
polar axis is along the \( x \)-axis with north pole, \( N \), at the origin
(Fig. 18). Construct the cylinder and cone of revolution obtained
by rotating the rectangle \( NABS \) and the triangle \( NCS \) about the \( x \)-axis.
Now cut from the three solids thin vertical slices (assumed to be
flat discs) at distance \( x \) from \( N \) and of thickness \( \Delta x \). The formulas
for the volumes of these slices are as follows: For the cylinder,
\[ \pi r^2 \Delta x. \] For the cone we show that since NC makes an angle of \(45^\circ\) with the x-axis, then the radius of the disc for the cone of revolution is equal to \(x\), thus the volume of this disc is given by \(\pi x^2 \Delta x\). For the sphere, consider the formula for the circle whose diameter is NS which is
\[
(x - r)^2 + y^2 = r^2
\]
or
\[
x^2 - 2rx + r^2 + y^2 = r^2
\]
and
\[
y^2 = x(2r - x)
\]
now since \(y\) is the radius of the spherical disc its volume is given by \(\pi x(2r - x) \Delta x\).

For ready reference we list them:

sphere: \(\pi x(2r - x) \Delta x\)
cylinder: \(\pi r^2 \Delta x\)
cone: \(\pi x^2 \Delta x\).

Next we hang at \(T\) the slices from the sphere and the cone, where \(TN = 2r\). Taking moments of these 2 discs about \(N\) and summing we have

\[
[\pi x(2r - x) \Delta x + \pi x^2 \Delta x]2r.
\]
We pause here briefly and define what is meant by taking moments about a point. In physics a moment of a volume about a point is "the product of the volume and the perpendicular distance from the point to the vertical line passing through the centroid of the volume."

Since the discs are located at T, the distance is 2r. So we have

$$
\sum \mu = (2 \pi x \Delta x - \pi x^2 \Delta x + \pi x^2 \Delta x)2r
= 4 \pi r^2 x \Delta x
$$

This is four times the moment of the cylindrical disc located distance x from the origin. Summing a large number of the slices together we have the volume of the three solids. It must be noted that by the time of Archimedes formulas for the volumes of a cone and a cylinder were known to be respectively $\frac{1}{3} \pi r^2 h$ and $\pi r^2 h$. For our particular figure these would be respectively $\frac{8 \pi r^3}{3}$ and $2 \pi r^3$, thus we have

$$
2r(V_{sphere} + V_{cone}) = 4r(V_{cylinder})
= 2r(V_{sphere} + 8 \pi \frac{r^3}{3}) = 8 \pi \frac{r^3}{3}
= 2r(V_{sphere}) = 8 \pi \frac{r^3}{3} - \frac{16 \pi r^3}{3}
= \frac{8 \pi r^3}{3}
V_{sphere} = \frac{8 \pi \frac{r^3}{3}}{3} \cdot \frac{1}{2r}
= \frac{4 \pi r^3}{6r}
= \frac{2 \pi r^3}{3}
$$

All of these herculean accomplishments are made almost mentally unfathomable when we recall that Archimedes had no convenient system of numbers, no symbolic algebra, and above all no pencil and paper to aid him in his labor.

He finally died when, as usual, lost in thought over a mathematical diagram he had sketched in the sand, he was surprised at
work by a pillaging Roman soldier after Syracuse had finally fallen under the Roman heel. He ordered the soldier to stand clear of his diagrams, whereupon the soldier slew the unarmed old geometer with his sword. The philosopher A.N. Whitehead observed the soldier was only doing what he had been taught to do—kill. "No Roman ever died in contemplation over a geometrical diagram."³⁸

The next mathematician of note was Apollonius (ca 262-206 B.C.). His chief claim to fame was *Conic Sections*, which earned him the title "The Greek Geometer". He originated the terms ellipse, parabola, hyperbola for the plane figures generated when a plane cuts a double cone. The terms were borrowed from the early Pythagorean terminology of application of areas. Recall that this refers to the placing or applying the base of a rectangle along a line segment. If the applied base fell short, exactly coincided with or exceeded the segment, it was referred to as an ellipse, a parabola, or an hyperbola respectively. Figure 19 illustrates the idea: Let AB and AR be the x and y axes of analytic geometry, and A be a vertex of a conic. P is any point on the conic and Q is the foot of the perpendicular from P to the major axis, AB. On
the $y$-axis mark off a distance equal to the latus rectum or $p$, of the conic. Apply to the segment $AR$, a rectangle with $AQ$ as one side and an area equal to $(PQ)^2$. According as to whether the application falls short, coincides with, or exceeds $AR$, Apollonius calls it an ellipse, parabola, or hyperbola respectively. Using analytic geometry, if we call the coordinates of $P$, $(x,y)$, then the curve is an ellipse if $y^2 < px$. He also showed in rhetorical and geometric language, the equivalent of $y^2 = px \pm \frac{2}{d}$, where $d$ is the length of the diameter through $A$ for the ellipse and hyperbola depending upon whether the sign between the terms was negative or positive respectively.

Apollonius gave an exhaustive treatment to conics, much more than is usually covered in a course in analytic geometry. Had it not been for Apollonius' development of the conics, Kepler might have never discovered the laws of planetary motion by showing that the path of every planet is an ellipse, Newton might not have formulated his theory of universal gravitation and Einstein might never have challenged that theory with a new one. But what of the other conics? The hyperbola is used in electronic navigation, the equilateral hyperbola $(xy = k)$ is inverse variation, the parabola is the path traced by a projectile such as an arrow or bullet. A parabola is in a sense an infinite ellipse with a second focus at infinity. Automobile headlights and reflective telescopes are parabolic shaped.\textsuperscript{39}

Greek Trigonometry

We close our discussion of the Greek contribution to geometry with a brief investigation of the origins of trigonometry. A Babylonian cuneiform tablet (Plimpton 322) contains a very complete
table of secants. The Babylonian astronomers of the 4th and 5th centuries B.C. accumulated a mass of observational data and passed much of it on to the Greeks. This was the origin of spherical trigonometry. The Greek astronomer Hipparchus (ca 130 B.C.), according to second hand sources made a number of extremely accurate calculations. For example, he is credited with calculating the mean lunar month to within one second of its presently accepted value. He originated the terms of latitude and longitude as applied to the earth.

Most important, he provided the foundation of a table given by Claudius Ptolemy relating the lengths of chords of all central angles of a given circle by half degree intervals from $0.5^\circ$ to $180^\circ$. Reflecting Babylonian influence, Ptolemy divided the radius into 60 parts and expressed his chords of angles in terms of parts or fractional parts of the radius. For example

\[
\text{chord (crd)} \ 36^\circ = 37\frac{1}{60} \cdot 15''
\]

meaning that the chord of $36^\circ$ is equal to $37/60$ of the radius plus $1/60$ of a first part plus $55/3600$ of a second part. We can see from Figure 20 that a table of chords is equivalent to a table of trigonometric sines, for

![Fig. 20](image-url)
\[
\sin \alpha = \frac{AM}{OA} = \frac{\text{AB}}{\text{diameter of circle}} = \frac{\text{crd } 2\alpha}{120}
\]

Ptolemy's table gives the sines of the angles from \(1/4^0\) to \(90^0\) in 15 minute intervals.

Another mathematician of repute was Menelaus who defined a spherical triangle and proved the spherical analogies of many of Euclid's propositions for plane geometry in his work *Sphaerica*.

The last geometer we note is Hero of Alexandria. He worked in the twilight years of the great Greek intellectual period, approximately 200 A.D. One of his best known works is the well-known formula for the area of a triangle in terms of its sides. He also originated the modern iterative method for approximating the square root of any non-square integer which is commonly used today in computers and is also taught in schools. If \(n = ab\), then \(n\) is approximated by \(\frac{a + b}{2}\). Let this be \(a_1\), then

\[
a_2 = \frac{a_1 + n/a_1}{2}
\]

\[
a_3 = \frac{a_2 + n/a_2}{2}
\]

and so on until we have the accuracy required.

Thus we come to the end of the glorious Greek mathematical period. The book is closed on the most intellectually advanced civilization certainly in the ancient world and probably throughout history. The great library of Alexandria had been sacked and burned by Roman soldiers around 300 A.D. Hypatia, the daughter of Theon of Alexandria, the first woman mathematician, was barbarously murdered by fanatical Christians in 415 A.D., who objected to her teachings of Neoplatonic philosophy and her failure to embrace Christianity. Thus ended the golden age of Greece. Following its demise almost nothing of original value was accomplished in mathe-
matics or in any other field for nearly 700 years. The period
known as the Dark Ages now unfolds.

Geometry in India, Arabia, and the Orient

Outside of Greece and those areas under Greek intellectual
influence, geometry was a science of observation and measurement,
the idea of a deductive proof was meaningless, even incomprehensi-
ble generally.

Hindus were interested in constructing very accurate altars
to honor their gods. They found the Pythagorean theorem fairly
early, possibly by 3000 B.C., which precedes any Sumerian record
of it. They were aware of several specific applications, namely
3, 4, 5 and 5, 12, 13 but cared not about a formal proof. Other
problems which interested them were the construction of a square
equal to the sum or difference of 2 given squares. The first is
easily handled using the Pythagorean theorem, the second using a
simple construction as follows. Assume we want

\[ x^2 = a^2 - b^2 \]

or

\[ x^2 + b^2 = a^2 \]

where \( a \) and \( b \) are known lengths. Lay off a segment equal to \( b \),
then at one end construct a right angle, from the other end draw
a circle whose radius is equal to \( a \). The point where the circle
intersects the perpendicular gives the length of \( x \). The next prob-
lem is that of constructing a rectangle upon a given segment equal
to a given square. Their solution is clumsy. The Greek mean
proportional construction is much better.

The next Hindu mathematician mentioned is Arzabhata (6th
century A.D.) who wrote a poem called the "Ganita". In it he
gives correct formulas for the area of a triangle and for the area
of a circle as half the circumference multiplied by half the diameter. He also gives a remarkably accurate value for \( \pi \): "add \( 4 \) to 100, multiply by 8, add 62,000, the result is approximately the circumference of a circle of which the diameter is 20,000." This gives

\[
\pi = \frac{62,832}{20,000} = \frac{3927}{1250} = 3.1416, \frac{42}{1250}
\]

an amazingly accurate value.

Next we have Brahmagupta (628 A.D.) who is reputed to be the most famous of the Hindu mathematicians. His work seems to be a curious mixture of truth and fiction. He gives a formula for the area of a tetragon (quadrilateral) which bears strong resemblance to Heron's formula for the area of a triangle in terms of its sides \( a, b, c, d \).

\[
F = \sqrt{(s-a)(s-b)(s-c)(s-d)}
\]

where \( 2s = a + b + c + d \)

This formula is incorrect in general and only true if the tetragon is cyclic (inscribed in a circle). He makes no mention of this restriction. He also deduced in a way which is not at all clear the correct formula for the volume of a frustum of a pyramid. Possibly he had heard of it. We know that it was used by both the ancient Egyptians and the Chinese.

The one remaining Hindu writer is Bhaskara who lived approximately 1200 A.D. His work is essentially a reworking of Brahmagupta's work. He did contribute a clever geometric proof of the Pythagorean theorem. We shall have more to say about the contributions of the Hindus to the art of computing and to algebra.

_**Oriental Geometry**_

The reason we know little of early Chinese mathematics is this:
In 213 B.C., the Emperor ordered the burning of all books in the kingdom. Undoubtedly some were hidden and thus saved, but certainly a great many were lost. One important work that survived is the Chiu-Chang Suan Chu or Arithmetic in Nine Sections. Three of the nine sections deal with measurement of plane figures, volumes, and right triangles. The value of \( \pi = 3 \), however, was a weakness.\(^3\)

**Arabic Geometry**

The contribution of the Arabs to geometry was to preserve the Greek classics. A great debt of gratitude is owed them for translating and saving such Greek works as we do have.

One of the most important Arabic geometers was Nasir Ed-din who investigated Euclid's parallel postulate and influenced Saccheri (1733) who started his work on non-Euclidean geometry as a result of his knowledge of the thinking of Nasir Ed-din. (We shall look presently at the contribution of Saccheri.) The Arab is also credited with an original proof of the Pythagorean theorem.

Moslem writers can also be credited with contributions of trigonometric tables and use of all six trigonometric functions. It is interesting to note briefly that the Latin equivalent "sinus" of the Arabic word "jaib" is a derivation of the word meaning chord half. Thus the origin of our word "sine".\(^4\)

Like the Hindus, the Arabs regarded themselves primarily as astronomers and their mathematics was developed to serve that end. This is certainly contrary to the Greek idea of mathematics as an end in itself. In Greece, also, the study of mathematics had been "democratic", that is, open to all non-slaves who cared to study it. Only the favored castes of Hindus were allowed to study it.
Non-Euclidian Geometries

"God geometrizes according to Euclid's Elements" was a popular theme in 1821 when Carl F. Gauss wrote a short account of a new geometry he had discovered by assuming that the sum of the angles of a triangle were less than 180°. Thus we begin an analysis of the repercussions felt through the space and time of mathematical history by Euclid's 5th postulate, "that if a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which the angles are less than two right angles.

Along with attempts to prove the 5th came a number of substitutes for the 5th. Among the most famous is one which is commonly used in secondary schools as the "parallel postulate" revised by John Playfair.

"Through a point outside a given line one and only one line can be drawn parallel to a given line." There are a number of equivalents which can be shown to be equivalent and allow alternate notions of a "given line," such as a number of equivalent alternate notions of a given line.

Debate over not to speak of deduction from and allowance of these notions is often confused with the need for the Proposition 29, Book I. Indeed, it takes the bristle, self-external aspects of the four predecessors and the need for the Proposition 29, Book I, to prove this and falling, simply added it as a 5th postulate because a theorem, meet on that side on which the angles are less than two right angles.

Two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which the angles are less than two right angles.

Repercussions fell through the space and time of mathematics like a bolt of lightning less than 180°. Thus we begin an analysis of the theme in (2) when Carl F. Gauss wrote a short account of a new geometry he had discovered by assuming that the sum of the angles of a triangle is likewise less than 180°.
where equidistant.

It is easily seen that the parallel postulate implies that the angle sum of a triangle is equal to 180° as that proof is part of high school plane geometry.

To prove that (1) above implies Playfair's postulate, we need two lemmas which are consequences of (1).

Lemma 1: An exterior angle of a triangle is equal to the sum of the two opposite interior angles. (Fig. 21)

\[
\angle 4 + \angle 3 + \angle 1 = 180^\circ
\]

or

\[
\angle 3 + \angle 1 = (180 - \angle 2) = \angle 1.
\]

Lemma 2: Through a given point P, there can always be drawn a ray making with a given line p, an angle less than any given angle, however small. (Fig. 22)

Proof: (1) From P draw PA₁ perpendicular to p. Lay off on p, A₁A₂ = PA₁ and draw PA₂. Then \(\angle A₁A₂P = \theta\).

\[
\theta = 45^\circ = \frac{180}{2^2}.
\]

Fig. 21

Fig. 22
(3) Lay off \( \triangle A_2A_3 = PA_2 \) and draw \( PA_3 \).

(4) \( \angle 1 = \angle 2 = \frac{9}{2} = \frac{180}{2} \).

(5) Continuing to lay off segments along \( p \) in this manner leads to a triangle \( \triangle PA_{n-1}A_n \) in which

(6) \( \angle A_nPA_{n-1} = \angle A_{n-1}A_nP = \frac{180}{2^n} \), where \( n \) is any positive integer greater than one.

(7) By the postulate of Archimedes, there exists a number, \( k \), such that

\[ k\alpha > \pi \quad \text{(where } \pi \text{ radians } = 180^\circ) \]

and

\[ \alpha > \frac{\pi}{k}. \]

(8) Now choose \( n \) (defined above) such that

\[ 2^n > k \]

and we have

\[ \alpha > \frac{\pi}{2^n} \]

or

\[ \frac{\pi}{2^n} < \alpha. \]

It can now be proven that if the sum of the angles of a triangle is always equal to two right angles, through any point outside a given line exactly one parallel can be drawn to any given line.\(^{15}\)

**Proof:** (Figure 21). Let \( P \) be the given point and \( p \) be the given line. Draw \( PA_1 \) perpendicular to \( p \) and at \( P \) draw \( PB \) perpendicular to \( PA_1 \). By Euclid's proposition 28 (proved without the 5th) "two lines perpendicular to a third line are parallel", \( PB \) parallel \( p \). Consider any line through \( P \) intersecting \( p \) such as \( PA_3 \). Since
\[ \angle 1 + \angle A_1PA_3 = \angle 3 + \angle A_1PA_3 = \frac{\pi}{2} \]

Therefore PB is the only line through P which does not cut p, for, no matter how small an angle a line through P makes with PB, there are other lines through P, by lemma 2, making smaller angles with PB and cutting p. The line below PB must, by Pasch's Axiom (Page 37), also cut p.

There were numerous ill-fated attempts to prove the 5th postulate for over 2000 years. In 1733, one of the most scientific investigations was published by a Jesuit Priest named Girolamo Saccheri, called Euclides ob omni naevo vindicatus (Euclid Freed of Every Flaw). Saccheri constructed a figure known today as the Saccheri Quadrilateral (Figure 23), in which the base angles A and B are right angles and the sides are congruent. It can be easily shown using simple congruence theorems not dependent upon the parallel postulate that the summit angles, C and D, are congruent. There are then three distinct and mutually exclusive possibilities:

- (1) C and D are equal right angles,
- (2) C and D are equal acute angles,
- (3) C and D are equal obtuse angles.

Saccheri's plan was to show indirectly that either (2) or (3) would
lead to a contradiction.\textsuperscript{147} It was unfortunate for him that he so obstinately clung to his conviction of the total inviolability of the Euclidean system; otherwise he might have been credited with the discovery of non-Euclidean geometry. After developing many of the new fundamental theorems of non-Euclidean geometry, Saccheri simply eliminated the acute angle hypothesis by speaking of parallels intersecting at infinity and then lumping the "points at infinity" with real points to reason that Euclid's 5th follows immediately. He thought he had eliminated the obtuse angle hypothesis by assuming the infinity of a straight line, neglecting the possibility that a straight line could be a closed curve.\textsuperscript{148}

Instead of being remembered for his contributions to the extension of geometry, Saccheri died in obscurity, little regarded by his countrymen. We now know that the acute angle hypothesis along with the basic set of Euclidean postulates is as consistent as Euclidean geometry with the hypothesis of the right angle.\textsuperscript{149}

The first to reach definitive conclusions, from assuming the falsity of the Playfair form of the 5th postulate, was Gauss in 1828. He, like many of his predecessors, was afraid of public ostracism and so refrained from publishing his findings. In 1831, Janos Bolyai, an Hungarian army officer, and Nikolai Lobachevsky, a Russian mathematician at the University of Kazan, published very nearly the same conclusions almost simultaneously but completely independently. Their work was based upon postulating more than one parallel through an external point to a given line. What follows is a very brief analysis of a few basic results which follow from the Lobachevskian assumption.\textsuperscript{50}

This new geometry was given the name "hyperbolic geometry" by
Felix Klein. We will make full use of all Euclidean geometry, except for the fifth postulate and all propositions which follow from. The new fifth postulate is the following:

"If \( p \) is any line and \( P \) is a point not on \( p \), then at least two distinct lines can be drawn through \( P \) which do not intersect \( p \)."

**Definition of Parallel Lines:** In Figure 2h, let \( Q \) be the foot of the perpendicular from \( P \) to \( p \). Rotate \( PQ \) about \( P \) in a counterclockwise manner until \( PQ \) no longer intersects \( p \). Assume that \( PN \) is the first such line. Then rotate \( PQ \) clockwise, let \( PM \) be that line. Thus it is said we have two parallels to line \( p \). Any lines which may fall in the interior of angle \( M'PN \) are called ultra-parallels.

We can now define a special type of triangle called an asymptotic triangle. For this purpose it is convenient to consider parallel lines as having a common "point at infinity" or "ideal point", symbolized by \( \Omega \).

**Definition:** Let \( A\Omega \) and \( B\Omega \) be parallels (not ultra-parallels) cut by transversal \( AB \), then triangle \( AB\Omega \) is an asymptotic triangle (Fig.25).
Theorem 1: If two asymptotic triangles $AB\Omega$ and $A'B'O'$ have $\angle B = \angle B'$ and $AB = A'B'$, then $\angle A = \angle A'$ (Figure 26).

Proof: Assume $\angle A > \angle A'$, then draw $AD$ making $\angle 2 = \angle A'$. This line must intersect $B\Omega$ in a real point $D$. On $B'O'$ lay off $B'D' = BD$. Now Euclidean $\triangle ABD \cong \triangle A'B'D'$. Therefore $\angle 2 = \angle A' = \angle 1$ which is a contradiction. In a like manner we could show that $\angle A < \angle A'$ which leads to an absurdity.

We have thus shown that the summit angles 1 and 2 in Figure 26 are equal. We shall call these angles of parallelism for distance $d = PQ$ (Fig.24) and refer to it by the Lobachevskian notation $\pi(d)$. We thus have by Theorem 1 $\pi(d) = \pi(d')$ or angles of parallelism for equal distances are equal.

Theorem 2: Angles of parallelism are acute (Fig.27).

Proof: We consider two cases:

(1) assume that $\angle 1 = \angle 2 = 90$. Then we could have only one parallel since it would violate Euclid's 28th proposition proved without his fifth postulate that two lines perpen-
Fig. 27

dicular to the same line are parallel and MPN would be one straight line violating the axiom of parallelism.

(2) assume the angles are obtuse: then PQ perpendicular to PR, and PR parallel to p would lie within the minimum parallels PN and PM, and intersect the given line, p. This is also a contradiction. Therefore, the angles of parallelism are acute.

There are several further results which could be easily shown: in an asymptotic triangle ABΩ, the exterior angle at A, χ₁, is greater than the interior angle at B (Fig. 28). An immediate corollary of this is the fact that the angle of parallelism decreases as the distance increases, i.e., if d > d', then π(d) < π(d').

Fig. 28

We shall, however, prove one major property of hyperbolic geometry. This was incidentally arrived at by Saccheri, before he rejected all
his efforts.

**Theorem 3:** The sum of the angles of a triangle is less than a straight angle.

**Proof:**

1. Let \( \text{GBCH} \) (Fig. 29) be a Saccheri quadrilateral with angles \( G \) and \( H \) which are equal and acute. Let \( E \) and \( D \) be midpoints of sides \( AB \) and \( AC \) respectively; \( AF \perp DE \).

2. We have \( \triangle GBD \cong \triangle ADF \) and \( \triangle AFE \cong \triangle ECH \) by (ASA).

\[
\angle 1 + \angle 2 + \angle 3 + \angle 4 < \angle 5 + \angle 3 + \angle 4 + \angle 6 < 180^\circ.
\]

![Fig. 29](image)

Other interesting properties are:

1. If three angles are equal respectively to three angles of another triangle, the two triangles are congruent.

2. Two triangles having the same angle sum are equal in area and conversely.

3. Area \( \triangle ABC = k^2(\pi - A - B - C) \).

4. The area of a triangle increases as the sum of its angles decreases and has a finite upper bound of \( k^2\pi \).

The next significant event occurred in 1854 when Bernhard Riemann announced his non-Euclidean geometry based on the elliptic angle of parallelism. We recall that Saccheri was able to show that
the summit angle of his quadrilateral could not be obtuse by assum-
ing that straight lines are infinite in length. Riemann, in a
lecture delivered in 1854, distinguished between the unboundedness
and infinitude of lines in space. He showed that on the surface of
a sphere it was possible to have a line (a great circle) which was
boundless (endless) but finite in length. Thus we have a new geom-
etry based on the following axiom: two lines in space always inter-
sect. Some of the more significant properties of elliptic geometry
are:

(1) All lines perpendicular to the same line p meet in a point
at a constant distance q from p. The point P is called the
pole of line p. And p is called the polar of P.

(2) All lines have the same finite length. A line is reentrant;
i.e., if we start at point A and traverse the line in
either direction we return to A.

(3) Two lines enclose an area.

(4) In a Saccheri quadrilateral the summit angles are equal
and obtuse.

(5) The sum of the angles of a triangle is greater than a
straight angle.

(6) The area of a triangle is proportional to its excess; i.e.,
area $\Delta ABC = k^2 (A + B + C - \pi)$. (Note: the k in the
formula above is the radius where our model is a sphere.)

One issue remains—what kinds of surfaces can serve as models
for non-Euclidean geometries? We have already seen that the surface
of a sphere where lines are defined as great circles is a model for
elliptic or Riemannian geometry, and spherical geometry was well-
known to the ancient Greeks. A model for hyperbolic geometry is not
quite so obvious. One commonly used is that of a pseudo-sphere, a curved surface of constant negative curvature (concave). It is generated by a tractrix (Fig. 30). The lines of a tractrix are asymptotic to the given line.

![Fig. 30](image)

Another model was suggested by Felix Klein. This consists of the interior points of a circle whose lines are chords of a Euclidean circle excluding the endpoints on the circle. It is easily shown that each of the first four Euclidean postulates and the hyperbolic axiom of parallelism are verified (Fig. 31).

![Fig. 31](image)

(1) Two points determine exactly one line.

(2) Two distinct lines intersect in at most one point.

(3) Given line \( l \) and point \( P \), there are an infinite number of lines, e.g., \( l_1, l_2 \) which do not intersect \( l \).
We define distance in the following unique way: \[ d(P, Q) = \ln \left( \frac{QS}{PT} \right) \left( \frac{PS}{QT} \right) \], then we prove

\[ PQ + QR = PR \quad \text{or} \quad d(P, Q) + d(Q, R) = d(P, R) \]

\[
\ln \left( \frac{(QS)(PT)}{(PS)(QT)} \right) + \ln \left( \frac{(QT)(RS)}{(QS)(RT)} \right) = \ln \left( \frac{(PT)(RS)}{(PS)(RT)} \right)
\]

\[
\ln \left[ \frac{(QS)(PT)}{(PS)(QT)} \right] \left( \frac{(QT)(RS)}{(QS)(RT)} \right) = \ln \left( \frac{(PT)(RS)}{(PS)(RT)} \right)
\]

It can be shown, using the above definition of length of a segment, that the hyperbolic length of a line is infinite. (See Appendix B for the problem.)

An interesting experiment showed that hyperbolic geometry can be used to describe space for people of normal vision. Subjects were asked to place light sources \( a, b, c \) at what they perceived to be the midpoints of the sides of a triangle formed by three fixed light sources \( A, B, C \) in a horizontal plane at eye level. The results (Figure 32) generally indicated triangles with sides that curved inward. This might provide some evidence of hyperbolic geometry describing binocular visual space.

![Fig. 32](attachment://image.png)

Projective Geometry

That the renowned artist Leonardo da Vinci inspired a form of
geometry may seem incongruous to a high school student; until he recognizes that to the realistic artist of the sixteenth century the concept of projection is paramount. Using the eye as the center of perspectivity the image on canvas can be thought of as a projection of the original onto a plane. Thus was born the idea of projective geometry which is a study of those properties which are invariant under a projection.

Lengths of segments, sizes of angles and shape of figures change under projection within the plane and between two non-parallel planes. There are, however, many features such as the type of figure, i.e., triangle, quadrilateral, closed curve, etc., and the incidence of a point and a line which remains unchanged under projection.

Several prominent mathematicians in widely separated eras contributed to projective geometry. Until the early nineteenth century it was simply a part of Euclidean geometry without its own undefined notions, definitions, postulates, and theorems. Three of the most important theorems were developed and proven independently of any other projective properties. The first two theorems of projective geometry were stated and proved by Pappus (4th century A.D.).
The first (Fig. 33) concerns the invariance of the cross ratio of four points in order, A, B, C, D. This is defined as a ratio of two ratios:

\[ x = \frac{CA}{CB} : \frac{DA}{DB} \]

**Cross Ratio Theorem:** If four concurrent rays (a pencil of lines) are cut by two transversals giving corresponding ranges A, B, C, D and A’, B’, C’, D’, then their cross ratios are equal:

\[ x = \frac{CA}{CB} : \frac{DA}{DB} = \frac{C'A'}{C'B'} : \frac{D'A'}{D'B'} \]

**Proof:** (Figure 33). Recall the area of a triangle is equal to one-half the product of any base and the altitude to that base. The area is also given by half the product of any two sides and the sine of the included angle:

- Area \( \triangle OCA = \frac{1}{2}h(CA) = \frac{1}{2}(OA)(OC) \sin \gamma_{COA} \)
- Area \( \triangle OCB = \frac{1}{2}h(CB) = \frac{1}{2}(OB)(OC) \sin \gamma_{COB} \)
- Area \( \triangle ODA = \frac{1}{2}h(DA) = \frac{1}{2}(OA)(OD) \sin \gamma_{DOA} \)
- Area \( \triangle ODB = \frac{1}{2}h(DB) = \frac{1}{2}(OB)(OD) \sin \gamma_{DOB} \)

It follows that

\[ \frac{CA}{CB} : \frac{DA}{DB} = \frac{CA}{CB} : \frac{DA}{DB} = \frac{(OA)(OC) \sin \gamma_{COA}}{(OB)(OC) \sin \gamma_{COB}} \cdot \frac{(OB)(OD) \sin \gamma_{DOB}}{(OA)(OD) \sin \gamma_{DOA}} \]

\[ = \frac{\sin \gamma_{COA}}{\sin \gamma_{COB}} \cdot \frac{\sin \gamma_{DOB}}{\sin \gamma_{DOA}} \]

Since the cross ratio of four points depends only upon the angles at \( O \) formed by the pencil of lines through the four points, it must be concluded that the cross ratio of any other set of four points formed by the intersection of this pencil with any other transversal must remain invariant under the projection through \( O \).
If the four points are projected under a parallel projection (Figure 34), the proof is part of elementary high-school geometry: "If three or more parallels are intersected by two transversals, the segments formed are in the same ratio."

![Diagram of points A, B, C, D and their projections A', B', C', D']

**Proof:**

(1) \[
\frac{a + b}{b} = \frac{a' + b'}{b'} \quad \text{and} \quad \frac{a + b + c}{b + c} = \frac{a' + b' + c'}{b' + c'}.
\]

Now

\[
\frac{CA}{DA} = \frac{a + b}{b} \quad \text{and} \quad \frac{DB}{DA} = \frac{a + b + c}{b + c}
\]

and

\[
\frac{C'A'}{D'A'} = \frac{b' + a'}{b'} \quad \frac{C'B'}{D'B'} = \frac{b' + a' + b' + c'}{b' + c'}.
\]

By (1) we have

\[
\frac{a + b}{b} \div \frac{a + b + c}{b + c} = \frac{a' + b'}{b'} \div \frac{a' + b' + c'}{b' + c'},
\]

which proves the invariance of cross ratio for the parallel projection.

The second theorem known as "Pappus' Theorem" states, "If A, B, C and A', B', C' are triples of distinct points on two distinct coplanar lines, the points of intersection of AB' and A'B, of AC' and A'C, and of BC' and B'C are collinear" (Fig. 35).
This may be proven quite easily using Figure 36 and an Euclidean method familiar to every high school geometry student.

**Proof:** Assume $P = B'C \cdot BC'$ (intersection of $B'C$ and $BC'$) and $Q = BA' \cdot B'A$ are ideal points or that $B'C$ parallel to $BC'$ and $B'A$ parallel to $BA'$. It must be shown that $R = A'C \cdot AC'$ is ideal and thus is collinear on the ideal line with $P$ and $Q$. This may be shown by proving $AC'$ parallel to $A'C$.

Since $A'B \parallel AB'$, we have

\[
\frac{b}{a + x} = \frac{b + y}{a + x + r},
\]

and similarly

\[
\frac{a}{b + y} = \frac{a + x}{b + y + c}.
\]
From these two equations we have

\[
\frac{a}{b+y} \cdot \frac{a+x}{b} = \frac{a+x}{b+y+s} \cdot \frac{a+x+r}{b+y}
\]

or

\[
\frac{a}{b} \cdot \frac{a+x}{b+y} = \frac{a+x+r}{b+y+s} \cdot \frac{a+x}{b+y}
\]

and

\[
\frac{a}{b} = \frac{a+x+r}{b+y+s}
\]

which implies that \( A'C \parallel AC' \).

The next event in the development of projective geometry was the publication in 1639 of a little noticed, but revolutionary new treatise on the conic sections written by Gerard Desargues. Yet, it did not create a stir in the mathematical world of the early seventeenth century. Rene Descartes, two years previously, had published his famous work on analytic geometry. He had invented a method to attack classical geometric problems by establishing a fundamental relationship between the real numbers (algebra) and the Euclidean plane (geometry). Thus, mathematicians were too busy with a new tool to be concerned with what seemed to be a rehash of Pappus. Rediscovered in the 19th century, Desargues' book inspired Gergonne, Poncelet, and von Staubt, who systematized and axiomitized projective geometry as a branch of geometry. It is in this little book that we find Desargues' fundamental "two triangle" theorem: "If two triangles, either coplanar or in distinct planes, are so situated that lines joining pairs of corresponding vertices are concurrent, then the points of intersection of pairs of corresponding sides are collinear, and conversely" (Fig. 37).

If we assume the triangles are in different planes, the proof is not difficult. Assume triangles ABC and A'B'C' are perspective from point O. Assume further that triangle ABC and triangle A'B'C'
lie respectively in distinct planes P and P'. Since lines AA' and BB' are coplanar, then AB and A'B' must likewise be coplanar and thus intersect in some point C" which must, by the second assumption above, lie in both P and P'. Similarly BB' and CC' are coplanar as are BC and B'C' which must intersect at some point A". A" must also lie in both P and P'. Likewise for AC and A'C' which intersect at some point B". Since all three must simultaneously lie on both P and P', they must lie in the line of intersection of the two planes. Hence the triangles are perspective from a line.56

One contemporary of Desargues was much impressed and influenced by his work. He was Blaise Pascal (1623-1662) a French mathematical genius. Pascal, unfortunately, was frail and sickly and often was tormented by a tendency to be neurotically religious. He died very prematurely at the age of 39. His brief life consisted of several stages during which he would devote himself zealously to mathematics. During one of these periods he cooperated with Fermat to lay the foundations of the mathematical theory of probability. Suddenly in 1654 he ceased all this activity when after being plucked from the jaws of death by an apparent miracle. He interpreted this as divine
objection to mathematics and vowed to devote himself to religious meditation. Later, when suffering from a toothache, his pain was relieved by contemplation of geometrical ideas. He then assiduously reapplied himself to mathematics.

Pascal's contributions to our examination of projective geometry is a generalization of Pappus' theorem in which the six vertices of the Pappus configuration are inscribed in a conic. If connections of the vertices are in a prescribed order, then the intersections of extensions of opposite sides are collinear. In Figure 38 below, two possible connection orders are illustrated. We can connect the numbered vertices in numerical order and extend opposite sides, e.g., (1,2) and (4,5) to intersect in X; (2,3) and (5,6) to intersect in Y; and (3,4) and (6,1) to intersect in Z. This is illustrated in black. If we connect the vertices in 1, 3, 5, 2, 6, 4 order (shown in red), we find opposite sides are (1,3) • (2,6) which = X'; (3,5) • (6,4) = Y'; and (5,2) • (4,1) = Z'. A good exercise at this point might
be to let the class discover the rule for determining opposite sides. This figure is what Pascal referred to as "mystic hexagram"; he proved no less than 400 theorems on conic sections as corollaries of this great theorem. His methods were entirely synthetic in the manner of Desargues.

The projectivity of Pascal's theorem can be further illustrated in nature by considering a cone of light emanating from a point. We pass a piece of clear glass on which is drawn a mystic hexagram through the cone. If another sheet of glass is passed through the cone so that the shadow of the hexagram falls upon it, the shadow will be another mystic hexagram with its three points of intersection of opposite pairs of sides lying in a straight line. This line will be the shadow (projection) of the 3-point line in the original hexagram. This means that Pascal's Theorem is invariant under conical projection.58

Projective geometry lay dormant until 1813 when Jean Victor Poncelet, an officer in Napoleon's Russian campaign, commenced his extensive work while a prisoner in Russia. Until this time the theorems of Pappus, Pascal, and Desargues were erratic blocks in Euclidean geometry.

Following a suggestion made 200 years earlier, by Johann Kepler that a new point be added to every Euclidean line called a "point at infinity" or an ideal point where all parallel lines will intersect, Poncelet took a significant step forward. Now the notion of parallelism fades and we are left with only incidence relations. We also extend the Euclidean plane by adding an ideal line consisting of all ideal points.

Poncelet studied those relationships which remain invariant
under projection, such as the harmonic ratio. His attempt to relate imaginary points and real points essentially failed, but he was able to establish ground work for showing how any conic can be projected into a circle, or two conics with a common ideal chord can be simultaneously projected into two circles. He also showed how the concept of a conic is invariant under projection, but that individual conics such as a circle can be projected into a parabola or another conic.\(^6\)

Poncelet was swiftly followed by a number of capable geometers including Jakob Steiner who introduced the powerful principle of duality by which the common notions and relations of projective geometry are paired. For example, a point is the dual of a line, collinearity is the dual of coincidence, and drawing a line through a point is the dual of locating a point on a line. Once the duals of all axioms have been proven, we obtain the theorem and its dual by proving the theorem. We can illustrate this by stating the dual of Pappus' Theorem: If two triples of lines are concurrent with two distinct points, then the lines joining opposite vertices are concurrent. P and Q are two distinct points concurrent with lines a, b, c and a', b', c' respectively. Connect opposite vertices (i.e., 1,4; 2,5; 3,6) and extend the three lines to be concurrent in point O (Figure 39).\(^6\)

![Fig.39](image-url)
Analytic Geometry

The next main area of geometry which warrants our investigation is Analytic Geometry. The foundations of analytic geometry were laid centuries before Descartes. The Babylonians used paired points and lines in surveying. Apollonius used the geometrical equivalents of cartesian equations to derive the bulk of his geometry of the conic sections. Nicole Oresme represented certain laws by graphing a dependent variable (latitude) against the independent one (longitude) which was permitted to augment by small increments. The analytic geometry we know today so highly dependent upon algebraic symbolism, originated with Rene Descartes and Pierre de Fermat, French mathematicians of the seventeenth century. 62

Descartes was born in 1596 near Tours, France. He attended the Jesuit school at La Flèche until he was 16 at which time he studied mathematics for a time, soldiered for several years, then commenced wandering about Western Europe. He finally settled in Holland and there he spent twenty of the most productive years of his life studying philosophy, mathematics, and science. It was during this time, between 1629 and 1649, that he published a lengthy book entitled A Discourse on the Method of Rightly Conducting the Reason and Seeking Truth in the Sciences. This book had three appendices, the last of which was called La Geometrie. In these 100 pages his contributions to analytic geometry appear. Basically the differences between Descartes' system and Greek geometry was that to the Greeks a variable, \( y \), represented a length of a segment; a square, \( x^2 \), the area of a square; the product of 2 variables, the area of a rectangle; and the product of 3 variables, the volume of a rectangular solid. Beyond this the Greeks could not venture. To Descartes, however, the
variable, $x$, denoted the fourth term in the proportion $\frac{1}{x} = \frac{x}{x^2}$. As
such, $x^2$ can be represented as a length when $x$ is known. Using a
unit length one can represent any power of a variable, for example
$x^3$ would be $\frac{1}{x} = \frac{x^2}{x^3}$. Descartes thus marks off $x$ on a given line and
then a length $y$ at a fixed angle to this and constructs a point
whose $x$'s and $y$'s satisfy a given relation. For example, if we have
the relations $y = x^2$, then for each value of $x$ we are able to con­
struct the corresponding $y$ as the fourth term of the above proportion.

The second part of the appendix deals with Descartes' laborious
method of constructing tangents to curves. With the invention of
differential calculus it became obsolete.\[63\]

While Descartes' appendix laid the foundation for a revolu­tion­
ary method, it remained for Leibnitz and others to add the terminol­
ogy such as coordinate, ordinate, linear, and quadratic. They smooth
out its rough spots and frost it with the simplicity and usefulness­
with which we can interpret complex geometric problems.

Descartes was not the only person at this time interested in
this idea. About the same time Pierre de Fermat began to investigate
the loci resulting from various types of equations which were ex­
pressed in the syncopated algebraic style of Viete. (By contrast,
Descartes examined a locus and then traced its equation analytically,
but in a much more modern symbolism.)

The idea of polar coordinates was introduced by Jakob Bernoulli
in 1691.\[61\] This is a system in which a point is located by its
angle with respect to a ray called a polar and its distance from the
origin of the ray. Polar coordinates are much more convenient than
Cartesian coordinates for certain types of curves, such as spirals
and other curves based on trigonometric functions.
Analytic geometry has led to a number of further developments. Among these has been the concept of n-dimensionality in which a point in n-space is represented as an n-tuple of coordinates: \( P = (x_1, x_2, \ldots, x_n) \) where in the case of real space, \( x_1, x_2, \ldots, x_n \) represent real numbers.

In 1829, Julius Plucker noted that the fundamental element in geometry need not be a point, but can be any geometric entity such as a straight line or a circle, etc. He designated a line by naming its intercepts with respect to a Cartesian frame of reference, provided the line does not pass through the origin. This is the analytic geometry of "line coordinates". He made the double interpretation of a pair of coordinates as either point coordinates or line coordinates and of a linear equation as either the equation of a line or the equation of a point. If instead of points of straight lines, we chose circles as our basic element we would need an ordered triple of numbers to determine one element completely. For example: an ordered pair to determine the center and a third to designate the radius.

**Topology**

"Where is in?", "Where is out?", "How many sides has a certain shape?", "What is a knot?"—those are fundamental questions to be answered by a study of the newest branch of mathematics called topology. It studies the common shapes of ordinary geometry such as a triangle, rectangle, polygon, circle, all of which are topologically equivalent. They separate the plane into two definite regions. When two configurations are topologically equivalent, they are said to be homeomorphic.

One of the fundamental theorems of topology is the Jordan Curve Theorem: "A simple closed curve divides the plane into exactly two
regions." This is a deceptively simple theorem whose statement anyone can readily understand. Most high school students would say, "well so what, everybody knows that." It was extremely difficult to prove with the rigorous thoroughness which has come to characterize mathematics in the last hundred years. Camille Jordan (1838-1922) was a first rate mathematician and an influential teacher. His original "proof" had a number of logical gaps in it, and he himself realized it did not meet his own high standards. It took the labors of other mathematicians to make it completely acceptable from a mathematical point of view.

Let us try to explain why it should be so difficult to prove, although we shall not here give its proof. Because there is a simple rule for determining whether a given point is inside or outside of a simple closed curve, the theorem is obvious. The rule is, start at P and travel along a given ray; if the ray intersects the curve an even number of times the point is outside, if the number of intersections are odd, the point is inside (see Fig.40).

For some labyrinthian curves whose boundaries cross and recross each other an infinite number of times, it would be impossible to
count the number of intersections a ray makes with a boundary. Herein lies the difficulty. This theorem has now been proven rigorously for the plane.

Let us consider the topological properties of some familiar objects. For example, picture a hollow ball; clearly it divides space. Now puncture a hole in its skin; it no longer separates space in two disjoint subsets. Next imagine a flat disk; it does not divide space but it has two sides and one edge. Third, consider a rectangular strip; it still has two sides but now it has two edges instead of one. The strip and the disk are both homeomorphic to a sphere with one hole. If both sides were painted a different color and both ends of the strip were joined to form a band so that the colors match, we would then have the topological equivalent of the sphere with two holes. Let us go one step further and take our unpainted strip, give it a half twist and glue the ends together. If we were to paint now we could use only one color since the band now has only one side and one edge. This is what is known as a Möbius strip, named after A.F. Möbius (1790-1868). What if we were to cut all three: the strip, the band, and the Möbius strip down their lengths? The strip and the band would fall apart into two separate pieces. The Möbius strip remains connected and no amount of stretching, pulling, etc., short of cutting will change it. When two configurations can be changed from one into the other and maintain their homeomorphic quality, we say they have been "deformed".

What does a topologist mean by a knot? He certainly does not mean something a Boy Scout or a sailor learns to tie from a long piece of line; for any knot which is tied can be untied and thus is homeomorphic to the line from which it was tied. A topologist's knot, on the other hand, is a knot which cannot be untied. The most
famous of these is the trefoil, two of which are shown below.

Fig.41

Nothing short of cutting can change trefoil #1 into trefoil #2. This brings us to the toughest problem facing topologists today—classifying different kinds of knots by their invariants. One method which works for many knots is associating each one with a certain surface, the edges of which can be arranged so that they trace out that particular knot. For example, a Möbius strip with three half twists would be associated with the trefoil. A general method has not yet been found which works for all topological surfaces and figures.

A number of important terms have come to be associated with topology: manifold, homeomorphic, deformable, and homotopic. A surface is said to be a manifold when the "neighborhood" of a point (roughly speaking, the area adjacent to a point) on the surface can be mapped one-to-one onto the interior of a circle. For example, a sphere is a manifold because the neighborhood of any point can be deformed by flattening it out. A mathematical cone of two nappes (an hourglass) is not a manifold because the neighborhood surrounding the vertex (with the exception of the vertex) cannot be flattened out to form a circle and its interior, since all non-vertex points are mapped two into one on the circle. The two terms, homeomorphic and deformable need close attention. Strictly speaking the two are
not exactly the same. An example will illustrate. We said that the circle and the polygon could be deformed (stretched or shaped) to form another. We also said that the trefoil (Figure 41) and the circle were not topologically equivalent since no amount of stretching or shaping could transform one into the other. However, both pairs of figures are homeomorphic since by homeomorphic we mean that a one-to-one correspondence can be made so that all points close together on one figure will be mapped close together on the second. We see that the trefoil and the circle are homeomorphic. To make the point perfectly clear let us say that two figures which can be deformed into one another are topologically equivalent (homeomorphic) but that the converse is false, i.e., that two topologically equivalent figures are not necessarily deformable.

This leads to another important concept in topology, that of homotopic. "A curve which (remaining within the surface) can be deformed continuously into another curve is said to be 'homotopic' to the latter." Thus the trefoil and the circle are not homotopic in the plane, but are in 3-dimensional space.67

The general classification of figures is the big question in topology today. We can classify certain types of figures as a sphere, as a torus (doughnut) with one hole, as a torus with two holes, as a torus with m-holes, etc. We must make use of topological invariants to assist us in classifying. We are looking for a number which corresponds to each of the classifications enumerated above. We want one number for the sphere, the polyhedron, and an analogous number for the torus and other figures which are topologically equivalent to it, etc.

We mentioned above when we were discussing the types of polyhedra
that a simple formula holds true for all different types of polyhedra. For instance, \( v - e + f = 2 \) (where \( v \) is the number of vertices, \( e \) is the number of edges, and \( f \) is the number of faces in a polyhedron) is true for all polyhedra. Imagine, if you will, taking a polyhedron and inflating it so that it forms a sphere with its vertices, faces, and edges still marked on the surface. The above formula still holds true. The number 2 has come to be called the Euler characteristic. In Fig. 42, notice that \( 16 - 32 - 16 = 0 \), or 0 is its Euler characteristic. If the polyhedron illustrated were inflated

![Fig. 42](image)

it would form a torus. Doing likewise with a nonsimple polyhedron with 2 such holes, we would find its characteristic would be \( 2\gamma - \gamma + 18 = -2 \). Continuing on we would find the set, 2, 0, -2, -4, ..., of Euler characteristics. In topology it is proved that if the number of holes in a torus is \( p \), its Euler characteristic is \( 2 - 2p \). The number \( p \) is called the genus of a surface. We can classify these surfaces by two numbers: by their genus or the set 0, 1, 2, ...; and by the Euler characteristic, 2, 0, -2, -4, ...

Topology originated in an attempt to solve a perplexing question. In the seventeenth century German town of Königsberg, seven bridges crossing the Pregel River connected the four parts of the town (Fig. 43a). The problem was, could one traverse all bridges without
crossing any one bridge more than once. Euler said that it could not be done. He transformed the problem into one involving vertices and edges (Fig.43b). Notice in Figure 43a that there are four areas

![Fig.43(a)](image)

![Fig.43(b)](image)

where a person might stand before crossing a bridge; we call these vertices. There are seven bridges to be crossed; we call these edges. Figure 43b shows four vertices and seven edges or lines joining them. A vertex is odd if there are an odd number of edges joining it. Euler then proved the following four important theorems, in answering the question:

1. In any graph the number of odd vertices is even.
2. A graph that has no odd vertices can be described unicursally (by a circuit or re-entrant route that starts at any vertex and returns to that vertex).
3. A graph which has exactly two odd vertices can be described unicursally by a route starting at one of the odd vertices and ending at the other. (Unicursally means traveling each edge once and only once.)
4. A graph which has more than two odd vertices cannot be described unicursally.

The Königsberg Bridge problem exhibits four odd vertices; thus it
cannot be described unicursally. Several excellent applications of these theorems are found in an area of topology called graph theory.

As we can see, topology, or rubber sheet geometry as it is often called, encompasses a wide variety of questions. A few of the areas of applied mathematics which are in the field of topology are electrical network theory, linear programming, and game theory.\textsuperscript{69}

Klein's Erlangen Programme

In concluding our study of the development of geometry, we examine a system by which various branches of geometry can be codified or classified. This system has been given the name "Erlangen Programme" after the location of a lecture given in 1872 by Felix Klein at the University of Erlangen, Germany. In this lecture Klein outlined a system under which we could define geometry: "A geometry is the study of those properties of figures which remain invariant under a given group of transformations."\textsuperscript{70} Let us explain what is meant by the three terms invariant, transformation, and group, for these are essential to an understanding of the definition.

Invariant means remaining unchanged. Transformation means subjecting a figure to a particular change: a rotation or translation in the plane, a proportional shrinking, a projection from one plane to another, or a stretching or a bending in a haphazard manner. For example, in Euclidean metric geometry, the concepts of length, area, and midpoint of a segment remain invariant when a figure such as a triangle is rotated or moved about the plane. In projective geometry the size and shape for figures will be changed, but a triangle will remain a triangle, collinear points will remain collinear points. The cross ratio of points examined above is an aspect of projective geometry since it remains unchanged under projection. In topology we examined properties which remained unchanged when a figure
underwent a stretching or an inflation and found that the location of a point inside or outside of a figure remained invariant under a topological transformation or homeomorphism. Another invariant of topology is the Euler characteristic of a particular type of surface.

The "Klein transformation" is defined in terms of a one-to-one mapping of a set S onto itself. A mapping is simply a correspondence which associates the elements of two sets. A one-to-one mapping of two sets is a correspondence in which each element of either set is associated with a unique element in the other set. It can easily be seen that a one-to-one mapping between the elements of finite sets (a permutation) is necessarily "onto" since each element in the image set must have a pre-image.

For infinite sets this is not necessarily true. Consider the following examples: (1) where each even whole number is mapped onto its half and each odd whole number is mapped onto zero.

\[
\begin{array}{cccccccc}
0 & 1 & 2 & 3 & 4 & 5 & 6 & \ldots \\
0 & 0 & 1 & 0 & 2 & 0 & 3 & \ldots \\
\end{array}
\]

Here every whole number will appear in the lower row so that the mapping is actually onto the set of whole numbers. But since many elements are mapped onto one element, namely zero, it is certainly not one-to-one. On the other hand, consider the following one-to-one mapping where each whole number is associated with its double.

\[
\begin{array}{cccccccc}
0 & 1 & 2 & 3 & 4 & 5 & 6 & \ldots \\
0 & 2 & 4 & 6 & 8 & 10 & 12 & \ldots \\
\end{array}
\]

The lower row contains only the positive even integers and is thus one-to-one into, but not onto, itself. If a Klein transformation is a permutation of an infinite set, it must be defined as a mapping that is both one-to-one and onto itself.\textsuperscript{71}
The third of the terms used in the definition is group. A group is a set of elements (Klein transformations) which satisfy four specific properties. Consider as an example the set of translations of points of the plane. A translation is a one-to-one mapping of the set of points of the plane onto itself. The properties are:

1. **Closure**: A sum of two translations is obviously equivalent to the effect of a single translation. "Sum" (+) is thus a binary operation on the set of translations (Fig. 44).

2. **Associativity**: Sum (+) is an associative operation. For example, Figure 45 indicates that the sum of translation #1 and #2 followed by #3 has the same effect (arrives at the same point) as translation #1 followed by the sum of translations #2 and #3.

3. **Inverse**: Each translation has an inverse in the form of an opposite translation whose sum is the identity translation.

4. **The Identity** translation of each point is the mapping which maps each point of the plane into itself.

It can be shown that the set of rigid motions (transformations)
in the plane, namely rotations and translations (sliding in a given direction) form a group. Thus plane Euclidean geometry under Klein's system is a study of those properties which remain invariant under rigid motion transformations. Projective geometry is a study of those properties which remain invariant under a projection or a series of projections. 72

This famous systemization of geometries was an attempt to categorize the varieties of geometry known in 1872. It has since revolutionized the treatment of geometry. In certain details, namely the group of transformations aspect, it is not applicable in general to certain quite abstract modern geometries. The idea of invariance and transformation are still important, however, in all these newer geometries.

One of these was inaugurated in 1906 by the French mathematician Maurice Fréchet who sought to establish a "general analysis, a calculus of relations between two sets whose elements are completely abstract in nature." 73 Fréchet's starting point was Hilbert space, which is defined as a countably infinite (Chapter III) analogue of a Euclidean space of 1, 2, 3, ..., n dimensions.

Consider an elementary example: In Euclidean space of two dimensions, the function \( y = x^2 \) has as its domain the real number continuum of the x-axis, a Euclidean one-dimensional space. A function of two variables, \( z = y - x^2 \) has as its domain the Euclidean plane of two dimensions. Theoretically we can extend this concept to n-dimensional space, where

\[
 y = g(x_1, x_2, x_3, \ldots, x_n)
\]

represents a function g whose domain is Euclidean space of n-dimensions, and n is a finite positive integer.
Hilbert generalized this still further by considering
\[ y = g(x_1, x_2, x_3, \ldots, x_n, \ldots) \]
where the domain is a space with a countable infinity of dimensions. A "point" of Hilbert space is an infinite sequence symbolized by
\[ (x_1, x_2, x_3, \ldots, x_n, \ldots) \]
Such a sequence can be thought of as representing a function, where the \( x_i \) are the Fourier coefficients of the function. Thus a Hilbert space can be considered a function space, a space whose points are in a one-to-one correspondence with the set of functions of a certain type.

To return to the geometry of Frechét: his general analysis is a study of functions in which neither the independent or dependent variable must be a number. In fact, the independent variable does not even have to be a function. Thus one is no longer dealing with a function of a curve, a surface, or even with a function in the Hilbert sense, but merely a function of a completely abstract entity. This is one of the reasons that the group aspect of Klein's Programme is not, in general, applicable.\(^7\)

Another of the modern geometries whose set of transformations is not necessarily a group is associated with Einstein's General Theory of Relativity (1916). The General Theory is based on Riemannian differential geometry. An analysis of the details of this relativistic geometry are well beyond the scope of this investigation. Suffice it to say, in the words of J.H.C. Whitehead, the nephew of the noted mathematical philosopher Alfred North Whitehead, while discussing the transition from pre-relativity to post-relativity geometry:\(^8\)

"Let us briefly recall the state of geometry just before and just after the discovery, in 1916, of general
relativity. During the period between this time and 1870 ideas concerning the foundations of geometry had been dominated by Felix Klein's Erlanger Programme. Riemann's philosophical concept of geometry had been ignored by most geometers though the analytical side of his theory had been extensively developed, notably by Ricci, and Levi-Civita. After the discovery of general relativity, which was based on Riemannian Geometry, it was realized that the Erlanger Programme was no longer adequate as a general description of geometry. The first person to understand the mathematical implications of this was Weyl, when he introduced generalized affine, projective, and conformal geometries, whose relation to their classical counterparts was analogous to that of Riemannian to Euclidean geometry."

This chapter has taken us far afield. We have traced the significant aspects of the growth of geometry from its beginnings as man's attempt to measure his earth, through Euclid's postulational structure to aspects of Riemann's differential geometry as applicable to Einstein's General Theory of Relativity. We have seen it grow from the one to the many. From its Euclidean root, geometry has taken many very different forms.

In the next two chapters we will trace the development of algebra and the development of mathematical logic and set theory.
FOOTNOTES

2. Ibid., p. 168.
5. Ibid., problem study 2-3, p. 42.
6. Ibid., p. 31.
7. Ibid., p. 41.
10. Ibid., p. 50.
11. Ibid., p. 51.
12. Ibid., pp. 51-52.
13. Ibid., p. 57.
17. Statement from Ibid., p. 65. The solution is mine.
18. A suggested problem is to have a student look up the proof of this and give a report to the class.
19. A suggested problem is to have the members of the class divide into teams of two or three students and construct models of the polyhedra. A reference is: Hartley, Miles C., Patterns of Polyhedra, Revised Edition, Ann Arbor, Michigan, privately printed by Edwards Brothers, 1957.

24. Ibid., p. 132.

25. Ibid., pp. 127-137.

26. Class members will be asked to prove this identity.

27. Eves, op. cit., p. 96.

28. Ibid., pp. 88-89.


30. Ibid., pp. 77-78.

31. See Appendix B for Euclid's proof of this.


35. Tuller, op. cit., pp. 21-22.


37. Ibid., pp. 154-156.

38. Quotation from Ibid., p. 33.


41. Ibid., p. 158.


43. Ibid., pp. 17-18.

44. Eves, op. cit., p. 196.

45. The hypothesis need not be so strong, for Legendre proved that if there exists one such triangle, then the sum of the angles of every triangle is equal to two right angles.

This point will be discussed at length in Chapter III where the properties of axiom systems are considered.

The whole brief discussion of the axioms and properties of hyperbolic geometry is in: Tuller, *op. cit.*, pp. 94-12 in much more detail.

---


51. Ibid., pp. 15-16.


55. Ibid., p. 188.


62. Ibid., pp. 281-286.

63. Ib. Id., p. 286.

64. Ibid., pp. 286-288 has a good general discussion of developments in Analytic Geometry.


67. Ibid., pp. 607-609.

68. Ibid., p. 617.

69. Reid, *op. cit.*, p. 188.
72. Ibid., pp. 427-428.
73. Ibid., p. 572.
74. Ibid., pp. 572-576.
CHAPTER II: THE DEVELOPMENT OF ALGEBRA

We cannot investigate the origins of algebra unless we go back into the dim recesses of history and investigate the origins of numbers; for, basically there were two origins of mathematics—numbers and form. Number led eventually to the concept of algebra and form led to geometry. With the birth of Analytic Geometry in the 16th century the two main streams finally converged to form the gigantic river of modern mathematics.1

We have no idea when man actually began to count objects. We do know that ancient peoples used the parts of their body in an ordered sequence to keep track of their flocks. For example, the natives of Muralog Island in the Torres Strait between New Guinea and Australia counted up to five by raising one by one the fingers of the left hand starting with the left little finger and proceeding to the thumb. They then expressed the numbers six to ten by touching and naming in order the left wrist, elbow, shoulder, breast, and sternum (left rib cage). The numbers from eleven to nineteen were indicated in inverse order on the right side starting with the right rib cage and proceeding through the little finger on the right hand. The words used to name the numbers were the actual parts of the body touched.2

It certainly was a monumental advance when words were developed to answer the question, how many? We thus have the origin of the word "cardinal number". We have early documents from China, India, Mesopotamia, and Egypt to indicate that the concept of cardinality was already well developed by the time those documents were written.

The next stage of advancement was when man realized that a
system of repetition was necessary if he were to avoid the colossal job of counting all his counters. Petroglyphs in caves dating back to the Middle Stone Age show men with ten fingers extended lined up in a row. This was the birth of the base system of counting although there is no positional value assigned. An investigation of the counting systems of primitive peoples indicates a use of various number bases such as 2, 3, 4, 5, 10, 12, 20, and even 60. The base 10 system was chosen, not surprisingly, far more often than any other. Some natives of Africa count a, oa, aa, aa-aa, aa-aa-a, aa-aa-aa, which is of course a base two system. Some American Indians used a base twenty counting system as did the Mayans, who had a well developed calendar. The ancient Babylonians used a sexagesimal (base 60) system.

This brings us to a consideration of a "best or preferred" system to use as a number base. Pure mathematicians have advocated a prime number base because they thought it would lead to greater uniformity in arithmetic procedures. Others have argued for a base twelve or duodecimal system since fractions would be much simpler with so many different factors—2, 3, 4, and 6. A simple problem illustrates:

\[
\frac{1}{3} = \frac{3}{12} = .3_{\text{twelve}} \\
\frac{1}{4} = .1_{\text{twelve}} \\
\frac{1}{2} = .6_{\text{twelve}}, \text{etc.}
\]

A colonial American mathematician, Hugh Jones, advocated the use of an octal system since, he argued, arithmetic was too complicated for "women and youths and often troublesome to the best artists."
One can see that octal fractions one-quarter, one-eighth, etc., are merely powers of one-half. Further, eight is a perfect cube, as is half of eight a perfect square. A giant among mathematicians, Gottfried Wilhelm Leibnitz, even argued for a binary system. Its advantages seem to be two: first, only two symbols need to be learned, and second, the addition and multiplication tables would be much simpler for young children to learn. But consider the disadvantage of pure length; our number 1000 would be 1,111,101,000 in binary. We know of course that computers use a binary system since representation inside the machine is by switching circuits which are either energized or not energized. Further, punched cards easily represent binary numerals since a hole punched represents one and a hole left unpunched, zero. This facilitates input of information into the machine.

The next important development, and certainly a major step, was the invention of place value. This was one of the most important contributions of the Babylonians to arithmetic. Algebraically it would appear

\[ a_n b^n + a_{n-1} b^{n-1} + \ldots + a_1 b + a_0 b^0 + a_{-1} b^{-1} + \ldots + a_{-m} b^{-m} \]

where \( 0 \leq a_i \leq b-1 \) and where \( b \) is the base. We represent this positionally

\[ a_n a_{n-1} a_{n-2} \ldots a_1 a_0 a_{-1} a_{-2} \ldots a_{-m} \]

understanding that juxtaposition of the \( a_i \)'s does not mean multiplication. Unfortunately no zero was used, especially in early writings. This made reading of some of the cuneiform tablets difficult since the meaning had to be determined by context.

Our number system has its origin in the old Hindu Arabic system. Details of its origin are clouded in mythology. We must patch
together bits and pieces of information from numerous sources. The use of symbols which have at least some resemblance to those we know today were found on a stone obelisk in India about 250 B.C. No symbol for zero was used and positional notation was also not in evidence at this time. Different symbols for 10 and 100 were used. Later the mathematics of the Hindus was expressed verbally. For example, 327 would appear as "three hundreds two tens seven." Since no place holding capacity of zero was necessary in this style it was not developed. The influence of the Greeks, Babylonians, and Chinese is unclear except the Greeks adopted the Babylonian sexagesimal system to express their fractions. Thus we have our present division of a circle as was seen in Chapter I.

Final development of the decimal system probably took place between the fourth and seventh centuries A.D. The system was in existence in Bagdad about 800 A.D. and was adopted by the Arabs with full recognition of its Hindu origins. As the Arabs moved across Africa and into Spain they carried it with them. A twelfth century translation of a book written by al-Khowarizmi about 825, entitled Liber Algorismi, brought the system to the western world. (Interestingly, "algoritmi dixit", according to al-Khowarizmi, is the origin of the modern word "algorithm" for a computational process.) The Liber Abaci of Fibonacci (1202) introduced the new system to Europe. The forms of the digit symbols became standardized by the invention of printing in the fifteenth century.

It is interesting to note that the modern Arabic numerals are not the same as the Hindu-Arabic form we recognize today. Simon Stevin, a Dutch scientist, published in 1585, La Disma, a book explaining the rules for and use of decimal fractions. His
notation, however, differs from ours. He wrote 5912 for 5.912. John Napier suggested a comma or period for the decimal point. Even today there is no general agreement: 3.25 (American notation) appears as 3•25 in Britain and 3,25 in France and Germany.\footnote{7}

It is instructive at this point to examine how early Hindu computations would have been accomplished. We must remember that the writing space was small, probably less than a foot square, and the medium was a thin white paint or a fine red dust; thus erasures of digits were made as soon as their purpose was served. Consider as an example the addition of $3\frac{4}{5}$ and $4\frac{8}{8}$.

\[
\begin{array}{c}
8 \ 3 \\
7 \ 2 \ 3 \\
3 \ 4 \ 5 \\
4 \ 8 \ 8 \\
\end{array}
\]

We can see they added from left to right and where they would have erased we have slashed out. The answer appears at the top: 833.

There is a method of multiplication using a lattice setup which was used by the Arabs and probably borrowed from the Hindus. For example, $135 \times 12$:

\[
\begin{array}{cccc}
1 & 3 & 5 \\
\hline
1 & 3 & 5 & 1 \\
2 & 6 & 1 & 0 & 2 \\
\end{array}
\]

Note that additions were performed diagonally and no carrying is necessary in multiplication because of diagonal cell division.\footnote{8}

Division among the Hindus was not considered a difficult and tedious operation as it was among European scholars as late as the fifteenth and sixteenth centuries. The Hindu method can be illustrated by the following example:

Divide 1620 by 12:
(1) Place the divisor below the dividend. 

\[
\begin{array}{c|c}
\text{1620} & \text{1} \\
\hline
\text{12} & \\
\end{array}
\]

Beginning at the far left, 16 is divided by 12. The quotient, 1, is placed in a separate line of quotients. The 16 is scratched out and the remainder, 4, substituted in its place.

(2) Rewrite and shift the divisor one place to the right. Divide again and write the quotient, 13, in the line of quotients to the right of the previous quotient.

\[
\begin{array}{c|c}
\text{1620} & \text{13} \\
\hline
\text{12} & \\
\end{array}
\]

(3) Proceed in this manner, eliminating the 12 and replacing it by the remainder, 6. Division yields the last quotient, 135, placed in the line of quotients.

(4) The 60 is rubbed out leaving the quotient, 135, and no remainder.

If figures are not obliterated and successive steps written one below the other, this process resembles closely the current method of long division taught in grade schools. When the use of paper became common in western Europe, this procedure became the galley method where digits were crossed out and new ones written above:

\[
\begin{array}{c|c}
\text{1} & \\
\hline
\text{1620} & \\
\text{122} & \\
\text{1} & \\
\hline
\text{1620} & \\
\text{122} & \\
\text{1} & \\
\end{array}
\]
Babylonian and Egyptian Numeration Systems

The Babylonian sexagesimal numeration system was developed prior to 2000 B.C. The symbol  stood repetitiously for units up to 9 and  stood for 10. Combinations of both as needed were used for numbers 11 to 59, using the positional notation to denote multiples of powers of 60. Using this system we read the following number  as

\[2 \times (60)^3 + 35 \times (60)^2 + 12 \times (60) + 31\]

which equals 560,551.

Fractions were represented by \( \frac{30}{60} = 1;30 = 1\frac{1}{2} \).

The origins of the use of a symbol for zero is obscure. There seems to be evidence of Hindu or Babylonian origin for the use of a symbol to indicate missing powers of the base (ten or sixty). Early Babylonian tablets indicate zero by leaving a space between groups of symbols. In later tablets, those dated during the last three centuries B.C., the symbol \( \Rightarrow \) was used for zero, inside numerical grouping and not at the end. The earliest Hindu symbol for zero was a heavy dot. Its use dates back to the third or fourth century A.D.

Ptolemy's tables in the Almagest (150 A.D.) include the symbol \( \overline{0} \) to indicate "no parts" of a given unit. Later, approximately 500 A.D., Greek texts used the initial letter, \( \omicron \), in the Greek word "ouden" meaning nothing.\(^{10}\)

Ancient Greeks used particular symbols to indicate powers of ten: \( \triangle = 10 \), \( \Pi = 100 \), \( \chi = 1000 \), \( \omicron = 10,000 \). 500 would have been represented \( \Xi \) where \( \Xi \equiv 5 \). Later the Ionic system, which came into use about 200 B.C., used twenty-seven letters of the Greek
alphabet in addition to three symbols from the old Phoenician alphabet. Multiples of 1,000 through 9,000 were represented by use of the first nine letters signifying the first nine numerals preceded by a small accent made to indicate thousands. For example: 1,000 = \( \alpha \), 2,000 = \( \beta \). Larger numerals were represented in a similar manner; 20,000 = \( \gamma \), or literally two ten-thousands.11

The use of a terminal zero as in 50 originated with the Hindus, shortly after 900 A.D.

The Egyptian numeration system dates from approximately 3400 B.C. in hieroglyphic form. It was a base ten system with no zero. The numerals one through nine were represented by vertical slashes; for example:

\[
\begin{align*}
1 & = | \\
2 & = || \\
10 & = \n \\
50 & = \n \n \n \\
100 & = \c \\
2000 & = \g \g
\end{align*}
\]

Later when the use of papyrus as a writing medium was developed the hieratic, or a more scriptive form, began to appear. Fractions, in ancient Egypt, were almost all unit fractions. In this system,

\[
\frac{\g}{\n n n} = \frac{1}{3}.
\]

Addition and subtraction are comparatively simple operations in any system, whereas multiplication and division tend to be more complicated. Multiplication was primarily a doubling process using the fact that any number can be expressed in binary form as sums of powers of two. A sample problem illustrates this: Find the product of 26 and 33.
Since $26 = 16 + 8 + 2$, double 33, result 66.

Then double 66, result 132; double 132, result

$8 \times 33 = 264$; double 264 which gives $16 \times 33 = 528$. Add only the starred multiples of 33. This gives the answer 858.

Division is similar. The divisor is successively doubled up to the point where the next doubling would exceed the dividend. For example, divide 753 by 26:

Since $753 = \frac{1}{16} \times 337$

\[
\begin{array}{cccc}
1 & 26 & = \frac{1}{16} & + 208 + 129 \\
2 & 52 & = \frac{1}{16} & + 208 + 104 + 25 \\
\times 4 & 104 & \text{x26} & \text{x26} \\
\times 8 & 208 & \text{x26} & \text{x26} \\
\times 16 & 416 & \text{x26} & \text{x26} \\
\end{array}
\]

The quotient is found by adding the underlined multiples of 26.\(^{12}\)

Ancient methods of calculation are instructive for several reasons. They are interesting in themselves, they serve to acquaint the student with the structure underlying our arithmetic operations, and they demonstrate the relationship between the binary system and the more familiar base ten system. With these goals in mind, several examples should be assigned to be carried out in the fashion of old.

Algebra in Babylonia and Egypt

As noted in Chapter I, our knowledge of Babylonian mathematics is gleaned painstakingly from translation of clay tablets. One appreciates just how painstaking this is, by considering two tablets in which the marks are hardly recognizable. (It would be useful at this time to pass out copies of Problems C and D, and work through with the class the solutions of these problems.) This example proves that these people had a specific working knowledge of the
Pythagorean relationship for right triangles. Problem D explains how the Babylonians would have solved a rectangle given the ratio between the length and the width and the measures of the diagonal. (See Appendix E for a copy of this problem.)

A remarkable tablet known as Plimpton 322 (ca 1900 B.C.) indicates to Babylonian scholars such as Otto Neugebauer, that these people were familiar with the parametric representation of Pythagorean triples, \((a, b, \text{ and } c)\): it gives:

\[
\begin{align*}
a &= 2uv \\
b &= u^2 - v^2 \\
c &= u^2 + v^2
\end{align*}
\]

where \(u\) and \(v\) are relatively prime and \(u > v\). These equations were used by later Arabic algebraists about 2000 years later to produce all primitive Pythagorean triples. (Primitive means that the triples are not integral multiples of the same number. For example \((3,4,5)\) is primitive, whereas \((6,8,10)\) is not.)

Accomplished algebraists that they were, the Babylonians had also developed in words the equivalent of our quadratic formula. Apparently they derived it by a geometric method of completing the square later used by the Greeks. Figure 1 shows how the quadratic \(x^2 + 2x = 4\) is represented geometrically. The square in the lower
left corner of the larger square has an area of $x^2$ while the unshaded rectangles must have areas each equal to $x$, total $2x$. The equation indicates that the unshaded area is four. We need an area of $1 \times 1 = 1$ to complete the larger square. If the shaded and unshaded areas are added together, the total area is five. This meant to the Greeks, and to the Babylonians before them, that the side of the larger square must have been $\sqrt{5}$ units, thus the side of the smaller square is $\sqrt{5} - 1$. This was the only solution which would have been recognized until negative numbers were understood and the solution of the problem was not dependent upon pictures.\textsuperscript{15}

The Babylonians were adept at solving certain special kinds of cubics. Before examining their methods, it must be pointed out that a great many of their tablets consisted of tables of all kinds and descriptions. Their calculations (especially multiplication and division) were carried out by use of these tables. Our elementary school children memorize the multiplication tables; Babylonian students simply consulted a tablet similar to the one recently unearthed and on display in the Brussels Museum collection. It gives products of 7, 10, 12, 16, 24 by 2, 3, ..., 9, 10, 20, ..., 50. Division was accomplished by use of extensive tables of inverses of numbers. Thus seven divided by two would be seven times one-half. The tables listed only regular reciprocals, ones whose denominators were factors of 60.\textsuperscript{16}

One can conjecture how they might have attempted to solve a cubic equation such as $x^3 + x^2 = 10$. Referring to tables of squares and cubes, they might have constructed a new table such as Fig. 2. If, for example, they wanted $x^3 + x^2 = 36$, they would just read the answer directly, three. If the sum they needed did not appear in
the table, they would use linear interpolation to obtain an approximation. If they wanted $x^3 + x^2 = 10$ as in the example above, they would approximate the answer to be $1;48$ or $1.8.17$.

One asks what sort of practical problems lead to a quadratic equation. According to the Dutch scholar, Professor Hans Freudenthal, Mesopotamian prices were not our unit prices at all, rather the inverse was used. Prices of barley might be quoted at five sacks per shekel. Consider a problem which might have arisen under this system. Suppose a merchant expected to sell 120 sacks of barley at market, for a total profit of ten shekels. What is the most he can pay the farmer if his selling price would be two sacks per shekel less than his cost? Using modern algebraic symbolism:

$$x = \text{the cost to merchant in sacks per shekel}$$
$$x - 2 = \text{the selling price to customer in sacks per shekel}$$
$$\frac{120}{x} = \text{cost in shekels}$$
$$\frac{120}{x - 2} = \text{selling price in shekels}$$
$$\frac{120}{x - 2} - \frac{120}{x} = 10.$$  

which it is easily seen leads to a quadratic in $x.18$

Our knowledge of Egyptian mathematics is based on two major
sources: the Moscow Papyrus and the Rhind or Ahmes Papyrus. Most of the problems in the papyri require little more calculation than is necessary for solving a linear equation. This was generally solved by what later became known as "false position". For example, to solve
\[ x + \frac{x}{7} = 2h \]
assume any convenient value for \( x \), say \( x = 7 \), then
\[ x + \frac{x}{7} = 8 \]
but since the answer must be twenty-four, our value of seven must be multiplied by three. The correct answer is twenty-one. False position is an application of the proportionality of sides of similar figures.

A strange problem (No. 79) is in the Rhind Papyrus. The following sets of data appeared:

<table>
<thead>
<tr>
<th>Houses</th>
<th>7</th>
<th>7^1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cats</td>
<td>39</td>
<td>7^2</td>
</tr>
<tr>
<td>Mice</td>
<td>343</td>
<td>7^3</td>
</tr>
<tr>
<td>Heads of Wheat</td>
<td>2401</td>
<td>7^4</td>
</tr>
<tr>
<td>Hekat Measures</td>
<td>16807</td>
<td>7^5</td>
</tr>
<tr>
<td>Total</td>
<td>19607</td>
<td></td>
</tr>
</tbody>
</table>

Its meaning puzzled anthropologists and archeologists until the historian Moritz Cantor in 1907, recognized it as the ancient forebear of the familiar old English children's rhyme:

"As I was going to St. Ives
I met a man with seven wives;
Every wife had seven sacks;
Every sack had seven cats;
Every cat had seven kits;
Kits, cats, sacks, wives,
How many were going to St. Ives?"

Greek Mathematics: 600 B.C. to 400 A.D.

The penchant for proof so typical of Greek mathematics probably has its origins in Greek philosophy. The early philosophers, Thales,
Anaximander, and Anaximenes, took great pains to establish their principles. There was conflict between their divergent viewpoints, and their followers had to resort to logic to refute one another. One may suppose that when Greek mathematicians were challenged, they resorted to logic to establish their positions.

We also know that these ancient inhabitants of the northern shores of the Mediterranean Sea had a penchant for traveling. The mathematics of Babylonia and Egypt was theirs for the taking. Along with astronomical tables and the algebraic and geometric skills from Mesopotamia, they came home with a fondness for mysticism and mythology. Indeed it was the Pythagoreans who defined the whole world in terms of its mathematical harmony, and were also influenced by mysticism. This fondness for the supernatural prompted the mathematical historian, Eric T. Bell, to make the comment:

"The sixth century before Christ was the time, and Greece the place, for human beings to reject once and for all the pernicious number mysticism of the East. Instead, Pythagoras and his followers eagerly accepted it all as the celestial revelation of higher mathematical harmony. Adding vast masses of sheer numerological nonsense of their own to an already enormous bulk, they transmitted this ancient superstition to the golden age of Greek thought, which passed it on in the first century A.D. ..."20

A discussion of Greek contributions to algebra must be preceded by explaining the Greek distinction between the terms "logistic" and "arithmetic". Logistic meant the science of computing or calculating. This is the meaning we give today to the term arithmetic. Arithmetic in the old sense meant the study of number theory. The study of arithmetic dominated the Pythagorean mind. The followers of Pythagoras disdained the practice of logistics. This may explain why the Greeks never developed an efficient numeration system. The explanation also exists that they recognized a base ten positional
number system would soon make the common people computationally competitive. This would democratize mathematics and thereby diminish the tight hold of the aristocratic members of the brotherhood on the economic lives of the people. It is also true that the precedent for this was set in Babylonia where only a small, exclusive group of logists handled computations and thus maintained influence and power. On the other hand, it must be remembered that the Pythagoreans were not logistically inclined. They cared little for economic and trade problems, preferring their "ivory tower" concentrations on arithmetic. Whatever the reason, they did launch the "golden age" of mathematics, and discovered many fascinating number properties.

One of these was the amicable or friendly numbers. Two numbers are said to be friendly if each is the sum of the proper divisors of the other. The smallest known pair are 220 and 284. Several prominent mathematicians spent some time searching for such pairs. Fermat, in 1636, announced the second pair to be 17,296 and 18,416. Euler launched a systematic search and, in 1747, published a list of more than 60 pairs.

Other mystical numbers are perfect, deficient, and abundant. These are classified by whether the sum of a number's proper divisors is equal to, less than, or greater than itself. $6 = 1 + 2 + 3$ is a perfect number as is $28 = 1 + 2 + 4 + 7 + 14$. Euclid proved (Book IX, Proposition 36) that if $2^n - 1$ is prime, then $2^{n-1}(2^n - 1)$ is perfect. Euler then proved that every even perfect number must be of the form $2^{n-1}(2^n - 1)$. One of the unsolved problems of number theory is the existence or non-existence of an odd perfect number. If such exists, it has no less than thirty-six digits.

The geometrical nature of Pythagorean mathematics would account
for the fact that figurate numbers would originate with these people. These numbers: triangular, square, pentagonal, etc., are numbers which can be represented with dots in the shape specified (Fig. 3).

\begin{figure}[h]
\centering
\begin{tabular}{cccc}
1 & 3 & 6 & 10 \\
4 & 9 & 16 \\
5 & 12 & 22 \\
\end{tabular}
\caption{Fig. 3}
\end{figure}

A number of theorems can be illustrated with figurate numbers. Three are shown:\cite{ref22}

(1) A square number can be represented as the sum of two successive triangular numbers (Fig. 4).

\begin{figure}[h]
\centering
\begin{tabular}{cccc}
\textbullet & \textbullet & \textbullet & \textbullet \\
\textbullet & \textbullet & \textbullet & \textbullet \\
\textbullet & \textbullet & \textbullet & \textbullet \\
\textbullet & \textbullet & \textbullet & \textbullet \\
\textbullet & \textbullet & \textbullet & \textbullet \\
\end{tabular}
\caption{Fig. 4}
\end{figure}
(2) The nth pentagonal number is equal to \( n \) plus three times the \((n-1)\)st triangular number (Fig. 5).

\[
\text{Fig. 5}
\]

(3) The sum of any number of consecutive odd integers, starting with 1 is a perfect square (Fig. 6).

\[
\text{Fig. 6}
\]

An algebraic proof of the irrationality of \( \sqrt{2} \), the stumbling block in Pythagorean geometry follows. Assume that \( \sqrt{2} = \frac{a}{b} \) where \( a, b \) are relatively prime integers, \( b \neq 0 \). Then we have

\[
a = b \sqrt{2}
\]

and

\[
a^2 = 2b^2
\]

hence \( a^2 \) must be an even number; let \( a = 2c \)

then

\[
4c^2 = 2b^2
\]

and

\[
2c^2 = b^2
\]

thus \( b^2 \) is also even. But this is impossible since \( a \) and \( b \) were taken to be relatively prime. The conclusion is that \( \sqrt{2} \neq \frac{a}{b} \).
Since it has been shown that non-rational numbers exist, let us examine a definition of a real (rational or non-rational) number such that all reals can be paired in a one-to-one correspondence with the set of points on the familiar number line. It is easily seen how such a correspondence can be set up with the integers by simply agreeing on a unit distance and a point to be paired with zero. Using the familiar Euclidean construction to divide a segment into any number of equal parts, a one-to-one correspondence can be set up with the rational numbers at least theoretically. For example: divide the segment, AB, into five equal parts (Figure 7).

Segment CB is one-fifth of AB since parallel lines divide transversals proportionally. Construct a length equal to the diagonal of a unit square, and thus set up a correspondence with $\pm m\sqrt{2}$ where $m$ is any natural number. Similarly, using the square root spiral (see Chapter I) a correspondence can be set up with all square root mag-
nitudes. But is this the end of the line? Where is \( \pi \) or a number such as \( \frac{5}{\sqrt{5}} + \frac{3}{\sqrt{2}} \)? The answer depends on how we define a real number. First it must be understood that all rational numbers can be expressed as a non-terminating, repeating decimal. By division any fraction can be changed to its decimal equivalent, and conversely we can change every repeating decimal to a rational expression. An illustration appears below.

It can also be shown that \( .3\overline{9} = 4.0 \), a terminating decimal:

<table>
<thead>
<tr>
<th>Repeating decimal to rational</th>
<th>( 0.3\overline{9} = 4.0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N = 0.3\overline{12} )</td>
<td>( 0.3\overline{9} = 0.3 + \overline{0.0\overline{9}} )</td>
</tr>
<tr>
<td>( 100N = 31.\overline{312} )</td>
<td>( = 0.3 + x )</td>
</tr>
<tr>
<td>( 99N = 30.\overline{9} )</td>
<td>( x = 0.\overline{09} )</td>
</tr>
<tr>
<td>( N = \frac{30.\overline{9}}{99} )</td>
<td>( 10x = 0.\overline{9} )</td>
</tr>
<tr>
<td>( = \frac{309}{990} )</td>
<td>( 100x = 9.\overline{9} )</td>
</tr>
<tr>
<td>( = \frac{103}{330} )</td>
<td>( = 9.0 + 0.\overline{9} )</td>
</tr>
<tr>
<td></td>
<td>( = 9.0 + 10x )</td>
</tr>
<tr>
<td></td>
<td>( 90x = 9 )</td>
</tr>
<tr>
<td></td>
<td>( x = .\overline{1} )</td>
</tr>
<tr>
<td></td>
<td>( 0.\overline{39} = 0.3 + x )</td>
</tr>
<tr>
<td></td>
<td>( = 0.3 + 0.\overline{1} )</td>
</tr>
<tr>
<td></td>
<td>( = 0.\overline{4} )</td>
</tr>
</tbody>
</table>

By the same token irrational numbers can be defined as non-repeating, non-terminating decimals. Therefore since the real numbers consist of the union of all rationals and irrationals, we may define a real number as the set of all non-terminating decimals.

We now prove the existence of a one-to-one correspondence between the real numbers and the points on the number line. Every real number can be represented by a non-terminating decimal, say for
It is readily seen that every real number is part of a nested interval; that is, it lies between two rationals. The interval \((3,4)\) contains \((3.6,3.7)\) which contains \((3.67,3.68)\) which contains \((3.678,3.679)\) which contains... . Each succeeding interval contains at least one point on the number line; then no matter how small our interval becomes, it always contains at least one point. Furthermore, that point is unique. Since there must be some finite distance between two distinct points, the interval can be made smaller, and ultimately, smaller than any distance between two points. Conversely, given any point \(P\) on the number line, if \(P\) corresponds to the end of an interval the proof is finished. If, however, \(P\) lies within some interval, the left endpoints of successive intervals give the corresponding decimal expansion. For example, suppose \(P\) were within \((1.36078,1.36079)\), then \(P\) corresponds to \(1.36078...\). Thus each point corresponds to a unique decimal expansion of a non-terminating real number.\(^{23}\) This formulation is due to Georg Cantor (1845-1918). Eudoxus (370 B.C.), Richard Dedekind and Georg Cantor (in 1872) established the continuum theory of numbers; see Chapter III.

This contrasts with the discretist's theory of a predecessor and successor to every number. The theory is attributed to the Pythagoreans. Some historians point to a continuing battle between discretists and continuum advocates. It seems reasonable to recognize that the two merely complement each other since both lead in different ways to the building of mathematical knowledge. The discretist generally is a number theorist, an algebraist, or a logician; whereas the latter are usually geometers, analysts, and mathematical physicists. All contribute their share!
As would be expected, we know very little of the early life of Eudoxus. It is told, however, how he argued against the rigid rules of the Platonists in regard to geometric constructions and advocated the use of curves other than the sanctified circle (compass) and straight line (straight edge). Plato prevailed, however. Being unpopular in Athens, Eudoxus returned to Asia Minor and opened his own academy in Cyzicus.

A comparison of Eudoxus' theory of proportionals with Dedekind's treatment of irrational numbers illustrates how mathematics is at once ever old, ever new. The theory of proportionals resolved the scandal and heresy created by the discovery of the irrationality of \( \sqrt{2} \). Eudoxus extended the Pythagorean theory of proportions to include both commensurable and incommensurable magnitudes.

According to the Pythagorean theory, if two segments are in the ratio of, for example, 9 to 5 or \( \frac{a}{b} = \frac{2}{5} \), then five segment "a's" would equal nine segment "b's" or \( 5a = 9b \). They would say that two magnitudes were "in the same ratio", i.e., \( \frac{a}{b} = \frac{c}{d} \), if for all integers \( m \) and \( n \), \( ma = nb \) implies \( mc = nd \). Using modern notation, this is simply the following:

\[
mc = nd \quad \text{and} \quad ma = nb
\]

implies

\[
\frac{c}{d} = \frac{n}{m} \quad \text{and} \quad \frac{a}{b} = \frac{n}{m}
\]

or

\[
\frac{c}{d} = \frac{n}{m} = \frac{a}{b}
\]

for some values of \( m \) and \( n \). This is what is meant by commensurable numbers.

The problem, however, is that a ratio such as \( \sqrt{2} : 1 \) is not expressible as \( m\sqrt{2} = n \). There was need for an explanation, which
Eudoxus furnished. He said

"Magnitudes are said to be in the same ratio, the first to the second and the third to the fourth, when, if any equimultiple whatever be taken of the first and third, and any equimultiple whatever of the second and fourth, the former equimultiples alike exceed, are alike equal to, or are alike less than the latter equimultiples taken in corresponding order."^{24}

What this means in modern notation is,

\[
\frac{a}{b} = \frac{c}{d}, \text{ if for all integers } m \text{ and } n, \text{ we have }
\]

\[
ma > nb \quad \text{and} \quad mc > nd
\]

or

\[
ma = nb \quad \text{and} \quad mc = nd
\]

or

\[
ma < nb \quad \text{and} \quad mc < nd.
\]

Another way of expressing this is: if for some choice of \(m\) and \(n\), one of the above three conditions is invalidated, then \(\frac{a}{b} \neq \frac{c}{d}.\)\(^{25}\)

The theory proposed by Dedekind, known as the "Dedekind Cut", is essentially the same as the Eudoxian Theory. Dedekind reasoned that if rational numbers were represented as points on the number line, then each rational number represents a "cut" in the line so that if a rational number, \(n\), is included in the right part of the line, say set \(B\), it would be the smallest element in set \(B\) and there would be no largest element in set \(A\). The reason for this is the density property of rational numbers: between any two rationals there exists another rational. For example, the rational 0.6 represents a cut such that if 0.6 is contained in set \(B\), then it would be the smallest element in \(B\), and \(A\) would have no largest element. At this point, Dedekind, desiring continuity on the number line, postulated the existence of a real number corresponding to every cut in the line. If for some cut \((A,B)\), there was no maximum value of \(A\) and no minimum value in \(B\), there would be a gap or hole in the line. He then defined such a gap or hole which he called a cut to be an
irrational number. 26

Greek Number Theory

Books VII, VIII, and IX of Euclid's Elements discuss elementary number theory or arithmetic. This concerns the properties of integers, called by Gauss "the Queen of mathematics". 27 One of the linkages between ancient and modern mathematics, it offers excellent illustrations of the concepts of abstract algebra.

This section deals with an examination and proof of several of the more important and well known theorems. The first is Euclid's proof of the infinitude of primes. This theorem has been praised for its elegance and its rigor. Assume, to the contrary, that there is a largest prime; call it P. Form the product of all primes up to and including P, add one to the result, and

\[ N = (p_1 \cdot p_2 \cdot p_3 \cdot \ldots P) + 1. \]

Now N is not divisible by any of the primes less than or equal to P, since such division would produce a remainder of one. Now N can be one of two types of numbers:

(1) it is prime, thus larger than P and we are finished.

(2) it is composite and divisible only by a prime larger than P. We are also finished.

Therefore there can be no largest prime.

The fundamental theorem of arithmetic states that "every composite number greater than one can be factored into a product of primes in a unique way, apart from the order of the factors". The proof given below is not the historic one found in the Elements, but it is a little shorter and easier to follow.

Assume that there exists a positive integer which can be factored into a product of primes in two different ways, and let m
the smallest such integer:

\[ m = p_1 p_2 \cdots p_r \quad \text{and} \quad m = q_1 q_2 \cdots q_s \]

where \( p_1 \) and \( q_1 \) are primes. Further, suppose the factors have been rearranged so that \( p_1 < p_2 < p_3 < \cdots < p_r \) and \( q_1 < q_2 < \cdots < q_s \).

Now \( p_1 \neq q_1 \), for if it were equal, we could divide the right member of (1) by \( q_1 = p_1 \) and have an integer smaller than \( m \). But this is impossible since we have taken \( m \) to be the smallest possible such integer. Then either \( p_1 < q_1 \) or \( p_1 > q_1 \). Suppose \( p_1 < q_1 \) (if \( p_1 > q_1 \), we interchange letters in the proof below) and form

\[ m' = m - (p_1 q_2 q_3 \cdots q_s) \]

Substituting for \( m \) the two expressions of equation (1) \( m' \) can be written in either of the forms

(3) \[ m' = (p_1 p_2 \cdots p_r) - (p_1 q_2 q_3 \cdots q_s) = p_1(p_2 p_3 \cdots p_r - q_2 q_3 \cdots q_s) \]

(h) \[ m' = (q_1 q_2 \cdots q_s) - (p_1 q_2 q_3 \cdots q_s) = (q_1 - p_1)(q_2 q_3 \cdots q_s) \]

Since \( p_1 < q_1 \), it follows that \( m' \) is positive. By (2) \( m' < m \) and by the hypothesis the prime factorization of \( m' \) is unique, aside from order of factors. Now by (3) \( p_1 \) is a factor of \( m' \), therefore from (h) it is a factor of either \( (q_1 - p_1) \) or \( (q_2 q_3 \cdots q_s) \). But since all \( q_s \) are primes distinct from \( p_1 \), the latter is impossible. Thus \( p_1 \) is a factor of \( (q_1 - p_1) \) and for some integer, \( h \),

\[ (q_1 - p_1) = p_1 h \quad \text{or} \quad q_1 = p_1 (h + 1) \]

But this implies that \( p_1 \) is a factor of \( q_1 \) which is contrary to the fact that \( q_1 \) is a prime. Thus our initial assumption that \( m \) is factorable in a non-unique way is false.28

The Fundamental Theorem of Arithmetic does not hold for all sets of numbers.29 Consider the set \( S \) of all even natural numbers under the operation of multiplication. It is easy to verify that this set is closed. Now consider the subset of \( S \) consisting of all numbers
not factorable into two or more members of the set, such as 2, 6, 10, 12, 18, ..., are called pseudoprimes. All other members of S can be factored such that all factors are also members of S. But this factorization is not unique, for example,

\[60 = 6 \cdot 10\]
\[= 2 \cdot 30\]

and

\[130 = 2 \cdot 66\]
\[= 6 \cdot 22.\]

An important corollary to this theorem states that "if a prime p were a factor of a product ab, then p must be a factor of either a or b."

The last of the famous theorems from Greek number theory is known as Euclid's Algorithm. It is a process for finding the greatest common divisor of two integers.

First establish the division algorithm which states that if a is any integer and b is any integer greater than zero, we can find an integer q such that

\[a = bq + r\]

where \(0 \leq r < b\). Now a is either a multiple of b implying that

\[\text{(1)} \quad a = bq\]

or a lies between two successive multiples of b,

\[\text{(2)} \quad bq < a < b(q + 1) = bq + b.\]

By the left inequality in (2)

\[a - bq = r > 0,\]

from the second,

\[a - bq = r \leq b\]

or

\[0 \leq r < b\]

as required.

This algorithm gives us readily the Euclidean Algorithm. It
must be shown that \( a = bq + r \), implies that every common divisor of 
a and \( b \) is also a common divisor of \( b \) and \( r \), and conversely, or 
symbolically
\[
(1) \quad (a, b) = (b, r).
\]
Let \( u \) be a common factor of \( a \) and \( b \), thus
\[
a = su \quad \text{and} \quad b = tu.
\]
Now if \( u \) divides \( a \) and \( b \) it must also divide \( r \), since
\[
r = a - bq = su - qt\ u = (s - tq)\ u.
\]
Also any number \( v \) which divides \( b \) and \( r \) can be written
\[
b = s'v \quad \text{and} \quad r = t'v.
\]
But \( v \) also divides \( a \), since
\[
a = bq + r = s'vq + t'\ v = (s'q + t')v.
\]
Therefore any common divisor of \( a \) and \( b \) is also a common divisor of 
\( b \) and \( r \), and conversely.30

An example: Find the greatest common divisor of 630 and 875.

We have 
\[
875 = 630 \cdot 1 + 245
\]
and using the property proven above, equation (1)
\[
(875, 630) = (630, 245).
\]
Now 
\[
630 = 245 \cdot 2 + 140
\]
and using equation (1) again
\[
(875, 630) = (630, 245) = (245, 140).
\]
Continuing,
\[
245 = 140 \cdot 1 + 105
\]
also
\[
140 = 105 \cdot 1 + 35
\]
and
\[
105 = 35 \cdot 3 + 0.
\]
So we have
(105, 35) = (35, 0) = 35.

One of the most intriguing questions of number theory down through the ages is that of developing a formula which will generate all primes. No such formula has been found. Several formulas have been proposed, such as \( n^2 - n + 41 \) and \( n^2 - 79n + 1601 \), but both yield composites eventually. The former gives primes for \( n = 1, 2, \ldots, 40 \), but a composite for \( n = 41 \). The latter is a good generator up to \( n = 79 \). Both these examples emphasize the critical necessity for rigorous proof of all conjectures.

Fermat, the great number theorist, conjectured that \( F_n = 2^{2^n} + 1 \) would produce only primes for \( n \), an integer. This supposition was shown erroneous in 1732, by Leonhard Euler, who proved that \( 2^{32} + 1 \) is composite with 641 as a divisor.

The next step in the long search for all possible primes was taken by Karl Gauss (1795). He considered the density of primes among the integers 1, 2, \ldots, \( n \), represented by the ratio \( A_n/n \), where \( A_n \) is the number of primes less than or equal to the natural number, \( n \). Observe the following table:

<table>
<thead>
<tr>
<th>( n )</th>
<th>( A_n/n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10^3</td>
<td>0.1680</td>
</tr>
<tr>
<td>10^6</td>
<td>0.0785</td>
</tr>
<tr>
<td>10^9</td>
<td>0.0508</td>
</tr>
</tbody>
</table>

Further, Gauss discovered that the values of \( A_n/n \) approached asymptotically the values of \( 1/\ln n \) where \( \ln n \) is the area under the hyperbola \( xy = 1 \) between \( x = 1 \) and \( x = n \). Consider the following table:

<table>
<thead>
<tr>
<th>( n )</th>
<th>( A_n/n )</th>
<th>( 1/\ln n )</th>
<th>( A_n/n )</th>
<th>( 1/\ln n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10^3</td>
<td>0.1680</td>
<td>0.1450</td>
<td>1.159</td>
<td>1.159</td>
</tr>
</tbody>
</table>
He conjectured on the basis of this data that the values of $\frac{A_n}{n}$ approaches one as $n$ becomes large without bound.\textsuperscript{32}$

The proof of this theorem was well beyond the realm of analysis at the time Gauss made his hypothesis. Nearly one hundred years passed before the first proof was given by Charles Hadamard and Charles J. Poussin. The Hadamard-Poussin proof was made possible when the Russian probability theorist Pafnuti Tchebycheff proved that the ratio lies between two positive constants $a$ and $A$. The demonstration was simplified several times until the most recent one in 1949 by Atle Selberg.\textsuperscript{33}

There are a number of fascinating conjectures in number theory which have never been proven. One of these is known as Fermat's Last Theorem. He had investigated the existences of solutions of $x^n + y^n = z^n$ for $n$, a natural number greater than two. In the margin of a book he was reading he made the famous remark:

"By contrast it is impossible to separate a cube into two cubes, a fourth power into two fourth powers or in general any power of the same degree. I have found a truly marvelous proof of this theorem but this margin is too narrow to contain it."\textsuperscript{34}

Mathematicians ever since have struggled with this famous theorem, but none has succeeded in a complete proof. A number of people have thought they found one, including Louis Cauchy, the well known French analyst of the early nineteenth century, only to discover they had assumed the Fundamental Theorem of Arithmetic holds for generalizations of the integers.

It can be shown that it is only necessary to establish it for primes greater than 2. To see why, consider the following: Fermat's
Theorem has no solutions for \( n \) equal to any higher power of two, for example \( n = 8 \), then we would have \((x^2)^4 + (y^2)^4 = (z^2)^4\) or \( a^4 + b^4 = c^4 \). But then the squares of \( x, y, \) and \( z \) would satisfy \( a^4 + b^4 = c^4 \) which is impossible because Fermat proved it for \( n = 4 \). For \( n > 4 \) and composite then we would have \( n \) expressible as a unique product of primes. Thus one is back to proving it for primes. Recent research has established the theorem for all primes less than \( 2^{3617} - 1 \). Some logicians have even advanced the theory that the conjecture is undecidable, that is, it can neither be proven true nor false.35

Modern methods of analysis, such as the study of the continuum, have been brought to bear on problems of integral number theory, the discrete. An excellent case in point is Gauss's Prime Number Theorem. This indicates firmly that the future of mathematics depends on both concepts proceeding hand in hand in a complementary fashion to open new vistas.

**Diophantine Equations**

An evolution of algebra next comes to the work of a giant among Greek algebraists. He did a great deal of work in number theory and algebra, but is probably best known for his syncopation of algebra. Prior to the time of Diophantus (ca 250 A.D.), all algebra was rhetorical; in other words, all arguments and reasoning was in words without using symbols of any kind. Diophantus originated the practice of using a shorthand in which the symbols are derived from the words which they represent. For example, his symbol for the unknown was, \( \delta \), a combination of the first two Greek letters \( \alpha \) and \( \rho \) of the Greek word arithmos. His symbolism for \( x^2 \) is \( \Delta \gamma \), the first two letters of the word dunamis meaning power. \( x^3 \) is \( \kappa \gamma \), or the first two letters of the Greek word kubos for cube. Numbers were simply
the Greek letters. Thus \( x^3 + 13x^2 + 5x \) would have appeared as
\[
\gamma^3 \alpha^2 \gamma \gamma \alpha
\]
or literally as "unknown cubed, unknown squared 13, unknown 5".\textsuperscript{36}

A Diophantine Equation is an algebraic equation in one or more unknowns with integer coefficients for which only integral or rational solutions are sought. Elementary algebra texts are replete with examples of these equations. Such equations may have no solutions, a finite number, or an infinite number. Using the Euclidean Algorithm a condition for integral solutions of the Diophantine Equation \( ax + by = c \) is that \( c \) be a multiple of \( a \) and \( b \), where \( a, b, c \) are integers. We must first show that if \( d = (a,b) \) (the greatest common divisor of \( a \) and \( b \)), then \( d = ka + \lambda b \) for a proper choice of integers \( k \) and \( \lambda \).

An example: if \( h = (12, 8) \) we have \( h = 12(1) + 8(-1) \). From the Euclidean algorithm we have:

\[
\begin{align*}
a &= bq_1 + r_1 \\
b &= r_1q_2 + r_2 \\
r_1 &= r_2q_3 + r_3 \\
r_2 &= r_3q_4 + r_4 \\
\end{align*}
\]

\[
\begin{align*}
r_{n-2} &= r_{n-1}q_n + r_n \\
r_{n-1}^n &= r_{n}q_{n+1} + 0
\end{align*}
\]
hence that \( r_n \) is the greatest common divisor of \( (a,b) \). From the first equation above we have

\[
r_1 = a - q_1b
\]
so that \( r_1 \) can be written in the form \( k_1a + \lambda_1b \) where in this case \( k_1 = 1 \) and \( \lambda_1 = -q_1 \). From the second we have

\[
r_2 = b - q_2r_1
\]
= b - q_2(k_1 + \ell_1 b) \\
= (-q_2k_1)a + (1 - q_2\ell_1)b \\
= k_2a + \ell_2b \\

repeating this process we have eventually \\
\[ r_n = ka + \ell b \]

for some \( k \) and \( \ell \).

Thus we know that if \( d = (a, b) \), then for some \( k, \ell \), \( d = ka + \ell b \).

Now \( c = ax + by \) will have integral solutions \( x = k \) and \( y = \ell \), if \( c = d \). We can also proceed from here to find all conditions under which \( ax + by = c \) will have integral solutions. If \( c \) is any multiple of \( d \), then \( c = dg \); thus we obtain from above \\
\[ a(kg) + b(\ell g) = dg = c \]

so that \( x = x^* = kg \) and \( y = y^* = \ell g \). We have shown that \( ax + by = c \) has integral solutions if and only if \( c \) is a multiple of \( d = (a, b) \).

An example: \( 3x + 6y = 22 \) has no integral solutions since \( 3 = (6, 3) \) does not divide 22. 37

As indicated in Chapter 1, the Greek mathematical contribution ended with the death of Hypatia, who was a student of, and much influenced by, the work of Diophantus. Under Roman rule, the arts and the intellectual climate in the Mediterranean generally languished. Then came the Arab conquest of North Africa in the late seventh century following the flight of Muhammad from Mecca to Medina. The star and crescent waved from India in the east to Spain in the west. The Arab caliphs encouraged learning, and Babylonian, Greek, and Hindu mathematics entered the Arab world.

Hindu algebraic contributions are from three scholars: Aryabhata, Bhaskara, and Brahmagupta. The Hindus operated freely with irrationals, and used syncopated algebra much like Diophantus. They admitted
negative solutions and often gave both formal roots to quadratics. Their method of solving a quadratic is similar to our modern method of completing the square. This problem is from Brahmagupta. Shown in three columns is a literal translation, modern notation, and a generalization.\footnote{38}

The equation is \( x^2 - 10x = -9 \) which was symbolized by Brahmagupta as

\[
yavlyayalo\]

\( ya \) is the unknown, \( v \) means squared, and the dot above a number indicates a negative number.

\[
x^2 - 10x = -9 \quad ax^2 + bx = c
\]

Here absolute number, \( 9 \) multiplied by \( 1 \) the \( \text{[coefficient of the]} \) square \( 9 \) made \( 16 \)

\[
-9 + \left[ \frac{-10}{2} \right]^2 = 16 \quad ca + \left[ \frac{b}{2} \right]^2
\]

of which the square root \( 4 \), less half the \( \text{[coefficient of the]} \) unknown \( 9 \);

\[
\sqrt{16} - \left[ \frac{-10}{2} \right] = 9 \quad \sqrt{ca + \left[ \frac{b}{2} \right]^2} - \frac{b}{2}
\]

and divided by the \( \text{[coefficient of the]} \) square \( 1 \) yields the value of the unknown \( 9 \).

\[
\frac{9}{1} = 9 \quad \sqrt{ca + \left[ \frac{b}{2} \right]^2} - \frac{b}{2} = x
\]

or

\[
x = -b + \sqrt{b^2 + 4ac} \quad 2a
\]

Note that the last equation on the right is equivalent to the modern quadratic formula, where \( c \) is the equivalent of our \( -c \).

(Replace the \( c \) inside the radical with \( -c \) and the familiar \( \sqrt{b^2 - 4ac} \) results.)
Bhaskara looked for the solution of $x^2 = -1$ in the domain of real numbers and labeled it impossible "as in the nature of things, a negative is not a square; it has therefore no square root" (Mahavira).\(^{39}\)

**Arabic Algebra**

The most important contribution of the Arab mathematicians is their translation and preservation of the Greek classics. To these they added much Hindu algebra including the solution of quadratics. A problem appearing in al-Khwarizmi's work "Al-jabr" (c825 A.D.) is virtually the same as the problem above from Brahmagupta. It is, however, interesting to note the origin of our word "algebra" from the title of al-Khwarizmi's book which means the "science of translation and cancellation." Omar Khayam (c1100 A.D.) is also remembered for his geometric solution of cubic equations by intersecting the circle and the rectangular hyperbola to find positive solutions.\(^{40}\)

**European Algebra - 500 to 1150 A.D.**

The Dark Ages settled upon Europe about 500 A.D. and with it began the disappearance of intellectual pursuits. The period is marked with much violence and intense religious fanaticism. Except for a few translations, mathematically the period is stagnant.

The next algebraist of note to appear was Leonardo of Pisa, better known as Fibonacci. In 1202 he published his famous *Liber Abaci*. The following problem appeared in that book: How many pairs of rabbits can be produced from a single pair in a year if every month each pair begets a new pair which from the second month on becomes productive. This problem gives rise to the well-known Fibonacci sequence: 1, 1, 2, 3, 5, 8, ..., $x$, $y$, $x+y$, ... Fibonacci also published in 1225 an outstanding work on indeterminate analysis,
entitled Liber Quadratorum. In searching for a solution to the cubic equation, \(x^3 + 2x^2 + 10x = 20\), Fibonacci investigated the quadratic surd \(\sqrt{a + \sqrt{b}}\). He came within a step of discovering a formula for the general cubic in terms of its coefficients when he examined the following surd:

\[ x = \sqrt{a + \sqrt{b}} + \sqrt{a - \sqrt{b}}. \]

Unfortunately he did not take the next logical step which was to find the sum of two cube roots; or

\[
(1) \quad x = 3\sqrt{a + \sqrt{b}} + 3\sqrt{a - \sqrt{b}}
\]

\[
(2) \quad x^3 = \left[(a + \sqrt{b})^{1/3} + (a - \sqrt{b})^{1/3}\right]^3
\]

\[
(3) \quad = a + \sqrt{b} + 3(a + \sqrt{b})^{2/3}(a - \sqrt{b})^{1/3} + 3(a + \sqrt{b})^{1/3} \cdot (a - \sqrt{b})^{2/3} + a - \sqrt{b}
\]

\[
(4) \quad = 2a + 3(a + \sqrt{b})^{1/3}(a - \sqrt{b})^{1/3}[a + \sqrt{b}]^{1/3} + (a - \sqrt{b})^{1/3}
\]

\[
(5) \quad = 2a + 3(a^2 - b)^{1/3} \cdot x
\]

\[
(6) \quad = p + qx,
\]

where \(p = 2a\) and \(q = 3(a^2 - b)^{1/3}\). But this implies, solving these simultaneous equations for \(a\) and \(b\):

\[
a = \frac{p}{2}
\]

and

\[
q = 3\left(\frac{p^2}{4} - b\right)^{1/3}
\]

\[
q^3 = 27\left(\frac{p^2}{4} - b\right)
\]

or

\[
b = \left(\frac{p}{2}\right)^2 - \left(\frac{q}{3}\right)^3.
\]

This is Cardano's solution to the incomplete cubic published in the Ars Magna (1545).

There is an interesting story regarding this solution. As was the custom, mathematical contests were held in which certain problems
were submitted for prize solution. In one such contest Antonio Fiore challenged Tartaglia, the stammerer, to a duel in solving cubic equations because he disbelieved Tartaglia's claim to have found a solution. Tartaglia, at the last minute, discovered a solution and won the contest. Girolamo Cardano, a friend of Tartaglia, learning of this, wrote and asked for an explanation of his method. After much discussion Tartaglia agreed to share his idea because he thought Cardano could help get him a job teaching mathematics. Tartaglia claims to have gotten a pledge of secrecy from Cardano, but this claim was disputed by a third party. At any rate, Cardano soon modified and extended these suggestions and published them in his Ars Magna, giving full credit to Tartaglia. Tartaglia flew into a rage at being tricked, but soon had to flee for his life.\footnote{122}

With a clever substitution, a complete cubic of the form
\[ax^3 + bx^2 + cx + d = 0,\]
where \(a, b, c, d\) are not zero, can be reduced to a cubic without quadratic term, where \(b = 0\). The method is shown below.

If in the equation
\[a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n = 0\]
x = \(z - \frac{a_1}{na_0}\) is substituted, the resulting equation in \(z\) will be without an \((n-1)\)st degree term.\footnote{123}

We have
\[
\begin{align*}
& a_0\left[z - \frac{a_1}{na_0}\right]^n + a_1\left[z - \frac{a_1}{na_0}\right]^{n-1} + \ldots + a_n = 0 \\
& a_0\left[z^n + \frac{n}{1}\left(-\frac{a_1}{na_0}\right)z^{n-1} + \ldots\right] + a_1\left[z^{n-1} + \frac{n}{1}\left(-\frac{a_1}{na_0}\right)z^{n-2} + \ldots\right] \\
& + a_n = 0
\end{align*}
\]
\[ a_0 z^n - a_1 z^{n-1} + \cdots + a_{n-2} z^2 - a_{n-1} z + a_n = 0 \]
\[ a_0 z^n + \left[ \frac{n-1}{2} \left( \frac{a_1^2}{a_0} \right) - \frac{a_1}{a_0} \right] z^{n-2} + \cdots + a_n = 0. \]

An example of this procedure is:
\[ y^3 - 3y^2 - 2y - 2 = 0 \]

Let \( y = x - \frac{3}{3(1)} \) or \( y = x + 1 \)

and this gives
\[ (x + 1)^3 - 3(x + 1)^2 - 2(x + 1) - 2 = 0 \]

which when simplified gives
\[ x^3 - 5x - 6 = 0. \]

Let us illustrate Cardano's procedure given above for solving cubic equations to find \( x \) in the equation \( x^3 - 5x - 6 = 0 \).

It becomes
\[ x^3 = 5x + 6 \]
where \( q = 5 \) and \( p = 6 \)

one root is \( x = \sqrt[3]{a + \sqrt{b}} + \sqrt[3]{a - \sqrt{b}} \)

where \( a = \frac{p}{2}; \ b = \left[ \frac{p}{2} \right]^2 - \left[ \frac{q}{3} \right]^3 \)

hence \( a = \frac{6}{2} = 3; \ b = 9 - \frac{125}{27} = \frac{243}{27} - \frac{125}{27} = \frac{118}{27} \)

\[ x = \sqrt[3]{3\sqrt[3]{118}} + \sqrt[3]{\frac{118}{27}}. \]

Referring to the original complete cubic above its roots will be found by adding one to those just found.

Cardano's pupil, Ferrari, soon followed with a method for solving a general quartic by reducing it to a product of a linear and a cubic factor. Both solutions are given in the Ars Magna.

It is instructive to mention at this point that the other roots are not given because negative and complex roots were unacceptable.
at this time.

There are better methods for solving general cubics today, namely those given by Newton and Horner. But Cardano's solution was given here for several reasons. First because of its historical importance as the first such method provided. Second, it (and the quartic solution which soon followed) marked the end of classical algebra. Cardano proved his method geometrically as the Greeks had done before him.

The solution of the quartic led mathematicians to search for 200 years for coefficient solutions of the quintic and higher degree equations. The search was ended by Nels Abel in 1826. He proved there exists no root of a fifth degree or higher equation, which is expressible in radicals in terms of its coefficients. In 1832, Galois showed that a polynomial equation is solvable if and only if its group is solvable. Before his very untimely death at age twenty, Galois launched modern abstract algebra by defining the notion of a group. We shall investigate later some of the advanced notions of abstract algebra.

Probability and Statistics

No treatment of the development of algebra would be complete without an investigation of the laws of chance, better known as probability theory. First of all, the subject has occupied the talents and energies of some of the most prominent mathematicians, men like Fermat, Pascal, DeMoivre, and many others. Secondly, its concepts are integral to the solutions of a number of complex problems. One such problem is the proof of the prime number theorem (p. 25). Finally its practical applications are limitless. The earliest known work on probability was Cardano's gambler's handbook
which was not published until almost 100 years after his death. In spite of this, most mathematical historians attribute the origins of probability theory to Pascal and Fermat who communicated often over the solutions to problems. One may conjecture as to why the laws of chance were not studied earlier. Probably it was due to the fact that gaming in any form was strongly discouraged by the church as being sinful. The church throughout history has been a dominant force in the lives of people. Possibly Cardano feared condemnation from this institution if he made known his investigation.

The origin of the science of probability is generally conceded to be the so-called problem of the points which concerned an interrupted game of chance between two players. The problem had been discussed by Pacioli in his *Suma* (1494). Cardano and Tartaglia also worked on it, but no real advances were made until Pascal became interested in 1654 and communicated his thoughts to Fermat. It is known that player A needs two more points and player B needs three more points to win. The question is how should the purse be divided at the time of interruption, assuming equally skilled players. Obviously a maximum of four more trials will decide the game one way or the other. Let "a" mean that A wins and "b" indicate that B wins. On the next trial and on each succeeding trial there are two possibilities, thus $2 \times 2 \times 2 \times 2 = 16$ total possibilities. The question is, in how many trial possibilities will an "a" appear at least twice? Turning the question around—on how many of the trials will an "a" appear only once or less? There are five such trials since "a" can appear once 4 different ways and not appear at all in one way. Thus there are 5 favorable for B and $16 - 5 = 11$ for A. Therefore the purse should be divided 11 to 5 for A.
Another problem was proposed to Pascal by Chevalier de Mère, a flamboyant French gambler but not a mathematician. He wanted to know the probability of rolling a double six on twenty-four rolls of two die. Pascal's and Fermat's solution to this was the following: there are thirty-six possible outcomes in the first roll, thirty-six in the second or $36^2$ in two rolls, thus there must be $(36)^{24}$ possible outcomes in twenty-four rolls. Now the probability of a double six means on each roll there are thirty-five unfavorable outcomes, thus $35^{24}$ unfavorable. The probability is the ratio of unfavorable outcomes to the total outcomes or $\frac{35^{24}}{36^{24}} = 0.51$. This means there is a probability of .49 that a double six will appear at least once.

In a book called The Doctrine of Chances (1738), Abraham DeMoivre gave a number of problems. The important rule of multiplication of successive probabilities is illustrated by one of these problems: (Problem 15, page 71) Any number of things $a$, $b$, $c$, $d$, $e$, $f$ being given, out of which three are taken as it happens; to find the probability that 'a' shall be the first taken, 'b' the second, and 'c' the third.

Solution: The probability of 'a' occurring in the first trial is $\frac{1}{6}$, assuming 'a' is taken then the probability is $\frac{1}{5}$ of taking 'b' on the second; the probability of 'c' is thus $\frac{1}{4}$ assuming 'a' and 'b' are taken. Thus the probability of $a$, $b$, and $c$ in that order is seen to be $\frac{1}{6} \times \frac{1}{5} \times \frac{1}{4} = \frac{1}{120}$.

The Growth of Modern Abstract Algebra

In this final section three topics will be investigated: (1) the invention and use of imaginary and complex numbers, (2) the beginnings of "new", as opposed to classical, algebra and the invention of
quaternions and finally, (3) the structure of abstract algebra.

Square roots of negative quantities had been occurring with increasing frequency during the development of algebra, but it was not until Rafael Bombelli (ca 1550) that this mysterious quantity was elevated to the status of "number". Since they occurred rarely in the works of Heron and Diophantus, negative square roots were generally disregarded. Cardano called them "fictitious". It was Bombelli who gave them the name "imaginary" numbers, a derogatory term on a par with irrational and negative numbers. He defined operations on all even roots (square root, fourth root, etc.) of negative numbers and incorporated the imaginaries to the reals. He also defined arithmetic operations on the set of imaginary numbers. In 1748, Euler introduced the use of the letter 'i' to represent the imaginary unit. Gauss coined the term complex number and the English physicist, William Rowan Hamilton (1805-1865), introduced the use of the ordered pair notation (a,b) for what Gauss had written a + bi. Wessel in 1798, Argand in 1806, and Gauss in 1813, defined complex numbers by graphical representation whereby the x-axis is identified as the axis of reals and the y-axis the axis of imaginaries. Thus every point in the Argand or Gaussian plane represents a complex number. Gauss defined multiplication of complex numbers as (a,b) • (c,d) = (ac-bd, ad+bc). It must be noted that the definitions of the operations with complex numbers are set up with only one thought in mind: Will these new operations preserve the consistency of the laws of the number system—associative, commutative, and distributive.

An important property of complex numbers is illustrated in Figure 8. We define |z| to be the modulus of z = (a,b) or the length of the vector in the Argand plane. Thus by the Euclidean distance
formula: $|z| = \sqrt{a^2 + b^2}$ one can show that

$$|z_1 + z_2| \leq |z_1| + |z_2|,$$

or geometrically, the side of a triangle is always less than or equal to the lengths of the other two sides. A complex number can also be represented in polar coordinate form (Figure 9): If $r$ is the radius of a circle whose center is the origin, then any point on the circle has coordinates $(r \cos \theta, r \sin \theta)$. Alternatively,

$$z = x + yi = r(\cos \theta + i \sin \theta).$$

Let us now investigate the multiplication of complex numbers in polar form, where

$$z = r(\cos \theta + i \sin \theta)$$

and

$$z_1 = r_1(\cos \phi + i \sin \phi).$$

Now

$$z z_1 = rr_1 (\cos \theta \cos \phi + i \sin \phi \cos \theta + i \sin \theta \cos \phi - \sin \theta \sin \phi)$$

$$= rr_1 [(\cos \theta \cos \phi - \sin \theta \sin \phi) + i(\sin \phi \cos \theta + \sin \theta \cos \phi)]$$

$$= rr_1 [\cos(\theta + \phi), \sin(\theta + \phi)].$$

In other words the product of two complex numbers is the product of their radii and the sum of their angles.

With this background, DeMoivre's formula can be deduced. Let
\[ z = z_1, \ r = r_1 \text{ and } x = \theta = \phi, \text{ then,} \]
\[ z^2 = r^2(\cos 2x + i\sin 2x). \]
Multiplying \( z^2 \) by \( z \),
\[ z^3 = r^3(\cos 3x + i\sin 3x). \]
Assume true for \( z^n \), i.e. that
\[ z^n = r^n(\cos nx + i\sin nx) \]
and for \( z^{n+1} \), we have
\[ z^{n+1} = z^n \cdot z \]
\[ = r^n(\cos nx + i\sin nx) \cdot r(\cos x + i\sin x) \]
\[ = r^{n+1} \left[ (\cos nx \cos x - \sin nx \sin x) + \right. \]
\[ \left. (\sin nx \cos x + \cos nx \sin x) \right] \]
\[ = r^{n+1} \left[ \cos(n+1)x + i\sin(n+1)x \right]. \]

This theorem was derived by Abraham DeMoivre (1667-1754). He was a French mathematician who fled the Protestant persecutions in France and settled in England. DeMoivre is known not only for the above application of complex numbers, but more for his outstanding work in the Theory of Probability, including the foundations of actuarial mathematics in his Annuities Upon Lives. In his Miscellanea Analytica, he is credited with deriving the probability integral \( \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \) and the formula for the curve of normal distribution,
\[ y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \]
where \( \sigma \) is the standard deviation and \( \mu \) is the mean of a statistical distribution. DeMoivre also is credited with so-called Stirling's Formula which for very large \( n \) is
\[ n! \approx (2\pi n)^{1/2} \cdot e^{-n} \cdot n^n. \]
This formula is useful in evaluation of factorials of very large
Before leaving the subject of complex numbers, an important use of DeMoivre's Theorem, that of finding roots of complex numbers, can be illustrated. The roots of a number are clearly solutions to the equation \( z^n = N \). If \( z = R(\cos nx + i\sin nx) \), and \( N = r(\cos y + i\sin y) \) then

\[
R^n(\cos nx + i\sin nx) = r(\cos y + i\sin y)
\]

which will be true if and only if

\[
R = \sqrt[n]{r}
\]

and \( nx = y + 2\pi k \), \( k = \{0,1,2,\ldots,n-1\} \)
or

\[
x = (y + 2\pi k)/n.
\]

Thus if \( z = r(\cos x + i\sin x) \) is a complex number its \( n^{\text{th}} \) roots will be given by

\[
\sqrt[n]{r} \left[ \cos \frac{k2\pi}{n} + i\sin \frac{k2\pi}{n} \right], \ k = \{0,1,\ldots,n-1\}.
\]

Example: Find the square roots of \( i \):

\[
i = 1(\cos \frac{\pi}{2} + i\sin \frac{\pi}{2})
\]

(1) \[
i^{\frac{1}{2}} = \sqrt{1}(\cos \frac{\pi/2 + 2(0)\pi}{2} + i\sin \frac{\pi/2 + 2(0)\pi}{2})
= \sqrt{1}(\cos \frac{\pi}{4} + i\sin \frac{\pi}{4}) = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}
\]

(2) \[
i^{\frac{1}{2}} = \sqrt{1}(\cos \frac{\pi/2 + 2(1)\pi}{2} + i\sin \frac{\pi/2 + 2(1)\pi}{2})
= \sqrt{1}(\cos \frac{5\pi}{4} + i\sin \frac{5\pi}{4}) = -\frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}
\]

Applying the properties of the roots of a quadratic equation to these roots \( x_1 + x_2 = -\frac{b}{a} = 0 \) and \( x_1x_2 = \frac{c}{a} = -i \)

thus \( x^2 + \frac{b}{a}x + \frac{c}{a} = 0 \)

and \( x^2 - i = 0 \).

An extension of the notion of complex numbers as ordered pairs of real numbers is the concept of quaternions as an ordered fourtuple
of real numbers. This is a larger system which contains the complex numbers as a proper subset. More important than that is the fact that these new numbers no longer satisfy all properties of classical numbers. We are at the threshold of the great beyond, about to leap into what has become the burgeoning land of "modern" algebra.

William Rowan Hamilton (1805-1865) is credited by most mathematicians as being the father of modern algebra. Hamilton was a child prodigy who had mastered Latin, Greek, and Hebrew by the age of three, French and Italian by the age of eight, and a half dozen other languages including Persian and Sanscrit by age thirteen. He was pronounced "first mathematician of the day" when he was seventeen. He thought of himself primarily as a physicist, not a pure mathematician. But the dividing line between theoretical physics and pure mathematics is a hazy one indeed! He was appointed a full professor of astronomy at Trinity College, Dublin, at the age of twenty-two over several highly credentialed applicants. In 1834, at the age of twenty-eight, he reached the pinnacle of success in physics with the publication and immediate verification of two new laws of light termed "Internal and External Conical Refraction." After this, for the remainder of his life he was torn between two unfortunates, a sickly, feeble wife, and an excessive weakness for alcohol.

The story of the invention of quaternions, whereby Hamilton became the father of modern algebra illustrates the conception and birth of a new idea.

For Hamilton, a complex number was not "one number", but an ordered pair of real numbers. He defined addition and multiplication on these ordered pairs and showed that they satisfied all six of the field properties:
A field, $F$, is a system consisting of a set $S$ of elements and two binary operations, $(+, \cdot)$ which satisfy the following properties:

1. If $a$ and $b$ are two elements of $F$, $a+b$ and $a\cdot b$ are uniquely determined elements of $F$.
2. The operations $(+, \cdot)$ are commutative.
3. The operations $(+, \cdot)$ are associative.
4. There exist elements 0 and 1 of $F$ such that for every $a$ in $F$, $a + 0 = a = 0 + a$ and $a \cdot 1 = a = 1 \cdot a$.
5. If $a$ is an element of $F$, then for every $a$, there is a unique element, $-a$, such that $a + (-a) = 0$. Also for every element $a \neq 0$ of $F$, there exists a unique element, $a^{-1}$, such that $a \cdot a^{-1} = 1$. $-a$ and $a^{-1}$ are respectively inverse elements for the binary operations $+$ and $\cdot$.
6. For $a, b, c$ in $F$, $a \cdot (b + c) = a \cdot b + a \cdot c$ (distributive property of multiplication over addition).

As can be readily seen these are nothing more than the properties of ordinary real numbers. The following are Hamilton's definitions of addition and multiplication of complex numbers:

$$(a,b) + (c,d) = (a+c, b+d)$$
and

$$(a,b) \cdot (c,d) = (ac-bd, ad+bc).$$

To illustrate these properties, consider three complex numbers $(a,b), (c,d), (e,f)$, where $a, b, c, \ldots, f$ are real numbers.

1. Closure:  
   $$(a,b) + (c,d) = (a+c, b+d)$$
   $$(a,b) \cdot (c,d) = (ac-bd, ad+bc)$$

2. Commutative:  
   $$(a,b) + (c,d) = (a+c, b+d) = (c+a, d+b)$$
   $$= (c,d) + (a,b)$$
   $$(a,b) \cdot (c,d) = (ac-bd, ad+bc)$$
\[= (ca-db, da+cb)\]
\[= (c,d) \cdot (a,b)\]

(3) Associative: 
\[(a,b) + [(c,d) + (e,f)] = (a,b) + (c+e, d+f)\]
\[= [(a+c)+e, (b+d)+f]\]
\[= [(a,b) + (c,d)] + (e,f)\]

Similarly for multiplication.

(4) The additive identity is \((0,0)\) since
\[(a,b) + (0,0) = (a+0, b+0) = (a,b)\]
The multiplicative identity is \((1,0)\) since
\[(a,b) \cdot (1,0) = (a-0, 0+b) = (a,b)\]

(5) Additive inverse: 
\[(a,b) + (-a,-b) = (0,0)\] implies that
\[-(a,-b)\] is the inverse of \((a,b)\)

Multiplicative inverse:
\[(a,b) \cdot (x,y) = (1,0).\] Solving we find
\[x = \frac{a}{a^2+b^2} \text{ and } y = -\frac{b}{a^2+b^2}\]
therefore the multiplicative inverse of
\[(a,b)\] is \(\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)\).

(6) Distributive property:
\[(a,b) \cdot [(c,d) + (e,f)] \equiv (a,b) \cdot (c,d) + (a,b) \cdot (e,f)\]
\[(a,b) \cdot (c+e, d+f) \equiv (ac-bd, ad+bc) + (ae-bf, af+be)\]
\[[(a(c+e)-b(d+f), a(d+f)+b(c+e)] \equiv (ac-bd+ae-bf, ad+bc+af+be)\]
\[(ac+ae-bd-bf, ad+af+bc+be) \equiv (ac+ae-bd-bf, ad+af+bc+be).\]

Hamilton, being a physicist, was concerned that there was no definition of a quotient of two vectors. Vectors, in Hamilton's day,
could be added and multiplied. Thus he set out to find a vector \( \mathbf{w} \)
such that given two vectors \( \mathbf{u} \) and \( \mathbf{v}, \mathbf{v} \neq 0, \mathbf{w} \mathbf{v} = \mathbf{u} \). Vector \( \mathbf{w} \) is
formed in the following way. Consider the three-dimensional rectangular axes and unit vectors \( i, j, k \) coinciding with the \( x, y, \) and \( z \)
axes respectively. The vectors \( \mathbf{u} \) and \( \mathbf{v} \) in space emanating from the
origin determine a plane \( X \). Vectors \( \mathbf{u} \) and \( \mathbf{v} \) are separated by an
angle \( a_3 \). Next consider what real numbers will be necessary to con­
vert vector \( \mathbf{v} \) into \( \mathbf{u} \). Hamilton found that it took rotations through
three angles and multiplication of \( \mathbf{v} \) by a scalar \( a_0 \) to transform \( \mathbf{v} \)
into \( \mathbf{u} \). The vector of transformation he called \( \mathbf{w} \), where

\[
\mathbf{w} = a_0 + a_1 i + a_2 j + a_3 k.
\]

The four coefficients \( (a_0, a_1, a_2, a_3) \) are real numbers to which he
gave the name quaternion.\(^49\) He defined addition of quaternions such
that all field properties were satisfied. But trying to define
multiplication of these numbers, he encountered a problem: try as he
might, for fifteen years he struggled, he could not find a way to
multiply these numbers so that all the field properties were satis­
fied. Multiplication of quaternions was not commutative. For hours,
days, weeks on end, he would isolate himself in his study, avoiding
all contacts, and even many meals in the hope that he could find the
elusive key. If suddenly he appeared at the breakfast table, his
wife and son would inquire, "have you succeeded in multiplying
quaternions yet?" We must realize, if all this seems silly, that
Hamilton was doing for algebra what Saccheri had done, but not real­
ized he had done, and what Gauss, Lobachevsky, and Bolyai had actual­
ly done just a few years before for geometry. It is very difficult
to break with tradition, whether that tradition be Euclid's parallel
postulate or the fundamental laws of real numbers.
Finally, in 1843, he resigned himself and released the revolutionary but logically consistent algebra of quaternions in which all properties except the commutative property of multiplication are satisfied. This was the key that would open the floodgates. Soon mathematics was filled with all sorts of algebras, many of which violated one or more of the inviolatable laws. Arthur Cayley invented the algebra of matrices in which multiplication is not commutative. There were Jordan and Lie algebras in which the associative property is violated.

But we are getting off the track. We have not indicated how ordinary complex numbers can be considered a subset of the hypercomplex numbers. A real number \(a\) is written \((a, 0, 0, 0)\) and the complex number \(3 - 2i = (3, -2, 0, 0)\). Since the hypercomplex numbers are merely an extension of the complex where \(i^2 = -1\), Hamilton defined the "units" \(j\) and \(k\), such that \(j^2 = -1\) and \(k^2 = -1\). One day while walking with his wife he carved the following multiplication table in the stone of a bridge near Dublin.\(^5\)

\[
\begin{array}{cccc}
. & 1 & i & j & k \\
1 & 1 & i & j & k \\
i & i & -1 & k & -j \\
j & j & -k & -1 & i \\
k & k & j & -1 & -1 \\
\end{array}
\]

Consider an example of quaternion multiplication. Let \(w = 1 + 2i + 3j + 4k\) and \(w' = 2 + i + 5k\) and we find

\[
ww' = (1 + 2i + 3j + 4k)(2 + i + 5k) \\
= -20 + 20i + 10k
\]

and

\[
w'w = (2 + i + 5k)(1 + 2i + 3j + 4k) \\
= -20 - 10i + 12j + 16k,
\]
from which it is clear that in general, for two quaternions, \( w \) and \( w' \), \( ww' \neq w'w \). However, for the special case where \( w' = a_0 + 0i + 0j + 0k \), a scalar, then it is obvious that \( ww' = w'w = a_0 w \).

It could be shown in a manner similar to that demonstrated for the complex numbers that quaternions satisfy the remaining field properties except commutative multiplication. As an example, we demonstrate that each quaternion \( w \) possesses an inverse \( w^{-1} \) such that \( ww^{-1} = w^{-1}w = 1 \). Express \( w = a_0 + v \) and \( v = a_1 i + a_2 j + a_3 k \). Then \( v \) is called the vector part of \( w \) and \( \overline{w} = a_0 - v \) is called the conjugate of \( w \). A simple computation will show that

\[
(1) \quad \overline{ww} = \overline{w}w = p = a_0^2 + a_1^2 + a_2^2 + a_3^2,
\]

called the norm of \( w \). Now \( p \) is a positive real number, and, by (1) we have

\[
\frac{1}{p} \overline{w} = \frac{1}{p} w = \frac{1}{p} \overline{w} \cdot w.
\]

In the previous example where \( w = 1 + 2i + 3j + 4k \),

\[
w^{-1} = \frac{1}{30} (1 - 2i - 3j - 4k).
\]

This corresponds closely with the inverse of a complex number.51

There is another surprising result. A corollary to the fundamental theorem of algebra states that classical algebra permits a maximum of two roots for any quadratic. Then over the complex numbers, \( x^2 + 1 = 0 \) has exactly two solutions, \( \pm i \). Over the quaternions, we have an infinite number of roots, for example \( \pm i; \pm j; \pm k; \pm \sqrt{2}(j+k); \ldots \) are only a few such. It can easily be verified that this is so. Such verification is left to the student.

Furthermore, for \( w = a_0 + a_1 i + a_2 j + a_3 k \), a necessary and sufficient condition for \( w^2 + 1 = 0 \) is \( a_0 = 0 \) and \( a_1^2 + a_2^2 + a_3^2 = 1 \). We assume first that \( w^2 + 1 = 0 \)

\[
(a_0 + a_1 i + a_2 j + a_3 k)^2 + 1 = 0
\]
and

\[ a_0^2 + 2a_1 a_0 i + 2a_2 a_0 j + 2a_3 a_0 k - a_1^2 - a_2^2 - a_3^2 + 1 = 0 \]

or

\[ a_0^2 - a_1^2 - a_2^2 - a_3^2 + 2a_0 (a_1 i + a_2 j + a_3 k) = -1. \]

This implies that \( 2a_0 (a_1 i + a_2 j + a_3 k) = 0 \) since the remaining four terms are real numbers. But the above implies further that either \( a_0 = 0 \) or \( a_1 i + a_2 j + a_3 k = 0 \) or both. If \( a_1 i + a_2 j + a_3 k = 0 \), then \( a_1 = a_2 = a_3 = 0 \) which is impossible because then \( a_0^2 = -1 \) which is also impossible since \( a_0 \) is real. Thus it must be that \( a_0 = 0 \) which leads to the immediate conclusion that \( a_1^2 + a_2^2 + a_3^2 = 1. \)

Conversely, assuming that \( a_0 = 0 \) and \( a_1^2 + a_2^2 + a_3^2 = 1 \), we have

\[(a_1 i + a_2 j + a_3 k)^2 + 1 = 0 \]

or

\[-a_1^2 - a_2^2 - a_3^2 + 1 = 0 \]

and the statement is proven.

The beginning of algebra as a postulational deductive system did not begin until the period just prior to Hamilton. It was gradually developed by members of the British Algebraic School, among whom were Augustus DeMorgan and George Boole. We shall investigate more of the work of the latter in Chapter III.

The "abstract" point of view demands that a pure science such as "an algebra" or "a geometry" be founded on undefined terms, postulates, definitions, and theorems much as Euclid did for geometry.

We can expand on the concept of hypercomplex numbers to one of order, or dimension, \( n \), and write the ordered \( n \)-tuple \((a_1, a_2, \ldots, a_n)\) or also as a linear combination \(a_1 e_1 + a_2 e_2 + \ldots + a_n e_n\), where \( e_i \) are to be considered vectors of unit length in our \( n \)-dimensional space, and the coefficients, \( a_j \), are numbers of some field, called a
scalar field. A scalar is a non vector, usually a real number. The product of a scalar and a given vector $\vec{v}$ is a linear multiple of the given vector. It amounts to a stretching or shrinking of the vector.

This idea was formulated by Hermann Grassmann (1809-1879), published in 1844 one year after Hamilton's quaternions were revealed.

Addition and scalar multiplication are defined, as would be expected, by simply adding corresponding coefficients and using the distributive law. For example, addition of two vectors is defined by

$$(a_1 e_1 + a_2 e_2) + (b_1 e_1 + b_2 e_2) = (a_1 + b_1) e_1 + (a_2 + b_2) e_2,$$

and scalar multiplication of a scalar and a vector is simply

$$a(b_1 e_1 + b_2 e_2) = ab_1 e_1 + ab_2 e_2$$

for a scalar, $a$, and a vector, $b_1 e_1 + b_2 e_2$. However, vector multiplication is more complicated. We illustrate with an algebra of order two:

$$(a_1 e_1 + a_2 e_2)(b_1 e_1 + b_2 e_2) = a_1 b_1 e_1^2 + a_1 b_2 e_1 e_2 + a_2 b_1 e_2 e_1 + a_2 b_2 e_2^2.$$}

Notice that since $a_1$ and $b_1$ are scalars they commute in the second and third terms. The product must be another hypercomplex number and therefore a linear combination of $e_1$ and $e_2$. A definition of multiplication of these unit vectors appears in the following table:

<table>
<thead>
<tr>
<th></th>
<th>$e_1$</th>
<th>$e_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_1$</td>
<td>$e_1$</td>
<td>$e_2$</td>
</tr>
<tr>
<td>$e_2$</td>
<td>$e_2$</td>
<td>$-e_1$</td>
</tr>
</tbody>
</table>

The product is thus: $a_1 b_1 e_1 + a_1 b_2 e_2 + a_2 b_1 e_2 + a_2 b_2 e_1$. The way we have defined the table $e_1 = 1$ and $e_2 = i$ in the ordinary complex number field. Notice also that the multiplication is commutative for order two.

Three types of associative linear algebras are: nilpotent,
matrix, and division. Define multiplication in accordance with the following table:

<table>
<thead>
<tr>
<th></th>
<th>e₁</th>
<th>e₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>e₁</td>
<td>0</td>
<td>e₁</td>
</tr>
<tr>
<td>e₂</td>
<td>e₁</td>
<td>e₂</td>
</tr>
</tbody>
</table>

Notice that $e₁^2 = 0$. Element $e₁$ is called a "divisor of zero" since a product of two non-zero elements is zero. The element $e₁$ is called "nilpotent" because one of its powers is zero. Such an element will be called "properly nilpotent" if its product with every hypercomplex number of the algebra is nilpotent. It can be shown that $e₁$ is properly nilpotent:

Consider $e₁(a₁e₁ + a₂e₂) = a₁e₁^2 + a₂e₁e₂ = 0 + a₂e₁ = a₂e₁$

Now $(a₂e₁)^2 = a₂^2e₁^2 = a₂^2(0) = 0$.

Therefore $e₁$ is properly nilpotent. It is possible for some of the elements of an algebra to be properly nilpotent and others not.

We mentioned that Arthur Cayley invented matrices. A matrix is a rectangular or square array of numbers. All operations can be defined and it can be shown that matrix multiplication, where defined, like the hypercomplex algebra of Hamilton, is non-commutative.

Matrices have been found very useful in many applications, one of which is solving systems of equations.

The last type of hypercomplex algebra is division algebra. It can be shown that the algebra defined above does not possess a division operation. For example consider the following problem:

$$\frac{6e₁ + 2e₂}{2e₁}$$

If some hypercomplex number were the answer, say
\[ xe_1 + ye_2 = \frac{6e_1 + 2e_2}{2e_1} \]

then we would have

\[ 2e_1(xe_1 + ye_2) = 6e_1 + 2e_2 \]

by definition of division as the inverse of multiplication

and

\[ 2xe_1^2 + 2ye_1e_2 = 6e_1 + 2e_2 \]

\[ 0 + 2ye_1 = 6e_1 + 2e_2. \]

But this implies that \( 2 = 0 \), which is impossible.

There is an important structure theorem proven by G. Frobenius (1849-1917) and by Charles Saunders Pierce (1839-1914). This theorem states that if the scalar or coefficient field is the system of real numbers, then the complex numbers of order two, the quaternions of order four, and the real numbers themselves (an algebra of order one) constitute the only associative division algebras. The theorem is a tremendous time saver for it eliminates laborious examination of many different associative division algebras.\(^5\)

This discussion of abstract algebra has been but a glimpse at the tremendous work being done by twentieth century algebraists. Another structure of abstract algebra not mentioned here is the group, whose properties were examined in Chapter I where the Klein geometry model was discussed. The group is the basic structure, or simplest algebraic system. Expanding the number of properties, the ring, commutative ring, the integral domain, the ideal, and the field can be defined.

I have tried to trace the development of algebra from its beginnings as man's attempts to catalog his belongings through its geometric manifestations during the Greek civilization, through its syncope to final symbolization, unto abstract fruition in the nine-
teenth and twentieth centuries.

My work has convinced me of one paramount thing: one cannot appreciate mathematical accomplishment apart from its history. I only hope my reader can glimpse, even partially, the beauty, and the majesty that is mathematics!
FOOTNOTES


4. Ibid., p.10.


6. Ibid., pp.146-149.


16. Ibid., pp.11-12.

17. Ibid., pp.91-92.

18. Ibid., p.86.

27. Ibid., p. 497.
32. Ibid., pp. 27-30.
34. Ibid., pp. 507-508.
35. Ibid., p. 509.
41. Lecture Notes, Mathematics 520.
42. Details assimilated and combined from Kramer, op. cit., p. 97 and Eves, op. cit., p. 217.
43. Problems in Eves, op. cit., p. 236, problems #8-11h; the proof is mine.

45. Problem in Eves, op. cit., pp. 292-293; the solution is mine.


47. Eves, op. cit., p. 356.


 CHAPTER III
THE DEVELOPMENT OF MATHEMATICAL LOGIC AND SET THEORY

The title of this chapter is an umbrella term which covers four related topics: (1) an examination of a postulational system together with an example; (2) David Hilbert's plan for mathematics and its shortcomings; (3) the development of Cantor's theory of sets; (4) George Boole's universal language as applied to logic and switching circuits.

(1) A Postulational System

Descartes is responsible for holding mathematical proof as the model for logical proof. He looked to Euclidean proof as his example. The origins of proof are much earlier than the age of the founder of the school at Alexandria. Postulational systems have matured considerably since the early morning musings of the founder of modern philosophy. Now mathematical proof looks to undefined or primitive terms, postulates giving meaning to the terms by statements about them, and theorems or conclusions which follow rigorously according to the laws of logic.

The ancient Babylonians developed their mathematics from empirical evidence, measurement, intuition, and guessing. So far as we know they made no attempt to, or saw any reason to, demonstrate deductively the truth of formulas they used. If a certain formula appeared to obtain the same answer which could be obtained by measurement, it was considered correct. Why then, if something seems to be true and a formula works in a number of trials, must we insist upon a formal proof? There are three important reasons, stated by Bell: 1

(1) Deductive reasoning is the only means yet devised for isolating and examining hidden assumptions, and for following the subtle implications of hypotheses which may be less factual than they seem.

(2) Modern technological precision demands accurate rules derived from certified math-
ematics. The magnitude of an error might be much larger than originally assumed thus ruining a product.

(3) The evolution of technology depends upon the constant improvement and enlargement of mathematical knowledge. This can only come from precise verification on the basis of known facts. Applications are quickly discovered for mathematical theories.

The methodology that is mathematical proof had its origins in the first half of the sixth century B.C. Thales of Miletus was the first to give a semblance of verification to such mathematical facts as, the base angles of isosceles triangles are equal. He was soon followed by Pythagoras and the Pythagorean brotherhood.

The question occurs, why did the Greeks, instead of the Babylonians or Egyptians, recognize the necessity for proof? Several possible explanations come to mind. First, the Greeks were primarily concerned with explaining the meaning of the universe and of man's role in the larger scheme of things. Second, there were outstanding philosophers who thought about these things, notably Socrates and later Plato. They put their minds to the problem of unifying the diversity they observed. Third, the value of mathematical logic was quickly recognized as fertile training ground for the mind, and among the wealthy Greeks intellectual pursuits were the pinnacle of life.

We can enumerate two from many achievements of the Greeks. The first is the explicit recognition that proof by deductive reasoning offers a foundation for mathematics. The second was the firm belief that the key to understanding the nature of things was mathematics, and that the language of mathematics was "most adequate for idealizing the complexity of nature into comprehensible simplicity". The best work was done by Euclid of Alexandria (ca 365-275 B.C.) He collected all of the known mathematics and systematized it into what became the world standard for 2200 years. He was the first to organize the structure under the familiar postulational umbrella.
Let us now examine a postulational system. The building blocks of any scientific subject are undefined or primitive terms. Since all definitions must use accepted terms to characterize that which is presently being defined, it is obvious that the whole process must start somewhere. Thus those terms upon which the structure is built must be left undefined in the strictest mathematical sense.

Next come the assumptions or unproved propositions which we call postulates. Actually Euclid grouped these into two categories: (1) postulates or assumptions which characterize the relationship among the elements of the structure, which can be given the name "primitive terms", and (2) axioms or assumptions considered to be "common knowledge". Modern practice relegates both of these to one group called postulates. The obvious question to the uninitiated or non-mathematician is "how do you know your postulates are true?" One answer is that their truth is not absolute. A better explanation is that the "truth" of a postulate is not a pertinent issue, for their purpose in an abstract sense is merely to describe relationships among our primitive terms. For example, suppose we are creating a system and our primitive terms are "nodes" and "poles" and our first postulate is the statement, "Every node contains at least two poles". The truth of this statement is irrelevant.

We are now ready to examine the properties which characterize a postulational system. The most important and absolutely essential of these properties is that the system be consistent. This means that our postulates must not allow the deduction of two theorems which are contradictory. The problem here is how does one guarantee that his system is consistent? We obviously cannot simply start proving theorems and, if we do not arrive at a contradictory pair, then assume we have consistency. We need a criteria for consistency. The one which mathematicians usually accept is a model or framework for the undefined terms within which, or relative to which, all the postulates are true statements. An excellent example of this is the Klein circle
model for the consistency of the postulates of hyperbolic geometry. The question which might still haunt us, is can we not establish in some absolute sense independent of a particular model the consistency of our axiom system? This is a question which has concerned the greatest mathematical minds of the recent era. It was not settled once and for all until Kurt Gödel, a prominent German mathematician, proved in 1931, that it was impossible to prove the absolute consistency of a formal (completely abstract) system. 3

Consistency is the only attribute which is a "must have" for any system. There are others which are important: For example, how do we know whether or not we have assumed too much, that we have postulated something which could be proven? The answer is found in the concept of independence. Redundancy in a set of axioms is abhorrent to the mathematician. If one or more of the axioms are deducible from the remainder they are said to be dependent. High school geometry postulates a number of statements which are actually provable. Because of mathematical immaturity on the part of high school sophomores, some propositions are assumed which would be too difficult for them to prove at that stage of their development. There are several notable examples: corresponding angles of parallel lines are congruent; if three sides of one triangle are congruent to three sides of another, the triangles are congruent; if two sides and the included angle of one triangle are congruent respectively to two sides and the included angle of another, the two triangles are congruent, etc. Independence is an ideal. If an axiom is independent of the remainder of axioms, it is proven by finding a physical representation, a model, of the primitive terms in which all axioms are true except the one in question, which is false. An excellent case in point is the almost 2000 year struggle to prove Euclid's fifth postulate. It was not until the consistency of the non-Euclidean geometries had been established that the independence of the Euclidean fifth could finally be established. The models exhibited for hyperbolic
geometry, notably Klein's boundaryless circle, used all Euclidian postulates except the fifth.

The next logical question is whether we have assumed too little. Do we need more postulates to prove everything we might wish to prove about our system. Again we are drawn to Euclid for an example.

In view of the fact that Euclid did not utilize the fifth or parallel postulate until proposition 29 of book I, it is conjectured that originally he planned to use only the first four postulates. Each of these is briefly expressed. Then, faced with the 29th, he discovered that he had postulated too little. Hence, he worked out a cumbersome lengthy postulate to establish a criterion for intersecting lines. Then he could attack the 29th which discusses parallel lines.

Completeness is that property of an axiom system which signals that we have enough postulates. This requires proof. Hence we define a system to be complete if it is possible to deduce either a proof of every proposition about the primitive terms or the negative of that proposition.

The next useful attribute of a system is that of categoricalness because it limits the models of a system to uniformity. We say that an axiom system is categorical if every pair of models of that system are isomorphic—the elements of two models are in a one-to-one correspondence to one another and all relations among the elements are preserved. Whether or not we demand categoricalness for our system depends upon the purpose for which our system is being devised. If we want our axiomatic system to formalize just one clearly delimited theory, such as Euclidean geometry of the real number system, then we would demand that it possess essentially one framework of interpretation. An example of a consistent but non-categorical axiom system is the theory of groups. It has a wide variety of examples and thus a wide range of application.

What follows is an abstract, finite geometry as an illustration of the way the properties defined for a postulational
system can be applied in practice. The postulates of a finite geometry limits the number of elements to a definite number. 6

Let \( S \) be a set of undefined elements, \( A, B, C, \ldots \), and let \( S \) have certain undefined subsets called \( m \)-sets. The relations "belonging to a set" and "containing an element" are undefined. We postulate the relationship between an element and an \( m \)-set in the following way:

\[
\begin{align*}
\text{P1} & \quad \text{Two elements determine an \( m \)-set. (existence)} \\
\text{P2} & \quad \text{There is not more than one \( m \)-set containing } A \text{ and } B. \quad \text{(uniqueness)} \\
\text{P3} & \quad \text{Any two \( m \)-sets have at least one element of } S \text{ in common.} \\
\text{P4} & \quad \text{There exists at least one \( m \)-set.} \\
\text{P5} & \quad \text{Every \( m \)-set contains at least three elements of } S. \\
\text{P6} & \quad \text{All elements of } S \text{ do not belong to the same \( m \)-set.} \\
\text{P7} & \quad \text{No \( m \)-set contains more than three elements of } S.
\end{align*}
\]

To verify consistency we offer the following model where the letters represent elements and the vertical columns represent \( m \)-sets, (Fig. 1).

\[
\begin{array}{cccccc}
A & B & C & D & E & F & G \\
B & C & D & E & F & G & A \\
D & E & F & G & A & B & C
\end{array}
\]

FIG 1

It is easily seen that each of the postulates are true. For example: \( \text{P1} \) is obviously true; for \( \text{P2} \), notice that not more than one vertical column contains any two elements; for \( \text{P3} \), taking the columns in pairs we notice that there is exactly one element in common; \( \text{P4}, \text{P5}, \text{P6}, \text{P7} \) are obviously true.

In order to verify independence we must offer seven models in which exactly one postulate is falsified. Note: \( \text{P1}' \) means \( \text{P1} \) is false, etc. It is left to the student to verify the truth of the remaining axioms in each case.

\[
\begin{align*}
\text{P1}' & \quad A & A & B & C & \text{There is no \( m \)-set containing } B \text{ and } E \text{ or} \\
& & B & E & E & D & A \text{ and } D, \text{ etc.} \\
& & C & F & G & E & \\
\text{P2}' & \quad A & A & A & B & \text{A and } B \text{ are contained in two \( m \)-sets.} \\
& & B & B & C & C & \\
& & C & D & D & D
\end{align*}
\]
P3' 1 2 3 4 5 6 7 8 9 10 11 12 M-sets 11 and 12
   A A A A B B C C C C D G have no elements
   D B E F E D F D E F E H in common.
   C G I H H I G H G I F I
P4' Consider a single element, A. The remainder of the postulates are true vacuously, that is, the remainder of the postulates are not exhibited, thus not false.

P5' A B C
    B C A
P6' A
    B
    C
P7' The real projective plane.
    We can prove categoricalness after proving some theorems.
In practice one might prove two for the class and leave the remainder to them.
    First, observe that the principle of duality for the system can be established by proving as theorems the duals of each of the axioms. P1 and P3 are duals of each other when m-set and element, "containing an element" and "belonging to a set" are interchanged.

Theorem 1: (Dual of P2) Two distinct m-sets have only one element in common.
Proof: Assume A and B are each members of m-set #1 and m-set #2, but this violates P2. Therefore our assumption is false and the theorem is proved.

Theorem 2: (Dual of P4) There exists at least one element in S.
Proof: By P4 there exists at least one m-set and by P5 it must contain at least three elements. Therefore we certainly have one element in S.
Theorems 3, 4, and 5 are left to the student.

Theorem 3: (Dual of P5) Every element belongs to at least three m-sets.

Theorem 4: (Dual of P6) No element is contained in all m-sets.

Theorem 5: (Dual of P7) No element belongs to more than three m-sets.

Theorem 6: There exist exactly seven elements in S.
Proof: By P4, P5, P6 there exist at least four elements, A, B, C, D. Now let D be the element not on m-set ABC. Now by P1, A and D; B and D; C and D determine three new m-sets which contain exactly three elements E, F, and G. Assume an eighth element H exists, but then A and H would determine a new m-set violating Theorem 1 unless there exists a ninth element J, but then AHJ would violate P3.

Theorem 7: (Dual of Th 6) There exist exactly seven m-sets.

Proof: True by duality.

Since we are limited to seven m-sets and seven elements, any model for this system will be isomorphic to the consistency model offered above. This proves our system is categorical.

(2) David Hilbert

Metamathematics is "a theory which studies the properties of formal axiomatic systems". Thus the consistency, independence, and completeness properties are part of the realm of metamathematics. David Hilbert originated the "formalist" school of thought in which consistency is the essential problem. Hilbert and his followers attempted to prove the absolute consistency of all formal systems of mathematics. They succeeded in proving that the additive arithmetic of integers was absolutely consistent. Recall that the relative consistency of non-Euclidean geometry was established as a result of Beltrami's pseudo-sphere and Klein's boundaryless circle. Hilbert and his followers showed that if Euclidean geometry was consistent, then so was non-Euclidean. Hilbert wanted to stop this "buck-passing" by proving, once and for all, that all classical mathematics is absolutely consistent and complete. Then in 1931, Kurt Gödel delivered the fatal blow to all these efforts by proving that "any formal system equivalent to the Russell-Whitehead Principia Mathematica must necessarily be incomplete". Gödel showed that there must be undecidable statements in any system and that consistency is one of those undecidable propositions. His proof is classic and revolutionary in the sense that Abel and Galois...
were revolutionary in proving the impossibility of solving quintic and higher degree equations in terms of the coefficients. Since his proof is so important his ideas are outlined in Appendix C.

(3) Cantor's Theory of the Infinite

Georg Cantor was a first rate mathematician who was left in a second class university because his teacher, Leopold Kronecker, violently disagreed with his theory. Kronecker had achieved the position to which Cantor so devoutly aspired, a professorship at the University of Berlin, the prestige German university in the middle of the nineteenth century. It would be tantamount to a professorship at, say, Harvard today.

Cantor was born in 1845, a Jew in a Christian household. His father converted to Protestantism, his mother was a Roman Catholic by birth. His father, whom he loved and respected had insisted for sometime that Georg be an engineer. Georg not wanting to disappoint his father submitted, much against his will and his own mental well being. This is one of the things which psychologists point to when they try to explain his emotional and mental instability in later life. Finally his father withdrew his objections and allowed his son to pursue mathematics for which the young man was extremely grateful.

Cantor's thinking shook the mathematical world. First of all he attacked the very foundations of analysis at the root of which are questions about continuity, limits, and convergence. The importance of his work is better appreciated by returning to ancient Greece.

The philosopher Zeno proposed four famous paradoxes. Two of these sought to destroy the concept of divisibility. 9

(1) The dichotomy: There is no motion because that which is moved must arrive at the middle of its course before it arrives at the end. Of course it must traverse half of the half before it reaches the half, and so ad infinitum. How is it possible to reach an infinite number of positions in a finite time?
(2) The Achilles: Achilles can run 1000 yards in the same time the turtle can run 100 yards. The turtle starts 1000 yards ahead of Achilles. Achilles will never overtake the slower moving turtle. Achilles covers the 1000 yards, the turtle is still 100 yards ahead, when he covers those 100 yards, the turtle has covered another 10 and so on.

Both of these paradoxes state the same principle: There are an infinite number of positions to be reached in a finite length of time. Therefore no distance can be traversed and Achilles will never overtake the turtle, since at each position, the turtle is always one-tenth of the previous increment ahead of Achilles. The table illustrates the issue:

<table>
<thead>
<tr>
<th>Position</th>
<th>Achilles</th>
<th>Turtle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>2</td>
<td>1000</td>
<td>1100</td>
</tr>
<tr>
<td>3</td>
<td>1100</td>
<td>1110</td>
</tr>
<tr>
<td>4</td>
<td>1110</td>
<td>1111</td>
</tr>
<tr>
<td>5</td>
<td>1111</td>
<td>1111.1</td>
</tr>
<tr>
<td>6</td>
<td>1111.1</td>
<td>1111.11</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

It is clear that the concept of "infinite" must be defined.

Cantor attacked the notion of infinity with the ingenuity and the clarity of a genius, although some Kronecker adherents will dispute his clarity. Cantor formalized the process of counting by substituting the concept of a one-to-one correspondence or matching. He stated that two sets have the same number of elements if a one-to-one correspondence can be set up such that when we have finished counting there are no elements left over in either set. He showed that the cardinal number, that is, the number of elements in the set of square integers is equal to the cardinality of the positive integers.

\[
1 \ 2 \ 3 \ 4 \ 5 \ 6 \ ... \ n \ ...
\]

\[
1 \ 4 \ 9 \ 16 \ 25 \ 36 \ n^2 \ ...
\]

If the set of squares are removed from the set of integers, some integers remain yet the sets have the same cardinality. A proper subset has the same number of elements as the whole set from which it was taken!
Consider the number of points on a small segment. We can prove (Cantor style) that it has the same number of points as a larger segment (Fig. 2).

Let ray PR rotate counterclockwise from PC' to PD'. For every point of intersection between C and D, say A, there will be a corresponding point A' on C'D'. Hence there is a one-to-one correspondence between the points on CD and C'D' which is longer. Notice without loss of generality we can simply apply one of the segments on the other so that there will be a deficiency between their lengths.

Cantor then defined an infinite set as one which can be put in a one-to-one correspondence with the set of natural numbers. Recalling Zeno's Dichotomy: the paradox occurred because a moving point will occupy an infinite number of positions, namely \( \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots \).

We see the correspondence:

\[
\begin{array}{ccccccc}
1 & 2 & 3 & 4 & 5 & 6 & \ldots & n \\
\frac{1}{2} & \frac{1}{2^2} & \frac{1}{2^3} & \frac{1}{2^4} & \frac{1}{2^5} & \frac{1}{2^6} & \ldots & \frac{1}{2^n} \\
\end{array}
\]

Similarly for the other problem, Achilles and the turtle:

<table>
<thead>
<tr>
<th>Achilles</th>
<th>0</th>
<th>1000</th>
<th>1100</th>
<th>1110</th>
<th>\ldots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turtle</td>
<td>1000</td>
<td>1100</td>
<td>1110</td>
<td>1111</td>
<td>\ldots</td>
</tr>
<tr>
<td>Natural Numbers</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>\ldots</td>
</tr>
</tbody>
</table>
Every child knows how to prove there is no largest natural number. We say that the natural numbers form a denumerable infinity. They certainly are countable, limited only by physical time. Thus any set which can be put into a one-to-one correspondence with the natural numbers is also denumerable. We can now see how Cantor solved Zeno's paradoxes. If Achilles caught up with the turtle, the places the turtle had been would be only part of the places Achilles had been, thus a part would be equal to the whole, which is countable by Cantor's definition. The cardinality of the set of rational numbers is denoted by the symbol $\aleph_0$, read "aleph null", the first letter in the Hebrew alphabet.

The next task is to show that the set of all rational numbers is a denumerable infinity. This is a shocking bit of mathematical wizardry to the "man in the street", especially when one tries to explain what density of the rationals means; i.e., that between any two rationals there is always another. How is such a correspondence constructed? We make a list of all rational numbers in order of some numerator and denominator, being careful not to repeat a fraction whose equivalent has been used, such as $1/2$ and $2/4$.

\[
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 \\
1/1 & 1/2 & 2/1 & 1/3 & 3/1 & 1/4 & 2/3 \\
8 & 9 & 10 & 11 & 12 & 13 & 14 \\
3/2 & 4/1 & 1/5 & 5/1 & 1/6 & 2/5 & 5/2 \\
15 & 16 & 17 & 18 & 19 & 20 & \text{etc.} \\
3/4 & 4/3 & 6/1 & 1/7 & 3/5 & 5/3 \\
\end{array}
\]

Additional surprises are observed when the arithmetic of these new numbers is investigated. What is $1 + \aleph_0$? The table is illustrative. The cardinality of the set of even integers is $\aleph_0$.

\[
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 \ldots n \\
2 & 4 & 6 & 8 & 10 & 12 & 14 \ldots 2n. \\
\end{array}
\]
An additional element has been added to $\mathcal{N}_0$, and the result is still $\mathcal{N}_0$. Consider the cardinality of the odd integers, it is certainly $\mathcal{N}_0$. Similarly the cardinality of the even natural numbers is $\mathcal{N}_0$. The cardinality of the union of these disjoint subsets of the naturals is $\mathcal{N}_0 + \mathcal{N}_0$, but that is simply the cardinality of the natural numbers, hence we have

$$\mathcal{N}_0 + \mathcal{N}_0 = \mathcal{N}_0$$

$$2 \mathcal{N}_0 = \mathcal{N}_0$$

Consider the following table:

<table>
<thead>
<tr>
<th></th>
<th>1/1</th>
<th>2/1</th>
<th>3/1</th>
<th>4/1</th>
<th>5/1</th>
<th>6/1</th>
<th>7/1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1/2</td>
<td>2/2</td>
<td>3/2</td>
<td>4/2</td>
<td>5/2</td>
<td>6/2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1/3</td>
<td>2/3</td>
<td>3/3</td>
<td>4/3</td>
<td>5/3</td>
<td>6/3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1/4</td>
<td>2/4</td>
<td>3/4</td>
<td>4/4</td>
<td>5/4</td>
<td>6/4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1/5</td>
<td>2/5</td>
<td>3/5</td>
<td>4/5</td>
<td>5/5</td>
<td>6/5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1/6</td>
<td>2/6</td>
<td>3/6</td>
<td>4/6</td>
<td>5/6</td>
<td>6/6</td>
<td></td>
</tr>
</tbody>
</table>

It can be seen that the cardinality of each column is $\mathcal{N}_0$ and there are certainly $\mathcal{N}_0$ different columns, therefore

$$\mathcal{N}_0 \cdot \mathcal{N}_0 = \mathcal{N}_0^2 = \mathcal{N}_0$$

Multiplying by $\mathcal{N}_0$ both sides, we have

$$\mathcal{N}_0^3 = \mathcal{N}_0^2 = \mathcal{N}_0$$

Assume inductively

$$\mathcal{N}_0^k = \mathcal{N}_0$$

Then

$$\mathcal{N}_0 \cdot \mathcal{N}_0^k = \mathcal{N}_0$$

$$\mathcal{N}_0^{k+1} = \mathcal{N}_0^2 = \mathcal{N}_0$$

One tends to suspect after such strange happenings that all sets are countable! Not so! It can be shown that the set of reals between 0 and 1 is uncountable. Suppose it is countable. Write a list purporting to be in a one-to-one correspondence with the naturals. Then construct a different real number not equal to any in the list. For the tenths place, select a digit different from the tenths place of the first number on the list and also different from 0 and 9, since 0 might lead to .000... = 0 and .9 might lead to .99 = 1, a rational number pos-
sibly on the list. For the hundredths place, select a digit different from 9 and different from the hundredths place of the second number on the list. For the thousandths place, pick a digit different from nine and from the thousandths place of the third number on the list, and so on. When this process is continued, we will obtain a real number different from any on the list in at least one decimal digit. Thus the cardinality of reals between zero and one must be greater than \( \mathbb{N}_0 \). It cannot be less or it would be a natural number. We call this new number \( C \). It represents a continuum (see Chapter II) as opposed to a discrete set.

The two new numbers, \( \mathbb{N}_0 \) and \( C \), are called transfinite numbers. Failing to find any transfinite number, \( T \), such that \( \mathbb{N}_0 < T < C \), Cantor hypothesized that none exist. This assumption has come to be known as the "continuum hypothesis".

There are other results of Cantorian Set Theory. Let \( R \) be the set of all rationals and \( I \) be the set of real numbers in the interval \((0,1)\); \( (x,y) \) = the set of all numbers between \( x \) and \( y \) exclusive of \( x \) and \( y \); and \( [x,y] \) = the set of all numbers between \( x \) and \( y \) including \( x \) and \( y \). We have

\[
I = R + R'
\]

where \( R' \) is the set of irrationals. Further, let

\[
E = \left\{ \sqrt{x}, \frac{1}{2}\sqrt{x}, \frac{1}{3}\sqrt{x}, \ldots, \frac{1}{n}\sqrt{x}, \ldots \right\}
\]

where \( x \) is any irrational in \( R' \). Now \( F = R + E \) is a countable set since it is the sum of two countable sets. Let \( E' \) be the set of irrationals in \((0,1)\) after the numbers in \( E \) are removed. Now

\[
I = (R + E) + E'
\]

and

\[
R' = E + E'.
\]

Now since \( E \) and \( F \) are countable and can be paired one-to-one and similarly \( E' \) can be paired with itself, \( I \) and \( R' \) can be paired one-to-one. Hence they both must have the same cardinal number \( C \). We thus have a subset and its parent set having the same cardinality once more, thus,
\[ \mathfrak{c} + c = c + \mathfrak{c} = c, \]

or

\[ c - \mathfrak{c} = c, \]

and

\[ c - \mathfrak{c} - \mathfrak{c} = c - \mathfrak{c} = c. \]

Observe that if \( C \) is the cardinality of \((0,1)\) it is also the cardinality of \([0,1]\), a corollary of the previous result.

Similarly by pairing every real number on the interval \((0,1)\) with its double and every real of the interval \((0,2)\) with its half, we have that \( C \) is the cardinality of \((0,2)\). Extending this idea we can show that \( C \) is the cardinality for all reals in the interval \((0,n)\) where \( n \) is any natural number. Symbolically

\[ n \cdot \mathfrak{c} = \mathfrak{c}. \]

By pairing \( x \) with \( 1+x \), one can establish a one-to-one correspondence between the reals of \((0,1)\) and \((1,2)\). Similarly, by pairing \( x \) with \((1/2)x - (1/6)\) we can pair the reals of \((0,1)\) with the reals of \((-1/6), (1/12)\). A one-to-one correspondence can be established between the reals of \((0,1)\) and the reals of any finite interval. Thus we have

\[ [0,1] \cup (1,2] \cup (2,3] \cup ... \]

is the set of all non-negative reals and

\[ (0,-1] \cup (-1,-2] \cup (-2,-3] \cup ... \]

is the set of all negative reals. It can be seen that for each set

\[ (1) \]

\[ \mathfrak{c} + \mathfrak{c} + \mathfrak{c} \ldots = \mathfrak{c} \cdot \mathfrak{c} \]

thus \( 2 \cdot \mathfrak{c} \cdot \mathfrak{c} \) is the cardinality of both sets or all real numbers.

Figure 3 shows how we can establish a one-to-one correspondence between all reals on the continuum \((0,1)\) and all real numbers.

In Figure 3, given any point \( P \) on the continuum \((0,1)\), it can be paired with another point \( P' \) on the semicircle which is projected through the center of perspectivity, \( A \), to \( P' \). Thus the cardinal number of all reals numbers is \( C \) and we have

\[ \mathfrak{c} \cdot \mathfrak{c} = \mathfrak{c} \]

by \((1)\) above.
Now consider the unit square (Fig. 4). Pairing a single real number on (0,1) with the ordered pair within the square on (0,1) in the following manner. Write the ordered pair in infinite decimal form; i.e., \( \left( \frac{1}{2}, \frac{\pi}{4} \right) = (.5000..., .78539...) \) and form a single decimal by choosing alternating digits from each of the two numbers to form a single number. (.5000..., .78539...) is mapped into .5708050309... Therefore we have
\[ C \cdot C = C^2 = C. \]

By a similar argument, \( C^3 = C. \) In general, it can be shown that \( C^n = C. \)

The next question is: are there any cardinals greater than \( C. \) The answer is affirmative. Consider the set consisting of all subsets of a given set. It can easily be seen that there are \( 2^n \) subsets for a set containing \( n \) elements. The proof follows: In forming a subset there are two choices for each element, either it is an element of the subset or it is not. Thus for \( n \) members there are \( 2^n \) possibilities in all. Further, it can be proven that \( 2^n > n \) for all \( n, \) finite or transfinite. Thus
$2^{\aleph_0} > \aleph_0$

and

$2^c > c$.

It can be shown in addition that $2^{\aleph_0} = c$. From the first inequality above,

$c > \aleph_0$.

We have the continuing inequality, where $S = 2^c$:

$\aleph_0 < c < S$.

Continuing the magnitude, $S_1 = 2^S > S$ and $S_1 > S_1$ and so on, we have an increasing series of cardinals where $10^1$:

$\aleph_0 < c < S < S_1 < S_2 < \ldots$.

An interesting chain of activity has taken place since Cantor first proposed his Theory of Sets. A number of mathematicians, including Ernst Zermelo (1871-1956) and Abraham Fraenkel, found the need to place Cantor's theory on a more rigorous foundation. A contradiction had developed in 1895-96 which Cantor was unable to explain.

Before we can consider this paradox we must investigate further Cantor's theory of transfinite ordinals. For this purpose, several definitions are necessary: partial ordering, order-preserving isomorphism, ordinal number, and well-ordered.

A relation $R$ in a set $X$ is called a partial ordering in $X$ if and only if the following three attributes are satisfied:

(1) $R$ is reflexive--$xRx$ for all $x$ in $X$.
(2) $R$ is asymmetric--for all $x,y$ in $X$, if $xRy$, then it is false that $yRx$.
(3) $R$ is transitive--if $xRy$ and $yRz$, then $xRz$.

Obviously examples of a partial ordering relation are $\leq$ (less than or equal to) for real numbers and the inclusion relation $\subseteq$ (is a subset of) for a collection of sets. There are, likewise, a number of examples of relations which fail to be partial orderings. One of these, "is a brother or sister of" fails to meet the second criteria--if John is a brother of Joe, then clearly Joe is a brother of John.

To make clear the idea of an order preserving isomorphism,
consider an example. A power set of a set of three elements has eight members, namely, the original set containing three elements, three subsets each containing two elements, three subsets each containing one element, and finally the empty set. This power set is partially ordered by inclusion. The set \{1, 2, 3, 5, 6, 10, 15, 30\} whose members are divisors of thirty is partially ordered by the relation "is a multiple of". 30 is a multiple of all eight elements; 15, 10, and 6 are each multiples of four elements; 5, 3, and 2 are each multiples of two elements, while 1 is a multiple of only itself.

These two sets are obviously not equal, but it can be seen that they are indistinguishable as far as their partial ordering is concerned. To appreciate this, consider the following dot and segment diagram (Fig. 5):

Fig. 5

where each dot represents an element, and a higher dot is connected to a lower dot by a downward path of one or more segments if and only if the two are related by the partial ordering relation. For example, in connection with our two sets, the highest dot would represent the improper subset in the first case and 30 in the second. Similarly, the next three lower dots represent the two element subsets in the former case and the elements 15, 10, and 6 in the latter case, etc. We thus define an isomorphism between two partially ordered sets if

(1) there is a one-to-one correspondence, \(x \mapsto f(x)\), between the elements of the two sets, and

(2) the partial ordering relation is order preserving—symbolically, \(x \mathbin{R} y\) implies \(f(x) \mathbin{R'} f(y)\).
The likeness which was observed between the collection of subsets of a three element set and the set of divisors of thirty, with their respective partial ordering relations, is an isomorphism.

A least member of a set \( X \) relative to a partial ordering relation is a \( y \) in \( X \) such that \( yRx \) for all \( x \) in \( X \). With these definitions we can define a partially ordered set \( X \) with an ordering relation \( R \) to be well-ordered if and only if each non-empty subset has a least member. A familiar example of a well-ordered set is the set of natural numbers. On the other hand, the set of rational numbers in order of size is not well-ordered since it contains many subsets with no least member, i.e., the set of rationals greater than four. It should be noted, however, that the set of rational numbers ordered according to the scheme on page 158 does indeed form a well-ordered set since any countable subset can be put into a one-to-one correspondence with the natural numbers, which are well-ordered.

Cantor required an ordinal number to characterize all sets which have the same cardinal number (the same number of elements) and which are isomorphic with respect to order. If we consider the set of non-negative integers arranged in order of increasing size, each number after the first is the ordinal number of the set of numbers preceding it. An ordinal number is often defined as the order type of a well-ordered set. For example, five is the order type of the set \( \{0, 1, 2, 3, 4\} \).

Cantor visualized extending the set of finite ordinals so that the extended set would be well-ordered. He created a set of transfinite ordinals and gave the first such ordinal the symbol \( \omega \). If the enlarged set of ordinals is to be well-ordered and if there are to be ordinals after \( \omega \), then there must be a "first" ordinal immediately following \( \omega \). He symbolized this next ordinal by \( \omega + 1 \), conceiving the following well-ordered set, where each ordinal gives the order type of the ordered set preceding it.
We are now in a position to consider the paradox which plagued Cantor and caused a number of mathematicians including Bertrand Russell, Zermelo, and Fraenkel to see the need to make Cantorian set theory more rigorous. Cantor believed the infinite set of all ordinals arranged according to increasing magnitude to be well-ordered. Thus it must be possible to create a new ordinal immediately following those in the set under consideration and representing the order type of that class. This is the contradiction: Consider the class of all ordinals, how can there be a new ordinal outside this class, but not part of "the all"?

In the fourth century B.C. Eubulides formulated the famous liar paradox: "What I am now saying is a lie". The paradox occurs because the sentence, if true, must by what it says be false, and if false, must by its statement be true. There is another paradox, formulated by Russell, called the barber paradox: In a certain town the barber shaves all those men and only those men who do not shave themselves. Who shaves the barber? If he is one of those who do not shave themselves, he must be shaved by the barber (himself) and hence he does shave himself, a paradoxical situation. If, on the other hand, it is claimed that he is someone who shaves himself, then he cannot be shaved by the town barber, which contradicts the fact stated by the original statement. These "all" paradoxes provided the motivation for Zermelo and others to provide set theory with a firm
foundation which would exclude dilemmas like those above.

The idea expressed by these men was to restrict the broad definition of a set as any collection whatever. Russell (1908) advocated a theory of types by which individual elements are of type 0, classes of individual elements are of type 1, sets of classes of type 2, etc. Anything of type \( n \) may be an element of a class of type \( n+1 \), so the latter class will contain as members only entities of the same type. Russell explains Cantor's dilemma by saying that a class cannot be a member of itself, it can only be an element of a class of classes.

In 1904, Zermelo proved the well-ordering theorem which states that every set can be well-ordered. But in order to prove it, he assumed what has become known as Zermelo's axiom of choice. It asserts that "for any class of non-overlapping classes, there exists at least one class which contains one representative element from each class of the aggregate of classes."\(^{12}\) Reduced to a finite situation this axiom can be compared to the election of representatives from each district (a set of people) to form a council. Thus each district is a set of individuals and the set of districts become a class of non-overlapping classes.

An objection to Zermelo's axiom of choice arose from the fact that the well-ordering theorem, proved by the axiom of choice implies that the set of real numbers between 0 and 1 inclusive can be well-ordered. No one has been able to exhibit such a well-ordering. Attempts to find fault with Zermelo's proof led to a verification that the well-ordering theorem is equivalent to the axiom of choice.

In 1938, Kurt Gödel proved that Cantor's continuum hypothesis is consistent with the Zermelo-Fraenkel set theory axioms.\(^{13}\) That is, one can take the continuum hypothesis as an additional postulate of set theory. It then remained for Paul J. Cohen in 1963 to prove that the Cantorian hypothesis was independent of the other set theory postulates much as Beltrami exhibited a mo-
model of a pseudosphere (see page 60) to prove that Euclid's fifth postulate is independent of the remaining four. Cohen's method is very technical and subtle. He established a model of non-Cantorian set theory in which the continuum hypothesis is negated while the remaining postulates hold true. Thus we are free to assume or deny as we wish the Cantorian hypothesis.

(3) **Boolean Algebra**

When we say that theorems are deduced from postulates on the basis of the laws of logic, what we are saying is, there must exist a body of rules which determine whether or not a theorem has been properly proved. This leads us to Boolean Algebra.

In 1854, George Boole, a member of the British school of postulational algebraists, published a book of profound originality. He won immediate praise and recognition. The book, *An Investigation of the Laws of Thought, on which are Founded the Mathematical Theories of Logic and Probabilities*, organized the laws of thought into an algebraic framework. What proceeds is a short history of logic followed by an analysis of the Boolean algebra of abstract set theory and its relationship to the algebra of propositions.

Formal logic began with Aristotle who wrote three works dealing with logic. Primarily, he originated the idea of a syllogism which is an argument consisting of two premises and a conclusion, for example:

\[
\begin{align*}
P1: & \text{ All men are mortal} \\
P2: & \text{ All Greeks are men} \\
C: & \text{ Therefore, all Greeks are mortal.}
\end{align*}
\]

Modern logic began with a work entitled *Ars Magna*, written by Ramon Lully (1235-1315). It contained mystical formulas which were completely unintelligible to later logicians. The next work of note was Leibnitz' *De Arte Combinatoria* in 1666. Leibnitz' motivation was similar to that of Lully: an attempt to establish "a general technique by which all reasoning can be reduced to mere calculation." But Leibnitz' formulation was much more capable of being interpreted and developed by logi-
cians who followed. There were others who followed. Giuseppe Peano (1858-1932) devised a complicated symbolic system in which eventually all mathematics was written in terms of his language. His was a monumental five volume work started in 1894. Apparently ignorant of what Leibnitz had done, Boole formulated his ideas in Laws of Thought (1854).

According to Bertrand Russell’s Principia Mathematica symbolic logic consists of three parts (1) the algebra of classes—set theory; (2) the propositional calculus—syllogisms, etc.; and (3) the calculus of relations which uses such statements as "for all", "for some", i.e., quantifiers.

An algebra of sets, $\mathcal{Q}$, is a non-empty collection of subsets of $U$ such that if $A, B \in \mathcal{Q}$, then $A \cup B, A \cap B \in \mathcal{Q}$, and if $A \in \mathcal{Q}$, then $\overline{A} \in \mathcal{Q}$. An example of this algebra of sets is the power set $P(U)$ consisting of the set of all subsets of $U$. In further references a set of sets will be termed a "class".

A Boolean algebra, one model of which is an algebra of sets, is a six-tuple $\langle B, U, \cap, ' , 0, 1 \rangle$, where $B$ is a set, $U$ is a binary operation which we shall call "union" in $B$, $\cap$ is a binary operation we will call "intersection" in $B$, ‘ is a unary operation in $B$, 0, and 1 are distinct elements of $B$, and the following axioms are satisfied:

1. Each binary operation is associative: for all $a, b, c \in B$, $a \cup (b \cup c) = (a \cup b) \cup c$ and $a \cap (b \cap c) = (a \cap b) \cap c$.

2. Each binary operation is commutative: for all $a, b \in B$, $a \cup b = b \cup a$ and $a \cap b = b \cap a$.

3. Each binary operation distributes over the other: for all $a, b, c \in B$, $a \cup (b \cap c) = (a \cup b) \cap (a \cup c)$ and $a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$.

4. For all $a$ in $B$, $a \cup 0 = a$ and $a \cap 1 = a$.

5. For each $a$ in $B$, there exists a unary operation ‘, (called complement of a) such that $a \cup a' = 1$ and $a \cap a' = 0$.

Using suitable models it can be shown that both parts of axioms (2) through (5) above are independent. The associativity...
axiom can be proven using these remaining axioms and Theorem I which follows.

The principle of duality is applicable to Boolean Algebras when the following interchanges are made: $\cap$ for $\cup$, $\cup$ for $\cap$, $0$ for $1$, and $1$ for $0$. It is evident that each axiom is a dual pair of statements. This means that if $T$ is any theorem of Boolean Algebra that the dual of $T$ is also a theorem, the steps in its proof being duals of the steps of the proof of $T$. In short, we obtain two theorems for one proof.

**Theorem I:** In each Boolean algebra, the following hold:

1. The elements $0$ and $1$ are unique.

   **Proof:** Assume $0_1$ and $0_2$ are elements of $B$ such that $a \cup 0_1 = a$ and $a \cup 0_2 = a$ for all $a$.

   Let $a = 0_2$ and then $0_2 \cup 0_1 = 0_2$ also.

   $0_1 \cup 0_2 = 0_1$, but $0_1 \cup 0_2 = 0_2 \cup 0_1$, therefore $0_1 = 0_2$. By duality it follows that $1$ is also unique.

2. Each element has a unique complement.

   **Proof:** Assume that $a_1'$ and $a_2'$ are both complements of $a$. Then

   $a_1' = a_1' \cup 0$

   $= a_1' \cup (a \cap a_2')$

   $= (a_1' \cup a) \cap (a_1' \cup a_2')$

   $= 1 \cap (a_1' \cup a_2')$

   $a_2' = a_1' \cup a_2'$;

   Similarly $a_2' = a_1' \cup a_2'$

   thus $a_1' = a_2'$.

3. For each element $a$, $(a')' = a$.

   **Proof:** By definition $a \cup a' = 1$ and $a \cap a' = 0$.

   But $a' \cup a = 1$ and $a' \cap a = 0$, this implies $(a')' = a$.

4. $0' = 1$.

   **Proof:** $1 = a \cup a'$

   $= 0 \cup 0'$
(5) For each element $a$, $a \cup a = a$ and $a \cap a = a$ (Idempotent Law).
Proof: $a \cup a = (a \cup a) \cap 1$
\[= (a \cup a) \cap (a \cup a')\]
\[= a \cup (a \cap a')\]
\[= a \cup 0\]
\[= a.\]

(6) For each element $a$, $a \cup 1 = 1$ and $a \cap 0 = 0$.
Proof: $1 = (a \cup a')$
\[= (a \cup a) \cup a'\]
\[= a \cup (a \cup a')\]
\[= a \cup 1.\]

(7) For all $a$ and $b$, $a \cup (a \cap b) = a$ and $a \cap (a \cup b) = a$ (Law of Absorption).
Proof: $a = a \cup 0$
\[= a \cup (0 \cap b)\]
\[= (a \cup 0) \cap (a \cup b)\]
\[= a \cap (a \cup b).\]

(8) For all $a$ and $b$, $(a \cup b)' = a' \cap b'$ and $(a \cap b)' = a' \cup b'$ (De Morgan's Laws).
Proof: (1) Let $x \in (a \cup b)'$, then $x \notin a \cup b$.
But this implies $x \notin a$ and $x \notin b$, then $x \in a'$ and $x \in b'$, which implies $x \in a' \cap b'$.

(2) Let $x \in a' \cap b'$, then $x \in a'$ and $x \in b'$. But this implies $x \notin a$ and $x \notin b$ which implies $x \notin a \cup b$ and thus $x \in (a \cup b)'$.
Therefore by (1) and (2) we have $(a \cup b)' = a' \cap b'$.
It is now possible to prove that the associative law is a consequence of axioms (2) through (5) and theorems I(1) through I(8), excluding I(6), and thus is dependent. It will be shown in the following way: Let \( x = a \cup (b \cup c) \) and \( y = (a \cup b) \cup c \). Further it will be shown that (1) \( a \cap x = a \cap y \) and (2) that \( a' \cap x = a' \cap y \) and thus (3) that \( x = y \).

**Proof:**

1. \( a \cap x = a \cap [a \cup (b \cup c)] \)
   \[ = (a \cup a) \cup [a \cap (b \cup c)] \]
   \[ = a \cup [a \cap (b \cup c)] \]
   \[ = a, \text{ by the law of absorption.} \]

2. \( a \cap y = a \cap [(a \cup b) \cup c] \)
   \[ = a \cap (a \cup (a \cap (b \cup c))) \]
   \[ = a \cup (a \cap (b \cup c)) \]
   \[ = a, \text{ by the law of absorption.} \]

3. \( a' \cap x = a' \cap [a \cup (b \cup c)] \)
   \[ = (a' \cap a) \cup a' \cap (b \cup c) \]
   \[ = 0 \cup a' \cap (b \cup c) \]
   \[ = a' \cap (b \cup c). \]

4. \( a' \cap y = a' \cap [(a \cup b) \cup c] \)
   \[ = a' \cap (a \cup (a' \cap (b \cup c))) \]
   \[ = a' \cap (a \cup (a' \cap b) \cup (a' \cap c)) \]
   \[ = a' \cap (b \cup c) \]

(3) Now this implies that

\[ (a \cap x) \cup (a' \cap x) = (a \cap y) \cup (a' \cap y) \]
\[ (a \cup a') \cap x = (a \cup a') \cap y \]
\[ 1 \cap x = 1 \cap y \]
\[ x = y. \]

A model of Boolean algebra other than the algebra of sets may be illustrated by defining \( \langle S, \land, \lor, ', 0, 1 \rangle \) where set \( S \) is the set of all single declarative statements which can be judged to be either true or false; \( \land \) is the conjunction "and" corresponding to intersection; \( \lor \) is the disjunction "and/or" corresponding to union; \( ' \) means negative; 0 and 1 mean false and true respectively. The previous set of axioms together with parts of Theorem I can be arrayed in a format known in logic as
a truth table. This display can be set up to illustrate relationships between these operations. The definitions of this new model will be given in truth table form.

Let a and b be elements of S:

<table>
<thead>
<tr>
<th>Truth Table for Negation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Truth Table for Conjunction</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Truth Table for Disjunction</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>

All of the laws of the theory proven above can be verified in this model by showing that both sides have the same truth tables. An example or two will suffice to illustrate this point. DeMorgan's Law in the theory above, \((a \land b)' = a' \lor b'\), becomes in logic notation \((a \land b)' = a' \lor b'\), as shown in the following two truth tables. Columns are formed using applicable truth table definitions of the respective operations \(\lor, \land\), and negation.

| a || b || a \land b || (a \land b)' |
|----|----|-----------|----------|
| 1  | 1  | 1         | 0        |
| 1  | 0  | 0         | 1        |
| 0  | 1  | 0         | 1        |
| 0  | 0  | 0         | 1        |

| a || b || a' || b' || a' \lor b' |
|----|----|----|----|----------|
| 1  | 1  | 0  | 0  | 0        |
| 1  | 0  | 0  | 1  | 1        |
| 0  | 1  | 1  | 0  | 1        |
| 0  | 0  | 1  | 1  | 1        |

Since the last columns of the two truth tables are exactly the same, the conclusion is that they are logically equivalent forms.

If the last column of a truth table is all ones, the table illustrates what is called in logic a tautology -- a logical
form which is always true regardless of the truth value of its components. For example, the law of excluded middle is symbolically expressed as \( a \lor a' \). (In the preceding model this was axiom (4): \( a \lor a' = 1 \).

\[
\begin{array}{c|c|c}
 a & a' & a \lor a' \\
\hline
 1 & 0 & 1 \\
 0 & 1 & 1 \\
\end{array}
\]

Another example is known in the classical logic of Aristotle and Leibnitz as the law of contradiction, \((a \land a')'\); verbally -- a statement and its negative cannot both be true. Its truth table is given below:

\[
\begin{array}{c|c|c|c}
 a & a' & a \land a' & (a \land a')' \\
\hline
 1 & 0 & 0 & 1 \\
 0 & 1 & 0 & 1 \\
\end{array}
\]

One logical connective which has not been mentioned heretofore is that of the conditional, "if \( a \), then \( b \)", symbolized by \( a \rightarrow b \). It is an unnecessary connective in the strictest sense in that it is possible to define \( \rightarrow \) in terms of \( \lor \) and \( ' \) or in terms of \( \land \) and \( ' \), but verbal expressions using the conditional can be made much simpler. In other words, the last column of the truth table defining "if \( a \), then \( b \)" is the same as the last column of the truth table for \((a \land b')'\) which is in turn the same as the table for \( b \lor a' \). The truth table for \( a \rightarrow b \) is

\[
\begin{array}{c|c|c}
 a & b & a \rightarrow b \\
\hline
 1 & 1 & 1 \\
 1 & 0 & 0 \\
 0 & 1 & 1 \\
 0 & 0 & 1 \\
\end{array}
\]

which is seen to be equivalent to the tables below for the two alternative logical forms.

\[
\begin{array}{c|c|c|c|c|c|c|c}
 a & b & b' & a \land b' & (a \land b')' & a & b & a' & b \lor a' \\
\hline
 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\
 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\
 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\
\end{array}
\]

The simplicity which this new connective introduces to verbal expressions is illustrated by the following. "If it
rains, I will stay home", is logically equivalent to "It is not the case that it is raining and I will not stay home", or to the sentence "I will stay home or it will not rain".

Two important forms related to the conditional may now be examined. These are the contrapositive and the converse. If the initial statement is \( a \rightarrow b \), then the contrapositive is symbolized by \( b' \rightarrow a' \) and the converse is \( b \rightarrow a \). It can be easily shown from truth tables that the contrapositive is always equivalent to the conditional while the converse is not. Consider this example from elementary geometry: If a quadrilateral is a square, then it is a rectangle is true; whereas its converse, if a quadrilateral is a rectangle then it is a square, is not always true.

Of significant importance is the biconditional. It is verbalized as "if and only if", and symbolized

\[ a \leftrightarrow b = (a \rightarrow b) \land (b \rightarrow a) \]

The truth table illustrates:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>( a \rightarrow b )</th>
<th>( b \rightarrow a )</th>
<th>( (a \rightarrow b) \land (b \rightarrow a) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

In 1938 Claude E. Shannon proved that the "switching circuit" algebra is isomorphic with the truth table algebra of propositions. In Shannon's interpretation a closed switch in which current flows is considered in state "1" and an open switch in which no current flows is in state "0". The ideas are illustrated in figure 6 which pictures two switches arranged in series and in parallel. For electricity to flow through switches in series all switches must be closed. On the other hand electricity will flow through switches in parallel if only one switch is closed. Symbolically series is equivalent to \( A \cap B = 1 \) and parallel is equivalent to \( A \cup B = 1 \). Switches in a circuit may operate so that two or more open or close simultaneously. In that case the diagram will represent the switch with the same letter, say A. Further, A and A' represent different switches
which operate so as to be controlled in opposite states. For example, if $A$ is in state "1", then $A'$ will be in state "0".21

![Diagram][1]

**FIG. 6**

The distributive law is illustrated in figure 7. Notice that fewer switches will accomplish the same result.

![Diagram][2]

**FIG. 7**

A complicated circuit can be reduced using the laws of Boolean algebra to $A \cup B$, illustrated in figure 8:

1. $A \cup (A \cup B) \cap C \cup B \cap (C' \cup A')$
   
   which by the distributive law becomes

2. $A \cup (A \cap C) \cup (B \cap C') \cup (B \cap A')$
   
   and by the law of absorption and commuting becomes

3. $A \cup (B \cap C) \cup (C' \cup B) \cup (B \cap A')$
   
   which by the distributive law becomes

4. $A \cup B \cap (C \cup C') \cup (B \cap A')$
   
   and using axioms 4 and 5 this is

5. $A \cup B \cup (B \cap A')$

   and using the law of absorption once again

6. $A \cup B$.

This problem illustrates the power of Boolean algebra to reduce a complicated circuit to a single parallel circuit.
We have examined the realm of the mathematical philosopher—metamathematics and seen how Hilbert's grandiose plan met with failure when scrutinized by Gödel. We have investigated the nature of and the need for mathematical proof. We have seen how modern mathematics can be applied to explain ancient paradoxes. Finally we have studied the development of a universal language of logical thought which attempted to reduce reasoning to a few simple calculations.

Mathematics is indeed ever old, ever new, ever dynamic. Often it is harnessed by man as his servant and just as often it becomes his master when he struggles in vain against its frustrations, its paradoxes, and its unsolved problems.

It is my fervent hope that by studying its history the student will gain an appreciation for mathematics, its beauty, its purpose and be challenged by its frustrations.
FOOTNOTES


2. Ibid., p. 55.


8. The Principia Mathematica by Alfred North Whitehead and Bertrand Russell (1910-1913) is considered by many mathematicians to be the outstanding modern contribution to the algebra of reasoning and a universal scientific language. (Kramer, op. cit., p. 101.)


10. Ibid., chapter 24, pp. 577-589.


13. For a full explanation of Zermelo-Fraenkel set theory see Stoll, op. cit., chapter 7, pp. 89ff.


15. See reference in footnote 3 above.


18. These are simply familiar names given to an abstract concept.

19. The suggestion for proof of the associative property is in Stoll, op. cit., p. 253, problem 2.3. The proof is mine.

20. The axioms and statement of Theorem I are in Ibid., chapter 6, pp. 248-252.

BIBLIOGRAPHY


APPENDIX A
Lesson Plan
The Other Contributions of Pythagoras and the Pythagoreans

Preface: The high school student of mathematics is quite familiar with "The Pythagorean Theorem", for its usefulness is excelled by few, if any, other theorem. Rare indeed is the student who comprehends the magnitude of the other contributions to man's knowledge made by Pythagoras and his followers. It is the purpose of this lesson to partially fill that void. Pythagoras was indeed an intellectual giant and really quite a remarkable man.

Behavioral Objectives
1. The students will be able to list at least five specific contributions to the field of mathematics, not including the famous theorem.
2. Students will be able to list one major contribution to musical theory, and one contribution to astronomy.
3. The student will be able to explain in a short paragraph the Pythagorean Philosophy and how it is related to numbers.

Instructional Materials
1. format--lecture
2. materials
   a. lecture notes
   b. lecture outline to be dittoed and passed out to each student to fill in missing points and ideas from the lecture. The ditto outline would contain only the information which is numbered with Roman numerals and non-parenthesized arabic numerals.

Lecture Outline
I The beginnings
   1. birth of Pythagoras
a. Pythagoras was born sometime between 580 and 568 B.C., possibly on the island of Samos in the Aegean Sea. His teachings reflect many different thoughts from widely dispersed areas—Babylon, India, Egypt, etc. Thus it is speculated that he must have traveled extensively. He also found time to study under the premier mathematician at that time, Thales.

b. Later he migrated to Crotona on the southeastern coast of Italy, called Magna Grecia, and founded a school which attracted well-to-do aristocratic young men. All instruction was done orally since there was a dearth of writing material in Greece at this time. Great emphasis was placed on scientific inquiry and study. Mathematics became the vehicle by which man and his relationship to the universe could be understood.

2. formation of the brotherhood

a. A brotherhood was soon formed among the students in this school. It was based on the principles of secrecy, simplicity of life, purity, abstinence, strict obedience to rules of the order, and scholarly inquiry.

3. life as a Pythagorean in the 5th century B.C.

a. The society which came to be known as the Pythagoreans was divided into groups: the probationers, called listeners, and the more senior members called mathematicians. Rather strange maxims were strictly enforced. For example one was strictly forbidden to eat beans or touch a hot coal with a piece of iron.

b. The ideas of these young aristocrats soon dominated the politics of the time, but ran afoul of the democratic beliefs of the common people.
They led a revolt in 501 B.C. murdering many followers and burning their dwellings. They drove Pythagoras himself away and scattered his followers. The master died of unknown causes soon afterward. The brotherhood thrived, however, and spread throughout the known world.

II Contributions

1. to mathematics
   a. to geometry
      (1) systematized the rules of mathematics into a rigid structure. They actually covered the bulk of material of Euclid's *Elements*, Books 1, 2, 4, 6, and probably much of 3.
      (2) proved that the sum of the angles of a triangle equal two right angles. Euclid's ideas on parallelism must have been known to prove this theorem.
      (3) proved that the plane space about a point was exactly filled by six regular triangles, four squares, or three regular hexagons.
      (4) constructed a polygon equal in area to one of lesser sides.
      (5) constructions of regular polyhedra: pyramids, cubes, tetrahedrons, etc.
   b. to number theory
      (1) figurate numbers, triangle, square, oblong numbers.
      (2) developed formula,

\[ m^2 + \left( \frac{1}{2}(m^2 - 1) \right)^2 = \left( \frac{1}{2}(m^2 + 1) \right)^2 \]

where \( m \) is an odd integer, for generating pythagorean triples. It was later shown that all triples are not generated using the formula.
(3) discovered the irrationality of \( \sqrt{2} \). This created quite a scandal to the extent that the discovery was kept a guarded secret. These numbers could not be expressed as a ratio of two integers, and thus were not commensurable quantities and violated the Pythagorean idea that all magnitudes could be expressed as rational numbers. It remained for Eudoxus about 370 B.C. to re-define proportion and thus put an end to the scandal.

2. to musical theory—harmony
   a. harmony of a stretched string
      (1) a string will vibrate at a certain pitch. If only half is allowed to vibrate the pitch is an octave higher. Allow \( \frac{2}{3} \) of it to vibrate and the pitch will rise a fifth above the one produced by the whole length. For example, if the whole string produces middle C, then half will produce C one octave higher. \( \frac{2}{3} \) will produce G (since five lines and spaces are traversed) and \( \frac{3}{4} \) will yield a pitch which is a fourth higher or F.
      (2) the intellectual appeal of simplicity to the Pythagoreans made the simpler ratios, \( \frac{1}{2} \), \( \frac{2}{3} \), and \( \frac{3}{4} \), come to be regarded as superior to the more complex ones. The octave, the fifth and the fourth are called "perfect consonances".
      (3) fascinatingly enough to the Pythagoreans, the same ratios kept appearing in other matters; the four seasons, two tides, etc.
   b. the "harmony of the spheres"
      (1) The Pythagoreans believed that bodies
moving in spheres produced sounds unheard by ordinary man and a higher pitch was emitted by bodies traveling at a higher speed.

(2) They also supposed the distances between the planets and the ratios of the speeds of these planets relative to the earth to be harmonically determined.

3. to astronomy
   a. earth and universe were spherical in shape.
   b. sun, moon, and planets have a movement of their own independent of daily rotation. They considered that the planets moved in spherical orbits.

III Philosophy
1. based on numbers
   a. observed in numbers many resemblances to "things that exist and are coming into being."
   b. numbers were elements of all things.

2. theory of opposites
   a. assumed as orderly world.
   b. opposites attract.

   (1) life characterized by change, but change after an orderly pattern in which contrary forces by their interaction produced all the variety of the visible world.

   c. set up a table of ten opposites
      (1) limited/unlimited (6) at rest/in motion
      (2) odd/even (7) straight/curved
      (3) unity/plurality (8) light/darkness
      (4) right/left (9) good/evil
      (5) male/female (10) square/oblong

3. relationship to numbers
   a. correspondence between all nature and numbers
      (1) number is the element "in" things.
(2) justice and soul are modifications of number; justice is a number in its evenness and its oddness.

(3) musical consonances governed by number.

(4) the decade (10 years) had a mystical significance, thus they asserted that the planets also numbered ten, but since there were only nine visible, they created the "counter-earth" which was always unseen as the tenth.
REFERENCES


APPENDIX B

The Fundamental Theorem of Algebra

Albert Girard (1595-1632) is generally regarded as the first to formulate the Fundamental Theorem of Algebra in 1629. It states that every algebraic equation of degree n with complex coefficients has exactly n roots, including complex roots. Prior to this, in 1600, Peter Rothe, a Nuremberg mathematician, stated it as "equations have at most as many roots as their degree indicates". Rothe's statement made no mention of the effect of coefficients. In 1637, Descartes, taking complex roots into account, admitted that an equation can have as many roots as the degree indicates.

The first attempt at proof of this theorem was made by D'Alembert in 1748. His proof was hardly convincing as he did not prove the existence of a root, but merely showed the form it must take. The next attempt was made by Euler in 1749. His proof stood as the best attempt until Gauss in 1799, in his doctoral dissertation, pointed out the discrepancies. He then proceeded to demonstrate in a rigorous fashion that every rational algebraic function of one variable can be decomposed into real factors of the first or second degree. Gauss's proof is extremely long and very complicated. Using De-Moivre's Theorem, however, a concise and somewhat topological proof to the theorem can be given. The statement of the theorem as it will be proven states that every polynomial equation

\[ f(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1z + a_0 = 0, \]

with complex coefficients, has at least one complex root, \( z = x + yi \), such that \( f(z) = 0 \).

Proof: (1) Preliminary discussion

(a) Let \( z = x + yi = r(\cos \theta + i \sin \theta) \), where \( r = \sqrt{x^2 + y^2} \).

(b) Now by De-Moivre's Theorem

\[ z^n = r^n(\cos n\theta + i \sin n\theta). \]

(c) Allow \( z \) to generate a circle of radius \( r \) about the origin, then \( z^n \) will generate \( n \) times, a concentric circle of
radius $r^n$ as $z$ describes its circle once. This follows directly from De Moivre's formula for integral values of $n$.

(d) Now $|z| = \text{the distance of point } z \text{ from } 0 \text{ and } |z| = r$.

(e) $|z - z'| = \text{distance between } z \text{ and } z' \text{ for } z' \text{ any complex number, } x' + iy'$.

(2) Proof: (a) Assume to the contrary, of that which we wish to prove, that the polynomial $f(z) = 0$ has no root, so that for all $z$, $f(z) \neq 0$.

(b) Allow $z$ to describe a closed curve $C$ in the $(x,y)$ plane and $f(z)$ will describe a closed curve $\Gamma$ which never passes through the origin, point $0$; since by original assumption $f(z) \neq 0$. (See Figure 1).

(c) Define the order of origin $0$ with respect to $f(z)$ for $C$ as "the net number of complete revolutions made by an arrow joining point $0$ to a point on $\Gamma = f(z)$ traced by $f(z)$ as $z$ traces curve $C$." Now $f(z)$ and $z$ will be joined by this arrow.

(d) Let the radius of circle $C$ be $t$.

(e) Define $\phi(t)$ as the order of point $0$ (definition above) with respect to $f(z)$ for circle $C$. 

\[ \text{Fig. 1} \]
(f) Now \( \phi(0) = 0 \), since if \( t = 0 \), we have a point circle and \( \Gamma \) becomes the point \( f(0) \neq 0 \).

(g) The order of \( \phi(t) \) depends continuously on \( t \) since \( f(z) \) is continuous (all polynomials are continuous functions). Now \( \phi(t) \) can only assume integral values since it represents the total number of complete revolutions.

(h) If we can show that \( \phi(t) = n \) for large \( t \), then we will have a contradiction for \( \phi(t) \) cannot pass continuously but only in discrete steps from 0 to \( n \).

(i) Choose \( t \) such that \( t > 1 \) and also that
\[
t > |a_0| + |a_1| + \cdots + |a_{n-1}|
\]

(j) Then for \( z \) on the circle \( C \) of radius \( t \), we have
\[
|f(z) - z^n| = \text{distance between } z^n \text{ and } f(z).
\]
\[
= |a_{n-1}z^{n-1} + \cdots + a_0|
\]
\[
\leq |a_{n-1}| |z|^{n-1} + |a_{n-2}| |z|^{n-2} + \cdots + |a_0|
\]
\[
= |a_{n-1}| t^{n-1} + |a_{n-2}| t^{n-2} + \cdots + |a_0|
\]
\[
= t^{n-1} \left( |a_{n-1}| + |a_{n-2}| + \cdots + \frac{|a_0|}{t^{n-1}} \right)
\]
\[
\leq t^{n-1} \left( |a_{n-1}| + |a_{n-2}| + \cdots + |a_0| \right)
\]
\[
< t^n
\]
\[
= |z^n|.
\]

Since the length of the arrow joining \( f(z) \) and \( z^n \) is always less than \( |z^n| \) (the radius of the circle formed by \( z^n \)), the two points are always separated from each other by a distance less than that from point \( 0 \) to point \( z^n \). This means that the curve traced out by \( f(z) \) may be continuously deformed, without ever passing through the origin, into the curve traced out by \( z^n \) by pushing each point \( f(z) \) along the arrow joining it to \( z^n \). Therefore the points \( f(z) \) and \( z^n \) must make the same number of
revolutions about point 0. Now the number of revolutions made by $z^n$ is $n$. Thus $\phi(t) = n$ for large $t$, and the proof is complete.\footnote{Richard Courant and Herbert Robbins, \textit{What is Mathematics?} (New York: Oxford University Press, 1941), pp. 269-271.}
APPENDIX C

A Proof
of
Gödel's Incompleteness Theorem
and
Inconsistency Corollary

Before an outline of Gödel's proof can be considered, one must examine what is known as Richard's Paradox, formulated by a Frenchman, Jules Richard in 1903. Suppose we list and number in order some properties of natural numbers, such as "even", "prime", etc. They are ordered by assigning to position one the definition with the fewest number of letters; to position two the definition which is next shortest; and so on. If two definitions have the same number of letters we order them alphabetically.

Now, perhaps a particular order number may possess the characteristic described in the definition to which it corresponds. For example, order number nineteen may correspond to the property "odd" and by contrast order number twelve may correspond to the property of being "divisible by nine" and certainly twelve does not have that property. Thus one would say that an order number like twelve which does not have the property being described is called Richardian, whereas an order number like nineteen, then, is not Richardian. Now the property of being Richardian can be assigned as order number, say n, in our sequence of definitions. Now the challenge: Is n Richardian? If yes, then n cannot possess the property to which it corresponds (i.e., being Richardian), hence it must not be Richardian. On the other hand, if the answer is no, then n is not Richardian, hence it must have the property to which it corresponds (being Richardian)--a paradox!

The significance of this paradox is disputed. Some mathematicians claim that "Richardian-ness" is not a mathematical
property of the natural numbers. Rather, they insist, it is a metamathematical property, hence does not belong on a list of mathematical properties.\textsuperscript{1}

Now what did Gödel do? He did not fall into the Richardian trap because he translated all metamathematical statements into arithmetic formulas which can be rightfully placed in such a list. Gödel constructed a formal system for elementary number theory, then assigned a different natural number (called a Gödel number) to each formula, proof, postulate, and theorem of his system. Gödel's numbering system maps the elements of his formalized number theory onto a proper subset of the natural numbers. Then, given a formula or a proof, its Gödel number can be found. Also, given a natural number one can determine whether or not it is a Gödel number. If it is a Gödel number it can be factored uniquely into primes (by the Fundamental Theorem of Arithmetic proven in Chapter II). This is called reconstituting the formula. Finally, metamathematical statements are mapped onto statements about ordinary arithmetic relations. Then these statements are represented by arithmetic symbols (numbers).

What follows will be an outline, non-rigorous and informal, of the main argument in his proof of the incompleteness theorem.

Gödel's procedure is this: A number is "troublesome" if and only if it cannot be proved to have the property to which it corresponds. Now Gödel considers only those formulas of his system which are open sentences in a single variable, \(x\), where the domain of \(x\) is the set of Gödel numbers. Those formulas can be ordered according to their increasing magnitude. This sequence is divided into two sets, \(T\) and \(T'\), depending on whether the elements are troublesome or non-troublesome Gödel numbers.

If the substitution of a Gödel number, \(k\), for \(x\) in the open sentence whose order number is \(k\) results in a formula which is unprovable, then it belongs in set \(T\). If the formula is provable, then \(k\) belongs to \(T'\).

The content of one of his open sentences would appear like this:
Now we substitute $n$ for $x$ in the open sentence to produce a proposition which is

$$G = n \text{ is a member of } T.$$  

Thus $G$ is a proposition whose formula is undecidable within the system. If we assumed $G$ were provable then it would be "true" and by its content reveal that $n$ is troublesome, thus the proposition $G$ is unprovable. If we assume $G$ is provable we are led to the conclusion that $G$ is unprovable, a contradiction! Next, suppose

$$\neg G = n \text{ is not a member of } T$$

is provable, hence "true". Then $n$ would belong to $T'$ and be a non-troublesome Gödel number. If we substitute it in the formula to which it corresponds, we find that $G$ is provable. Thus again we see that the assumption that "not $G$" is provable, leads to the conclusion that $G$ is provable. But this is impossible in a consistent system because a proposition $G$ and its negation $\neg G$ cannot both be provable.

In conclusion, the argument thus far shows that if a formal system is consistent neither $G$ nor $\neg G$ is provable. Hence $G$ is undecidable and the system is incomplete.

Arguing further, if neither $G$ nor $\neg G$ is provable, add either $G$ or $\neg G$ as a postulate, call it $G_0$, then by Gödel's method a new undecidable formula can be developed. We can add that formula or its negation, call it $G_1$, and we are led to the same argument. But we cannot ever, in a finite number of postulates, make the system complete. Some undecidable formula $F$ remains.

Finally, it must be shown that the corollary demonstrating inconsistency is true. Gödel did this by mapping the statement "the system is consistent" onto a formula $C$ within the system. Then he demonstrated that

$$F \text{ is not provable implies } C \text{ is not provable.}$$
But the statement "F is not provable" is true, since F is undecidable, thus the statement "C is not provable" is also true. We assumed the present system is consistent, but its consistency cannot be proven within the system.\(^2\)

---


APPENDIX D

Euclid's Proof of the Pythagorean Theorem

The oldest known written proof of the Pythagorean Theorem is believed attributable to Euclid himself in The Elements.

Given: Right triangle ABC, right angle at C

Prove: Area of sq. AHIB = Area sq. AGFC + Area sq. BDEC

Proof:
(1) BCED, AGFC, and AHIB are squares drawn on sides BC, AC, and AB respectively.
(2) Draw CX parallel to AH intersecting AB and HI in J and X respectively. Draw, also, BG and CH.
(3) $\angle HAC \cong \angle BAG$ since they are equal to $\angle BAC + \text{right } \angle s \triangle BAC$ and CAG.
(4) $AB \cong AH$ and $AC \cong AG$.
(5) $\triangle HAC \cong \triangle BAG$ by (SAS).
(6) Area of sq. AGFC = $2(Area \triangle BAG)$, since AC is the altitude...
and AG is the base of both.

(7) Area of rectangle AHXJ = 2(Area \angle HAC) for a similar reason as for (6).

(8) Area square AGFC = Area rectangle AHXJ.

(9) Draw AD and CI and prove in a similar manner that \angle CBI = \angle ABD, thus the Area of square BCED = Area of rectangle BIXJ.

(10) Area of square ABIH = Area of rectangle AHXJ + Area of rectangle BIXJ.

(11) Therefore Area square ABIH = Area square BCED + Area of square AGFC.
APPENDIX E

Babylonian Texts

The following two problems taken from Babylonian cuneiform texts are from sources indicated in footnote number 13, Chapter II. They illustrate the geometrical knowledge of these ancient people.

The first, Problem C, shows a tablet on which is drawn a square whose side is given as thirty units. The remainder of the wedges indicate the length of the diagonal. The second, Problem D, shows how the Babylonians would use their knowledge of the Pythagorean Theorem to find the width of a rectangle given its length and its diagonal. (See also Chapter II, pages 98 and 99.)
Problem D

One diagonal, 20 the area. I square the side. Multiply the square by the side, (again) multiply by the diagonal.

14.48.53.20. What are the length, the width and the diagonal? By you. Square 20 the area, 6.40 you get. This you retain. Again 14.48.53.20 square: 3.39.2(8.43.27.2h).26.40 you get.

(Take) half of 6.40, which had been retained, fraction (it): 3.20 you get. Square 3.20: 11.6.40 you get; 11.6.40 added to 3.39.2(8.43.27.2h).26.40 you get: 3.50.36.43.34.26.40. you get. What is the root of the square? 15.11.6.40 is the root of the square.

Subtract 3.20, the half of the square, from 15.11.6.40: 11.51.6.40 you get. What is the root of the square? 26.40 is the root of the square. Find the inverse of 26.40: 2.15 you get. Multiply 2.15 by 6.40 of the square which had been kept: (15) you get. What is the root of the square? 30 is the root of the square. 30 is the width.