Linear Algebra

A First Course with Applications to Differential Equations

Tom M. Apostol

Linear Algebra

Tom M. Apostol

Linear Algebra

A First Course, with Applications to Differential Equations

A Wiley-Interscience Publication
JOHN WILEY & SONS, INC.
New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

A NOTE TO THE READER
This book has been electronically reproduced from digital information stored at John Wiley & Sons, Inc. We are pleased that the use of this new technology will enable us to keep works of enduring scholarly value in print as long as there is a reasonable demand for them. The content of this book is identical to previous printings.

This text is printed on acid-free paper.

Library of Congress Cataloging in Publication Data:

Apostol, Tom M.
Linear algebra : a first course, with applications to differential equations / Tom M. Apostol.
p. cm.
"A Wiley-Interscience publication."
ISBN 0-471-17421-1 (cloth : acid-free paper)
1. Algebra, Linear. 2. Differential equations. I. Title.
QA184.A66 1997
512'.5—dc21 96-37131
CIP

Printed in the United States of America

10 9 8 7 6 5 4

To
Erica, Emily, and Caitlin Jane

CONTENTS

1. VECTOR ALGEBRA

2. APPLICATIONS OF VECTOR ALGEBRA TO ANALYTIC GEOMETRY

3. LINEAR SPACES

4. LINEAR TRANSFORMATIONS AND MATRICES

5. DETERMINANTS

6. EIGENVALUES AND EIGENVECTORS

7. EIGENVALUES OF OPERATORS ACTING ON EUCLIDEAN SPACES

8. APPLICATIONS TO LINEAR DIFFERENTIAL EQUATIONS

9. APPLICATIONS TO SYSTEMS OF DIFFERENTIAL EQUATIONS

10. THE METHOD OF SUCCESSIVE APPROXIMATIONS

PREFACE

For many years the author has been urged to develop a text on linear algebra based on material in the second edition of his two-volume *Calculus*, which presents calculus of functions of one or more variables, integrated with differential equations, infinite series, linear algebra, probability, and numerical analysis. To some extent this was done by others when the two *Calculus* volumes were translated into Italian and divided into three volumes,* the second of which contained the material on linear algebra. The present text is designed to be independent of the *Calculus* volumes.

To accommodate a variety of backgrounds and interests, this text begins with a review of prerequisites (Chapter 0). The review is divided into two parts: pre-calculus prerequisites, needed to understand the material in Chapters 1 through 7, and calculus prerequisites, needed for Chapters 8 through 10. Chapters 1 and 2 introduce vector algebra in n-space with applications to analytic geometry. These two chapters provide motivation and concrete examples to illustrate the more abstract treatment of linear algebra presented in Chapters 3 through 7.

Chapter 3 discusses linear spaces, subspaces, linear independence, bases and dimension, inner products, orthogonality, and the Gram-Schmidt process. Chapter 4 introduces linear transformations and matrices, with applications to systems of linear equations. Chapter 5 is devoted to determinants, which are introduced axiomatically through their properties. The treatment is somewhat simpler than that given in the author's *Calculus*. Chapter 6 treats eigenvalues and eigenvectors, and includes the triangularization theorem, which is used to deduce the Cayley-Hamilton theorem. There is also a brief section on the Jordan normal form. Chapter 7 continues the discussion of eigenvalues and eigenvectors in the setting of Euclidean spaces, with applications to quadratic forms and conic sections.

In Chapters 3 through 7, calculus concepts occur only occasionally in some illustrative examples, or in some of the exercises; these are clearly identified and can be omitted or postponed without disrupting the continuity of the text. This part of the text is suitable for a first course in linear algebra not requiring a calculus prerequisite. However, the level of presentation is more appropriate for readers who have acquired some degree of mathematical sophistication in a course such as elementary calculus or finite mathematics.

Chapters 8, 9, and 10 definitely require a calculus background. Chapter 8 applies linear algebra concepts to linear differential equations of order n, with special emphasis on

**Calcolo*, Volume primo: Analisi 1; Volume Secondo: Geometria; Volume Terzo: Analisi 2. Published by Editore Boringhieri, 1977.

equations with constant coefficients. Chapter 9 uses matrix calculus to discuss systems of differential equations. This chapter focuses on the exponential matrix, whose properties are derived by an interplay between linear algebra and matrix calculus. Chapter 10 treats existence and uniqueness theorems for systems of differential equations, using Picard's method of successive approximations, which is also cast in the language of contraction operators.

Although most of the material in this book was extracted from the author's *Calculus*, some topics have been revised or rearranged, and some new material and new exercises have been added.

This textbook can be used by first- or second-year students in college, and it can also be of interest to more mature individuals, who may have studied mathematics many years ago without learning linear algebra, and who now wish to learn the basic concepts without undue emphasis on abstraction or formalization.

TOM M. APOSTOL

California Institute of Technology

Linear Algebra

0

REVIEW OF PREREQUISITES

Part 1 of this chapter summarizes some pre-calculus prerequisites for this book—facts about real numbers, rectangular coordinates, complex numbers, and mathematical induction. Part 2 does the same for calculus prerequisites. Chapters 1 and 2, which deal with vector algebra and its applications to analytic geometry, do not require calculus as a prerequisite. These two chapters provide motivation and concrete examples to illustrate the abstract treatment of linear algebra that begins with Chapter 3. In Chapters 3 through 7, calculus concepts occur only occasionally in some illustrative examples, or in some exercises; these are clearly identified and can be omitted or postponed without disrupting the continuity of the text.

Although calculus and linear algebra are independent subjects, some of the most striking applications of linear algebra involve calculus concepts—integrals, derivatives, and infinite series. Familiarity with one-variable calculus is essential to understand these applications, especially those referring to differential equations presented in the last three chapters. At the same time, the use of linear algebra places some aspects of differential equations in a natural setting and helps increase understanding.

Part 1. Pre-calculus Prerequisites

0.1 Real numbers as points on a line

Real numbers can be represented geometrically as points on a straight line. A point is selected to represent 0 and another, to the right of 0, to represent 1, as illustrated in Figure 0.1. This choice determines the scale, or unit of measure. If one adopts an appropriate set of axioms for Euclidean geometry, then each real number corresponds to exactly one point on this line and, conversely, each point on the line corresponds to one and only one real

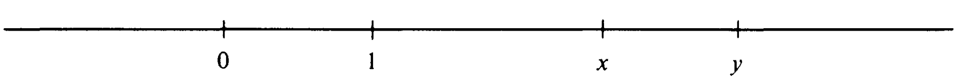

FIGURE 0.1 Real numbers represented geometrically on a line.

number. For this reason, the line is usually called the *real line* or the *real axis*. We often speak of the *point* x rather than the point corresponding to the real number x. The set of all real numbers is denoted by **R**.

If $x < y$, point x lies to the left of y as shown in Figure 0.1. Each positive real number x lies at a distance x to the right of zero. A negative real number x is represented by a point located at a distance $|x|$ to the left of zero.

0.2 Pairs of real numbers as points in a plane

Points in a plane can be represented by *pairs* of real numbers. Two perpendicular reference lines in the plane are chosen, a horizontal x axis and a vertical y axis. Their point of intersection, denoted by 0, is called the *origin*. On the x axis a convenient point is chosen to the right of 0 to represent 1; its distance from 0 is called the *unit distance*. Vertical distances along the y axis are usually measured with the same unit distance. Each point in the plane is assigned a pair of numbers, called its *coordinates*, which tell us how to locate the point. Figure 0.2 illustrates some examples. The point with coordinates $(3, 2)$ lies three units to the right of the y axis and two units above the x axis. The number 3 is called the *x coordinate* or *abscissa* of the point, and 2 is its *y coordinate* or *ordinate*. Points to the left of the y axis have a negative abscissa; those below the x axis have a negative ordinate. The coordinates of a point, as just defined, are called its *Cartesian coordinates* in honor of René Descartes (1596–1650), one of the founders of analytic geometry.

When a pair of numbers is used to represent a point, we agree that the abscissa is written first, the ordinate second. For this reason, the pair (a, b) is referred to as an *ordered pair*: the first entry is a, the second is b. Two ordered pairs (a, b) and (c, d) represent the same point if and only if we have $a = c$ and $b = d$. Points (a, b) with both a and b positive are said to lie in the *first quadrant*; those with $a < 0$ and $b > 0$ are in the *second quadrant*; those with $a < 0$ and $b < 0$ are in the *third quadrant*; and those with $a > 0$ and $b < 0$ are in the *fourth quadrant*. Figure 0.2 shows one point in each quadrant.

The procedure for locating points in space is analogous. We take three mutually perpendicular lines in space intersecting at a point (the origin). These lines determine three

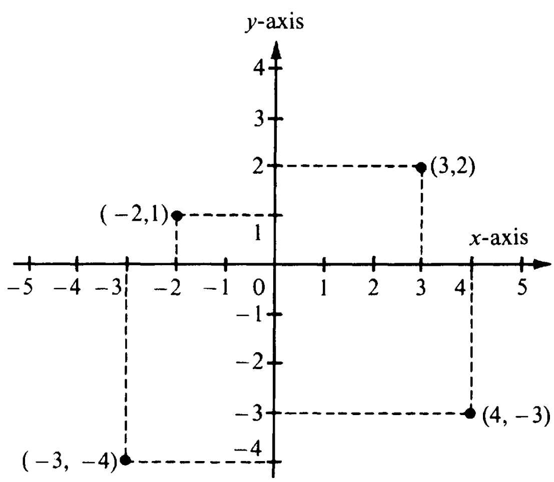

FIGURE 0.2 Points in the plane represented by pairs of real numbers.

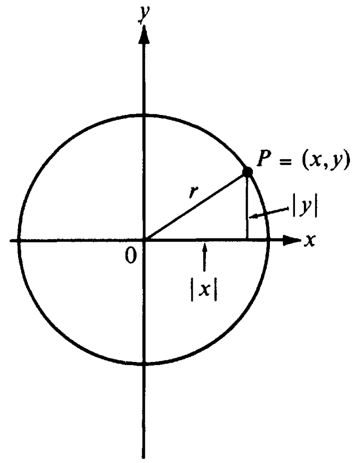

FIGURE 0.3 The circle represented by the Cartesian equation $x^2 + y^2 = r$.

mutually perpendicular planes, and each point in space can be completely described by specifying, with appropriate regard for signs, the distances from these planes. We shall discuss three-dimensional Cartesian coordinates in a later chapter; for the present we confine our attention to the two-dimensional case.

A geometric figure, such as a curve in the plane, is a collection of points satisfying one or more special conditions. By expressing these conditions in terms of the coordinates x and y we obtain one or more relations (equations or inequalitites) that characterize the figure in question. For example, consider a circle of radius r with its center at the origin, as shown in Figure 0.3.

Let (x, y) denote the coordinates of an arbitrary point P on this circle. The line segment OP is the hypotenuse of a right triangle whose legs have lengths $|x|$ and $|y|$ and, hence, by the theorem of Pythagoras, we have

$$x^2 + y^2 = r^2.$$

This equation, called a *Cartesian equation* of the circle, is satisfied by all points (x, y) on the circle and by no others, so the equation completely characterizes the circle. Points *inside* the circle satisfy the inequality $x^2 + y^2 < r^2$, while those *outside* satisfy $x^2 + y^2 > r^2$. This example illustrates how analytic geometry is used to reduce geometrical statements about points to algebraic relations about real numbers.

0.3 Polar coordinates

Points in a plane can also be located by using polar coordinates. This is done as follows. Let P be a point distinct from the origin. Suppose the line segment joining the origin to P has length $r > 0$ and makes an angle of θ radians with the positive x axis, as shown by the example in Figure 0.4. The two numbers r and θ are called *polar coordinates* of P. They are related to the rectangular coordinates x and y by the equations

(0.1) $$x = r\cos\theta, \qquad y = r\sin\theta.$$

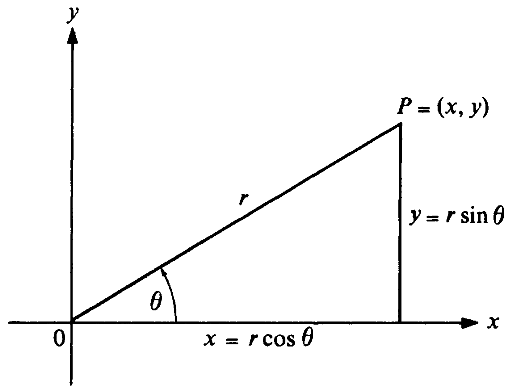

FIGURE 0.4 Polar coordinates.

The positive number r is called the *radial distance* of P, and θ is called *a polar angle.* We say *a* polar angle rather than *the* polar angle because if θ satisfies (0.1) so does $\theta + 2n\pi$ for any integer n. We agree to call all pairs of real numbers (r, θ) polar coordinates of P if they satisfy (0.1) with $r > 0$.

The radial distance r is uniquely determined by x and y: $r = \sqrt{x^2 + y^2}$, but the polar angle θ is determined only up to integer multiples of 2π.

When P is the origin, Eqs. (0.1) are satisfied with $r = 0$ and any θ. For this reason, we assign the radial distance $r = 0$ to the origin, and we agree that *any* real θ may be used as a polar angle.

Some curves are described more simply with polar coordinates rather than rectangular coordinates. For example, a circle of radius 2 with center at the origin has Cartesian equation $x^2 + y^2 = 4$. In polar coordinates the same circle is described by the simpler equation $r = 2$. The interior of the circle is described by the inequality $r < 2$, the exterior by $r > 2$.

0.4 Complex numbers

The quadratic equation $x^2 + 1 = 0$ has no solution in the real-number system because there is no real number whose square is negative. New types of numbers, called *complex numbers*, have been introduced to provide solutions to such equations.

As early as the 16th century, a symbol $\sqrt{-1}$ was introduced to provide solutions of the quadratic equation $x^2 + 1 = 0$. This symbol, later denoted by the letter i, was regarded as a fictitious or imaginary number, which could be manipulated algebraically like an ordinary real number, except that its square was -1. Thus, for example, the quadratic polynomial $x^2 + 1$ was factored by writing

$$x^2 + 1 = x^2 - i^2 = (x - i)(x + i),$$

and the solutions of the equation $x^2 + 1 = 0$ were exhibited as $x = \pm i$, without any concern regarding the meaning or validity of such formulas. Expressions such as $2 + 3i$ were called complex numbers, and they were used in a purely formal way for nearly 300 years before they were described in a manner that would be considered satisfactory by present-day standards.

Early in the 19th century, Carl Friedrich Gauss (1777–1855) and William Rowan Hamilton (1805–1865) independently and almost simultaneously proposed the idea of defining complex numbers as ordered pairs of real numbers (a, b) endowed with certain special properties. This idea is widely accepted today and is described in the next section.

0.5 Definition and algebraic properties of complex numbers

Complex numbers are defined as ordered pairs of real numbers, in the same way that we described the rectangular coordinates of points in the plane. The new feature is that we also define addition and multiplication so that we can perform algebraic operations on complex numbers.

DEFINITION. *If a and b are real numbers, the pair (a, b) is called a complex number, provided that equality, addition, and multiplication of pairs is defined as follows:*
(a) *Equality: $(a, b) = (c, d)$ means $a = c$ and $b = d$.*
(b) *Sum: $(a, b) + (c, d) = (a + c, b + d)$.*
(c) *Product: $(a, b)(c, d) = (ac - bd, ad + bc)$.*

The definition of equality states that (a, b) is to be regarded as an *ordered pair*. Thus, the complex number $(2, 3)$ is distinct from the complex number $(3, 2)$. The numbers a and b are called *components* of the complex number. The first component, a, is also called the *real part* of the complex number; the second component, b, is called the *imaginary part*.

Note that the symbol $\sqrt{-1}$ does not appear anywhere in this definition. Presently we shall introduce i as a particular complex number that has all the algebraic properties ascribed to the fictitious symbol $\sqrt{-1}$ introduced by the early mathematicians. However, before we do this we discuss basic properties of the operations just defined.

THEOREM 0.1. *Addition and multiplication of complex numbers satisfy the commutative, associative and distributive laws. That is, if x, y, and z are arbitrary complex numbers we have the following properties:*
Commutative laws: $x + y = y + x$, $\quad xy = yx$.
Associative laws: $x + (y + z) = (x + y) + z$, $\quad x(yz) = (xy)z$.
Distributive law: $x(y + z) = xy + xz$.

Proof. All these laws are easily verified directly from the definition of sum and product. For example, to prove the associative law for multiplication, we express x, y, z in terms of their components, say $x = (x_1, x_2)$, $y = (y_1, y_2)$, $z = (z_1, z_2)$ and note that

$$\begin{aligned} x(yz) &= (x_1, x_2)(y_1 z_1 - y_2 z_2, y_1 z_2 + y_2 z_1) \\ &= \big(x_1(y_1 z_1 - y_2 z_2) - x_2(y_1 z_2 + y_2 z_1), x_1(y_1 z_2 + y_2 z_1) + x_2(y_1 z_1 - y_2 z_2)\big) \\ &= \big((x_1 y_1 - x_2 y_2) z_1 - (x_1 y_2 + x_2 y_1) z_2, (x_1 y_2 + x_2 y_1) z_1 + (x_1 y_1 - x_2 y_2) z_2\big) \\ &= (x_1 y_1 - x_2 y_2, x_1 y_2 + x_2 y_1)(z_1, z_2) = (xy)z. \end{aligned}$$

The commutative and distributive laws may be similarly proved.

Further algebraic concepts, such as *zero*, *negative*, *reciprocal*, and *quotient*, analogous to those for real numbers, are defined as follows:

The complex number $(0, 0)$ is called the *zero* complex number. It is an identity element for addition because $(0, 0) + (a, b) = (a, b)$ for all complex numbers (a, b). Similarly, the complex number $(1, 0)$ is an identity for multiplication because

$$(a, b)(1, 0) = (a, b)$$

for all (a, b).

Since $(-a, -b) + (a, b) = (0, 0)$ we call the complex number $(-a, -b)$ the *negative* of (a, b) and we write $-(a, b)$ for $(-a, -b)$.

The *difference* $(a, b) - (c, d)$ of two complex numbers is defined to be the sum of (a, b) and the negative of (c, d).

Each nonzero complex number (a, b) has a *reciprocal* relative to the identity element $(1, 0)$, which we denote by $(a, b)^{-1}$. It is given by the ordered pair

(0.2) $$(a, b)^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right) \qquad \text{if } (a, b) \neq (0, 0),$$

and it has the property that $(a, b)(a, b)^{-1} = (1, 0)$. Note that $a^2 + b^2 \neq 0$ because $(a, b) \neq (0, 0)$.

The *quotient* $(a, b)/(c, d)$ of two complex numbers with $(c, d) \neq (0, 0)$ is defined to be the product $(a, b)(c, d)^{-1}$.

0.6 Complex numbers as an extension of real numbers

Let $\mathbf{C}$ denote the set of all complex numbers. Consider the subset $\mathbf{C}_0$ of $\mathbf{C}$ consisting of all complex numbers of the form $(a, 0)$, that is, all complex numbers with zero imaginary part. The sum or product of two members of $\mathbf{C}_0$ is again in $\mathbf{C}_0$. In fact we have

$$(a, 0) + (b, 0) = (a + b, 0) \qquad \text{and} \qquad (a, 0)(b, 0) = (ab, 0).$$

This shows that we can add or multiply two numbers in $\mathbf{C}_0$ by adding or multiplying the real parts alone. Or, in other words, with respect to addition and multiplication, the numbers in $\mathbf{C}_0$ act exactly as though they were real numbers. The same is true for subtraction and division because $-(a, 0) = (-a, 0)$, and $(b, 0)^{-1} = (b^{-1}, 0)$ if $b \neq 0$. For this reason, we make no distinction between the real number x and the complex number $(x, 0)$ whose real part is x. We agree to identify x and $(x, 0)$ and we write $x = (x, 0)$. In particular, we write $0 = (0, 0)$, $1 = (1, 0)$, $-1 = (-1, 0)$, and so on. Thus, we can regard the complex number system as an extension of the real number system.

This also makes sense geometrically. In a later section we will represent the complex number (x, y) by a point in the plane with Cartesian coordinates x and y; the subset $\mathbf{C}_0$ is represented geometrically by the points on the x axis.

0.7 The imaginary unit i

Complex numbers have some algebraic properties not possessed by real numbers. For example, the quadratic equation $x^2 + 1 = 0$, which has no solution among the real numbers, can now be solved with the use of complex numbers. In fact, the complex number $(0, 1)$ is a solution, because we have

$$(0, 1)^2 = (0, 1)(0, 1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0) = -1.$$

DEFINITION. *The complex number* $(0, 1)$ *is denoted by the symbol* i *and is called the imaginary unit.*

The imaginary unit has the property that its square is -1, $i^2 = -1$. Therefore the quadratic equation $x^2 + 1 = 0$ has the solution $x = i$. The reader can easily verify that $x = -i$ is another solution.

Now we can relate the ordered-pair idea with the notation used by the early mathematicians. First we note that the definition of multiplication gives us $(b, 0)(0, 1) = (0, b)$, and hence we have

$$(a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0)(0, 1).$$

Therefore if we write $a = (a, 0)$, $b = (b, 0)$, and $i = (0, 1)$, we get $(a, b) = a + bi$. In other words, we have proved the following:

THEOREM 0.2. *Every complex number* (a, b) *can be expressed in the form* $(a, b) = a + bi$.

This notation aids us in calculations involving addition and multiplication. For example, to multiply $a + bi$ by $c + di$, use the distributive and associative laws, and replace i^2 by -1. Thus,

$$(a + bi)(c + di) = ac - bd + (ad + bc)i,$$

which, of course, agrees with the definition of multiplication. Similarly, to compute the reciprocal of a nonzero complex number $a + bi$ we write

$$\frac{1}{a + bi} = \frac{a - bi}{(a + bi)(a - bi)} = \frac{a - bi}{a^2 + b^2} = \frac{a}{a^2 + b^2} - \frac{bi}{a^2 + b^2}.$$

This formula agrees with that given in (0.2).

With complex numbers we can solve not only the simple quadratic equation $x^2 + 1 = 0$, but also the more general quadratic equation $ax^2 + bx + c = 0$, where a, b, c are real and $a \neq 0$. By completing the square, we can write this quadratic equation in the form

$$\left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a^2} = 0.$$

If $4ac - b^2 \leq 0$, the equation has the real roots $(-b \pm \sqrt{b^2 - 4ac})/(2a)$. If $4ac - b^2 > 0$, the left member is positive for every real x and the equation has no real roots. In this case, however, there are two complex roots, given by the formulas

$$(0.3) \qquad r_1 = -\frac{b}{2a} + i\frac{\sqrt{4ac - b^2}}{2a} \quad \text{and} \quad r_2 = -\frac{b}{2a} - i\frac{\sqrt{4ac - b^2}}{2a}.$$

In 1799, Gauss proved that every polynomial equation of the form

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0,$$

where $a_0, a_1, \ldots, a_n$ are arbitrary real numbers, with $a_n \neq 0$, has a solution among the complex numbers if $n \geq 1$. Moreover, even if the coefficients $a_0, a_1, \ldots, a_n$ are complex, a solution exists in the complex-number system. This fact is known as the *fundamental theorem of algebra*. It shows that there is no need to construct numbers more general than complex numbers to solve polynomial equations with complex coefficients.

0.8 Exercises

1. If the product of two complex numbers is zero, prove that at least one of the factors is zero.
2. Prove that $x = i$ and $x = -i$ are the *only* solutions of the quadratic equation $x^2 + 1 = 0$.
3. Instead of the definition of multiplication given in Section 0.5, suppose that the product of two complex numbers is defined by the simpler equation $(a, b)(c, d) = (ac, bd)$, which is analogous to that for addition.
   (a) Show that this new product is commutative and associative and also satisfies the distributive law.
   (b) Give two reasons why you think this simpler definition is not appropriate for multiplying complex numbers.

0.9 Geometric interpretation. Modulus and argument

Because a complex number (x, y) is an ordered pair of real numbers, it can be represented geometrically by a point in a plane, or by an arrow extending from the origin to the point (x, y), as shown in Figure 0.5. In this context, the xy plane is often referred to as the complex plane. The x axis is called the real axis; the y axis is the imaginary axis. It is customary to use the words *complex number* and *point* interchangeably. Thus, we refer to the point z rather than the point corresponding to the complex number z.

The operations of addition and subtraction of complex numbers have a simple geometric interpretation. If two complex numbers z_1 and z_2 are represented by arrows from the origin to z_1 and z_2, respectively, then the sum $z_1 + z_2$ is determined by the *parallelogram law*. The arrow from the origin to $z_1 + z_2$ is a diagonal of the parallelogram determined by 0, z_1, and z_2, as illustrated by the example in Figure 0.6. The other diagonal is related to the difference of z_1 and z_2. The arrow from z_1 to z_2 is parallel to and equal in length to the arrow from 0 to $z_2 - z_1$; the arrow in the opposite direction, from z_2 to z_1, is related in the same way to $z_1 - z_2$.

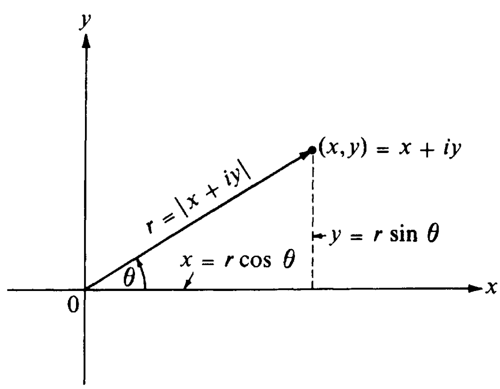

FIGURE 0.5 Geometric representation of the complex number $x + iy$.

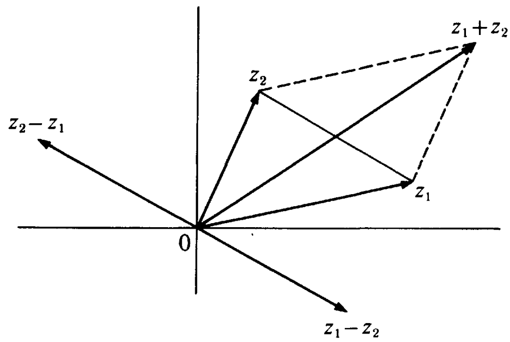

FIGURE 0.6 Addition and subtraction of complex numbers represented geometrically by the parallelogram law.

If $(x, y) \neq (0, 0)$ we can express x and y in polar coordinates,

$$x = r\cos\theta, \qquad y = r\sin\theta,$$

and we obtain

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

(See Figure 0.5.) The positive number r, which represents the distance of (x, y) from the origin, is called the *modulus* or *absolute value* of $x + iy$ and is denoted by $|x + iy|$. Thus, we have

$$|x + iy| = \sqrt{x^2 + y^2}.$$

The polar angle θ is call an *argument* of $x + iy$. We say *an* argument rather than *the* argument because for a given point (x, y) the angle θ is determined only up to multiples of 2π. Sometimes it is desirable to assign a unique argument to a complex number. This may be done by restricting θ to lie in a half-open interval of length 2π. The intervals $[0, 2\pi)$ and $(-\pi, \pi]$ are commonly used for this purpose. We shall use the interval $(-\pi, \pi]$ and refer to the corresponding θ in this interval as the *principal argument* of $x + iy$; we denote this θ by $\arg(x + iy)$. Thus, if $x + iy \neq 0$ and $r = |x + iy|$, we define $\arg(x + iy)$ to be the unique real θ satisfying the conditions

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad -\pi < \theta \leq \pi.$$

For the zero complex number we assign the modulus 0 and agree that any real θ may be used as argument.

Since the absolute value of a complex number z is simply the length of a line segment, it is not surprising to learn that it has the usual properties of absolute values of real numbers. For example,

$$|z| > 0 \text{ if } z \neq 0, \qquad \text{and} \qquad |z_1 - z_2| = |z_2 - z_1|.$$

Geometrically, the absolute value $|z_1 - z_2|$ represents the distance between the points z_1 and z_2 in the complex plane.

0.10 Complex conjugates

If $z = x + iy$, the *complex conjugate* of z is the complex number $\bar{z} = x - iy$. Geometrically, $\bar{z}$ represents the reflection of z through the real axis; it has the same real part, but the imaginary part has opposite sign. The definition of conjugate implies that $|\bar{z}| = |z|$ and that

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \qquad \overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}, \qquad \overline{z_1/z_2} = \overline{z_1}/\overline{z_2}, \qquad z\bar{z} = |z|^2.$$

Using these properties we find that $|z_1 z_2|^2 = z_1 z_2 \overline{z_1 z_2} = z_1 \overline{z_1} z_2 \overline{z_2} = |z_1|^2 |z_2|^2$ and hence

$$|z_1 z_2| = |z_1||z_2|. \tag{0.4}$$

Similarly, we find $|z_1/z_2| = |z_1|/|z_2|$ if $z_2 \neq 0$. The triangle inequality

$$|z_1 + z_2| \leq |z_1| + |z_2| \tag{0.5}$$

is also valid. To prove this we write

$$|z_1 + z_2|^2 = (z_1 + z_2)(\overline{z_1} + \overline{z_2}) = |z_1|^2 + |z_2|^2 + z_1\overline{z_2} + \overline{z_1} z_2.$$

Now observe that a complex number plus its conjugate is twice its real part; and since the real part of a complex number does not exceed its modulus, we have

$$z_1\overline{z_2} + \overline{z_1} z_2 \leq 2|z_1\overline{z_2}| = 2|z_1||z_2|.$$

Therefore

$$|z_1 + z_2|^2 \leq |z_1|^2 + |z_2|^2 + 2|z_1||z_2| = \left(|z_1| + |z_2|\right)^2,$$

from which we get the triangle inequality in (0.5).

If a quadratic equation with real coefficients has no real roots, its complex roots, given by (0.3), are conjugates. Conversely, if r_1 and r_2 are complex conjugates, say $r_1 = \alpha + i\beta$ and $r_2 = \alpha - i\beta$, where α and β are real, then r_1 and r_2 are roots of a quadratic equation with real coefficients. In fact,

$$r_1 + r_2 = 2\alpha \qquad \text{and} \qquad r_1 r_2 = \alpha^2 + \beta^2,$$

so

$$(x - r_1)(x - r_2) = x^2 - (r_1 + r_2)x + r_1 r_2$$

and the quadratic equation in question is

$$x^2 - 2\alpha x + \alpha^2 + \beta^2 = 0.$$

0.11 Exercises

1. Express each of the following complex numbers in the form $a + bi$.
   (a) $(1 + i)^2$.
   (b) $1/i$.
   (c) $1/(1 + i)$.
   (d) $(2 + 3i)(3 - 4i)$.
   (e) $(1 + i)/(1 - 2i)$.
   (f) $i^5 + i^{16}$.
   (g) $1 + i + i^2 + i^3$.
   (h) $\frac{1}{2}(1 + i)(1 + i^{-8})$.