CHAPTER 1

Review of Prerequisites

Every account of advanced engineering mathematics must rely on earlier mathematics courses to provide the necessary background. The essentials are a first course in calculus and some knowledge of elementary algebraic concepts and techniques. The purpose of the present chapter is to review the most important of these ideas that have already been encountered, and to provide for convenient reference results and techniques that can be consulted later, thereby avoiding the need to interrupt the development of subsequent chapters by the inclusion of review material prior to its use.

Some basic mathematical conventions are reviewed in Section 1.1, together with the method of proof by mathematical induction that will be required in later chapters. The essential algebraic operations involving complex numbers are summarized in Section 1.2, the complex plane is introduced in Section 1.3, the modulus and argument representation of complex numbers is reviewed in Section 1.4, and roots of complex numbers are considered in Section 1.5. Some of this material is required throughout the book, though its main use will be in Part 5 when developing the theory of analytic functions.

The use of partial fractions is reviewed in Section 1.6 because of the part they play in Chapter 7 in developing the Laplace transform. As the most basic properties of determinants are often required, the expansion of determinants is summarized in Section 1.7, though a somewhat fuller account of determinants is to be found later in Section 3.3 of Chapter 3.

The related concepts of limit, continuity, and differentiability of functions of one or more independent variables are fundamental to the calculus, and to the use that will be made of them throughout the book, so these ideas are reviewed in Sections 1.8 and 1.9. Tangent line and tangent plane approximations are illustrated in Section 1.10, and improper integrals that play an essential role in the Laplace and Fourier transforms, and also in complex analysis, are discussed in Section 1.11.

The importance of Taylor series expansions of functions involving one or more independent variables is recognized by their inclusion in Section 1.12. A brief mention is also made of the two most frequently used tests for the convergence of series, and of the differentiation and integration of power series that is used in Chapter 8 when considering series solutions of linear ordinary differential equations. These topics are considered again in Part 5 when the theory of analytic functions is developed.

The solution of many problems involving partial differential equations can be simplified by a convenient choice of coordinate system, so Section 1.13 reviews the theorem for the change of variable in partial differentiation, and describes the cylindrical polar and spherical polar coordinate systems that are the two that occur most frequently in practical problems.

Because of its fundamental importance, the implicit function theorem is stated without proof in Section 1.14, though it is not usually mentioned in first calculus courses.

1.1 Real Numbers, Mathematical Induction, and Mathematical Conventions

Numbers are fundamental to all mathematics, and real numbers are a subset of complex numbers. A real number can be classified as being an **integer**, a **rational** number, or an **irrational** number. From the set of positive and negative integers, and zero, the set of positive integers $1, 2, 3, \ldots$ is called the set of **natural numbers**. The rational numbers are those that can be expressed in the form m/n, where m and n are integers with $n \neq 0$. Irrational numbers such as π, $\sqrt{2}$, and $\sin 2$ are numbers that cannot be expressed in rational form, so, for example, for no integers m and n is it true that $\sqrt{2}$ is equal to m/n. Practical calculations can only be performed using rational numbers, so all irrational numbers that arise must be approximated arbitrarily closely by rational numbers.

Collectively, the sets of integers and rational and irrational numbers form what is called the set of all **real numbers**, and this set is denoted by **R**. When it is necessary to indicate that an arbitrary number a is a real number a shorthand notation is adopted involving the symbol $\in$, and we will write $a \in \mathbf{R}$. The symbol $\in$ is to be read "belongs to" or, more formally, as "is an element of the set." If a is not a member of set **R**, the symbol $\in$ is negated by writing $\notin$, and we will write $a \notin \mathbf{R}$ where, of course, the symbol $\notin$ is to be read as "does not belong to," or "is not an element of the set." As real numbers can be identified in a unique manner with points on a line, the set of all real numbers **R** is often called the **real line**. The set of all complex numbers **C** to which **R** belongs will be introduced later.

One of the most important properties of real numbers that distinguishes them from other complex numbers is that they can be arranged in numerical order. This fundamental property is expressed by saying that the real numbers possess the **order property**. This simply means that if $x, y \in \mathbf{R}$, with $x \neq y$, then

$$\text{either } x < y \quad \text{or} \quad x > y,$$

where the symbol $<$ is to be read "is less than" and the symbol $>$ is to be read "is greater than." When the foregoing results are expressed differently, though equivalently, if $x, y \in \mathbf{R}$, with $x \neq y$, then

$$\text{either } x - y < 0 \quad \text{or} \quad x - y > 0.$$

It is the order property that enables the graph of a real function f of a real variable x to be constructed. This follows because once length scales have been chosen for the axes together with a common origin, a real number can be made to correspond to a unique point on an axis. The graph of f follows by plotting all possible points $(x, f(x))$ in the plane, with x measured along one axis and $f(x)$ along the other axis.

absolute value

The **absolute value** $|x|$ of a real number x is defined by the formula

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0. \end{cases}$$

This form of definition is in reality a concise way of expressing two separate statements. One statement is obtained by reading $|x|$ with the top condition on the right and the other by reading it with the bottom condition on the right. The absolute value of a real number provides a measure of its magnitude without regard to its sign so, for example, $|3| = 3$, $|-7.41| = 7.41$, and $|0| = 0$.

Sometimes the form of a general mathematical result that only depends on an arbitrary natural number n can be found by experiment or by conjecture, and then the problem that remains is how to prove that the result is either true or false for all n. A typical example is the proposition that the product

$$(1 - 1/4)(1 - 1/9)(1 - 1/16)\ldots[1 - 1/(n+1)^2]$$
$$= (n+2)/(2n+2), \quad \text{for } n = 1, 2, \ldots.$$

This assertion is easily checked for any specific positive integer n, but this does not amount to a proof that the result is true for all natural numbers.

mathematical induction

A powerful method by which such propositions can often be shown to be either true or false involves using a form of argument called **mathematical induction**. This type of proof depends for its success on the order property of numbers and the fact that if n is a natural number, then so also is $n + 1$. The steps involved in an inductive proof can be summarized as follows.

Proof by Mathematical Induction

Let $P(n)$ be a proposition depending on a positive integer n.

STEP 1 Show, if possible, that $P(n)$ is true for some positive integer n_0.

STEP 2 Show, if possible, that if $P(n)$ is true for an arbitrary integer $n = k \geq n_0$, then the proposition $P(k + 1)$ follows from proposition $P(k)$.

STEP 3 If Step 2 is true, the fact that $P(n_0)$ is true implies that $P(n_0 + 1)$ is true, and then that $P(n_0 + 2)$ is true, and hence that $P(n)$ is true for all $n \geq n_0$.

STEP 4 If no number $n = n_0$ can be found for which Step 1 is true, or if in Step 2 it can be shown that $P(k)$ does not imply $P(k + 1)$, the proposition $P(n)$ is false.

The example that follows is typical of the situation where an inductive proof is used. It arises when determining the nth term in the Maclaurin series for $\sin ax$ that involves finding the nth derivative of $\sin ax$. A result such as this may be found intuitively by inspection of the first few derivatives, though this does not amount to a formal proof that the result is true for all natural numbers n.

EXAMPLE 1.1 Prove by mathematical induction that

$$d^n/dx^n[\sin ax] = a^n \sin(ax + n\pi/2), \quad \text{for } n = 1, 2, \ldots.$$

Solution The proposition $P(n)$ is that

$$d^n/dx^n[\sin ax] = a^n \sin(ax + n\pi/2), \quad \text{for } n = 1, 2, \ldots.$$

STEP 1 Differentiation gives

$$d/dx[\sin ax] = a \cos ax,$$

but setting $n = 1$ in $P(n)$ leads to the result

$$d/dx[\sin ax] = a\sin(ax + \pi/2) = a\cos ax,$$

showing that proposition $P(n)$ is true for $n = 1$ (so in this case $n_0 = 1$).

STEP 2 Assuming $P(k)$ to be true for $k > 1$, differentiation gives

$$d/dx\{d^k/dx^k[\sin ax]\} = d/dx[a^k \sin(ax + k\pi/2)],$$

so

$$d^{k+1}/dx^{k+1}[\sin ax] = a^{k+1}\cos(ax + k\pi/2).$$

However, replacing k by $k+1$ in $P(k)$ gives

$$\begin{aligned} d^{k+1}/dx^{k+1}[\sin ax] &= a^{k+1}\sin[ax + (k+1)\pi/2] \\ &= a^{k+1}\sin[(ax + k\pi/2) + \pi/2] \\ &= a^{k+1}\cos(ax + k\pi/2), \end{aligned}$$

showing, as required, that proposition $P(k)$ implies proposition $P(k+1)$, so Step 2 is true.

STEP 3 As $P(n)$ is true for $n = 1$, and $P(k)$ implies $P(k+1)$, it follows that the result is true for $n = 1, 2, \ldots$ and the proof is complete. ■

The **binomial theorem** finds applications throughout mathematics at all levels, so we quote it first when the exponent n is a positive integer, and then in its more general form when the exponent α involved is any real number.

Binomial theorem when n is a positive integer

If a, b are real numbers and n is a positive integer, then

$$(a+b)^n = a^n + na^{n-1}b + \frac{n(n-1)}{2!}a^{n-2}b^2 + \frac{n(n-1)(n-2)}{3!}a^{n-3}b^3 + \cdots + b^n,$$

binomial coefficient

or more concisely in terms of the **binomial coefficient**

$$\binom{n}{r} = \frac{n!}{(n-r)!r!},$$

we have

$$(a+b)^n = \sum_{r=0}^{n}\binom{n}{r}a^{n-r}b^r,$$

where $m!$ is the factorial function defined as $m! = 1\cdot 2\cdot 3\cdots m$ with $m > 0$ an integer, and $0!$ is defined as $0! = 1$. It follows at once that

$$\binom{n}{0} = \binom{n}{n} = 1.$$

The binomial theorem involving the expression $(a+b)^\alpha$, where a and b are real numbers with $|b/a| < 1$ and α is an arbitrary real number takes the following form.

General form of the binomial theorem when α is an arbitrary real number

If a and b are real numbers such that $|b/a| < 1$ and α is an arbitrary real number, then

$$(a+b)^\alpha = a^\alpha\left(1+\frac{b}{a}\right)^\alpha = a^\alpha\left(1+\frac{\alpha}{1!}\left(\frac{b}{a}\right)+\frac{\alpha(\alpha-1)}{2!}\left(\frac{b}{a}\right)^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}\left(\frac{b}{a}\right)^3+\cdots\right).$$

The series on the right only terminates after a finite number of terms if α is a positive integer, in which case the result reduces to the one just given. If α is a negative integer, or a nonintegral real number, the expression on the right becomes an infinite series that diverges if $|b/a| > 1$.

EXAMPLE 1.2 Expand $(3+x)^{-1/2}$ by the binomial theorem, stating for what values of x the series converges.

Solution Setting $b/a = \frac{1}{3}x$ in the general form of the binomial theorem gives

$$(3+x)^{-1/2} = 3^{-1/2}\left(1+\frac{1}{3}x\right)^{-1/2} = \frac{1}{\sqrt{3}}\left(1-\frac{1}{6}x+\frac{1}{24}x^2-\frac{5}{432}x^3+\cdots\right).$$

The series only converges if $|\frac{1}{3}x| < 1$, and so it is convergent provided $|x| < 3$. ■

Some standard mathematical conventions

Use of combinations of the $\pm$ and $\mp$ signs

The occurrence of two or more of the symbols $\pm$ and $\mp$ in an expression is to be taken to imply two separate results, the first obtained by taking the upper signs and the second by taking the lower signs. Thus, the expression $a \pm b\sin\theta \mp c\cos\theta$ is an abbreviation for the two separate expressions

$$a + b\sin\theta - c\cos\theta \quad \text{and} \quad a - b\sin\theta + c\cos\theta.$$

Multi-statements

multi-statement

When a function is defined sectionally on n different intervals of the real line, instead of formulating n separate definitions these are usually simplified by being combined into what can be considered to be a single **multi-statement**. The following example is typical of a multi-statement:

$$f(x) = \begin{cases} \sin x, & x < \pi \\ 0, & \pi \le x \le 3\pi/2 \\ -1, & x > 3\pi/2. \end{cases}$$

It is, in fact, three statements. The first is obtained by reading $f(x)$ in conjunction with the top line on the right, the second by reading it in conjunction with the second line on the right, and the third by reading it in conjunction with the third line on the right. An example of a multi-statement has already been encountered in the definition of the absolute value $|x|$ of a number x. Frequent use of multi-statements will be made in Chapter 9 on Fourier series, and elsewhere.

Polynomials

polynomials

A **polynomial** is an expression of the form $P(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$. The integer n is called the **degree** of the polynomial, and the numbers a_i are called its **coefficients**. The **fundamental theorem of algebra** that is proved in Chapter 14 asserts that $P(x) = 0$ has n **roots** that may be either real or complex, though some of them may be repeated. ($a_0 \neq 0$ is assumed.)

Notation for ordinary and partial derivatives

If $f(x)$ is an n times differentiable function then $f^{(n)}(x)$ will, on occasion, be used to signify $d^n f/dx^n$, so that

$$f^{(n)}(x) = \frac{d^n f}{dx^n}.$$

suffix notation for partial derivatives

If $f(x, y)$ is a suitably differentiable function of x and y, a concise notation used to signify partial differentiation involves using suffixes, so that

$$f_x = \frac{\partial f}{\partial x},\ f_{yx} = (f_y)_x = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y \partial x},\ f_{yy} = \frac{\partial^2 f}{\partial y^2}, \ldots,$$

with similar results when f is a function of more than two independent variables.

Inverse trigonometric functions

The periodicity of the real variable trigonometric sine, cosine, and tangent functions means that the corresponding general inverse trigonometric functions are *many valued*. So, for example, if $y = \sin x$ and we ask for what values of x is $y = 1/\sqrt{2}$, we find this is true for $x = \pi/4 \pm 2n\pi$ and $x = 3\pi/4 \pm 2n\pi$ for $n = 0, 1, 2, \ldots$. To overcome this ambiguity, we introduce the *single valued* inverses, denoted respectively by $x = \text{Arcsin } y$, $x = \text{Arccos } y$, and $x = \text{Arctan } y$ by restricting the domain and range of the sine, cosine, and tangent functions to one where they are either strictly increasing or strictly decreasing functions, because then one value of x corresponds to one value of y and, conversely, one value of y corresponds to one value of x.

In the case of the function $y = \sin x$, by restricting the argument x to the interval $-\pi/2 \leq x \leq \pi/2$ the function becomes a strictly increasing function of x. The corresponding single valued inverse function is denoted by $x = \text{Arcsin } y$, where y is a number in the domain of definition $[-1, 1]$ of the Arcsine function and x is a number in its range $[-\pi/2, \pi/2]$. Similarly, when considering the function $y = \cos x$, the argument is restricted to $0 \leq x \leq \pi$ to make $\cos x$ a strictly decreasing function of x. The corresponding single valued inverse function is denoted by $x = \text{Arccos } y$, where y is a number in the domain of definition $[-1, 1]$ of the Arccosine function and x is a number in its range $[0, \pi]$. Finally, in the case of the function $y = \tan x$, restricting the argument to the interval $-\pi/2 < x < \pi/2$ makes the tangent function a strictly increasing function of x. The corresponding single valued inverse function is denoted by $x = \text{Arctan } y$ where y is a number in the domain of definition $(-\infty, \infty)$ of the Arctangent function and x is a number in its range $(-\pi/2, \pi/2)$.

As the inverse trigonometric functions are important in their own right, the variables x and y in the preceding definitions are interchanged to allow consideration of the inverse functions $y = \text{Arcsin } x$, $y = \text{Arccos } x$, and $y = \text{Arctan } x$, so that now x is the independent variable and y is the dependent variable.

With this interchange of variables the expression $y = \arcsin x$ will be used to refer to any single valued inverse function with the *same* domain of definition as $\text{Arcsin } x$, but with a *different* range. Similar definitions apply to the functions $y = \arccos x$ and $y = \arctan x$.

Double summations

An expression involving a double summation like

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn} \sin mx \sin ny,$$

double summation

means sum the terms $a_{mn} \sin mx \sin ny$ over all possible values of m and n, so that

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn} \sin mx \sin ny = a_{11} \sin x \sin y + a_{12} \sin x \sin 2y$$
$$+ a_{21} \sin 2x \sin y + a_{22} \sin 2x \sin 2y + \cdots .$$

A more concise notation also in use involves writing the double summation as

$$\sum_{m=1,n=1}^{\infty} a_{mn} \sin mx \sin nx.$$

The signum function

signum function

The **signum function**, usually written $\text{sign}(x)$, and sometimes $\text{sgn}(x)$, is defined as

$$\text{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0. \end{cases}$$

We have, for example, $\text{sign}(\cos x) = 1$ for $0 < x < \pi/2$, and $\text{sign}(\cos x) = -1$ for $\pi/2 < x < \pi$ or, equivalently,

$$\text{sign}(\cos x) = \begin{cases} 1, & 0 < x < \frac{1}{2}\pi \\ -1, & \frac{1}{2}\pi < x < \pi. \end{cases}$$

Products

Let $\{u_k\}_{k=1}^{n}$ be a sequence of numbers or functions $u_1, u_2, \ldots$; then the product of the n members of this sequence is denoted by $\prod_{k=1}^{n} u_k$, so that

$$\prod_{k=1}^{n} u_k = u_1 u_2 \cdots u_n.$$

infinite product

When the sequence is infinite,

$$\lim_{n\to\infty} \prod_{k=1}^{n} u_k = \prod_{k=1}^{\infty} u_k$$

is called an **infinite product** involving the sequence $\{u_k\}$. Typical examples of infinite products are

$$\prod_{k=2}^{\infty}\left(1-\frac{1}{k^2}\right)=\frac{1}{2} \quad \text{and} \quad \prod_{k=1}^{\infty}\left(1-\frac{x^2}{k^2\pi^2}\right)=\frac{\sin x}{x}.$$

More background information and examples can be found in the appropriate sections in any of references [1.1], [1.2], and [1.5].

Logarithmic functions

the functions ln and Log

The notation $\ln x$ is used to denote the **natural logarithm** of a real number x, that is, the logarithm of x to the base e, and in some books this is written $\log_e x$. In this book logarithms to the base 10 are not used, and when working with functions of a complex variable the notation $\mathrm{Log}\, z$, with $z = re^{i\theta}$ means $\mathrm{Log}\, z = \ln r + i\theta$.

EXERCISES 1.1

1. Prove that if $a > 0, b > 0$, then $a/\sqrt{b} + b/\sqrt{a} \geq \sqrt{a} + \sqrt{b}$.

Prove Exercises 2 through 6 by mathematical induction.

2. $\sum_{k=0}^{n-1}(a + kd) = (n/2)[2a + (n-1)d]$ (sum of an arithmetic series).
3. $\sum_{k=0}^{n-1} r^k = (1 - r^n)/(1 - r) \quad (r \neq 1)$ (sum of a geometric series).
4. $\sum_{k=1}^{n} k^2 = (1/6)n(n+1)(2n+1)$ (sum of squares).
5. $d^n/dx^n[\cos ax] = a^n \cos(ax + n\pi/2)$, with n a natural number.
6. $d^n/dx^n[\ln(1+x)] = (-1)^{n+1}(n-1)!/(1+x)^n$, with n a natural number.
7. Use the binomial theorem to expand $(3 + 2x)^4$.
8. Use the binomial theorem and multiplication to expand $(1 - x^2)(2 + 3x)^3$.

In Exercises 9 through 12 find the first four terms of the binomial expansion of the function and state conditions for the convergence of the series.

9. $(3 + 2x)^{-2}$.
10. $(2 - x^2)^{1/3}$.
11. $(4 + 2x^2)^{-1/2}$.
12. $(1 - 3x^2)^{3/4}$.

1.2 Complex Numbers

Mathematical operations can lead to numbers that do not belong to the real number system $\mathbf{R}$ introduced in Section 1.1. In the simplest case this occurs when finding the roots of the quadratic equation

$$ax^2 + bx + c = 0 \quad \text{with } a, b, c \in \mathbf{R}, a \neq 0$$

by means of the quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

discriminant of a quadratic

The **discriminant** of the equation is $b^2 - 4ac$, and if $b^2 - 4ac < 0$ the formula involves the square root of a negative real number; so, if the formula is to have meaning, numbers must be allowed that lie outside the real number system.

The inadequacy of the real number system when considering different mathematical operations can be illustrated in other ways by asking, for example, how to find the three roots that are expected of a third degree algebraic equation as simple as $x^3 - 1 = 0$, where only the real root 1 can be found using $y = x^3 - 1$, or by seeking to give meaning to $\ln(-1)$, both of which questions will arise later.

Difficulties such as these can all be overcome if the real number system is extended by introducing the **imaginary unit** i defined as

$$i^2 = -1,$$

so expressions like $\sqrt{(-k^2)}$ where k a positive real number may be written $\sqrt{(-1)}\sqrt{(k^2)} = \pm ik$. Notice that as the real number k only *scales* the imaginary unit i, it is immaterial whether the result is written as ik or as ki.

The extension to the real number system that is required to resolve problems of the type just illustrated involves the introduction of **complex numbers**, denoted collectively by **C**, in which the general complex number, usually denoted by z, has the form

$$z = \alpha + i\beta, \quad \text{with } \alpha, \beta \text{ real numbers.}$$

real and imaginary part notation

The real number α is called the **real part** of the complex number z, and the real number β is called its **imaginary part**. When these need to be identified separately, we write

$$\text{Re}\{z\} = \alpha \quad \text{and} \quad \text{Im}\{z\} = \beta,$$

so if $z = 3 - 7i$, $\text{Re}\{z\} = 3$ and $\text{Im}\{z\} = -7$.

If $\text{Im}\{z\} = \beta = 0$ the complex number z reduces to a real number, and if $\text{Re}\{z\} = \alpha = 0$ it becomes a purely imaginary number, so, for example, $z = 5i$ is a purely imaginary number. When a complex number z is considered as a variable it is usual to write it as

$$z = x + iy,$$

where x and y are now real variables. If it is necessary to indicate that z is a general complex number we write $z \in \mathbf{C}$.

When solving the quadratic equation $az^2 + bz + c = 0$ with a, b, and c real numbers and a discriminant $b^2 - 4ac < 0$, by setting $4ac - b^2 = k^2$ in the quadratic formula, with $k > 0$, the two roots z_1 and z_2 are given by the complex numbers

$$z_1 = -(b/2a) + i(k/2a) \quad \text{and} \quad z_2 = -(b/2a) - i(k/2a).$$

Algebraic rules for complex numbers

Let the complex numbers z_1 and z_2 be defined as

$$z_1 = a + ib \quad \text{and} \quad z_2 = c + id,$$

with a, b, c, and d arbitrary real numbers. Then the following rules govern the arithmetic manipulation of complex numbers.

Equality of complex numbers

The complex numbers z_1 and z_2 are **equal**, written $z_1 = z_2$ if, and only if, $\text{Re}\{z_1\} = \text{Re}\{z_2\}$ and $\text{Im}\{z_1\} = \text{Im}\{z_2\}$. So $a + ib = c + id$ if, and only if,

$$a = c \quad \text{and} \quad b = d.$$

EXAMPLE 1.3

(a) $3 - 9i = 3 + bi$ if, and only if, $b = -9$.

(b) If $u = -2 + 5i$, $v = 3 + 5i$, $w = a + 5i$, then
$u = w$ if, and only if, $a = -2$ but $u \neq v$, and
$v = w$ if, and only if, $a = 3$. ■

Zero complex number

The **zero** complex number, also called the **null** complex number, is the number $0 + 0i$ that, for simplicity, is usually written as an ordinary zero 0.

EXAMPLE 1.4

If $a + ib = 0$, then $a = 0$ and $b = 0$. ■

Addition and subtraction of complex numbers

The **addition (sum)** and **subtraction (difference)** of the complex numbers z_1 and z_2 is defined as

$$z_1 + z_2 = \mathrm{Re}\{z_1\} + \mathrm{Re}\{z_2\} + i[\mathrm{Im}\{z_1\} + \mathrm{Im}\{z_2\}]$$

and

$$z_1 - z_2 = \mathrm{Re}\{z_1\} - \mathrm{Re}\{z_2\} + i[\mathrm{Im}\{z_1\} - \mathrm{Im}\{z_2\}].$$

So, if $z_1 = a + ib$ and $z_2 = c + id$, then

$$\begin{aligned} z_1 + z_2 &= (a + ib) + (c + id) \\ &= (a + c) + i(b + d), \end{aligned}$$

and

$$\begin{aligned} z_1 - z_2 &= (a + ib) - (c + id) \\ &= (a - c) + i(b - d). \end{aligned}$$

EXAMPLE 1.5

If $z_1 = 3 + 7i$ and $z_2 = 3 + 2i$, then the sum

$$z_1 + z_2 = (3 + 3) + (7 + 2)i = 6 + 9i,$$

and the difference

$$z_1 - z_2 = (3 - 3) + (7 - 2)i = 5i.$$ ■

Multiplication of complex numbers

The **multiplication (product)** of the two complex numbers $z_1 = a + ib$ and $z_2 = c + id$ is defined by the rule

$$z_1 z_2 = (a + ib)(c + id) = (ac - bd) + i(ad + bc).$$

An immediate consequence of this definition is that if k is a real number, then $kz_1 = k(a + ib) = ka + ikb$. This operation involving multiplication of a complex number by a real number is called *scaling* a complex number. Thus, if $z_1 = 3 + 7i$ and $z_2 = 3 + 2i$, then $2z_1 - 3z_2 = (6 + 14i) - (9 + 6i) = -3 + 8i$.

In particular, if $z = a + ib$, then $-z = (-1)z = -a - ib$. This is as would be expected, because it leads to the result $z - z = 0$.

In practice, instead of using this formal definition of multiplication, it is more convenient to perform multiplication of complex numbers by multiplying the bracketed quantities in the usual algebraic manner, replacing every product i^2 by -1, and then combining separately the real and imaginary terms to arrive at the required product.

EXAMPLE 1.6

(a) $5i(-4 + 3i) = -15 - 20i$.

(b) $(3 - 2i)(-1 + 4i)(1 + i) = (-3 + 12i + 2i - 8i^2)(1 + i)$

$= [(-3 + 8) + (12 + 2)i](1 + i) = (5 + 14i)(1 + i)$

$= 5 + 14i + 5i + 14i^2 = (5 - 14) + (5 + 14)i = -9 + 19i$. ■

Complex conjugate

If $z = a + ib$, then the **complex conjugate** of z, usually denoted by $\overline{z}$ and read "z bar," is defined as $\overline{z} = a - ib$. It follows directly that

$$\overline{(\overline{z})} = z \quad \text{and} \quad z\overline{z} = a^2 + b^2.$$

In words, the complex conjugate operation has the property that taking the complex conjugate of a complex conjugate returns the original complex number, whereas the product of a complex number and its complex conjugate always yields a real number.

If $z = a + ib$, then adding and subtracting z and $\overline{z}$ gives the useful results

$$z + \overline{z} = 2\mathrm{Re}\{z\} = 2a \quad \text{and} \quad z - \overline{z} = 2i\ \mathrm{Im}\{z\} = 2ib.$$

These can be written in the equivalent form

$$\mathrm{Re}\{z\} = a = \frac{1}{2}(z + \overline{z}) \quad \text{and} \quad \mathrm{Im}\{z\} = b = \frac{1}{2i}(z - \overline{z}).$$

Quotient (division) of complex numbers

Let $z_1 = a + ib$ and $z_2 = c + id$. Then the **quotient** z_1/z_2 is defined as

$$\frac{z_1}{z_2} = \frac{(ac + bd) + i(bc - ad)}{c^2 + d^2},\ z_2 \neq 0.$$

In practice, division of complex numbers is not carried out using this definition. Instead, the quotient is written in the form

$$\frac{z_1}{z_2} = \frac{z_1\overline{z}_2}{z_2\overline{z}_2},$$

where the denominator is now seen to be a real number. The quotient is then found by multiplying out and simplifying the numerator in the usual manner and dividing the real and imaginary parts of the numerator by the real number $z_2\overline{z}_2$.

EXAMPLE 1.7 Find z_1/z_2 given that $z_1 = (3 + 2i)$ and $z_2 = 1 + 3i$.

Solution

$$\frac{3+2i}{1+3i} = \frac{(3+2i)(1-3i)}{(1+3i)(1-3i)} = \frac{3-9i+2i-6i^2}{10} = \frac{9}{10} - \frac{7i}{10}.$$

Modulus of a complex number

The **modulus** of the complex number $z = a + ib$ denoted by $|z|$, and also called its **magnitude**, is defined as

$$|z| = (a^2 + b^2)^{1/2} = (z\overline{z})^{1/2}.$$

It follows directly from the definitions of the modulus and division that

$$|z| = |\overline{z}| = (a^2 + b^2)^{1/2},$$

and

$$z_1/z_2 = z_1\overline{z}_2/|z_2|^2.$$

EXAMPLE 1.8 If $z = 3 + 7i$, then $|z| = |3 + 7i| = (3^2 + 7^2)^{1/2} = \sqrt{58}$.

It is seen that the foregoing rules for the arithmetic manipulation of complex numbers reduce to the ordinary arithmetic rules for the algebraic manipulation of real numbers when all the complex numbers involved are real numbers. Complex numbers are the most general numbers that need to be used in mathematics, and they contain the real numbers as a special case. There is, however, a fundamental difference between real and complex numbers to which attention will be drawn after their common properties have been listed.

Properties shared by real and complex numbers

Let z, u, and w be arbitrary real or complex numbers. Then the following properties are true:

1. $z + u = u + z$. This means that the order in which complex numbers are added does not affect their sum.
2. $zu = uz$. This means that the order in which complex numbers are multiplied does not affect their product.

3. $(z + u) + w = z + (u + w)$. This means that the order in which brackets are inserted into a sum of finitely many complex numbers does not affect the sum.
4. $z(uw) = (zu)w$. This means that the terms in a product of complex numbers may be grouped and multiplied in any order without affecting the resulting product.
5. $z(u + w) = zu + zw$. This means that the product of z and a sum of complex numbers equals the sum of the products of z and the individual complex numbers involved in the sum.
6. $z + 0 = 0 + z = z$. This result means that the addition of zero to any complex number leaves it unchanged.
7. $z \cdot 1 = 1 \cdot z = z$. This result means that multiplication of any complex number by unity leaves the complex number unchanged.

Despite the properties common to real and complex numbers just listed, there remains a fundamental difference because, unlike real numbers, complex numbers have *no* natural order. So if z_1 and z_2 are any complex numbers, a statement such as $z_1 < z_2$ has no meaning.

EXERCISES 1.2

Find the roots of the equations in Exercises 1 through 6.

1. $z^2 + z + 1 = 0$.

2. $2z^2 + 5z + 4 = 0$.

3. $z^2 + z + 6 = 0$.

4. $3z^2 + 2z + 1 = 0$.

5. $3z^2 + 3z + 1 = 0$.

6. $2z^2 - 2z + 3 = 0$.

7. Given that $z = 1$ is a root, find the other two roots of $2z^3 - z^2 + 3z - 4 = 0$.

8. Given that $z = -2$ is a root, find the other two roots of $4z^3 + 11z^2 + 10z + 8 = 0$.

9. Given $u = 4 - 2i$, $v = 3 - 4i$, $w = -5i$ and $a + ib = (u + iv)w$, find a and b.

10. Given $u = -4 + 3i$, $v = 2 + 4i$, and $a + ib = uv^2$, find a and b.

11. Given $u = 2 + 3i$, $v = 1 - 2i$, $w = -3 - 6i$, find $|u + v|$, $u + 2v$, $u - 3v + 2w$, uv, uvw, $|u/v|$, v/w.

12. Given $u = 1 + 3i$, $v = 2 - i$, $w = -3 + 4i$, find uv/w, uw/v and $|v|w/u$.

1.3 The Complex Plane

Complex numbers can be represented geometrically either as *points*, or as *directed line segments* (*vectors*), in the **complex plane**. The complex plane is also called the **z-plane** because of the representation of complex numbers in the form $z = x + iy$. Both of these representations are accomplished by using rectangular cartesian coordinates and plotting the complex number $z = a + ib$ as the point (a, b) in the plane, so the x-coordinate of z is $a = \text{Re}\{z\}$ and its y-coordinate is $b = \text{Im}\{z\}$. Because of this geometrical representation, a complex number written in the form $z = a + ib$ is said to be expressed in **cartesian form**. To acknowledge the Swiss amateur mathematician Jean-Robert Argand, who introduced the concept of the complex plane in 1806, and who by profession was a bookkeeper, this representation is also called the **Argand diagram**.

cartesian representation of z

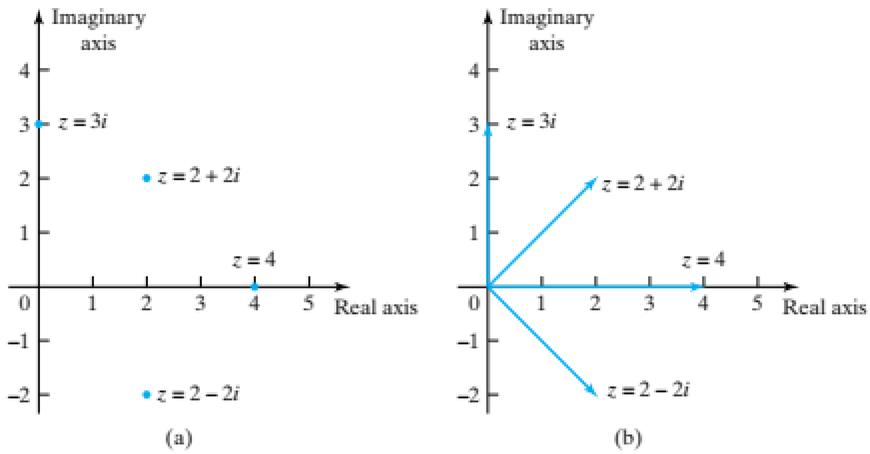

FIGURE 1.1 (a) Complex numbers as points. (b) Complex numbers as vectors.

For obvious reasons, the x-axis is called the **real axis** and the y-axis the **imaginary axis**. Purely real numbers are represented by points on the real axis and purely imaginary ones by points on the imaginary axis. Examples of the representation of typical points in the complex plane are given in Fig. 1.1a, where the numbers $4, 3i, 2+2i$, and $2-2i$ are plotted as points. These same complex numbers are shown again in Fig. 1.1b as directed line segments drawn from the origin (vectors). The arrow shows the *sense* along the line, that is, the direction from the origin to the tip of the vector representing the complex number. It can be seen from both figures that, when represented in the complex plane, a complex number and its complex conjugate (in this case $2+2i$ and $2-2i$) lie symmetrically above and below the real axis. Another way of expressing this result is by saying that a complex number and its complex conjugate appear as **reflections** of each other in the real axis, which acts like a mirror.

The addition and subtraction of two complex numbers have convenient geometrical interpretations that follow from the definitions given in Section 1.2. When complex numbers are added, their respective real and imaginary parts are added, whereas when they are subtracted, their respective real and imaginary parts are subtracted. This leads at once to the **triangle law** for addition illustrated in Fig. 1.2a, in which the directed line segment (vector) representing z_2 is translated without rotation or change of scale, to bring its base (the end opposite to the arrow) into coincidence with the tip of the directed line element representing z_1 (the end at which the arrow is located). The sum $z_1 + z_2$ of the two complex numbers is then represented by the directed line segment from the base of the line segment representing z_1 to the tip of the newly positioned line segment representing z_2.

triangle and parallelogram laws

The name *triangle law* comes from the triangle that is constructed in the complex plane during this geometrical process of addition. Notice that an immediate consequence of this law is that addition is *commutative*, because both $z_1 + z_2$ and $z_2 + z_1$ are seen to lead to the same directed line segment in the complex plane. For this reason the addition of complex numbers is also said to obey the **parallelogram law** for addition, because the commutative property generates the parallelogram shown in Fig. 1.2a.

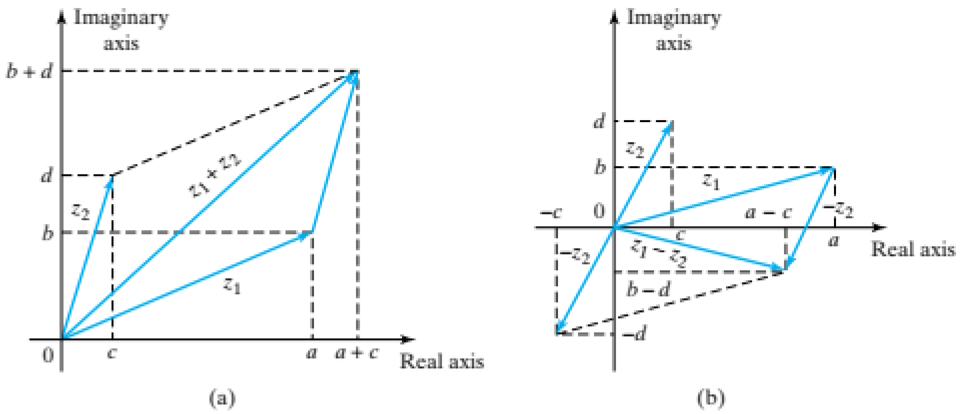

FIGURE 1.2 Addition and subtraction of complex numbers using the triangle/parallelogram law.

The geometrical interpretation of the subtraction of z_2 from z_1 follows similarly by adding to z_1 the directed line segment $-z_2$ that is obtained by reversing of the sense (arrow) along z_2, as shown in Fig. 1.2b.

It is an elementary fact from Euclidean geometry that the sum of the lengths of the two sides $|u|$ and $|v|$ of the triangle in Fig. 1.3 is greater than or equal to the length of the hypotenuse $|u+v|$, so from geometrical considerations we can write

$$|u+v| \leq |u| + |v|.$$

triangle inequality

This result involving the moduli of the complex numbers u and v is called the **triangle inequality** for complex numbers, and it has many applications.

An algebraic proof of the triangle inequality proceeds as follows:

$$\begin{aligned} |u+v|^2 &= (u+v)\overline{(u+v)} = u\overline{u} + v\overline{u} + u\overline{v} + v\overline{v} \\ &= |u|^2 + |v|^2 + (u\overline{v} + \overline{u\overline{v}}) \leq |u^2| + |v^2| + 2|u\overline{v}| \\ &= (|u| + |v|)^2. \end{aligned}$$

The required result now follows from taking the positive square root.

A similar argument, the proof of which is left as an exercise, can be used to show that $||u| - |v|| \leq |u+v|$, so when combined with the triangle inequality we have

$$||u| - |v|| \leq |u+v| \leq |u| + |v|.$$

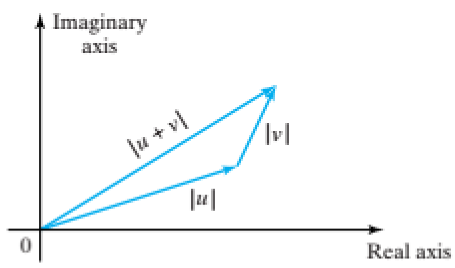

FIGURE 1.3 The triangle inequality.

EXERCISES 1.3

In Exercises 1 through 8 use the parallelogram law to form the sum and difference of the given complex numbers and then verify the results by direct addition and subtraction.

1. $u = 2 + 3i$, $v = 1 - 2i$.
2. $u = 4 + 7i$, $v = -2 - 3i$.
3. $u = -3$, $v = -3 - 4i$.
4. $u = 4 + 3i$, $v = 3 + 4i$.
5. $u = 3 + 6i$, $v = -4 + 2i$.
6. $u = -3 + 2i$, $v = 6i$.
7. $u = -4 + 2i$, $v = -4 - 10i$.
8. $u = 4 + 7i$, $v = -3 + 5i$.

In Exercises 9 through 11 use the parallelogram law to verify the triangle inequality $|u + v| \le |u| + |v|$ for the given complex numbers u and v.

9. $u = -4 + 2i$, $v = 3 + 5i$.
10. $u = 2 + 5i$, $v = 3 - 2i$.
11. $u = -3 + 5i$, $v = 2 + 6i$.

1.4 Modulus and Argument Representation of Complex Numbers

polar representation of z

When representing $z = x + iy$ in the complex plane by a point P with coordinates (x, y), a natural alternative to the cartesian representation is to give the *polar coordinates* (r, θ) of P. This polar representation of z is shown in Fig. 1.4, where

$$OP = r = |z| = (x^2 + y^2)^{1/2} \quad \text{and} \quad \tan\theta = y/x. \tag{1}$$

The radial distance OP is the **modulus** of z, so $r = |z|$, and the angle θ measured counterclockwise from the positive real axis is called the **argument** of z. Because of this, a complex number expressed in terms of the polar coordinates (r, θ) is said to be in **modulus–argument** form. The argument θ is indeterminate up to a multiple of 2π, because the polar coordinates (r, θ), and $(r, \theta + 2k\pi)$, with $k = \pm1, \pm2, \ldots$, identify the *same* point P. By convention, the the angle θ is called the **principal value** of the argument of z when it lies in the interval $-\pi < \theta \le \pi$. To distinguish the principal value of the argument from all of its other values, we write

$$\operatorname{Arg} z = \theta, \quad \text{when } -\pi < \theta \le \pi. \tag{2}$$

The values of the argument of z that differ from this value of θ by a multiple of 2π are denoted by $\arg z$, so that

$$\arg z = \theta + 2k\pi, \quad \text{with } k = \pm 1, \pm 2, \ldots. \tag{3}$$

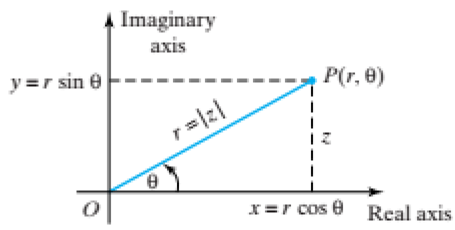

FIGURE 1.4 The complex plane and the (r, θ) representation of z.

The significance of the multivalued nature of arg z will become apparent later when the roots of complex numbers are determined.

The connection between the cartesian coordinates (x, y) and the polar coordinates (r, θ) of the point P corresponding to $z = x + iy$ is easily seen to be given by

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta.$$

modulus–argument representation of z

This leads immediately to the representation of $z = x + iy$ in the alternative *modulus–argument form*

$$z = r(\cos\theta + i\sin\theta). \tag{4}$$

A routine calculation using elementary trigonometric identities shows that

$$(\cos\theta + i\sin\theta)^2 = (\cos 2\theta + i\sin 2\theta).$$

An inductive argument using the above result as its first step then establishes the following simple but important theorem.

THEOREM 1.1

De Moivre's theorem

$$(\cos\theta + i\sin\theta)^n = (\cos n\theta + i\sin n\theta), \quad \text{for } n \text{ a natural number.}$$

■

EXAMPLE 1.9

Use de Moivre's theorem to express $\cos 4\theta$ and $\sin 4\theta$ in terms of powers of $\cos\theta$ and $\sin\theta$.

Solution The result is obtained by first setting $n = 4$ in de Moivre's theorem and expanding $(\cos\theta + i\sin\theta)^4$ to obtain

$$\cos^4\theta + 4i\cos^3\theta\sin\theta - 6\cos^2\theta\sin^2\theta - 4i\cos\theta\sin^3\theta + \sin^4\theta = \cos 4\theta + i\sin 4\theta.$$

Equating the respective real and imaginary parts on either side of this identity gives the required results

$$\cos 4\theta = \cos^4\theta - 6\cos^2\theta\sin^2\theta + \sin^4\theta$$

and

$$\sin 4\theta = 4\cos^3\theta\sin\theta - 4\cos\theta\sin^3\theta.$$

■

As the complex number $z = \cos\theta + i\sin\theta$ has unit modulus, it follows that all numbers of this form lie on the unit circle (a circle of radius 1) centered on the origin, as shown in Fig. 1.5.

Using (5), we see that if $z = r(\cos\theta + i\sin\theta)$, then

$$z^n = r^n(\cos n\theta + i\sin n\theta), \text{ for } n \text{ a natural number.} \tag{5}$$

The relationship between e^θ, $\sin\theta$, and $\cos\theta$ can be seen from the following well-known series expansions of the functions

$$e^\theta = \sum_{n=0}^{\infty}\frac{\theta^n}{n!} = 1 + \theta + \frac{\theta^2}{2!} + \frac{\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{\theta^5}{5!} + \frac{\theta^6}{6!} + \cdots;$$

$$\sin\theta = \sum_{n=0}^{\infty}(-1)^n\frac{\theta^{2n+1}}{(2n+1)!} = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots;$$

$$\cos\theta = \sum_{n=0}^{\infty}(-1)^n\frac{\theta^{2n}}{(2n)!} = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots.$$

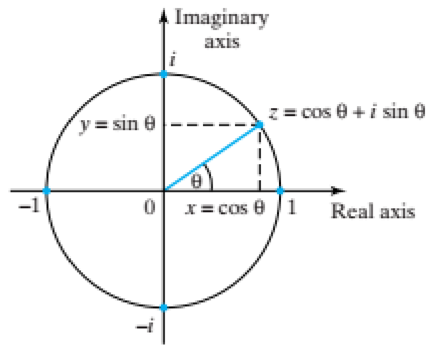

FIGURE 1.5 Point $z = \cos\theta + i\sin\theta$ on the unit circle centered on the origin.

By making a formal power series expansion of the function $e^{i\theta}$, simplifying powers of i, grouping together the real and imaginary terms, and using the series representations for $\cos\theta$ and $\sin\theta$, we arrive at what is called the real variable form of the

Euler formula

Euler formula

$$e^{i\theta} = \cos\theta + i\sin\theta, \quad \text{for any real } \theta. \tag{6}$$

This immediately implies that if $z = re^{i\theta}$, then

$$z^{\alpha} = r^{\alpha}e^{i\alpha\theta}, \quad \text{for any real } \alpha. \tag{7}$$

When θ is restricted to the interval $-\pi < \theta \leq \pi$, formula (6) leads to the useful results

$$1 = e^{i0}, \quad i = e^{i\pi/2}, \quad -1 = e^{i\pi}, \quad -i = e^{-i\pi/2}$$

and, in particular, to

$$1 = e^{2k\pi i} \quad \text{for } k = 0, \pm 1, \pm 2, \ldots.$$

The Euler form for complex numbers makes their multiplication and division very simple. To see this we set $z_1 = r_1e^{i\alpha}$ and $z_2 = r_2e^{i\beta}$ and then use the results

$$z_1z_2 = r_1r_2e^{i(\alpha+\beta)} \quad \text{and} \quad z_1/z_2 = r_1/r_2e^{i(\alpha-\beta)}. \tag{8}$$

These show that when complex numbers are multiplied, their moduli are *multiplied* and their arguments are *added*, whereas when complex numbers are divided, their moduli are *divided* and their arguments are *subtracted*.

EXAMPLE 1.10 Find uv, u/v, and u^{25} given that $u = 1 + i$, $v = \sqrt{3} - i$.

Solution $u = 1 + i = \sqrt{2}e^{i\pi/4}$, $v = \sqrt{3} - i = 2e^{-i\pi/6}$, so $uv = 2\sqrt{2}e^{i\pi/12}$, $u/v = (1/\sqrt{2})e^{i5\pi/12}$ while $u^{25} = (\sqrt{2}e^{i\pi/4})^{25} = (\sqrt{2})^{25}(e^{i\pi/4})^{25} = 4096\sqrt{2}(e^{i(6+1/4)\pi}) = 4096\sqrt{2}(e^{i6\pi})(e^{i\pi/4}) = 4096\sqrt{2}(e^{i\pi/4}) = 4096\sqrt{2}(1 + i)$. ■

To find the principal value of the argument of a given complex number z, namely Arg z, use should be made of the signs of $x = \text{Re}\{z\}$, and $y = \text{Im}\{z\}$ together with the results listed below, all of which follow by inspection of Fig. 1.5.

Signs of x and y	*Arg z = θ*
$x < 0,\ y < 0$	$-\pi < \theta < -\pi/2$
$x > 0,\ y < 0$	$-\pi/2 < \theta < 0$
$x > 0,\ y > 0$	$0 < \theta < \pi/2$
$x < 0,\ y > 0$	$\pi/2 < \theta < \pi$

EXAMPLE 1.11

Find $r = |z|$, Arg z, arg z, and the modulus–argument form of the following values of z.

(a) $-2\sqrt{3} - 2i$ (b) $-1 + i\sqrt{3}$ (c) $1 + i$ (d) $2 - i2\sqrt{3}$.

Solution **(a)** $r = \{(-2\sqrt{3})^2 + (-2)^2\}^{1/2} = 4$, $\text{Arg}\,z = \theta = -5\pi/6$ and $\arg z = -5\pi/6 + 2k\pi$, $k = \pm 1, \pm 2, \ldots$, $z = 4(\cos(-5\pi/6) + i\sin(-5\pi/6))$.

(b) $r = \{(-1)^2 + (\sqrt{3})^2\}^{1/2} = 2$, $\text{Arg}\,z = \theta = 2\pi/3$ and $\arg z = 2\pi/3 + 2k\pi$, $k = \pm 1, \pm 2, \ldots$, $z = 2(\cos(2\pi/3) + i\sin(2\pi/3))$.

(c) $r = \{(1)^2 + (1)^2\}^{1/2} = \sqrt{2}$, $\text{Arg}\,z = \theta = \pi/4$ and $\arg z = \pi/4 + 2k\pi$, $k = \pm 1, \pm 2, \ldots$, $z = \sqrt{2}(\cos(\pi/4) + i\sin(\pi/4))$.

(d) $r = \{(2)^2 + (-2\sqrt{3})^2\}^{1/2} = 4$, $\text{Arg}\,z = \theta = -\pi/3$ and $\arg z = -\pi/3 + 2k\pi$, $k = \pm 1, \pm 2, \ldots$, $z = 4(\cos(-\pi/3) + i\sin(-\pi/3))$. ■

EXERCISES 1.4

1. Expand $(\cos\theta + i\sin\theta)^2$ and then use trigonometric identities to show that
$$(\cos\theta + i\sin\theta)^2 = (\cos 2\theta + i\sin 2\theta).$$

2. Give an inductive proof of de Moivre's theorem
$$(\cos\theta + i\sin\theta)^n = (\cos n\theta + i\sin n\theta),$$
for n a natural number.

3. Use de Moivre's theorem to express $\cos 5\theta$ and $\sin 5\theta$ in terms of powers of $\cos\theta$ and $\sin\theta$.

4. Use de Moivre's theorem to express $\cos 6\theta$ and $\sin 6\theta$ in terms of powers of $\cos\theta$ and $\sin\theta$.

5. Show by expanding $(\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta)$ and using trigonometric identities that
$$(\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) = \cos(\alpha + \beta) + i\sin(\alpha + \beta).$$

6. Show by expanding $(\cos\alpha + i\sin\alpha)/(\cos\beta + i\sin\beta)$ and using trigonometric identities that
$$(\cos\alpha + i\sin\alpha)/(\cos\beta + i\sin\beta) = \cos(\alpha - \beta) + i\sin(\alpha - \beta).$$

7. If $z = \cos\theta + i\sin\theta = e^{i\theta}$, show that when n is a natural number,
$$\cos(n\theta) = \frac{1}{2}\left(z^n + \frac{1}{z^n}\right) \quad \text{and} \quad \sin(n\theta) = \frac{1}{2i}\left(z^n - \frac{1}{z^n}\right).$$
Use these results to express $\cos^3\theta\sin^3\theta$ in terms of multiple angles of θ. Hint: $\bar{z} = 1/z$.

8. Use the method of Exercise 7 to express $\sin^6\theta$ in terms of multiple angles of θ.

9. By expanding $(z + 1/z)^4$, grouping terms, and using the method of Exercise 7, show that
$$\cos^4\theta = (1/8)(3 + 4\cos 2\theta + \cos 4\theta).$$

10. By expanding $(z - 1/z)^5$, grouping terms, and using the method of Exercise 7, show that
$$\sin^5\theta = (1/16)(\sin 5\theta - 5\sin 3\theta + 10\sin\theta).$$

11. Use the method of Exercise 7 to show that
$$\cos^3\theta + \sin^3\theta = (1/4)(\cos 3\theta + 3\cos\theta - \sin^3\theta + 3\sin\theta).$$

In Exercises 12 through 15 express the functions of u, v, and w in modulus-argument form.

12. uv, u/v, and v^5, given that $u = 2 - 2i$ and $v = 3 + i3\sqrt{3}$.

13. uv, u/v, and u^7, given that $u = -1 - i\sqrt{3}$, $v = -4 + 4i$.

14. uv, u/v, and v^6, given that $u = 2 - 2i$, $v = 2 - i2\sqrt{3}$.

15. uvw, uw/v, and w^3/u^4, given that $u = 2 - 2i$, $v = 3 - i3\sqrt{3}$ and $w = 1 + i$.

16. Express $[(-8 + i8\sqrt{3})/(-1 - i)]^2$ in modulus–argument form.

17. Find in modulus–argument form $[(1 + i\sqrt{3})^3/(-1 + i)^2]^3$.

18. Use the factorization

$$(1 - z^{n+1}) = (1 - z)(1 + z + z^2 + \cdots + z^n) \qquad (z \neq 1)$$

with $z = e^{i\theta} = \exp(i\theta)$ to show that

$$\sum_{k=1}^{n} \exp(ik\theta) = \frac{\exp(in\theta) - 1}{1 - \exp(-i\theta)}.$$

19. Use the final result of Exercise 18 to show that

$$\sum_{k=1}^{n} \exp(ik\theta) = \frac{\exp[i(n + 1/2)\theta] - \exp(i\theta/2)}{\exp(i\theta/2) - \exp(-i\theta/2)},$$

and then use the result to deduce the **Lagrange identity**

$$1 + \cos\theta + \cos 2\theta + \cdots + \cos n\theta = 1/2 + \frac{\sin[(n + 1/2)\theta]}{2\sin(\theta/2)}, \quad \text{for } 0 < \theta < 2\pi.$$

1.5 Roots of Complex Numbers

It is often necessary to find the n values of $z^{1/n}$ when n is a positive integer and z is an arbitrary complex number. This process is called finding the ***n*th roots** of z. To determine these roots we start by setting

$$w = z^{1/n}, \quad \text{which is equivalent to } w^n = z.$$

Then, after defining w and z in modulus–argument form as

$$w = \rho e^{i\phi} \quad \text{and} \quad z = re^{i\theta}, \tag{9}$$

we substitute for w and z in $w^n = z$ to obtain

$$\rho^n e^{in\phi} = re^{i\theta}.$$

It is at this stage, in order to find all n roots, that use must be made of the many-valued nature of the argument of a complex number by recognizing that $1 = e^{2k\pi i}$ for $k = 0, \pm 1, \pm 2, \ldots$. Using this result we now multiply the right-hand side of the foregoing result by by $e^{2k\pi i}$ (that is, by 1) to obtain

$$\rho^n e^{in\phi} = re^{i\theta}e^{2k\pi i} = re^{i(\theta + 2k\pi)}.$$

Equality of complex numbers in modulus–argument form means the equality of their moduli and, correspondingly, the equality of their arguments, so applying this to the last result we have

$$\rho^n = r \quad \text{and} \quad n\phi = \theta + 2k\pi,$$

showing that

$$\rho = r^{1/n} \quad \text{and} \quad \phi = (\theta + 2k\pi)/n.$$

Here $r^{1/n}$ is simply the nth positive root of r: $\rho = \sqrt[n]{r}$.

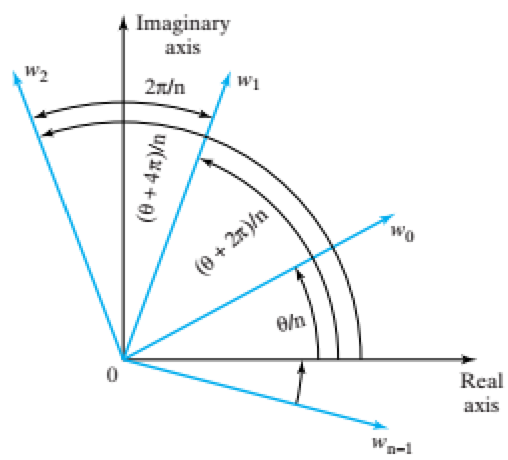

FIGURE 1.6 Location of the roots of $z^{1/n}$.

nth roots of a complex number z

Finally, when we substitute these results into the expression for w, we see that the n values of the roots denoted by $w_0, w_1, \ldots, w_{n-1}$ are given by

$$w_k = r^{1/n}\{\cos[(\theta + 2k\pi)/n] + i\sin[(\theta + 2k\pi)/n]\},\quad \text{for } k = 0, 1, \ldots, n-1. \tag{10}$$

Notice that it is only necessary to allow k to run through the successive integers $0, 1, \ldots, n-1$, because the period of the sine and cosine functions is 2π, so allowing k to increase beyond the value $n-1$ will simply repeat this *same* set of roots. An identical argument shows that allowing k to run through successive negative integers can again only generate the same n roots $w_0, w_1, \ldots, w_{n-1}$.

Examination of the arguments of the roots shows them to be spaced uniformly around a circle of radius $r^{1/n}$ centered on the origin. The angle between the radial lines drawn from the origin to each successive root is $2\pi/n$, with the radial line from the origin to the first root w_0 making an angle θ/n to the positive real axis, as shown in Fig. 1.6. This means that if the location on the circle of any one root is known, then the locations of the others follow immediately.

Writing unity in the form $1 = e^{i0}$ shows its modulus to be $r = 1$ and the principal value of its argument to be $\theta = 0$. Substitution in formula (10) then shows the n roots of $1^{1/n}$, called the ***n*th roots of unity**, to be

$$w_0 = 1,\quad w_1 = e^{i\pi/n},\quad w_2 = e^{i2\pi/n}, \ldots, w_{n-1} = e^{i(n-1)\pi/n}. \tag{11}$$

By way of example, the fifth roots of unity are located around the unit circle as shown in Fig. 1.7.

If we set $\omega = w_1$, it follows that the nth roots of unity can be written in the form

$$1, \omega, \omega^2, \ldots, \omega^{n-1}.$$

As $\omega^n = 1$ and $\omega^n - 1 = (\omega - 1)(1 + \omega + \omega^2 + \cdots + \omega^{n-1}) = 0$, as $\omega_1 \neq 1$ we see that the the nth roots of unity satisfy

$$1 + \omega + \omega^2 + \cdots + \omega^{n-1} = 0. \tag{12}$$

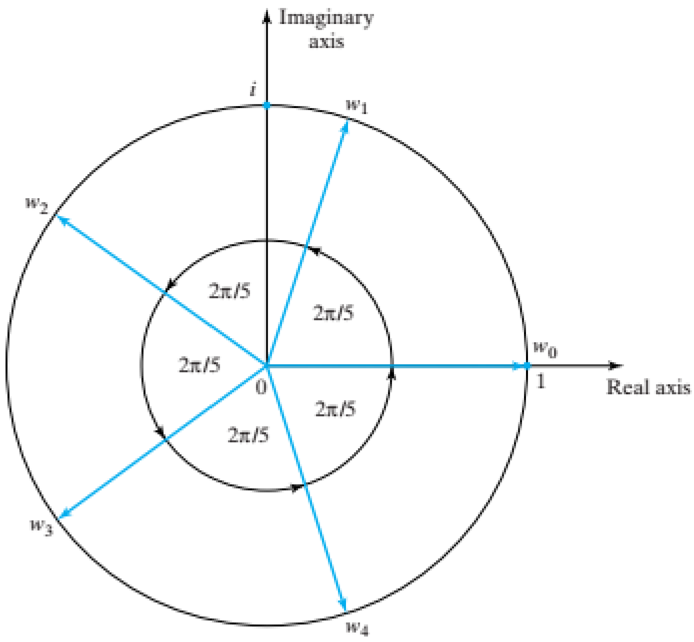

FIGURE 1.7 The fifth roots of unity.

This result remains true if ω is replaced by any one of the other nth roots of unity, with the exception of 1 itself.

EXAMPLE 1.12 Find $w = (1+i)^{1/3}$.

Solution Setting $z = 1 + i = \sqrt{2}e^{i\pi/4}$ shows that $r = |z| = \sqrt{2}$ and $\theta = \pi/4$. Substituting these results into formula (1) gives

$$w_k = 2^{1/6}\{\cos[(1/12)(1+8k)\pi] + i\sin[(1/12)(1+8k)\pi]\}, \quad \text{for } k = 0, 1, 2. \quad ■$$

The square root of a complex number $\zeta = \alpha + i\beta$ is often required, so we now derive a useful formula for its two roots in terms of $|\zeta|$, α and the sign of β. To obtain the result we consider the equation

$$z^2 = \zeta, \quad \text{where } \zeta = \alpha + i\beta,$$

and let Arg $\zeta = \theta$. Then we may write

$$z^2 = |\zeta|e^{i\theta},$$

and taking the square root of this result we find the two square roots z_- and z_+ are given by

$$\begin{aligned} z_\pm &= \pm|\zeta|^{1/2}e^{i\theta/2} \\ &= \pm|\zeta|^{1/2}\{\cos(\theta/2) + i\sin(\theta/2)\}. \end{aligned}$$

Now $\cos\theta = \alpha/|\zeta|$, but

$$\cos^2(\theta/2) = (1/2)(1+\cos\theta), \quad \text{and} \quad \sin^2(\theta/2) = (1/2)(1-\cos\theta),$$

so

$$\cos^2(\theta/2) = (1/2)(1 + \alpha/|\zeta|), \quad \text{and} \quad \sin^2(\theta/2) = (1/2)(1 - \alpha/|\zeta|).$$

As $-\pi < \theta \leq \pi$, it follows that in this interval $\cos(\theta/2)$ is nonnegative, so taking the square root of $\cos^2(\theta/2)$ we obtain

$$\cos(\theta/2) = \left(\frac{|\zeta| + \alpha}{2|\zeta|}\right)^{1/2}.$$

However, the function $\sin(\theta/2)$ is negative in the interval $-\pi < \theta < 0$ and positive in the interval $0 < \theta < \pi$, and so has the same sign as β. Thus, the square root of $\sin^2(\theta/2)$ can be written in the form

$$\sin(\theta/2) = \text{sign}\,(\beta)\left(\frac{|\zeta| - \alpha}{2|\zeta|}\right)^{1/2}.$$

Using these expressions for $\cos(\theta/2)$ and $\sin(\theta/2)$ in the square roots $z_\pm$ brings us to the following useful rule.

Rule for finding the square root of a complex number

Let $z^2 = \zeta$, with $\zeta = \alpha + i\beta$. Then the square roots z_+ and z_- of ζ are given by

$$z_+ = \left(\frac{|\zeta| + \alpha}{2}\right)^{1/2} + i\,\text{sign}\,(\beta)\left(\frac{|\zeta| - \alpha}{2}\right)^{1/2}$$
$$z_- = -\left(\frac{|\zeta| + \alpha}{2}\right)^{1/2} - i\,\text{sign}\,(\beta)\left(\frac{|\zeta| - \alpha}{2}\right)^{1/2}.$$

EXAMPLE 1.13 Find the square roots of (a) $\zeta = 1 + i$ and (b) $\zeta = 1 - i$.

Solution **(a)** $\zeta = 1 + i$ so $|\zeta| = \sqrt{2}$, $\alpha = 1$ and $\text{sign}(\beta) = 1$, so the square roots of $\zeta = 1 + i$ are

$$z_\pm = \pm\left\{\left(\frac{\sqrt{2}+1}{2}\right)^{1/2} + i\left(\frac{\sqrt{2}-1}{2}\right)^{1/2}\right\}.$$

(b) $\zeta = 1 - i$, so $|\zeta| = \sqrt{2}$, $\alpha = 1$ and $\text{sign}(\beta) = -1$, from which it follows that the square roots of $\zeta = 1 - i$ are

$$z_\pm = \pm\left\{\left(\frac{\sqrt{2}+1}{2}\right)^{1/2} - i\left(\frac{\sqrt{2}-1}{2}\right)^{1/2}\right\}.$$ ■

The theorem that follows provides information about the roots of polynomials with *real* coefficients that proves to be useful in a variety of ways.

THEOREM 1.2 **Roots of a polynomial with real coefficients** Let

$$P(z) = z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots a_{n-1} z + a_n$$

be a polynomial of **degree** n in which all the coefficients $a_1, a_2, \ldots, a_n$ are real. Then either all the n **roots** of $P(z) = 0$ are real, that is, the n **zeros** of $P(z)$ are all real, or any that are complex must occur in complex conjugate pairs.

Proof The proof uses the following simple properties of the complex conjugate operation.

1. If a is real, then $\overline{a} = a$. This result follows directly from the definition of the complex conjugate operation.
2. If b and c are any two complex numbers, then $\overline{b+c} = \overline{b} + \overline{c}$. This result also follows directly from the definition of the complex conjugate operation.
3. If b and c are any two complex numbers, then $\overline{bc} = \overline{b}\overline{c}$ and $\overline{b^r} = (\overline{b})^r$.

We now proceed to the proof. Taking the complex conjugate of $P(z) = 0$ gives

$$\overline{z}^n + \overline{a_1 z^{n-1}} + \overline{a_2 z^{n-2}} + \cdots + \overline{a_{n-1} z} + \overline{a_n} = 0,$$

but the a_r are all real so $\overline{a_r z^{n-r}} = \overline{a_r}\,\overline{z^{n-r}} = a_r \overline{z^{n-r}} = a_r (\overline{z})^{n-r}$, allowing the preceding equation to be rewritten as

$$(\overline{z})^n + a_1(\overline{z})^{n-1} + a_2(\overline{z})^{n-2} + \cdots + a_{n-1}\overline{z} + a_n = 0.$$

This result is simply $P(\overline{z}) = 0$, showing that if z is a complex root of $P(z)$, then so also is $\overline{z}$. Equivalently, z and $\overline{z}$ are both zeros of $P(z)$.

If, however, z is a real root, then $z = \overline{z}$ and the result remains true, so the first part of the theorem is proved. The second part follows from the fact that if $z = \alpha + i\beta$ is a root, then so also is $z = \alpha - i\beta$, and so $(z - \alpha - i\beta)$ and $(z - \alpha + i\beta)$ are factors of $P(z)$. The product of these factors must also be a factor of $P(z)$, but

$$(z - \alpha - i\beta)(z - \alpha + i\beta) = z^2 - 2\alpha z + \alpha^2 + \beta^2,$$

and the expression on the right is a quadratic in z with real coefficients, so the final result of the theorem is established. ■

EXAMPLE 1.14 Find the roots of $z^3 - z^2 - z - 2 = 0$, given that $z = 2$ is a root.

Solution If $z = 2$ is a root of $P(z) = 0$, then $z - 2$ is a factor of $P(z)$, so dividing $P(z)$ by $z - 2$ we obtain $z^2 + z + 1$. The remaining two roots of $P(z) = 0$ are the roots of $z^2 + z + 1 = 0$. Solving this quadratic equation we find that $z = (-1 \pm i\sqrt{3})/2$, so the three roots of the equation are 2, $(-1 + i\sqrt{3})/2$, and $(-1 - i\sqrt{3})/2$. ■

For more background information and examples on complex numbers, the complex plane and roots of complex numbers, see Chapter 1 of reference [6.1], Sections 1.1 to 1.5 of reference [6.4], and Chapter 1 of reference [6.6].

EXERCISES 1.5

In Exercises 1 through 8 find the square roots of the given complex number by using result (10), and then confirm the result by using the formula for finding the square root of a complex mumber.

1. $-1+i$.
2. $3+2i$.
3. i.
4. $-1+4i$.
5. $2-3i$.
6. $-2-i$.
7. $4-3i$.
8. $-5+i$.

In Exercises 9 through 14 find the roots of the given complex number.

9. $(1+i\sqrt{3})^{1/3}$.
10. $i^{1/4}$.
11. $(-1)^{1/4}$.
12. $(-1-i)^{1/3}$.
13. $(-i)^{1/3}$.
14. $(4+4i)^{1/4}$.

15. Find the roots of $z^3+z(i-1)=0$.

16. Find the roots of $z^3+iz/(1+i)=0$.

17. Use result (12) to show that

$$1+\cos(2\pi/n)+\cos(4\pi/n)+\cdots$$
$$+\cos[(2(n-1)\pi/n)]=0$$

and

$$\sin(2\pi/n)+\sin(4\pi/n)+\cdots+\sin[(2(n-1)\pi/n)]=0.$$

18. Use Theorem 1.1 and the representation $z=re^{i\theta}$ to prove that if a and b are any two arbitrary complex numbers, then $\overline{ab}=\overline{a}\overline{b}$ and $\overline{(a^r)}=(\overline{a})^r$.

19. Given $z=1$ is a zero of the polynomial $P(z)=z^3-5z^2+17z-13$, find its other two zeros and verify that they are complex conjugates.

20. Given that $z=-2$ is a zero of the polynomial $P(z)=z^5+2z^4-4z-8$, find its other four zeros and verify that they occur in complex conjugate pairs.

21. Find the two zeros of the quadratic $P(z)=z^2-1+i$, and explain why they do not occur as a complex conjugate pair.

1.6 Partial Fractions

Let $N(x)$ and $D(x)$ be two polynomials. Then a **rational function** of x is any function of the form $N(x)/D(x)$. The method of **partial fractions** involves the decomposition of rational functions into an equivalent sum of simpler terms of the type

$$\frac{P_1}{ax+b},\ \frac{P_2}{(ax+b)^2},\ldots \quad \text{and} \quad \frac{Q_1x+R_1}{Ax^2+Bx+C},\ \frac{Q_2x+R_2}{(Ax^2+Bx+C)^2},\ldots,$$

where the coefficients are all real together with, possibly, a polynomial in x.

The steps in the reduction of a rational function to its partial fraction representation are as follows:

STEP 1 Factorize $D(x)$ into a product of linear factors and quadratic factors with real coefficients with complex roots, called **irreducible** factors. This is the hardest step, and real quadratic factors will only arise when $D(x)=0$ has pairs of complex conjugate roots (see Theorem 1.2). Use the result to express $D(x)$ in the form

$$D(x)=(a_1x+b_1)^{r_1}\ldots(a_mx+b_m)^{r_m}\left(A_1x^2+B_1x+C_1\right)^{s_1}$$
$$\ldots\left(A_kx^2+B_kx+C_k\right)^{s_k},$$

where r_i is the number of times the linear factor (a_ix+b_i) occurs in the factorization of $D(x)$, called its **multiplicity**, and s_j is the corresponding multiplicity of the quadratic factor $(A_jx^2+B_jx+C_j)$.

STEP 2 Suppose first that the degree n of the numerator is *less* than the degree d of the denominator. Then, to every different linear factor $(ax + b)$ with multiplicity r, include in the partial fraction expansion the terms

$$\frac{P_1}{(ax+b)} + \frac{P_2}{(ax+b)^2} + \cdots + \frac{P_r}{(ax+b)^r},$$

where the constant coefficients P_i are unknown at this stage, and so are called **undetermined coefficients**.

partial fraction undetermined coefficients

STEP 3 To every quadratic factor $(Ax^2 + Bx + C)^s$ with multiplicity s include in the partial fraction expansion the terms

$$\frac{Q_1 x + R_1}{(Ax^2+Bx+C)} + \frac{Q_2 x + R_2}{(Ax^2+Bx+C)^2} + \cdots + \frac{Q_s x + R_s}{(Ax^2+Bx+C)^s},$$

where the Q_j and R_j for $j = 1, 2, \ldots, s$ are undetermined coefficients.

STEP 4 Take as the partial fraction representation of $N(x)/D(x)$ the sum of all the terms in Steps 2 and 3.

STEP 5 Multiply the expression

$$N(x)/D(x) = \text{Partial fraction representation in Step 4}$$

by $D(x)$, and determine the unknown coefficients by equating the coefficients of corresponding powers of x on either side of this expression to make it an identity (that is, true for all x).

STEP 6 Substitute the values of the coefficients determined in Step 5 into the expression in Step 4 to obtain the required partial fraction representation.

STEP 7 If $n \geq d$, use long division to divide the denominator into the numerator to obtain the sum of a polynomial of degree $n - d$ of the form

$$T_0 + T_1 x + T_2 x^2 + \cdots + T_{n-d} x^{n-d},$$

together with a remainder term in the form of a rational function $R(x)$ of the type just considered. Find the partial fraction representation of the rational function $R(x)$ using Steps 1 to 6. The required partial fraction representation is then the sum of the polynomial found by long division and the partial fraction representation of $R(x)$.

EXAMPLE 1.15 Find the partial fraction representations of

$$\text{(a) } F(x) = \frac{x^2}{(x+1)(x-2)(x+3)} \quad \text{and} \quad \text{(b) } F(x) = \frac{2x^3 - 4x^2 + 3x + 1}{(x-1)^2}.$$

Solution **(a)** All terms in the denominator are linear factors, so by Step 1 the appropriate form of partial fraction representation is

$$\frac{x^2}{(x+1)(x-2)(x+3)} = \frac{A}{x+1} + \frac{B}{x-2} + \frac{C}{x+3}.$$

Cross multiplying, we obtain

$$x^2 = A(x-2)(x+3) + B(x+1)(x+3) + C(x+1)(x-2).$$

Setting $x = -1$ makes the terms in B and C vanish and gives $A = -1/6$. Setting $x = 2$ makes the terms in A and C vanish and gives $B = 4/15$, whereas setting $x = -3$ makes the terms in A and B vanish and gives $C = 9/10$, so

$$\frac{x^2}{(x+1)(x-2)(x+3)} = \frac{-1}{6(x+1)} + \frac{4}{15(x-2)} + \frac{9}{10(x+3)}.$$

(b) The degree of the numerator exceeds that of the denominator, so from Step 7 it is necessary to start by dividing the denominator into the numerator longhand to obtain

$$\frac{2x^3 - 4x^2 + x + 3}{(x-1)^2} = 2x + \frac{3-x}{(x-1)^2}.$$

We now seek a partial fraction representation of $(3-x)/(x-1)^2$ by using Step 1 and writing

$$\frac{3-x}{(x-1)^2} = \frac{A}{x-1} + \frac{B}{(x-1)^2}.$$

When we multiply by $(x-1)^2$, this becomes

$$3 - x = A(x-1) + B.$$

Equating the constant terms gives $3 = -A + B$, whereas equating the coefficients of x gives $-1 = A$ so that $B = 2$. Thus, the required partial fraction representation is

$$\frac{2x^3 - 4x^2 + x + 3}{(x-1)^2} = 2x + \frac{1}{1-x} + \frac{2}{(x-1)^2}.$$ ■

An examination of the way the undetermined coefficients were obtained in (a) earlier, where the degree of the numerator is less than that of the denominator and linear factors occur in the denominator, leads to a simple rule for finding the undetermined coefficients called the "cover-up rule."

The cover-up rule

Let a partial fraction decomposition be required for a rational function $N(x)/D(x)$ in which the degree of the numerator $N(x)$ is less than that of the denominator $D(x)$ and, when factored, let $D(x)$ contain some linear factors (factors of degree 1).

Let $(x-\alpha)$ be a linear factor of $D(x)$. Then the unknown coefficient K in the term $K/(x-\alpha)$ in the partial fraction decomposition of $N(x)/D(x)$ is obtained by "covering up" (ignoring) all of the other terms in the partial fraction expansion, multiplying the remaining expression $N(x)/D(x) = K/(x-\alpha)$ by $(x-\alpha)$, and then determining K by setting $x = \alpha$ in the result.

To illustrate the use of this rule we use it in case (a) given earlier to find A from the representation

$$\frac{x^2}{(x+1)(x-2)(x+3)} = \frac{A}{x+1} + \frac{B}{x-2} + \frac{C}{x+3}.$$

We "cover up" (ignore) the terms involving B and C, multiply through by $(x+1)$, and find A from the result

$$\frac{x^2}{(x-2)(x+3)} = A$$

by setting $x = -1$, when we obtain $A = -1/6$. The undetermined coefficients B and C follow in similar fashion.

Once a partial fraction representation of a function has been obtained, it is often necessary to express any quadratic $x^2 + px + q$ that occurs in a denominator in the form $(x + A)^2 + B$, where A and B may be either positive or negative real numbers. This is called **completing the square**, and it is used, for example, when integrating rational functions and when finding inverse Laplace transforms.

completing the square

To find A and B we set

$$\begin{aligned} x^2 + px + q &= (x + A)^2 + B \\ &= x^2 + 2Ax + A^2 + B, \end{aligned}$$

and to make this an identity we now equate the coefficients of corresponding powers of x on either side of this expression:

(coefficients of x^2)	$1 = 1$ (this tells us nothing)
(coefficients of x)	$p = 2A$
(constant terms)	$q = A^2 + B$.

Consequently $A = (1/2)p$ and $B = q - (1/4)p^2$, and so the result obtained by completing the square is

$$x^2 + px + q = [x + (1/2)p]^2 + q - (1/4)p^2.$$

If the more general quadratic $ax^2 + bx + c$ occurs, all that is necessary to reduce it to this same form is to write it as

$$ax^2 + bx + c = a[x^2 + (b/a)x + c/a],$$

and then to complete the square using $p = b/a$ and $q = c/a$.

EXAMPLE 1.16

Complete the square in the following expressions:

(a) $x^2 + x + 1$.
(b) $x^2 + 4x$.
(c) $3x^2 + 2x + 1$.

Solution **(a)** $p = 1, q = 1$, so $A = 1/2$, $B = 3/4$, and hence

$$x^2 + x + 1 = (x + 1/2)^2 + 3/4.$$

(b) $p = 4,\ q = 0$, so $A = 2,\ B = -4$, and hence

$$x^2 + 4x = (x + 2)^2 - 4.$$

(c) $3x^2 + 2x + 1 = 3[x^2 + (2/3)x + 1/3]$ and so $p = 2/3$, $q = 1/3$, from which it follows that $A = 1/3$ and $B = 2/9$, so

$$3x^2 + 2x + 1 = 3\{(x + 1/3)^2 + 2/9\}.$$

Further information and examples of partial fractions can be found in any one of references [1.1] to [1.7].

EXERCISES 1.6

Express the rational functions in Exercises 1 through 8 in terms of partial fractions using the method of Section 1.6, and verify the results by using computer algebra to determine the partial fractions.

1. $(3x+4)/(2x^2+5x+2)$.
2. $(x^2+3x+5)/(2x^2+5x+3)$.
3. $(3x-7)/(2x^2+9x+10)$.
4. $(x^2+3x+2)/(x^2+2x-3)$.
5. $(x^3+x^2+x+1)/[(x+2)^2(x^2+1)]$.
6. $(x^2-1)/(x^2+x+1)$.
7. $(x^3+x^2+x+1)/\{(x+2)^2(x+1)\}$.
8. $(x^2+4)/(x^3+3x^2+3x+1)$.

Complete the square in Exercises 9 through 14.

9. x^2+4x+5.
10. x^2+6x+7.
11. $2x^2+3x-6$.
12. $4x^2-4x-3$.
13. $2-2x+9x^2$.
14. $2+2x-x^2$.

1.7 Fundamentals of Determinants

A **determinant** of **order** $\boldsymbol{n}$ is a single number associated with an array $\mathbf{A}$ of n^2 numbers arranged in n rows and n columns. If the number in the ith row and jth column of a determinant is a_{ij}, the determinant of $\mathbf{A}$, denoted by det $\mathbf{A}$ and sometimes by $|\mathbf{A}|$, is written

$$\det \mathbf{A} = |\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \tag{13}$$

It is customary to refer to the entries a_{ij} in a determinant as its *elements*. Notice the use of vertical bars enclosing the array $\mathbf{A}$ in the notation $|\mathbf{A}|$ for the *determinant* of $\mathbf{A}$, as opposed to the use of the square brackets in $[\mathbf{A}]$ that will be used later to denote the *matrix* associated with an array $\mathbf{A}$ of quantities in which the number of rows need not be equal to the number of columns.

The value of a first order determinant det $\mathbf{A}$ with the single element a_{11} is defined as a_{11} so that $\det[a_{11}] = a_{11}$ or, in terms of the alternative notation for a determinant, $|a_{11}| = a_{11}$. This use of the notation $|.|$ to signify a determinant should not be confused with the notation used to signify the absolute value of a number.

The second order determinant associated with an array of elements containing two rows and two columns is defined as

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \tag{14}$$

so, for example, using the alternative notation for a determinant we have

$$\begin{vmatrix} 9 & 3 \\ -7 & -4 \end{vmatrix} = 9(-4) - (-7)3 = -15.$$

Notice that **interchanging** two rows or columns of a determinant changes its sign.

We now introduce the terms *minor* and *cofactor* that are used in connection with determinants of all orders, and to do so we consider the third order determinant

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}. \tag{15}$$

minors and cofactors

The **minor** M_{ij} associated with a_{ij}, the element in the ith row and jth column of $\det \mathbf{A}$, is defined as the second order determinant obtained from $\det \mathbf{A}$ by deleting the elements (numbers) in its ith row and jth column. The **cofactor** C_{ij} of an element in the ith row and jth column of the $\det \mathbf{A}$ in (15) is defined as the **signed minor** using the rule

$$C_{ij} = (-1)^{i+j} M_{ij}. \tag{16}$$

With these ideas in mind, the determinant $\det \mathbf{A}$ in (15) is defined as

$$\begin{aligned}\det \mathbf{A} &= \sum_{j=1}^{3} a_{1j}(-1)^{1+j} \det M_{1j} \\ &= a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13}.\end{aligned}$$

If we introduce the cofactors C_{ij}, this last result can be written

$$\det \mathbf{A} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}, \tag{17}$$

and more concisely as

$$\det \mathbf{A} = \sum_{j=1}^{3} a_{1j}C_{1j}. \tag{18}$$

Result (18), or equivalently (17), will be taken as the definition of a third order determinant.

EXAMPLE 1.17 Evaluate the determinant

$$\begin{vmatrix} 1 & 3 & -3 \\ 2 & 1 & 0 \\ -2 & 1 & 1 \end{vmatrix}.$$

Solution

The minor $M_{11} = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = (1)(1) - (0)(1) = 1$, so the cofactor $C_{11} = (-1)^{(1+1)}M_{11} = 1$.
The minor $M_{12} = \begin{vmatrix} 2 & 0 \\ -2 & 1 \end{vmatrix} = (2)(1) - (0)(-2) = 2$, so the cofactor $C_{12} = (-1)^{(1+2)}M_{12} = -2$.
The minor $M_{13} = \begin{vmatrix} 2 & 1 \\ -2 & 1 \end{vmatrix} = (2)(1) - (1)(-2) = 4$, so the cofactor $C_{13} = (-1)^{(1+3)}M_{13} = 4$.

Using (17) we have

$$\begin{vmatrix} 1 & 3 & -3 \\ 2 & 1 & 0 \\ -2 & 1 & 1 \end{vmatrix} = (1)C_{11} + (3)C_{12} + (-3)C_{13} = (1)(1) + (3)(-2) + (-3)(4) = -17.$$

■

When expanded, (17) becomes

$$\det \mathbf{A} = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22},$$

and after regrouping these terms in the form

$$\det \mathbf{A} = -a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} + a_{22}a_{11}a_{33} - a_{22}a_{31}a_{13} - a_{23}a_{11}a_{32} + a_{23}a_{31}a_{12},$$

we find that

$$\det \mathbf{A} = a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23}.$$

Proceeding in this manner, we can easily show that det **A** may be obtained by forming the sum of the products of the elements of **A** and their cofactors in *any* row or column of det **A**. These results can be expressed symbolically as follows.

Expanding in terms of the elements of the *i*th row:

$$\det \mathbf{A} = a_{i1}C_{i1} + a_{i2}C_{i2} + a_{i3}C_{i3} = \sum_{j=1}^{3} a_{ij}C_{ij}. \tag{19}$$

Laplace expansion theorem

Expanding in terms of the elements of the *j*th column:

$$\det \mathbf{A} = a_{1j}C_{1j} + a_{2j}C_{2j} + a_{3j}C_{3j} = \sum_{i=1}^{3} a_{ij}C_{ij}. \tag{20}$$

Results (19) and (20) are the form taken by the **Laplace expansion theorem** when applied to a third order determinant. The extension of the theorem to determinants of any order will be made later in Chapter 3, Section 3.3.

EXAMPLE 1.18

Expand the following determinant (a) in terms of elements of its first row, and (b) in terms of elements of its third column:

$$|\mathrm{A}| = \begin{vmatrix} 1 & 2 & 4 \\ 1 & 0 & 2 \\ 1 & 2 & 1 \end{vmatrix}.$$

Solution **(a)** Expanding in terms of the elements of the first row requires the three cofactors $C_{11} = M_{11}$, $C_{12} = -M_{12}$, and $C_{13} = M_{13}$, where

$$M_{11} = \begin{vmatrix} 0 & 2 \\ 2 & 1 \end{vmatrix} = -4, \quad M_{12} = \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = -1, \quad M_{13} = \begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix} = 2,$$

so $C_{11} = (-1)^{(1+1)}(-4) = -4$, $C_{12} = (-1)^{(1+2)}(-1) = 1$, $C_{13} = (-1)^{(1+3)}(2) = 2$, and so

$$|\mathrm{A}| = (1)(-4) + (2)(1) + (4)(2) = 6.$$

(b) Expanding in terms of the elements of the third column requires the three cofactors $C_{13} = M_{13}$, $C_{23} = -M_{23}$, and $C_{33} = M_{33}$, where

$$M_{13} = \begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix} = 2, \quad M_{23} = \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = 0, \quad M_{33} = \begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} = -2,$$

so $C_{13} = (-1)^{(1+3)}(2) = 2$, $C_{23} = 0$, $C_{33} = (-1)^{(3+3)}(-2) = -2$ and so

$$|\mathrm{A}| = (4)(2) + (2)(0) + (1)(-2) = 6.$$ ■

Two especially simple *third order* determinants are of the form

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix} \quad \text{and} \quad \det \mathbf{A} = \begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

The first of these determinants has only zero elements below the diagonal line drawn from its top left element to its bottom right one, and the second determinant has only zero elements above this line. This diagonal line in every determinant is called the **leading diagonal**. The value of each of the preceding determinants is easily seen to be given by the product $a_{11}a_{22}a_{33}$ of the terms on its leading diagonal.

Simpler still in form is the third order determinant

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33},$$

whose value $a_{11}a_{22}a_{33}$ is again the product of the elements on the leading diagonal.

For another approach to the elementary properties of determinants, see Appendix A16 of reference [1.2], and Chapter 2 of reference [2.1].

EXERCISES 1.7

Evaluate the determinants in Exercises 1 through 6 (a) in terms of elements of the first row and (b) in terms of elements of the second column.

1. $\begin{vmatrix} 1 & 5 & 7 \\ 1 & -1 & 1 \\ 1 & 2 & 1 \end{vmatrix}$.

2. $\begin{vmatrix} 2 & 1 & -1 \\ 2 & 6 & -1 \\ 5 & 1 & -1 \end{vmatrix}$.

3. $\begin{vmatrix} 5 & 2 & 4 \\ 1 & 2 & 1 \\ 3 & 1 & 5 \end{vmatrix}$.

4. $\begin{vmatrix} -1 & 3 & 6 \\ 2 & 1 & 4 \\ -1 & 3 & 1 \end{vmatrix}$.

5. $\begin{vmatrix} 1 & 0 & -6 \\ 2 & 1 & 3 \\ 4 & 3 & 21 \end{vmatrix}$.

6. $\begin{vmatrix} 1 & 5 & -1 \\ 2 & 1 & -3 \\ -4 & 1 & 1 \end{vmatrix}$.

7. On occasion the elements of a matrix may be functions, in which case the determinant may be a function. Evaluate the *functional* determinant

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & \sin x & -\cos x \\ 0 & \cos x & \sin x \end{vmatrix}.$$

8. Determine the values of λ that make the following determinant vanish:

$$\begin{vmatrix} 3-\lambda & 2 & 2 \\ 2 & 2-\lambda & 0 \\ 2 & 0 & 4-\lambda \end{vmatrix}.$$

Hint: This is a polynomial in λ of degree 3.

9. A matrix is said to be **transposed** if its first row is written as its first column, its second row is written as its second column . . . , and its last row is written as its last column. If the determinant is $|\mathbf{A}|$, the determinant of $\mathbf{A}^{\mathrm{T}}$, the transpose matrix $\mathbf{A}$, is denoted by $|\mathbf{A}^{\mathrm{T}}|$. Write out the expansion of $|\mathbf{A}|$ using (17) and reorder the terms to show that

$$|\mathbf{A}| = |\mathbf{A}^{\mathrm{T}}|.$$

10. Use elimination to solve the system of linear equations

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

for x_1 and x_2, in which not both b_1 and b_2 are zero, and show that the solution can be written in the form

$$x_1 = D_1/|\mathbf{A}| \quad \text{and} \quad x_2 = D_2/|\mathbf{A}|, \quad \text{provided } |\mathbf{A}| \neq 0,$$

where $|\mathbf{A}|$ is the determinant of the matrix of *coefficients* of the system

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad D_1 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, \quad \text{and } D_2 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}.$$

Notice that D_1 is obtained from $|\mathbf{A}|$ by replacing its *first* column by b_1 and b_2, whereas D_2 is obtained from $|\mathbf{A}|$ by replacing its *second* column by b_1 and b_2. This is **Cramer's rule** for a system of two simultaneous equations. Use this method to find the solution of

$$x_1 + 5x_2 = 3$$
$$7x_1 - 3x_2 = -1.$$

11. Repeat the calculation in Exercise 10 using the system of equations

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3,$$

in which not all of b_1, b_2, and b_3 are zero, and show that provided $|\mathbf{A}| \neq 0$,

$$x_1 = D_1/|\mathbf{A}|, \quad x_2 = D_2/|\mathbf{A}|, \quad \text{and} \quad x_3 = D_3/|\mathbf{A}|,$$

where $|\mathbf{A}|$ is the determinant of the matrix of coefficients and D_i is the determinant obtained from $|\mathbf{A}|$ by replacing its ith column by b_1, b_2, and b_3 for $i = 1, 2, 3$. This is **Cramer's rule** for a system of three simultaneous equations, and the method generalizes to a system of n linear equations in n unknowns. Use this method to find the solution of

$$x_1 + 2x_2 - x_3 = 2$$
$$x_1 - 3x_2 - 2x_3 = -1$$
$$2x_1 + x_2 + 2x_3 = 1.$$

1.8 Continuity in One or More Variables

If the function $y = f(x)$ is defined in the interval $a \leq x \leq b$, the interval is called the **domain of definition** of the function. The function f is said to have a **limit** at a point c in $a \leq x \leq b$, written $\lim_{x \to c} f(x) = L$, if for every arbitrarily small number $\varepsilon > 0$ there exists a number $\delta > 0$ such that

$$|f(x) - L| < \varepsilon \quad \text{when } |x - c| < \delta. \tag{21}$$

This technical definition means that as x either increases toward c and becomes arbitrarily close to it, or decreases toward c and becomes arbitrarily close to it, so $f(x)$ approaches arbitrarily close to the value L. Notice that it is not necessary for $f(x)$ to be defined at $x = c$, or, if it is, that $f(c)$ assumes the value L. If $f(x)$ has a limit L as $x \to c$ and in addition $f(c) = L$, so that

$$\lim_{x \to c} f(x) = f(c) = L, \tag{22}$$

then the function f is said to be **continuous** at c. It must be emphasized that in this definition of continuity the limiting operation $x \to c$ must be true as x tends to c from *both* the left and right. It is convenient to say that x approaches c from the *left* when it increases toward c and, correspondingly, to say that x approaches c from the *right* when it decreases toward it.

continuity from the right

The function f is **continuous from the right** at $x = c$ if

$$\lim_{x \to c^+} f(x) = f(c), \tag{23}$$

continuity from the left

where the notation $x \to c^+$ means that x decreases toward c, causing x to tend to c from the *right*. Similarly, f is **continuous from the left** at $x = c$ if

$$\lim_{x \to c^-} f(x) = f(c), \tag{24}$$

continuity at $x = c$

where now $x \to c^-$ means that x increases toward c, causing x to tend to c from the *left*. The relationship among definitions (22), (23), and (24) is that f is continuous at the point c if

$$\lim_{x \to c^-} f(x) = \lim_{x \to c^+} f(x) = f(c). \tag{25}$$

When expressed in words, this says that f is continuous at $x = c$ if the limits of f as x tends to c from both the left and right exist and, furthermore, the limits equal the functional value $f(c)$.

continuous function

A function f that is continuous at all points of $a \leq x \leq b$ is said to be a **continuous function** on that interval. Graphically, a continuous function on $a \leq x \leq b$ is a function whose graph is unbroken but not necessarily smooth. A function f is said

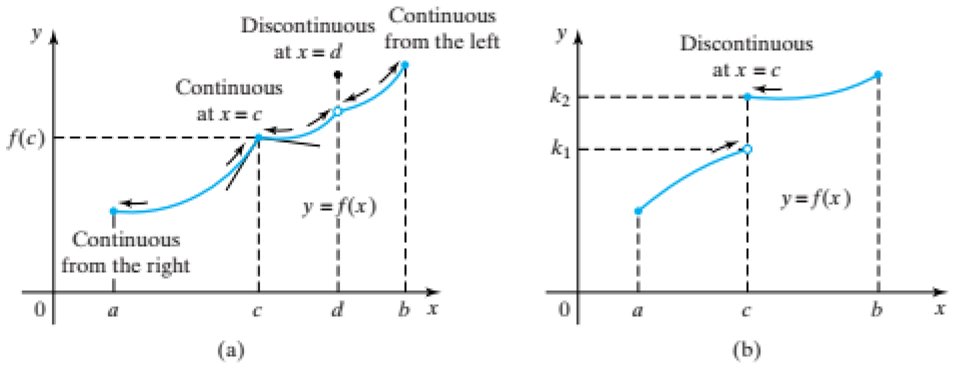

FIGURE 1.8 (a) A continuous function for $a < x < b$. (b) A discontinuous function.

smooth function

to be **smooth** over an interval if at each point of the graph the tangent lines to the left and right of the point are the same. Figure 1.8a shows the graph of a continuous function that is smooth over the intervals $a \le x < c$ and $c < x < b$ but has different tangent lines to the immediate left and right of $x = c$ where the function is *not* smooth. A function such as this is said to be **continuous and piecewise smooth** over the interval $a \le x \le b$.

continuous and piecewise smooth function

discontinuous function

A function f is said to be **discontinuous** at a point c if it is not continuous there. For a jump discontinuity we have

$$\lim_{x \to c^-} f(x) = k_1 \quad \text{and} \quad \lim_{x \to c^+} f(x) = k_2, \quad \text{but } k_1 \neq k_2. \tag{26}$$

A function f is said to have a **removable discontinuity** at a point c if $k_1 = k_2$ in (26), but $f(c) \neq k_1$, as at the point c_2 in Fig. 1.9.

An example of a discontinuous function is shown in Fig. 1.8b where a jump discontinuity occurs at $x = c$.

piecewise continuity

A function f is said to be **piecewise continuous** on an interval $a \le x \le b$ if it is continuous on a finite number of adjacent subintervals, but discontinuous at the end points of the subintervals, as shown in Fig. 1.9.

The notion of continuity of a function of several variables is best illustrated by considering a function $f(x, y)$ of the two independent variables x and y. The function f defined in some region of the (x, y)-plane D, say, is said to be **continuous**

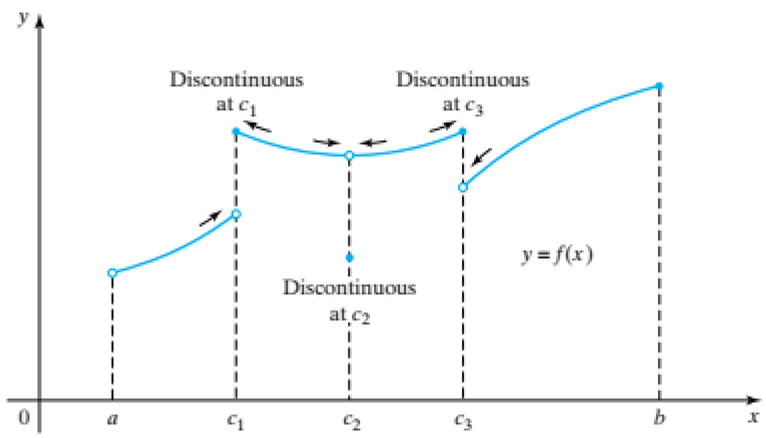

FIGURE 1.9 A piecewise continuous function.

at the point (a, b) in D if

continuity of $f(x, y)$

$$\lim_{x \to a, y \to b} f(x, y) = f(a, b), \tag{27}$$

and to be discontinuous otherwise.

In this definition of continuity, it is important to recognize that a general point P at (x, y) is allowed to tend to the point (a, b) in D along *any* path in the (x, y)-plane that lies in D. Expressed differently, f will only be continuous at (a, b) if the limit in (27) is independent of the way in which the point (x, y) approaches the point (a, b). When this is true for all points in D, the function f is said to be **continuous** in D.

discontinuity of $f(x, y)$

The function f is, for instance, **discontinuous** at (a, b) if

$$\lim_{x \to a, y \to b} f(x, y) = k, \quad \text{but } f(a, b) \neq k.$$

Sufficient for showing that a function f is discontinuous at a point (a, b) is by demonstrating that two *different* limiting values of f are obtained if the point P at (x, y) is allowed to tend to (a, b) along two *different* straight-line paths. This approach can be used to show that the function

$$f(x, y) = \frac{xy}{x^2 + a^2 y^2}$$

has no limit at the origin. If we allow the point P at (x, y) to tend to the origin along the straight line $y = kx$, with k an arbitrary constant, the function f becomes

$$f(x, kx) = \frac{k}{1 + a^2 k^2},$$

and it is seen from this that f is constant along each such line. However, the value of f on each line, and hence at the origin, depends on k, so f has no limit at the origin and so is discontinuous at that point, though f is defined and continuous at all other points of the (x, y)-plane.

An example of a function $f(x, y)$ that is continuous everywhere except at points along a curve Γ in the (x, y)-plane is shown in Fig. 1.10.

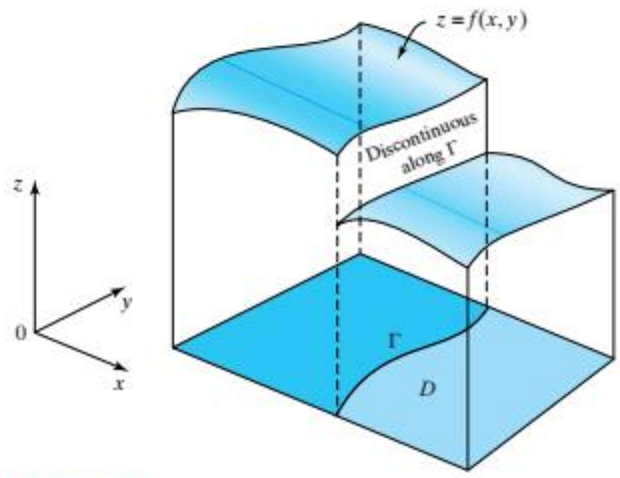

FIGURE 1.10 A function $f(x, y)$ continuous everywhere except at points on Γ.

The extension of these definitions to functions of n variables is immediate and so will not be discussed.

Discussions on continuity and its consequences can be found in any one of references [1.1] to [1.7].

1.9 Differentiability of Functions of One or More Variables

The function $f(x)$ defined in $a \le x \le b$ is said to be **differentiable** with the **derivative** $f'(c)$ at a point c inside the interval if the following limit exists:

$$\lim_{h\to 0} \frac{f(c+h) - f(c)}{h} = f'(c). \tag{28}$$

differentiability of $f(x)$

Here, as in the definition of continuity, for f to be differentiable at point c the limit must remain unchanged as h tends to zero through both positive and negative values. The function f is said to be **differentiable** in the interval $a \le x \le b$ if it is differentiable at every point in the interval. When f is differentiable at a point c with derivative $f'(c)$, the number $f'(c)$ is the gradient, or slope, of the tangent line to the graph at the point $(c, f(c))$. A function with a continuous derivative throughout an interval is said to be a **smooth** function over the interval. The function f will be said to be **nondifferentiable** at any point c where the limit in (28) does not exist.

left- and right-hand derivatives of $f(x)$

Even when a function f is nondifferentiable at a point, it is possible that a special form of derivative can still be defined to the left and right of the point if the requirement that the limit in (28) exists as $h \to 0$ through both positive and negative values is relaxed. The function f has a **right-hand derivative** at a if the limit

$$\lim_{h\to 0+} \frac{f(a+h) - f(a)}{h} \tag{29}$$

exists, and a **left-hand derivative** at b if the limit

$$\lim_{h\to 0-} \frac{f(b+h) - f(b)}{h}. \tag{30}$$

exists.

When c is a specific point, $f'(c)$ is a *number*, but when x is a variable, $f'(x)$ becomes a function. Left- and right-hand derivatives are illustrated in Fig. 1.11. An important consequence of differentiability is that **differentiability implies continuity**, but the converse is not true.

first order partial derivatives of $f(x, y)$

The **first order partial derivative with respect to x** of the function $f(x, y)$ of the two independent variables x and y at the point (a, b) is the number defined by

$$\lim_{h\to 0} \frac{f(a+h, b) - f(a, b)}{h}, \tag{31}$$

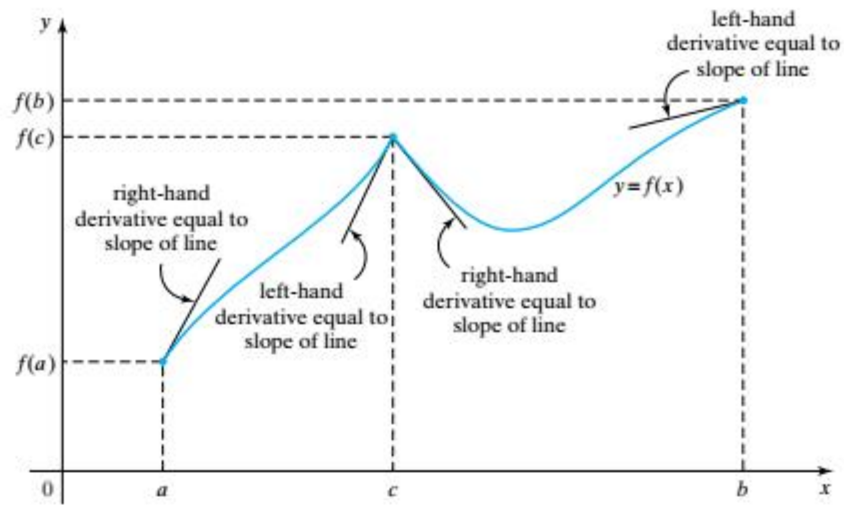

FIGURE 1.11 Left- and right-hand derivatives as tangent lines.

provided the limit exists. The value of this partial derivative is denoted either by $\partial f/\partial x$ at (a, b), or by $f_x(a, b)$. The corresponding partial derivative at a general point (x, y) is the function $f_x(x, y)$.

Similarly, the **first order partial derivative with respect to y** of the function $f(x, y)$ at the point (a, b) is the number defined by the limit

$$\lim_{k \to 0} \frac{f(a, b+k) - f(a, b)}{k}, \tag{32}$$

provided the limit exists. The value of this partial derivative is denoted either by $\partial f/\partial y$ at (a, b), or by $f_y(a, b)$. At a general point (x, y) this partial derivative becomes the function $f_y(x, y)$. Higher order partial derivatives are defined in a similar fashion leading, for example, to the **second order partial derivatives**

second order partial derivatives of $f(x, y)$

$$\partial^2 f/\partial x^2 = \partial/\partial x(\partial f/\partial x),\ \partial^2 f/\partial y^2 = \partial/\partial y(\partial f/\partial y),$$
$$\partial^2 f/\partial x\partial y = \partial/\partial y(\partial f/\partial x), \quad \text{and} \quad \partial^2 f/\partial y\partial x = \partial/\partial x(\partial f/\partial y).$$

A more compact notation for these same derivatives is

f_{xx}, f_{yy}, f_{xy}, and f_{yx}, so that, for example $f_{yx} = \partial^2 f/\partial y\partial x$ and $f_{yy} = \partial^2 f/\partial y^2$.

mixed partial derivatives

The derivatives f_{xy} and f_{yx} are called **mixed partial derivatives**, and their relationship forms the statement of the next theorem, the proof of which can be found in any one of references [1.1] to [1.7].

THEOREM 1.3

Equality of mixed partial derivatives Let f, f_x, f_{xy}, and f_{yx} all be defined and continuous at a point (a, b) in a region. Then

$$f_{xy}(a, b) = f_{yx}(a, b).$$

■

This result, given conditions for the *equality* of mixed partial derivatives, is an important one, and use will be made of it on numerous occasions as, for example, in Chapter 18 when second order partial differential equations are considered.

total differential

If $z = f(x, y)$, the **total differential** dz of f is defined as

$$dz = (\partial f/\partial x)\,dx + (\partial f/\partial y)\,dy, \tag{33}$$

where dz, dx, and dy are *differentials*. Here, a **differential** means a small quantity, and the differential dz is determined by (33) when the differentials dx and dy are specified. When $\partial f/\partial x$ and $\partial f/\partial y$ are evaluated at a specific point (a, b), result (33) provides a linear approximation to $f(x, y)$ near to the point (a, b). Although finite, the limits of the quotients of the differentials $dz \div dx$ and $dy \div dx$ as the differential $dx \to 0$ are such that they become the values of the derivatives dz/dx and dy/dx, respectively, at a point (x, y) where $\partial f/\partial x$ and $\partial f/\partial y$ are evaluated.

1.10 Tangent Line and Tangent Plane Approximations to Functions

tangent line approximation

Let $y = f(x)$ be defined in the interval $a \le x \le b$ and be differentiable throughout it. Then a **tangent line (linear) approximation** to f near a point x_0 in the interval is given by

$$y_T = f(x_0) + (x - x_0) f'(x_0). \tag{34}$$

This linear expression approximates the function f close to x_0 by the tangent to the graph of $y = f(x)$ at the point $(x_0, f(x_0))$.

This simple approximation has many uses; one will be in the Euler and modified Euler methods for solving initial value problems for ordinary differential equations developed in Chapter 19.

EXAMPLE 1.19

Find a tangent line approximation to $y = 1 + x^2 + \sin x$ near the point $x = \alpha$.

Solution Setting $x_0 = \alpha$ and substituting into (34) gives

$$y \approx 1 + \alpha^2 + \sin\alpha + (x - \alpha)(2\alpha + \cos\alpha) \text{ for } x \text{ close to } \alpha.$$

■

tangent plane approximation

Let the function $z = f(x, y)$ be defined in a region D of the (x, y)-plane where it possesses continuous first order partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$. Then a **tangent plane (linear) approximation** to f near any point (x_0, y_0) in D is given by

$$z_T = f(x_0, y_0) + (x - x_0) f_x(x_0, y_0) + (y - y_0) f_y(x_0, y_0). \tag{35}$$

This linear expression approximates the function f close to the point (x_0, y_0) by a plane that is tangent to the surface $z = f(x, y)$ at the point $(x_0, y_0, f(x_0, y_0))$. The tangent plane approximation in (35) is an immediate extension to functions of two variables of the tangent line approximation in (34), to which it simplifies when only one independent variable is involved.

Both of these approximations are derived from the appropriate Taylor series expansions of functions discussed in Section 1.12 by retaining only the linear terms.

EXAMPLE 1.20

Find the tangent plane approximation to the function $z = x^2 - 3y^2$ near the point $(1, 2)$.

Solution Setting $x_0 = 1$, $y_0 = 2$ and substituting into (35) gives

$$z \approx -11 + 2(x-1) - 12(y-2) \text{ for } (x, y) \text{ close to } (1, 2).$$

1.11 Integrals

indefinite and definite integrals

A differentiable function $F(x)$ is called an **antiderivative** of the function $f(x)$ on some interval if at each point of the interval $dF/dx = f(x)$. If $F(x)$ is any antiderivative of $f(x)$, the **indefinite integral** of $f(x)$, written $\int f(x)\,dx$, is

$$\int f(x)\,dx = F(x) + c,$$

where c is an arbitrary constant called the *constant of integration*. The function $f(x)$ is called the **integrand** of the integral. Thus, an indefinite integral is a function, and an antiderivative and an indefinite integral can only differ by an arbitrary additive constant.

The expression $\int_a^b f(x)\,dx$, called a **definite integral**, is a number and may be interpreted geometrically as the area between the graph of $f(x)$ and the lines $x = a$ and $x = b$, for $b > a$, with areas above the x-axis counted as positive and those below it as negative.

The relationship between definite integrals that are *numbers* and indefinite integrals that are *functions* is given in the next theorem, included in which is also the mean value theorem for integrals. See the references at the end of the chapter for proofs and further information.

THEOREM 1.4

Fundamental theorem of integral calculus and the mean value theorem for integrals If $F'(x)$ is continuous in the interval $a \le x \le b$, throughout which $F'(x) = f(x)$, then

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Another result is

$$\int_a^b f(x)\,dx = (b-a)\,f'(\xi),$$

if f is differentiable, where the number ξ, although unknown, lies in the interval $a < \xi < b$. In this form the result is called the **mean value theorem for integrals**.

An **improper integral** is a definite integral in which one or more of the following cases arises: (a) the integrand becomes infinite inside or at the end of the interval of integration, or (b) one (or both) of the limits of integration is infinite.

Types of Improper Integrals

Case (a)

convergence and divergence of improper integrals

If the integrand of an integral becomes infinite at a point c inside the interval of integration $a \leq x \leq b$ as shown in Fig. 1.12a, the improper integral is said to exist if the limits in (36) exist. When the improper integral exists it is said to **converge** to the (finite) value of the following limit:

$$\int_a^b f(x)\,dx = \lim_{h\to 0}\int_a^{c-h} f(x)\,dx + \lim_{k\to 0}\int_{c+k}^b f(x)\,dx. \tag{36}$$

In this definition $h > 0$ and $k > 0$ are allowed to tend to zero *independently* of each other. If, when the limit is taken, the integral is either infinite or indeterminate, the integral is said to **diverge**.

Cauchy principal value

Some integrals of this type diverge when h and k are allowed to tend to zero independently of each other, but converge when the limit is taken with $h = k$, in which case the result of the limit is called the **Cauchy principal value** of the integral. Integrals of this type arise frequently when certain types of definite integral are evaluated in the complex plane by means of contour integration (see Chapter 15, Section 15.5).

Case (b)

If a limit of integration in a definite integral is infinite, say the upper limit as shown in Fig. 1.12b, then, when it exists, the improper integral is said to **converge** to the value of the limit

$$\int_a^\infty f(x)\,dx = \lim_{R\to\infty}\int_a^R f(x)\,dx, \tag{37}$$

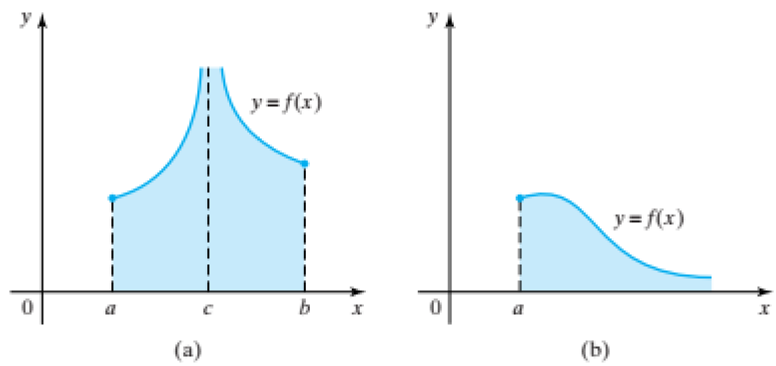

FIGURE 1.12 (a) $f(x)$ is infinite inside the interval of integration. (b) The interval of integration is infinite in length.

and the integral is **divergent** if the limit is either infinite or indeterminate. If both limits are infinite, the improper integral is said to **converge** to the value of the limit

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{R\to\infty, S\to\infty} \int_{-S}^{R} f(x)\,dx \tag{38}$$

when it exists, and the integral is said to be **divergent** if the limit is either infinite or indeterminate.

In (38) R and S are allowed to tend to infinity *independently* of each other. Integrals of this type also have Cauchy principal values if the foregoing process leads to divergence, but the integrals are convergent when the limit is taken with $R = S$. Integrals of this type also occur when certain real integrals are evaluated by means of contour integration (see Chapter 15, Section 15.5).

Elementary examples of convergent improper integrals of the types shown in (36) to (38) are

$$\int_0^1 \frac{x^p - x^{-p}}{x-1}\,dx = \frac{1}{p} - \pi \cot p\pi, \quad (p^2 < 1),$$

$$\int_0^{\infty} \exp(-x)\sin x\,dx = 1/2 \quad \text{and} \quad \int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \pi.$$

THEOREM 1.5

Differentiation under the integral sign — Leibniz' rule If $\xi(t)$, $\eta(t)$, $d\xi/dt$, $d\eta/dt$, $f(x,t)$, and $\partial f/\partial t$ are continuous for $t_0 \le t \le t_1$ and for x in the interval of integration, then

$$\frac{d}{dt}\int_{\xi(t)}^{\eta(t)} f(x,t)\,dx = \int_{\xi(t)}^{\eta(t)} \frac{\partial f(x,t)}{\partial t}\,dx + f(\eta(t),t)\frac{d\eta}{dt} - f(\xi(t),t)\frac{d\xi}{dt}. \quad \blacksquare$$

This theorem is used, for example, in Chapter 18 when discussing discontinuous solutions of a class of partial differential equations called *conservation laws*. Extensions of the theorem to functions of more variables are developed in Chapter 12, Section 12.3, where certain vector integral theorems are developed, and applications of the results of that section to fluid mechanics are to be found in Chapter 12, Section 12.4.

An application of Theorem 1.5 that is easily checked by direct calculation is

$$\frac{d}{dt}\int_{2t}^{t^2} (x^2+t)\,dx = \int_{2t}^{t^2} dx + (t^4+t)\cdot 2t - (4t^2+t)\cdot 2 = 2t^5 - 5t^2 - 4t.$$

A proof of Leibniz' rule can be found, for example, in Chapter 12 of reference [1.6].

1.12 Taylor and Maclaurin Theorems

THEOREM 1.6

Taylor's theorem for a function of one variable Let a function $f(x)$ have derivatives of all orders in the interval $a < x < b$. Then for each positive integer n and each x_0 in the interval

$$f(x) = f(x_0) + (x - x_0) f^{(1)}(x_0) + \frac{(x - x_0)^2}{2!} f^{(2)}(x_0) + \cdots + \frac{(x - x_0)^n}{n!} f^{(n)}(x_0) + R_{n+1}(x),$$

where $f^{(r)}(x) = d^r f/dx^r$, and the **remainder term** $R_{n+1}(x)$ is given by

$$R_{n+1}(x) = \frac{(x - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi),$$

for some ξ between x_0 and x. ■

Taylor polynomial

Taylor's theorem becomes the **Taylor series** for $f(x)$ when n is allowed to become infinite, and if the remainder term is neglected in Taylor's theorem the result is called the **Taylor polynomial approximation** to $f(x)$ of **degree** n. The Taylor polynomial of degree 1 is simply the tangent line approximation to f at x_0 given in (34).

Maclaurin's theorem

Taylor's theorem reduces to **Maclaurin's theorem** if $x_0 = 0$, and if we allow n to become infinite in Maclaurin's theorem, it becomes the **Maclaurin series** for $f(x)$.

A special case of Theorem 1.6 arises when Taylor's theorem is terminated with the term $R_1(x)$, corresponding to $n = 0$, because the result can be written

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\xi), \tag{39}$$

mean value theorem

with ξ between x_0 and x, and in this form it is called the **mean value theorem for derivatives** (see the last result of Theorem 1.4).

A Taylor series is an example of an infinite series called a **power series**, the general form of which is

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots. \tag{40}$$

In (40) the quantity x is a variable, the numbers a_i are the **coefficients** of the power series, the constant x_0 is called the **center** of the series, or the point about which the series is **expanded**, and unless otherwise stated, x, x_0, and the a_i are real numbers, so the power series is a function of x.

A power series is said to **converge** for a given value of x if the sum of the infinite series for this value of x is *finite*. If the sum is *infinite*, or is *not defined*, the power series will be said to **diverge** for that value of x. Power series converge in an interval $x_0 - R < x < x_0 + R$, where the number R is called the **radius of convergence** of the series. Expressions for R are derived in Section 15.1.

The interval $x_0 - R < x < x_0 + R$ is called the **interval of convergence** of the power series. A power series converges for all x inside the interval of convergence and diverges for all x outside it, and the series may, or may not, converge at the end points of the interval. The convergence properties of power series are shown diagramatically in Fig. 1.13, and results (40) and combining expressions for R with

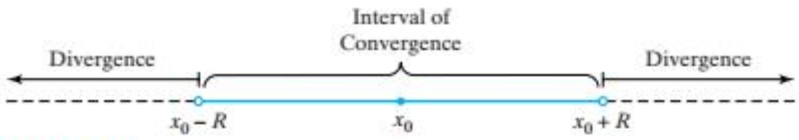

FIGURE 1.13 Interval of convergence of a power series with center x_0.

(40) gives the following theorem (see the references at the end of the chapter for real variable proofs of the following results and for more information).

THEOREM 1.7

Ratio test and *n*th root test for the convergence of power series The power series

$$\sum_{n=0}^{\infty} a_n(x - x_0)^n = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots$$

radius and interval of convergence

converges in the *interval of convergence* $x_0 - R < x < x_0 + R$, where the *radius of convergence* R is determined by either of the formulas

$$\text{(a) } R = 1/\lim_{n\to\infty} |a_{n+1}/a_n| \quad \text{or} \quad \text{(b) } R = 1/\lim_{n\to\infty} |a_n|^{1/n}.$$

The power series will diverge outside the interval of convergence, and its behavior at the ends of the interval of convergence must be determined separately. ■

A simple result on the convergence of a series that is often useful is the alternating series test. An **alternating series** is so named because the signs of successive terms of the series alternate in sign.

THEOREM 1.8

The alternating series test for convergence The alternating series $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ converges if $a_n > 0$ and $a_{n+1} < a_n$ for all n and $\lim_{n\to\infty} a_n = 0$. ■

The following theorem on the differentiation and integration of power series is often needed, and it is a real variable form of a result proved later in Chapter 15 when complex power series are studied.

THEOREM 1.9

Differentiation and integration of power series Let a power series have an interval of convergence $x_0 - R < x < x_0 + R$. Then the series may be differentiated and integrated term by term, and in each case the resulting series will have the same interval of convergence as the original series. In addition, within an interval of convergence common to any two power series, the series may be scaled by a constant and added or subtracted term by term and the resulting power series will have the same common interval of convergence. ■

The simplest form of Taylor's theorem for a function of two variables that finds many applications is given in the next theorem.

THEOREM 1.10

Taylor's theorem for a function of two variables Let $f(x, y)$ be defined for $a < x < b$ and $c < y < d$ and have continuous partial derivatives up to and including those of order 2. Then for x_0 and y_0 any points such that $a < x_0 < b$ and $c < y_0 < d$,

$$\begin{aligned} f(x, y) = {} & f(x_0, y_0) + (x - x_0) f_x(x_0, y_0) + (y - y_0) f_y(x_0, y_0) \\ & + \frac{1}{2!}\big[(x - x_0)^2 f_{xx}(x_0 + \xi, y_0 + \eta) + 2(x - x_0)(y - y_0) \\ & \times f_{xy}(x_0 + \xi, y_0 + \eta)(y - y_0)^2 f_{yy}(x_0 + \xi, y_0 + \eta)\big], \end{aligned}$$

where the numbers ξ and η are unknown, but ξ lies between x_0 and x and η lies between y_0 and y. ■

The group of second order partial derivatives in Theorem 1.10 forms the remainder term, and when these derivatives are ignored, the result reduces to the tangent plane approximation to $f(x, y)$ at the point (x_0, y_0) given in (35).

More information on Taylor's theorem and series can be found, for example, in reference [1.2].

1.13 Cylindrical and Spherical Polar Coordinates and Change of Variables in Partial Differentiation

Mathematical problems formulated using a particular coordinate system, such as cartesian coordinates, often need to be reexpressed in terms of a different coordinate system in order to simplify the task of finding a solution. When partial derivatives occur in the formulation of problems, it becomes necessary to know how they transform when a different coordinate system is used. The fundamental theorem governing the transformation of partial derivatives under a change of variables takes the following form (see the references at the end of the chapter for the proof of Theorem 1.11 and for more examples of its use).

THEOREM 1.11

Change of variables in partial differentiation Let $f(x_1, x_2, \ldots, x_n)$ be a differentiable function with respect to the n independent variables $x_1, x_2, \ldots, x_n$, and let the n new independent variables $u_1, u_2, \ldots, u_n$ be determined in terms of $x_1, x_2, \ldots, x_n$ by

$$x_1 = X_1(u_1, u_2, \ldots, u_n), \quad x_2 = X_2(u_1, u_2, \ldots, u_n), \ldots, \quad x_n = X_n(u_1, u_2, \ldots, u_n),$$

where $X_1, X_2, \ldots, X_n$ are differentiable functions of their arguments. Then, if as a result of the change of variables the function $f(x_1, x_2, \ldots, x_n)$ becomes the function $F(X_1, X_2, \ldots, X_n)$, and using chain rules we have

$$\begin{aligned} \frac{\partial F}{\partial u_1} &= \frac{\partial f}{\partial x_1}\frac{\partial X_1}{\partial u_1} + \frac{\partial f}{\partial x_2}\frac{\partial X_2}{\partial u_1} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial X_n}{\partial u_1} \\ \frac{\partial F}{\partial u_2} &= \frac{\partial f}{\partial x_1}\frac{\partial X_1}{\partial u_2} + \frac{\partial f}{\partial x_2}\frac{\partial X_2}{\partial u_2} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial X_n}{\partial u_2} \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{\partial F}{\partial u_n} &= \frac{\partial f}{\partial x_1}\frac{\partial X_1}{\partial u_n} + \frac{\partial f}{\partial x_2}\frac{\partial X_2}{\partial u_n} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial X_n}{\partial u_n}. \end{aligned} \tag{41}$$

■

To find higher order partial derivatives it is necessary to express the relationships between the *operations* of differentiation in the two coordinate systems, rather than between the actual derivatives themseves. This can be accomplished by rewriting the results of Theorem 1.11 in the form of **partial differential operators** as follows:

$$
\begin{aligned}
\frac{\partial}{\partial u_1} &\equiv \frac{\partial X_1}{\partial u_1}\frac{\partial}{\partial x_1} + \frac{\partial X_2}{\partial u_1}\frac{\partial}{\partial x_2} + \cdots + \frac{\partial X_n}{\partial u_1}\frac{\partial}{\partial x_n} \\
\frac{\partial}{\partial u_2} &\equiv \frac{\partial X_1}{\partial u_2}\frac{\partial}{\partial x_1} + \frac{\partial X_2}{\partial u_2}\frac{\partial}{\partial x_2} + \cdots + \frac{\partial X_n}{\partial u_2}\frac{\partial}{\partial x_n} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\frac{\partial}{\partial u_n} &\equiv \frac{\partial X_1}{\partial u_n}\frac{\partial}{\partial x_1} + \frac{\partial X_2}{\partial u_n}\frac{\partial}{\partial x_2} + \cdots + \frac{\partial X_n}{\partial u_n}\frac{\partial}{\partial x_n}.
\end{aligned}
\tag{42}
$$

When expressed in this form the relationships between the partial differentiation operations $\partial/\partial x_1, \partial/\partial x_2, \ldots, \partial/\partial x_n$ and $\partial/\partial u_1, \partial/\partial u_2, \ldots, \partial/\partial u_n$ become clear. This interpretation is needed when finding higher order partial derivatives such as $\partial^2 F/\partial u_2 \partial u_1$, because

$$\frac{\partial^2 F}{\partial u_2 \partial u_1} = \frac{\partial}{\partial u_1}\left(\frac{\partial F}{\partial u_2}\right) = \left(\frac{\partial X_1}{\partial u_1}\frac{\partial}{\partial x_1} + \frac{\partial X_2}{\partial u_1}\frac{\partial}{\partial x_2} + \cdots + \frac{\partial X_n}{\partial u_1}\frac{\partial}{\partial x_n}\right)\left(\frac{\partial F}{\partial u_2}\right).$$

An important combination of partial derivatives that occurs throughout physics and engineering is called the **Laplacian** of a function. When a twice differentiable function $f(x, y, z)$ of the cartesian coordinates x, y, and z is involved, the Laplacian of f, denoted by Δf and sometimes by $\nabla^2 f$, read "del squared f," takes the form

$$\Delta f = \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}. \tag{43}$$

Cylindrical Polar Coordinates (r, θ, z)

The cylindrical polar coordinate system (r, θ, z) is illustrated in Fig. 1.14, and its relationship to cartesian coordinates is given by

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = z, \quad \text{with } 0 \le \theta < 2\pi \text{ and } r \ge 0. \tag{44}$$

Spherical Polar Coordinates (r, ϕ, θ)

The spherical polar coordinate system (r, ϕ, θ) shown in Fig. 1.15 is related to cartesian coordinates by

$$
\begin{aligned}
x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\theta, \\
\text{with } 0 \le \theta \le \pi,\ 0 \le \phi < 2\pi.
\end{aligned}
\tag{45}
$$

The derivation of the formulas for the change of variables in functions of several variables can be found in any one of references [1.1] to [1.7], where cylindrical and

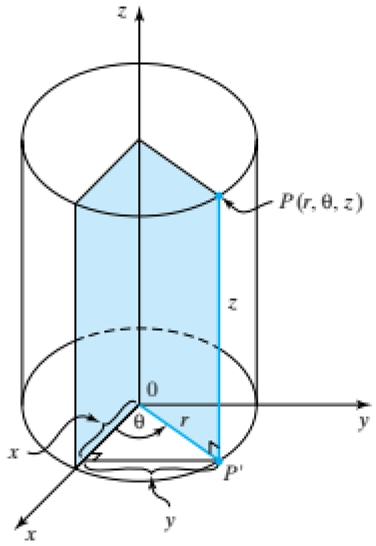

FIGURE 1.14 Cylindrical polar coordinates (r, θ, z).

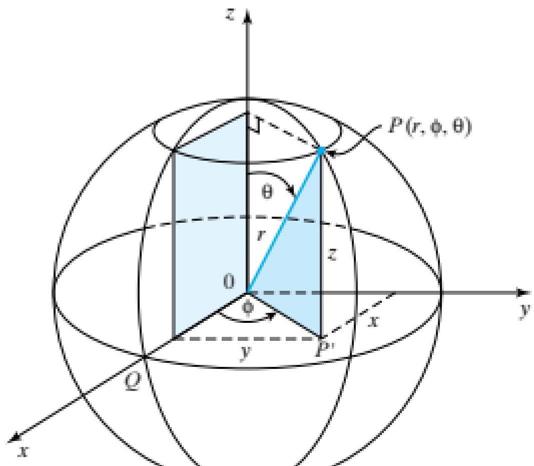

FIGURE 1.15 Spherical polar coordinates (r, ϕ, θ).

spherical polar coordinates are also discussed. Information on general orthogonal coordinate systems can be found in references [G.3] and [2.3].

EXERCISES 1.13

1. By making the change of variables $x = r\cos\theta$, $y = r\sin\theta$, $z = z$, in the function $f(x, y, z)$, when it becomes the function $F(r, \theta, z)$, show that in cylindrical polar coordinates

$$\frac{\partial F}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \sin\theta\frac{\partial f}{\partial y},$$
$$\frac{\partial F}{\partial \theta} = -r\sin\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y}, \quad \frac{\partial F}{\partial z} = \frac{\partial f}{\partial z}.$$

2. Use the results of Exercise 1 to show that in cylindrical polar coordinates the Laplacian

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \quad \text{becomes}$$
$$\Delta F = \frac{\partial^2 F}{\partial r^2} + \frac{1}{r}\frac{\partial F}{\partial r} + \frac{1}{r^2}\frac{\partial^2 F}{\partial \theta^2} + \frac{\partial^2 F}{\partial z^2},$$

and hence that an equivalent form of ΔF is

$$\Delta F = \frac{1}{r}\left[\frac{\partial}{\partial r}\left(r\frac{\partial F}{\partial r}\right) + \frac{1}{r}\frac{\partial}{\partial \theta}\left(\frac{\partial F}{\partial \theta}\right) + \frac{\partial}{\partial z}\left(r\frac{\partial F}{\partial z}\right)\right].$$

3. By making the change of variable $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$ in the function $f(x, y, z)$, when it becomes $F(r, \phi, \theta)$, show that in spherical polar coordinates

$$\frac{\partial F}{\partial r} = \sin\theta\cos\phi\frac{\partial f}{\partial x} + \sin\phi\sin\theta\frac{\partial f}{\partial y} + \cos\phi\frac{\partial f}{\partial z}$$
$$\frac{\partial F}{\partial \phi} = r\cos\phi\cos\theta\frac{\partial f}{\partial x} + r\cos\phi\sin\theta\frac{\partial f}{\partial y} - r\sin\phi\frac{\partial f}{\partial z}$$
$$\frac{\partial F}{\partial z} = -r\sin\phi\sin\theta\frac{\partial f}{\partial x} + r\sin\phi\cos\theta\frac{\partial f}{\partial y}.$$

4. Use the results of Exercise 3 to show that in spherical polar coordinates the Laplacian

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

becomes

$$\Delta F = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial F}{\partial r}\right) + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial^2 F}{\partial \phi^2}\right)$$
$$+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial F}{\partial \theta}\right).$$

1.14 Inverse Functions and the Inverse Function Theorem

In mathematics and its applications it is often necessary to find the inverse of a function $y = f(x)$ so x can be expressed in the form $x = g(y)$, and when this can be done the function g is called the **inverse** of f and is such that $y = f(g(y))$. When f is an arbitrary function its inverse is often denoted by f^{-1}, and this superscript notation is also used to denote the inverse of trigonometric functions so if, for example, $y = \sin x$, the inverse sine function is written $\sin^{-1}$, so that $x = \sin^{-1} y$. However, the notation $y = \arcsin y$ is also used with the understanding that the notations arcsin and $\sin^{-1}$ are equivalent.

A trivial example of a function whose inverse can be found unambiguously is $y = ax + b$, because provided $a \neq 0$ we can write $x = (y - b)/a$ for all x and y. This is not the case, however, when trigonometric functions are involved, because the function $y = \sin x$ will give a unique value of y for any given x, but given y there are infinitely many values of x for which $y = \sin x$. This and similar inverse trigonometric functions are considered in elementary calculus courses. There the multivalued nature of the inverse sine function is resolved by restricting it to make y lie in a specific interval chosen so that one y corresponds to one x and, conversely, one x corresponds to one y. This situation is described by saying that the relationship between x and y is **one-to-one**. Specifically, in the case of the sine function, this is accomplished by requiring that if $x = \sin y$, the inverse function $y = \text{Arcsin } x$ is restricted so its **principal value** lies in the interval $-\pi/2 \leq \text{Arcsin} x \leq \pi/2$, where the domain of definition of the inverse function is $-1 \leq x \leq 1$.

A different possibility that arises frequently is when x and y are related by an equation of the form $f(x, y) = 0$ from which it is impossible to extract either x as a function of y, or y as a function of x in terms of known functions. A typical example of this type is $f(x, y) = x^2 - 2y^2 - \sin xy$. To make matters precise, if x and y are related by an equation $f(x, y) = 0$, then if a function $y = g(x)$ exists such that $f(x, g(x)) = 0$, the function $y = g(x)$ is said to be defined **implicitly** by $f(x, y) = 0$.

Although it is often not possible to find the function $g(x)$, it is still necessary to know when, in a neighborhood of a point (x_0, y_0), given a value of x, a unique value of y can be found, sometimes only numerically. The *implicit function theorem* that follows is seldom mentioned in first calculus courses because its proof involves certain technicalities, but it is quoted here in the simplest possible form because of its fundamental importance and the fact that is it frequently used by implication.

THEOREM 1.12

The implicit function theorem Let $f(x, y)$ and $f_y(x, y)$ be continuous in a region D of the (x, y)-plane and let (x_0, y_0) be a point inside D, where $f(x_0, y_0) = 0$ and $f_y(x_0, y_0) \neq 0$. Then

(i) There is a rectangle R inside D containing (x_0, y_0) at all points of which there can be found a unique y such that $f(x, y) = 0$.

(ii) If the value of y is denoted by $g(x)$, then $y_0 = g(x_0)$, with $f(x, g(x)) = 0$, and $g(x)$ is continuous inside R.

(iii) If, in addition, $f_x(x, y)$ is continuous in D then $g(x)$ is differentiable in R and $g'(x) = -\frac{f_x(x,g(x))}{f_y(x,g(x))}$. ■

In general terms, the implicit function theorem gives conditions that ensure the existence of an inverse function that is continuous and smooth enough to be differentiable. The theorem has a more general form involving functions $f(x_1, x_2, \ldots, x_n)$ of n variables, though this will not be given here. The interested reader can find accounts of the implicit function theorem and some of its generalizations in references [1.4], [1.6], and [5.1].