CHAPTER 0

Preliminaries

1. Introduction

In this preliminary chapter we consider briefly some important concepts from calculus and algebra which we shall require for our study of differential equations. Many of these concepts may be familiar to the student, in which case this chapter can serve as a review. First the elementary properties of complex numbers are outlined. This is followed by a discussion of functions which assume complex values, in particular polynomials and power series. Some consequences of the Fundamental Theorem of Algebra are given. The exponential function is defined using power series; it is of central importance for linear differential equations with constant coefficients. The role that determinants play in the solution of systems of linear equations is discussed. Lastly we make a few remarks concerning principles of discovery, and methods of proof, of mathematical results.

2. Complex numbers

It is a fundamental fact about real numbers that the square of any such number is never negative. Thus there is no real x which satisfies the equation

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

We shall use the real numbers to define new numbers which include numbers which satisfy such equations.

A *complex number* z is an ordered pair of real numbers (x, y), and we write

$$z = (x, y).$$

If

$$z_1 = (x_1, y_1), \qquad z_2 = (x_2, y_2),$$

are two such numbers, we define z_1 to be *equal* to z_2, and write $z_1 = z_2$, if $x_1 = x_2$ and $y_1 = y_2$. The *sum* $z_1 + z_2$ is defined to be the complex number given by

$$z_1 + z_2 = (x_1 + x_2, y_1 + y_2).$$

If $z = (x, y)$, the *negative* of z, denoted by $-z$, is defined to be the number

$$-z = (-x, -y).$$

The *zero* complex number, also denoted by 0, is defined by

$$0 = (0, 0).$$

It is clear from these definitions that

(i) $z_1 + z_2 = z_2 + z_1$

(ii) $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$

(iii) $z + 0 = z$

(iv) $z + (-z) = 0$

for all complex numbers z, z_1, z_2, z_3.

The *difference* $z_1 - z_2$ is defined by

$$z_1 - z_2 = z_1 + (-z_2),$$

and we have

$$z_1 - z_2 = (x_1 - x_2, y_1 - y_2).$$

The *product* z_1z_2 is defined by

$$z_1z_2 = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1).$$

This definition appears curious at first, but we shall soon see a justification for it. It is easy to check that multiplication satisfies

(v) $z_1z_2 = z_2z_1$

(vi) $(z_1z_2)z_3 = z_1(z_2z_3)$

for all complex numbers z_1, z_2, z_3.

The *unit* complex number, with respect to multiplication, is the number $(1, 0)$ for we see that if $z = (x, y)$ is any complex number

$$z(1, 0) = (x, y)\,(1, 0) = (x, y) = z.$$

For this reason we denote the number $(1, 0)$ by just 1. Then we have

(vii) $z1 = z$

for all complex z.

If $z = (x, y) \neq (0, 0)$ there is a unique complex number w such that $zw = 1$ $(= (1, 0))$. Indeed, if $w = (u, v)$, where u, v are real, the equation $zw = 1$ says that

$$xu - yv = 1$$

$$yu + xv = 0.$$

These equations have the unique solution

$$u = \frac{x}{x^2 + y^2}, \qquad v = \frac{-y}{x^2 + y^2},$$

provided $x^2 + y^2 \neq 0$, which is equivalent to the assumption we made that $z \neq 0$. The number w, such that $zw = 1$, is called the *reciprocal* of z, and we denote it by z^{-1} or $1/z$. Thus

$$z^{-1} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2}\right), \qquad \text{if } z \neq 0.$$

Then

(viii) $zz^{-1} = 1, \quad \text{if} \quad z \neq 0.$

The *quotient* z_1/z_2 is defined when $z_2 \neq 0$ by

$$\frac{z_1}{z_2} = z_1z_2^{-1}, \qquad \text{if } z_2 \neq 0.$$

The interaction between addition and multiplication is given by the rule

(ix) $z_1(z_2 + z_3) = z_1z_2 + z_1z_3.$

The complex numbers of the form $(x, 0)$ are such that the negative and reciprocal of any such number have the same form, for

$$-(x, 0) = (-x, 0),$$

$$(x, 0)^{-1} = (x^{-1}, 0), \qquad \text{if } x \neq 0.$$

Moreover, the sum and product of two such numbers have the same form, since

$$(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0),$$

$$(x_1, 0)(x_2, 0) = (x_1x_2, 0).$$

The real numbers are in a one-to-one correspondence with the complex numbers of this form, the real number x corresponding to the complex number $z = (x, 0)$. Further, as we have just seen, the numbers corresponding to $-x, x^{-1}, x_1 + x_2, x_1x_2$ are just $-z, z^{-1}, z_1 + z_2, z_1z_2$, if $z_1 = (x_1, 0)$, $z_2 = (x_2, 0)$. For this reason it is usual to identify the complex number $(x, 0)$ with the real number x, and we write $x = (x, 0)$. [Notice that this agrees with our earlier identifications $0 = (0, 0)$, $1 = (1, 0)$.] In this sense, the complex numbers contain the real numbers. The properties (i)–(ix), which hold for complex numbers, are also valid for real numbers, and thus we see that we have succeeded in enlarging the set of real numbers without losing any of these algebraic properties. We have gained something also, since there are complex numbers z which satisfy the equation

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

One such number is the *imaginary unit* $i = (0, 1)$, as can be easily checked, and this provides one justification for our definition of multiplication.

If $z = (x, y)$ is a complex number, the real number x is called the *real part* of z, and we write $\operatorname{Re} z = x$; whereas y is called the *imaginary part* of z, and we write $\operatorname{Im} z = y$. Thus

$$z = (x, y) = x(1, 0) + y(0, 1) = x + iy = \operatorname{Re} z + i(\operatorname{Im} z).$$

Hereafter it will be convenient to denote a complex number (x, y) as $x + iy$.

It is clear that the complex numbers are in a one-to-one correspondence with the points of the (x, y)-plane, the complex number $z = x + iy$ corresponding to the point with coordinates (x, y). Then thought of in this way the x-axis is often called the *real axis*, the y-axis is called the *imaginary axis*, and the plane is called the *complex plane*.

If $z = x + iy$, its mirror image in the real axis is the point $x - iy$. This number is called the *complex conjugate* of z, and is denoted by $\bar{z}$. Thus $\bar{z} = x - iy$ if $z = x + iy$. We see immediately that

$$\bar{\bar{z}} = z, \overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2, \overline{z^{-1}} = (\bar{z})^{-1},$$

for any complex numbers z, z_1, z_2.

Introducing polar coordinates (r, θ) in the complex plane via

$$x = r\cos\theta, \quad y = r\sin\theta, \quad (r \geqq 0, 0 \leqq \theta < 2\pi),$$

we see that we may write

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

The *magnitude* of $z = x + iy$, denoted by $|z|$, is defined to be r. Thus

$$|z| = (x^2 + y^2)^{1/2} = (z\bar{z})^{1/2},$$

where the positive square root is understood. Clearly $|\bar{z}| = |z|$. Suppose z is real (that is, $\operatorname{Im} z = 0$). Then $z = x + i0$, for some real x, and

$$|z| = (x^2)^{1/2},$$

which is the magnitude of x considered as a real number. In addition the magnitude of a complex number obeys the same rules as the magnitude of a real number, namely:

$$|z| \geqq 0,$$

$$|z| = 0 \quad \text{if and only if} \quad z = 0,$$

$$|-z| = |z|,$$

$$|z_1 + z_2| \leqq |z_1| + |z_2|,$$

$$|z_1 z_2| = |z_1|\,|z_2|.$$

We show that $|z_1 + z_2| \leqq |z_1| + |z_2|$, for example. First we note that

$$\operatorname{Re} z \leqq |z|$$

for any complex number z. Then

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)\overline{(z_1 + z_2)} = |z_1|^2 + |z_2|^2 + z_1\bar{z}_2 + \bar{z}_1 z_2 \\ &= |z_1|^2 + |z_2|^2 + 2\operatorname{Re}(z_1\bar{z}_2) \\ &\leqq |z_1|^2 + |z_2|^2 + 2|z_1\bar{z}_2| \\ &= |z_1|^2 + |z_2|^2 + 2|z_1|\,|z_2| \\ &= (|z_1| + |z_2|)^2, \end{aligned}$$

from which it follows that $|z_1 + z_2| \leqq |z_1| + |z_2|$.

From the above rules one can deduce further that

$$||z_1| - |z_2|| \leqq |z_1 + z_2| \leqq |z_1| + |z_2|,$$

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}.$$

Geometrically we see that $|z_1 - z_2|$ represents the distance between the two points z_1 and z_2 in the complex plane.

EXERCISES

1. Compute the following complex numbers, and express in the form $x + iy$, where x, y are real:

(a) $(2 - i3) + (-1 + i6)$ (b) $(4 + i2) - (6 - i3)$

(c) $(6 - i\sqrt{2})(2 + i4)$ (d) $\dfrac{1 + i}{1 - i}$

(e) $|4 - i5|$ (f) $\operatorname{Re}(4 - i5)$

(g) $\operatorname{Im}(6 + i2)$

2. Express the following complex numbers in the form $r(\cos\theta + i\sin\theta)$ with $r \geqq 0$ and $0 \leqq \theta < 2\pi$:

(a) $1 + i\sqrt{3}$ (b) $(1+i)^2$

(c) $\dfrac{1+i}{1-i}$ (d) $(1+i)(1-i)$

3. Indicate graphically the set of all complex numbers z satisfying:

(a) $|z-2| = 1$ (b) $|z+2| < 2$
(c) $|\operatorname{Re} z| \leqq 3$ (d) $|\operatorname{Im} z| > 1$
(e) $|z-1| + |z+2| = 8$.

4. Prove that:
(a) $z + \bar{z} = 2\operatorname{Re} z$ (b) $z - \bar{z} = 2i\operatorname{Im} z$
(c) $|\operatorname{Re} z| \leqq |z|$ (d) $|z| \leqq |\operatorname{Re} z| + |\operatorname{Im} z|$

5. If r is a real number, and z complex, show that

$$\operatorname{Re}(rz) = r(\operatorname{Re} z), \qquad \operatorname{Im}(rz) = r(\operatorname{Im} z).$$

6. Prove that

$$||z_1| - |z_2|| \leqq |z_1 + z_2|.$$

(*Hint*: $z_1 = z_1 + z_2 + (-z_2)$, and $z_2 = z_1 + z_2 + (-z_1)$.)

7. Prove that

$$|z_1 + z_2|^2 + |z_1 - z_2|^2 = 2|z_1|^2 + 2|z_2|^2,$$

for all complex z_1, z_2.

8. If $|a| < 1$, what complex z satisfy $\dfrac{|z-a|}{|1-\bar{a}z|} \leqq 1$?

9. If n is any positive integer, prove that

$$r^n(\cos n\theta + i\sin n\theta) = [r(\cos\theta + i\sin\theta)]^n.$$

(*Hint*: Use induction.)

10. Use the result of Ex. 9 to find
(a) two complex numbers satisfying $z^2 = 2$,
(b) three complex numbers satisfying $z^3 = 1$.

3. Functions

Suppose D is a set whose elements are denoted by $P, Q, \cdots$, which are called the points of the set. Let R be another set. A *function* on D to R is a law f which associates with each point P in D exactly one point in R, which we denote by $f(P)$. The set D is called the *domain* of f. The point $f(P)$ is called the *value* of f at P. We can visualize the concept of a function as

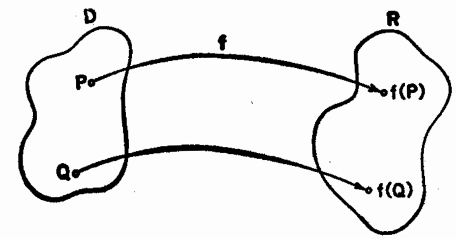

Figure 1

in Fig. 1, where each P in D is connected to a unique $f(P)$ in R by a string according to some rule. This rule, or what amounts to the same thing, the collection of all these strings, is the function f on D to R.

We say that two functions f and g are *equal*, $f = g$, if they have the same domain D, and $f(P) = g(P)$ for all P in D.

The idea of a function is very general, and is a fundamental one in mathematics. We shall consider some examples which are of importance for our study of differential equations.

(a) *Complex-valued functions.* If the set R which contains the values of f is the set of all complex numbers, we say that f is a *complex-valued function.* If f and g are two complex-valued functions with the same domain D, we can define their *sum* $f + g$ and *product* fg by

$$(f+g)(P) = f(P) + g(P),$$
$$(fg)(P) = f(P)g(P),$$

for each P in D. Thus $f + g$ and fg are also functions with domain D. If α is any complex number the function which assigns to each P in a domain D the number α is called a *constant function*, and is also denoted by α. Thus if f is any complex-valued function on D we have

$$(\alpha f)(P) = \alpha f(P)$$

for all P in D.

A *real-valued function* f defined on D is one whose values are real numbers. Such a function is a special case of a complex-valued function. Clearly the sum and product of two real-valued functions on D are real-valued functions. Real-valued functions are usually the principal object of study in first courses in calculus.

Every complex-valued function f defined on a domain D gives rise to two real-valued functions $\operatorname{Re} f$, $\operatorname{Im} f$ defined by

$$(\operatorname{Re} f)(P) = \operatorname{Re}[f(P)],$$
$$(\operatorname{Im} f)(P) = \operatorname{Im}[f(P)],$$

for all P in D. $\operatorname{Re} f$ and $\operatorname{Im} f$ are called the *real* and *imaginary* parts of f respectively and we have

$$f = \operatorname{Re} f + i \operatorname{Im} f.$$

Thus the study of complex-valued functions can be reduced to the study of pairs of real-valued functions. To obtain examples of complex-valued functions we must specify their domains.

(b) *Complex-valued functions with real domains.* Many of the functions we consider in this book have a domain D which is an interval I of the real axis. Recall that an *interval* is a set of real x satisfying one of the nine inequalities

$$a \leqq x \leqq b, \quad a \leqq x < b, \quad a < x \leqq b, \quad a < x < b,$$

$$a \leqq x < \infty, \quad -\infty < x \leqq b, \quad a < x < \infty, \quad -\infty < x < b,$$

$$-\infty < x < \infty,$$

where a, b are distinct real numbers. The calculus of complex-valued functions defined on real intervals is entirely analogous to the calculus of real-valued functions defined on intervals. We sketch the main ideas.

Suppose f is a complex-valued function defined on a real interval I. Then f is said to have the complex number L as a *limit* at x_0 in I, and we write

$$\lim_{x \to x_0} f(x) = L, \quad \text{or} \quad f(x) \to L, \qquad (x \to x_0),$$

if

$$|f(x) - L| \to 0, \quad \text{as} \quad 0 < |x - x_0| \to 0.$$

This means that given any $\epsilon > 0$ there is a $\delta > 0$ such that

$$|f(x) - L| < \epsilon, \quad \text{whenever} \quad 0 < |x - x_0| < \delta, \quad x \text{ in } I.$$

Note that here we are using the magnitude of complex numbers. Formally our definition is the same as that for real limits of real-valued functions. Because of this the usual rules for limits, and their proofs, are valid. In particular, if f and g are complex-valued functions defined on I such that for some x_0 in I

$$f(x) \to L, \quad g(x) \to M, \qquad (x \to x_0),$$

then

$$(f + g)(x) \to L + M, \quad (fg)(x) \to LM, \qquad (x \to x_0).$$

Suppose f has a limit $L = L_1 + iL_2$ at x_0, where L_1, L_2 are real. Then since

$$|(\operatorname{Re} f)(x) - L_1| = |\operatorname{Re}[f(x) - L]| \leqq |f(x) - L|,$$

and

$$|(\operatorname{Im} f)(x) - L_2| = |\operatorname{Im}[f(x) - L]| \leqq |f(x) - L|,$$

it follows that

$$(\operatorname{Re} f)(x) \to L_1, \quad (\operatorname{Im} f)(x) \to L_2, \qquad (x \to x_0).$$

Conversely, if $\operatorname{Re} f$ and $\operatorname{Im} f$ have limits L_1, L_2 respectively at x_0, then f will have the limit $L = L_1 + iL_2$ at x_0.

We say that a complex-valued function f defined on an interval I is *continuous* at x_0 in I if f has the limit $f(x_0)$ at x_0, that is,

$$|f(x) - f(x_0)| \to 0, \quad \text{as} \quad 0 < |x - x_0| \to 0.$$

Equivalently, f is continuous at x_0 if both $\operatorname{Re} f$ and $\operatorname{Im} f$ are continuous at x_0. We say f is continuous on I if it is continuous at each point of I. The sum and product of two functions which are continuous at x_0 are continuous there.

The complex-valued function f defined on an interval I is said to be *differentiable* at x_0 in I if the ratio

$$\frac{f(x) - f(x_0)}{x - x_0}, \qquad (x \neq x_0),$$

has a limit at x_0. If f is differentiable at x_0 we define its *derivative* at x_0, $f'(x_0)$, to be this limit. Thus, if $f'(x_0)$ exists,

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| \to 0, \quad \text{as} \quad 0 < |x - x_0| \to 0.$$

An equivalent definition is: f is differentiable at x_0 if both $\operatorname{Re} f$ and $\operatorname{Im} f$ are differentiable at x_0. The derivative of f at x_0 is given by

$$f'(x_0) = (\operatorname{Re} f)'(x_0) + i(\operatorname{Im} f)'(x_0).$$

Using these definitions one can show that the usual rules for differentiating real-valued functions are valid for complex-valued functions. For example, if f, g are differentiable at x_0 in I, then so are $f + g$ and fg, and

$$(f + g)'(x_0) = f'(x_0) + g'(x_0),$$

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

If f is differentiable at every x in an interval I, then f gives rise to a new function f' on I whose value at each x on I is $f'(x)$.

A complex-valued function f with domain the interval $a \leqq x \leqq b$ is said to be *integrable* there if both Re f and Im f are, and in this case we define its integral by

$$\int_a^b f(x)\,dx = \int_a^b (\operatorname{Re} f)(x)\,dx + i\int_a^b (\operatorname{Im} f)(x)\,dx.$$

Every function f which is continuous on $a \leqq x \leqq b$ is integrable there. This definition implies the usual integration rules. In particular, if f and g are integrable on $a \leqq x \leqq b$, and α, β are two complex numbers,

$$\int_a^b (\alpha f + \beta g)(x)\,dx = \alpha\int_a^b f(x)\,dx + \beta\int_a^b g(x)\,dx.$$

An important inequality connected with the integral of a continuous complex-valued function f defined on $a \leqq x \leqq b$ is

$$\left|\int_a^b f(x)\,dx\right| \leqq \int_a^b |f(x)|\,dx.^*$$

This inequality is valid if f is real-valued, and the proof for the case when f is complex-valued can be based on this fact. Let

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

If $F = 0$ the inequality is obvious. If $F \neq 0$, let

$$F = |F|u, \quad u = \cos\theta + i\sin\theta, \qquad (0 \leqq \theta < 2\pi).$$

Then $u\bar{u} = 1$, and we have

$$\left|\int_a^b f(x)\,dx\right| = \bar{u}\int_a^b f(x)\,dx = \operatorname{Re}\left[\bar{u}\int_a^b f(x)\,dx\right]$$

$$= \int_a^b \operatorname{Re}\,[\bar{u}f(x)]\,dx \leqq \int_a^b |f(x)|\,dx,$$

*By

$$\int_a^b |f(x)|\,dx$$

is meant the integral of the function $|f|$ given by $|f|(x) = |f(x)|$ for $a \leqq x \leqq b$. Thus a more appropriate notation would be

$$\int_a^b |f|(x)\,dx.$$

We shall use the former notation since it is commonly used, and there will be no chance of confusion.

since

$$\operatorname{Re}\,[\bar{u}f(x)] \leqq |\bar{u}f(x)| = |f(x)|.$$

As particular examples of complex-valued functions let

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

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

for all real x. Then

$$(\operatorname{Re} f)(x) = x + x^2, \qquad (\operatorname{Im} f)(x) = -x^2,$$

$$(f+g)(x) = x + 2x^2,$$

$$(fg)(x) = (1+i)x^3 + 2x^4,$$

$$f'(x) = 1 + (2-2i)x,$$

$$\int_0^1 f(x)\,dx = \int_0^1 x\,dx + (1-i)\int_0^1 x^2\,dx = \frac{5}{6} - \frac{i}{3}.$$

(c) *Complex-valued functions with complex domains.* We shall need to know a little about complex-valued functions whose domains consist of complex numbers. An example is the function f given by

$$f(z) = z^n,$$

for all complex z, where n is a positive integer.

Let f be a complex-valued function which is defined on some disk

$$D: \quad |z - a| < r$$

with center at the complex number a and radius $r > 0$. Much of the calculus for such functions can be patterned directly after the calculus of complex-valued functions defined on a real interval I. We say that f has the complex number L as a *limit* at z_0 in D if

$$|f(z) - L| \to 0, \quad \text{as} \quad 0 < |z - z_0| \to 0,$$

and we write

$$\lim_{z\to z_0} f(z) = L, \quad \text{or} \quad f(z) \to L, \quad (z \to z_0).$$

If f and g are two complex-valued functions defined on D such that for some z_0 in D

$$f(z) \to L, \quad g(z) \to M, \quad (z \to z_0),$$

then

$$(f+g)(z) \to L + M, \quad (fg)(z) \to LM, \quad (z \to z_0).$$

The proofs are identical to those for functions defined on real intervals.

The function f, defined on the disk D, is said to be *continuous* at z_0 in D if

$$|f(z) - f(z_0)| \to 0, \quad \text{as} \quad 0 < |z - z_0| \to 0.$$

It is said to be continuous on D if it is continuous at each point of D. The sum and product of two functions which are continuous at z_0 are continuous there. Examples of continuous functions on the whole complex plane are

$$f(z) = |z|, \qquad g(z) = z^2.$$

Let g be defined on some disk D_1 containing z_0, and let its values be in some disk D_2, where a function f is defined. If g is continuous at z_0, and f is continuous at $g(z_0)$, then "the function of a function" F given by

$$F(z) = f(g(z)), \qquad (z \text{ in } D_1), \tag{3.1}$$

is continuous at z_0. The proof follows the same lines as in calculus for real-valued functions defined for real x.

If f is defined on a disk D containing z_0 we say that f is *differentiable* at z_0 if

$$\frac{f(z) - f(z_0)}{z - z_0}, \qquad (z \neq z_0),$$

has a limit at z_0. If f is differentiable at z_0 its *derivative* at z_0, $f'(z_0)$, is defined to be this limit. Thus

$$\left| \frac{f(z) - f(z_0)}{z - z_0} - f'(z_0) \right| \to 0, \qquad \text{as } 0 < |z - z_0| \to 0.$$

Formally our definition is the same as that for the derivative of a complex-valued function defined on a real interval. For this reason if f and g are functions which have derivatives at z_0 in D then $f + g$, fg have derivatives there, and

$$\begin{aligned} (f+g)'(z_0) &= f'(z_0) + g'(z_0), \\ (fg)'(z_0) &= f'(z_0)g(z_0) + f(z_0)g'(z_0). \end{aligned} \tag{3.2}$$

Also, suppose f and g are two functions as given in (3.1), and that g is differentiable at z_0, whereas f is differentiable at $g(z_0)$. Then F is differentiable at z_0, with

$$F'(z_0) = f'(g(z_0))g'(z_0).$$

It is clear from the definition of a derivative that the function q defined by $q(z) = c$, where c is a complex constant, has a derivative which is zero everywhere, that is, $q'(z) = 0$. Also, if $p_1(z) = z$ for all z, then $p_1'(z) = 1$. Combining these results with the rules (3.2) we obtain the fact that every polynomial has a derivative for all z. A *polynomial* is a function p whose domain is the set of all complex numbers and which has the form

$$p(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n,$$

where $a_0, a_1, \cdots, a_n$ are complex constants. The rules (3.2) imply that for such a p

$$p'(z) = a_0 n z^{n-1} + a_1 (n-1) z^{n-2} + \cdots + a_{n-1}.$$

Thus p' is also a polynomial.

It is a rather strong restriction on a function defined on a disk D t demand that it be differentiable at a point z_0 in D. To illustrate this we not that the real-valued function f given by

$$f(x) = |x|,$$

for all real x, is differentiable at all $x \neq 0$. Indeed $f'(x)$ is $+1$ or -1 accor ing as x is positive or negative. However the continuous complex-valu function g given by

$$g(z) = |z|,$$

for all complex z, is not differentiable for any z. Suppose $z_0 = x_0 + y_0 i \neq$ for example, and let $z = x + yi$. Then for $z \neq z_0$

$$\frac{|z| - |z_0|}{z - z_0} = \frac{(x^2 + y^2)^{1/2} - (x_0^2 + y_0^2)^{1/2}}{(x - x_0) + i(y - y_0)}$$

$$= \frac{(x^2 + y^2) - (x_0^2 + y_0^2)}{[(x - x_0) + i(y - y_0)][(x^2 + y^2)^{1/2} + (x_0^2 + y_0^2)^{1/2}]}$$

If we let $|z - z_0| \to 0$ using z of the form $z = x_0 + yi$ (that is $y \to$ we see that

$$\frac{|z| - |z_0|}{z - z_0} \to \frac{y_0}{i(x_0^2 + y_0^2)^{1/2}}, \qquad ($$

whereas if we let $|z - z_0| \to 0$ using z of the form $z = x + y_0 i$ (th $x \to x_0$) we obtain

$$\frac{|z| - |z_0|}{z - z_0} \to \frac{x_0}{(x_0^2 + y_0^2)^{1/2}}.$$

The two limits (3.3) and (3.4) are different. However, in order that g be differentiable at z_0 we must obtain the same limit no matter how $|z - z_0| \to 0$. This shows that g is not differentiable at z_0.

(d) *Other functions.* Other types of functions which are important for our study of differential equations are usually combinations of the types discussed in (*b*), (*c*) above. Typical is a complex-valued function f which is defined for real x on some interval $|x - x_0| \leqq a$ (x_0 real, $a > 0$), and for complex z on some disk $|z - z_0| \leqq b$ (z_0 complex, $b > 0$). Thus the domain D of f is given by

$$D: \quad |x - x_0| \leqq a, \quad |z - z_0| \leqq b,$$

and the value of f at (x, z) is denoted by $f(x, z)$. Such a function f is said to be *continuous* at (ξ, η) in D if

$$|f(x, z) - f(\xi, \eta)| \to 0, \quad \text{as} \quad 0 < |x - \xi| + |z - \eta| \to 0.$$

There are two important facts which we shall need in Chap. 5 concerning such continuous functions. The first is that a continuous f on the D given above (*with the equality signs included*) is bounded, that is, there is a positive constant M such that

$$|f(x, z)| \leqq M,$$

for all (x, z) in D. This result is usually proved in advanced calculus courses. The second result relates to "plugging in" a complex-valued function ϕ into f. Suppose ϕ is a complex-valued function defined on

$$|x - x_0| \leqq a,$$

which is continuous there, and has values in $|z - z_0| \leqq b$. Then if f is continuous on D, the function F given by

$$F(x) = f(x, \phi(x)),$$

for all x such that $|x - x_0| \leqq a$, is continuous for such x.

A slightly more complicated type of complex-valued function f is one which is defined for real x and complex $z_1, \cdots, z_n$ on a domain

$$D: \quad |x - x_0| \leqq a, \quad |z_1 - z_{10}| + \cdots + |z_n - z_{n0}| \leqq b.$$

Here x_0 is real, $z_{10}, \cdots, z_{n0}$ are complex, and a, b are positive. The value of f at $x, z_1, \cdots, z_n$ is denoted by $f(x, z_1, \cdots, z_n)$. Continuity of f is defined just

as in the case of one z. Thus f is *continuous* at $\xi, \eta_1, \cdots, \eta_n$ in D if

$$|f(x, z_1, \cdots, z_n) - f(\xi, \eta_1, \cdots, \eta_n)| \to 0,$$

as

$$0 < |x - \xi| + |z_1 - \eta_1| + \cdots + |z_n - \eta_n| \to 0.$$

Such an f is bounded on D, and if $\phi_1, \cdots, \phi_n$ are n continuous complex-valued functions defined on $|x - x_0| \leqq a$, having the property that

$$|\phi_1(x) - z_{10}| + \cdots + |\phi_n(x) - z_{n0}| \leqq b$$

for all such x, then the function F given by

$$F(x) = f(x, \phi_1(x), \cdots, \phi_n(x))$$

for $|x - x_0| \leqq a$ is continuous there.

EXERCISES

1. Let $a = 2 + i3$, $b = 1 - i$. If for all real x

$$f(x) = ax + (bx)^2,$$

compute:

(a) $(\text{Re}\, f)(x)$ (b) $(\text{Im}\, f)(x)$

(c) $f'(x)$ (d) $\int_0^1 f(x)\, dx$

2. If for all real x

$$f(x) = x + ix^2, \qquad g(x) = \frac{x^2}{2},$$

compute:

(a) The function F given by $F(x) = f(g(x))$ (b) $F'(x)$

3. If a is a real-valued function defined on an interval I, and f is a complex-valued function defined there, show that

$$\text{Re}\,(af) = a(\text{Re}\, f), \qquad \text{Im}\,(af) = a(\text{Im}\, f).$$

4. Let $f(z) = z^2$ for all complex z, and let

$$u(x, y) = (\text{Re}\, f)(x + iy), \quad v(x, y) = (\text{Im}\, f)(x + iy).$$

(a) Compute $u(x, y)$ and $v(x, y)$.

(b) Show that $\dfrac{\partial u}{\partial x} = \dfrac{\partial v}{\partial y}, \quad \dfrac{\partial u}{\partial y} = -\dfrac{\partial v}{\partial x}.$

(c) Show that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \qquad \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0.$$

5. Let f be a complex-valued function defined on a disk

$$D: \quad |z| < r \qquad (r > 0),$$

which is differentiable there. Let

$$u(x, y) = (\text{Re } f)(x + iy), \qquad v(x, y) = (\text{Im } f)(x + iy).$$

Show that

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}, \tag{*}$$

for all $z = x + iy$ in D. (*Hint*: If $z_0 = x_0 + iy_0$ is in D, let $0 < |z - z_0| \to 0$, in the definition of $f'(z_0)$, through z of the form $z = x + iy_0$, and then of the form $z = x_0 + iy$, to obtain

$$f'(z_0) = \frac{\partial u}{\partial x}(x_0, y_0) + i\frac{\partial v}{\partial x}(x_0, y_0)$$

$$= \frac{\partial v}{\partial y}(x_0, y_0) - i\frac{\partial u}{\partial y}(x_0, y_0).$$

The equations (*) are called the *Cauchy-Riemann* equations.)

6. Let f be the complex-valued function defined on

$$D: \quad |x| \leqq 1, \qquad |z| \leqq 2,$$

(x real, z complex) by

$$f(x, z) = 3x^2 + xz + z^2,$$

and let ϕ be the function defined on $|x| \leqq 1$ by

$$\phi(x) = x + i.$$

(a) Compute the function F given by

$$F(x) = f(x, \phi(x)), \qquad (|x| \leqq 1).$$

(b) Compute $F'(x)$.

(c) Compute

$$\int_0^1 F(x)\, dx.$$

7. If r is a complex number, and

$$p(z) = (z - r)^n,$$

where n is a positive integer, show that

$$p(r) = p'(r) = \cdots = p^{(n-1)}(r) = 0, \qquad p^{(n)}(r) = n!.$$

4. Polynomials

We have defined a polynomial as a complex-valued function p whose domain is the set of all complex numbers and which has the form

$$p(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n,$$

where n is a non-negative integer, and $a_0, a_1 \cdots, a_n$ are complex constants. The highest power of z with non-zero coefficient which appears in the expression defining a polynomial p is called the *degree* of p, and written deg p. A *root* of a polynomial p is a complex number r such that $p(r) = 0$. A root of p is sometimes called a *zero* of p. We shall require, and assume, the following important result.*

Fundamental theorem of algebra. *If* p *is a polynomial such that* deg p $\geqq 1$, *then* p *has at least one root.*

This is a rather remarkable result, and justifies our introduction of the complex numbers. We have seen that not every polynomial with real coefficients (for example $z^2 + 1$) has a real root, but polynomials of degree greater than zero with complex coefficients always have a complex root. The remarkable fact is that we do not need to invent new numbers, which include the complex numbers, to guarantee a complex root.

We derive some consequences of this fundamental theorem.

Corollary 1. *Let* p *be a polynomial of degree* n $\geqq 1$, *with leading coefficient* 1 (*the coefficient of* z^n), *and let* r *be a root of* p. *Then*

$$p(z) = (z - r)q(z)$$

where q *is a polynomial of degree* n $-$ 1, *with leading coefficient* 1.

Proof. Let $p(z)$ have the form

$$p(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n,$$

* A proof can be found in G. Birkhoff and S. MacLane, *A survey of modern algebra*, New York, rev. ed., 1953, p. 107, and also in K. Knopp, *Theory of functions*, New York, 1945, p. 114.

and let c be any complex number. Then

$$p(z) - p(c) = (z^n - c^n) + a_1(z^{n-1} - c^{n-1}) + \cdots + a_{n-1}(z - c)$$
$$= (z - c)q(z),$$

where q is the polynomial given by

$$q(z) = z^{n-1} + cz^{n-2} + c^2z^{n-3} + \cdots + c^{n-1}$$
$$+ a_1(z^{n-2} + cz^{n-3} + \cdots + c^{n-2}) + \cdots + a_{n-1}.$$

Clearly deg $q = n - 1$ and q has leading coefficient 1. In particular if $c = r$, a root of p, then we have

$$p(z) = (z - r)q(z),$$

as desired.

If $n - 1 \geqq 1$, the polynomial q has a root, and this root is also a root of p by Corollary 1. Thus applying the Fundamental Theorem of Algebra n times, together with Corollary 1, we obtain

Corollary 2. *If* p *is a polynomial,* deg p = n $\geqq$ 1, *with leading coefficient* $a_0 \neq 0$, *then* p *has exactly* n *roots. If* $r_1, r_2, \cdots, r_n$ *are these roots, then*

$$p(z) = a_0(z - r_1)(z - r_2) \cdots (z - r_n). \tag{4.1}$$

Note that $a_0^{-1}p$ is a polynomial which has leading coefficient 1. We remark that the roots need not all be distinct. If r is a root of p, the number of times $z - r$ appears as a factor in (4.1) is called the *multiplicity* of r.

Theorem 1. *If* r *is a root of multiplicity* m *of a polynomial* p, deg p $\geqq$ 1, *then*

$$p(r) = p'(r) = \cdots = p^{(m-1)}(r) = 0,$$

and

$$p^{(m)}(r) \neq 0.$$

Proof. Let p have leading coefficient $a_0 \neq 0$, and degree $n \geqq m$. It follows from Corollary 2 that

$$p(z) = a_0(z - r)^m q(z), \tag{4.2}$$

where q is a polynomial of degree $n - m$, and $q(r) \neq 0$. Clearly $p(r) = 0$ by the definition of a root. Also

$$p'(z) = a_0 m(z - r)^{m-1}q(z) + a_0(z - r)^m q'(z),$$

and this implies that, if $m - 1 > 0$, $p'(r) = 0$. If $m = 1$ we have

$$p'(z) = a_0 q(z) + a_0(z - r)q'(z),$$

and thus $p'(r) = a_0 q(r) \neq 0$.

The general argument can be based on (4.2) and the formula

$$(fg)^{(k)} = f^{(k)}g + kf^{(k-1)}g' + \frac{k(k-1)}{2!}f^{(k-2)}g'' + \cdots + fg^{(k)} \tag{4.3}$$

for the k-th derivative of the product fg of two functions having k derivatives. Formula (4.3) can be established by induction. Applying (4.3) to the functions $f(z) = (z - r)^m$, $g(z) = q(z)$ in (4.2), we obtain

$$p^{(k)}(z) = a_0[m(m-1)\cdots(m-k+1)(z-r)^{m-k}q(z)$$
$$+ \text{(terms with higher powers of } (z - r) \text{ as a factor)}].$$

It is now clear that

$$p(r) = p'(r) = \cdots = p^{(m-1)}(r) = 0,$$

and

$$p^{(m)}(r) = a_0 m!\, q(r) \neq 0,$$

which is the desired result.

EXERCISES

1. Compute the roots, with multiplicities, of the following polynomials:
   (a) $z^2 + z - 6$ (b) $z^2 + z + 1$
   (c) $z^3 - 3z^2 + 4$ (d) $z^3 - (2 + i)z^2 + (1 + i2)z - i$
   (e) $z^4 - 3$

2. If r is such that $r^3 = 1$, and $r \neq 1$, prove that $1 + r + r^2 = 0$.

3. Let p be the polynomial given by

$$p(z) = a_0z^n + a_1z^{n-1} + \cdots + a_n,$$

with $a_0, a_1, \cdots, a_n$ all *real*. Show that

$$\overline{p(z)} = p(\bar{z}).$$

As a consequence show that if r is a root of p, then so is $\bar{r}$.

4. Prove that every polynomial of degree 3 with real coefficients has at least one real root.

Hence

$$\cos\theta + i\sin\theta = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \cdots$$

$$= \sum_{k=0}^{\infty} \frac{(i\theta)^k}{k!} = e^{i\theta}.$$

A consequence of (5.8) is that

$$e^{-i\theta} = \cos\theta - i\sin\theta, \tag{5.9}$$

since $\cos(-\theta) = \cos\theta$, and $\sin(-\theta) = -\sin\theta$. Using (5.8) and (5.9) we can solve for $\cos\theta$ and $\sin\theta$, obtaining

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2},$$

$$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

If z is a complex number with polar coordinates (r, θ), then

$$z = r(\cos\theta + i\sin\theta), \qquad (r \geq 0, 0 \leq \theta < 2\pi),$$

and we have, using (5.8),

$$z = re^{i\theta}. \tag{5.10}$$

Note that $|z| = r$, $|e^{i\theta}| = 1$ for every real θ. The relation (5.10) can be employed to find the roots of polynomials p of the form

$$p(z) = z^n - c, \tag{5.11}$$

where c is a complex constant. Suppose $c = |c|e^{i\alpha}$, where α is real, $0 \leq \alpha < 2\pi$, and $re^{i\theta}$ is a root. Then

$$r^n e^{in\theta} = |c|e^{i\alpha},$$

and taking magnitudes of both sides we see that

$$r^n = |c|, \quad \text{or} \quad r = |c|^{1/n},$$

where the positive n-th root is understood. Further

$$e^{in\theta} = e^{i\alpha}, \quad \text{or} \quad e^{i(n\theta-\alpha)} = 1.$$

There are exactly n distinct values of θ satisfying this relation and $0 \leq \theta < 2\pi$, namely, those for which

$$n\theta - \alpha = 2\pi k,$$

or

$$\theta = \frac{\alpha + 2\pi k}{n}, \qquad (k = 0, 1, \cdots, n-1).$$

Thus the roots $z_1, \cdots, z_n$ of the polynomial p in (5.11) are given by

$$z_{k+1} = |c|^{1/n} e^{i(\alpha+2\pi k)/n}$$

$$= |c|^{1/n}\left[\cos\left(\frac{\alpha + 2\pi k}{n}\right) + i\sin\left(\frac{\alpha + 2\pi k}{n}\right)\right], \quad (k = 0, 1, \cdots, n-1).$$

Geometrically we can describe the roots of p as follows. All roots lie on a circle about the origin with radius $|c|^{1/n}$. One root has an angle α/n with the real axis, if c has angle α with the real axis. The remainder of the roots are located by cutting the circle into n even parts, with the first cut being at the root at angle α/n.

As a particular example let us find the three cube roots of $4i$. Thus we want the roots of $z^3 - 4i$. Here $c = 4i$, and hence the cube roots will all have a magnitude of $|4i|^{1/3} = 4^{1/3}$. If we write $c = |c|e^{i\alpha}$, we see that $\alpha = \pi/2$ in this case. Thus the three cube roots of $4i$ are given by

$$z_1 = 4^{1/3}e^{i\pi/6}, \qquad z_2 = 4^{1/3}e^{i5\pi/6}, \qquad z_3 = 4^{1/3}e^{i9\pi/6},$$

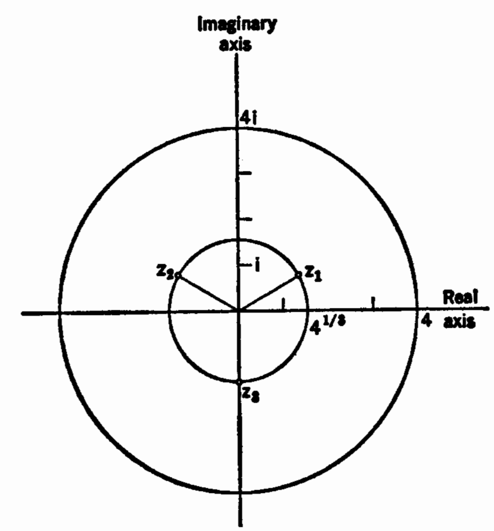

Figure 2. Three cube roots of $4i$

or since $\pi/6$ represents 30°,

$$z_1 = 4^{1/3}\left(\frac{\sqrt{3}}{2} + \frac{i}{2}\right), \qquad z_2 = 4^{1/3}\left(-\frac{\sqrt{3}}{2} + \frac{i}{2}\right),$$

$$z_3 = -4^{1/3}i.$$

These roots are sketched in Fig. 2.

EXERCISES

1. Find the three cube roots of 1.

2. Find the two square roots of i.

3. Find all roots of the polynomials:
(a) $z^3 + 24$ (b) $z^4 + i64$
(c) $z^4 + 4z^2 + 4$ (d) $z^{100} - 1$

4. If $z = x + iy$, where x, y are real, show that $|e^z| = e^x$. As a consequence show that there is no complex z such that $e^z = 0$.

5. If a, b, x are real show that:
(a) $\text{Re}\,[e^{(a+ib)x}] = e^{ax}\cos bx$ (b) $\text{Im}\,[e^{(a+ib)x}] = e^{ax}\sin bx$

6. (a) If $r = a + ib \neq 0$, where a, b are real, show that $(e^{rx})' = re^{rx}$.
(b) Using (a) compute:

(i) $\displaystyle\int_0^1 e^{rx}\,dx$

(ii) $\displaystyle\int_0^1 e^{ax}\cos bx\,dx$

(iii) $\displaystyle\int_0^1 e^{ax}\sin bx\,dx$

7. (a) If $\phi(x) = e^{rx}$, where r is a complex constant, and x is real, show that $\phi'(x) - r\phi(x) = 0$.
(b) If $\phi(x) = e^{iax}$, where a is a real constant, show that:
(i) $\phi'(x) - ia\phi(x) = 0$
(ii) $\phi''(x) + a^2\phi(x) = 0$

8. For what values of the constant r will the function ϕ given by $\phi(x) = e^{rx}$ satisfy

$$\phi''(x) + 3\phi'(x) - 2\phi(x) = 0$$

for all real x?

9. Let $a_k = k! + (i/k!)$. For what real x are the following series convergent?

(a) $\displaystyle\sum_{k=0}^{\infty} (\text{Re}\,a_k)x^k$ (b) $\displaystyle\sum_{k=0}^{\infty} (\text{Im}\,a_k)x^k$

(c) $\displaystyle\sum_{k=0}^{\infty} a_k x^k$

10. Consider the series

$$\sum_{k=0}^{\infty} z^k, \tag{*}$$

where z is complex.
(a) Show that the partial sum

$$s_n(z) = \sum_{k=0}^{n} z^k = \frac{1 - z^{n+1}}{1 - z},$$

if $z \neq 1$.
(b) Show that the series (*) converges absolutely for $|z| < 1$.
(c) Compute the sum $s(z)$ of the series (*) for $|z| < 1$.

6. Determinants

We shall need to know the connection between determinants and the solution of systems of linear equations. Suppose we have such a system of n equations

$$\begin{aligned} a_{11}z_1 + a_{12}z_2 + \cdots + a_{1n}z_n &= c_1 \\ a_{21}z_1 + a_{22}z_2 + \cdots + a_{2n}z_n &= c_2 \\ &\vdots \\ a_{n1}z_1 + a_{n2}z_2 + \cdots + a_{nn}z_n &= c_n, \end{aligned} \tag{6.1}$$

where the a_{ij} and c_i are given complex constants. The problem is to find complex numbers $z_1, \cdots, z_n$ satisfying these equations. Such a set of n numbers is called a *solution* of (6.1). We say that two solutions $z_1, \cdots, z_n$ and $z_1', \cdots, z_n'$ of (6.1) are *equal* if $z_1 = z_1', \cdots, z_n = z_n'$. If $c_1 = c_2 = \cdots = c_n = 0$ we say that the system is a *homogeneous system* of n linear equations, otherwise we say (6.1) is a *non-homogeneous* system. The determinant Δ of the coefficients in (6.1) is denoted by

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

5. Prove that if p is a polynomial, $\deg p \geq 1$, and r is a complex number such that

$$p(r) = p'(r) = \cdots = p^{(m-1)}(r) = 0, \qquad p^{(m)}(r) \neq 0,$$

then r is a root of p with multiplicity m. This is the converse of Theorem 1.

6. (a) Use the result of Ex. 5 to show that i is a root of the polynomial p given by

$$p(z) = z^5 + (2-3i)z^4 + (-1-6i)z^3 + (-6-5i)z^2 + (-6+2i)z + 2i,$$

and compute the multiplicity of i.
(b) Find the other roots of the polynomial p in (a).

7. Prove the formula (4.3). This can be written in the form

$$(fg)^{(k)} = f^{(k)}g + \binom{k}{1}f^{(k-1)}g' + \binom{k}{2}f^{(k-2)}g'' + \cdots + \binom{k}{l}f^{(k-l)}g^{(l)} + \cdots + fg^{(k)},$$

where

$$\binom{k}{l} = \frac{k!}{l!(k-l)!}$$

is a binomial coefficient. *Hint*: Use induction, and show that

$$\binom{k+1}{l} = \binom{k}{l-1} + \binom{k}{l}.$$

5. Complex series and the exponential function

If x is a real number, and e is the base for the natural logarithms, the number e^x exists, and

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \qquad (0! = 1),$$

where the series converges for all real x. Indeed, this series may be taken as the definition of e^x. We shall need to know what e^z is for *complex* z. One way is to *define* e^z by

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}. \tag{5.1}$$

Now we have to prove that this series converges for all complex z, and in fact there is the problem of defining what we mean by a convergent series with complex terms. The method is the same as that used to define convergent series with real terms.

A series

$$\sum_{k=0}^{\infty} c_k, \tag{5.2}$$

where all c_k are complex numbers, is said to be *convergent* if the sequence of partial sums

$$s_n = \sum_{k=0}^{n} c_k, \qquad (n = 0, 1, 2, \cdots),$$

tends to a limit s, as $n \to \infty$. That is, s is a complex number such that

$$|s_n - s| \to 0, \qquad (n \to \infty),$$

where the magnitude is the magnitude for complex numbers. If the series (5.2) is convergent, and $s_n \to s$, we call s the *sum* of the series, and write

$$s = \sum_{k=0}^{\infty} c_k.$$

If the series is not convergent we say that it is *divergent*.

The series (5.2) with complex terms c_k gives rise to two series with real terms, namely

$$\sum_{k=0}^{\infty} \operatorname{Re} c_k, \qquad \sum_{k=0}^{\infty} \operatorname{Im} c_k, \tag{5.3}$$

and it is not difficult to see that the series (5.2) is convergent with sum $s = \operatorname{Re} s + i \operatorname{Im} s$ if, and only if, the two real series in (5.3) are convergent with sums $\operatorname{Re} s$ and $\operatorname{Im} s$ respectively. In principle, therefore, the study of series with complex terms is the study of pairs of real series.

The series (5.2) is said to be *absolutely convergent* if the series

$$\sum_{k=0}^{\infty} |c_k| \tag{5.4}$$

is convergent. It can be shown that every absolutely convergent series is convergent. Since the series (5.4) has terms which are real and non-negative, any condition which implies the convergence of such series can be applied to guarantee the convergence of the series (5.2). One of the most important tests for convergence is the *ratio test*. One version of this is the following.

Ratio test. *Consider the series*

where the c_k *are complex. If* $|c_k| > 0$ *for all* k *beyond a certain positive integer, and*

$$\frac{|c_{k+1}|}{|c_k|} \to L, \qquad (k \to \infty), \tag{5.5}$$

then the series is convergent if $L < 1$, *and divergent for* $L > 1$.

Thus the series (5.2) is convergent if (5.5) is valid for an $L < 1$.

An immediate application of this result is to the series

$$\sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

Here $c_k = z^k/k!$ and

$$\frac{|c_{k+1}|}{|c_k|} = \left|\frac{z^{k+1}}{(k+1)!}\cdot\frac{k!}{z^k}\right| = \frac{|z|}{k+1} \to 0, \qquad (k \to \infty).$$

Thus this series converges for every z such that $|z| < \infty$, that is, for all complex z. Hence our definition (5.1) of e^z as the sum of this series makes sense. The function which associates with each z the complex number e^z is called the *exponential function*.

The series defining e^z is an example of a *power series*

$$\sum_{k=0}^{\infty} a_k(z - z_0)^k \tag{5.6}$$

about some point z_0, the a_k being complex. Many of the properties of a power series of the type

$$\sum_{k=0}^{\infty} a_k(x - x_0)^k,$$

where the a_k, x, x_0 are real, remain true for series of the form (5.6), and the proofs are identical. In particular, if a series (5.6) is convergent on a disk $D: |z - z_0| < r$ $(r > 0)$, then the function f defined by

$$f(z) = \sum_{k=0}^{\infty} a_k(z - z_0)^k, \qquad (z \text{ in } D),$$

has all derivatives in D, and these may be computed by differentiating term by term. Thus

$$f'(z) = \sum_{k=0}^{\infty} k a_k(z - z_0)^{k-1} = \sum_{k=1}^{\infty} k a_k(z - z_0)^{k-1},$$

where the last series converges in D. Applying this result to (5.1) we find that

$$(e^z)' = \sum_{k=1}^{\infty} k\frac{z^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{z^{k-1}}{(k-1)!} = \sum_{k=0}^{\infty} \frac{z^k}{k!} = e^z.$$

Another important property of the exponential function is that

$$e^{z_1+z_2} = e^{z_1}e^{z_2} \tag{5.7}$$

for every complex z_1, z_2. This can be proved by justifying the following steps

$$e^{z_1}e^{z_2} = \left(\sum_{k=0}^{\infty} \frac{z_1^k}{k!}\right)\left(\sum_{k=0}^{\infty} \frac{z_2^k}{k!}\right) = \sum_{k=0}^{\infty} c_k.$$

Here

$$\begin{aligned} c_k &= \sum_{n=0}^{k} \frac{z_1^{k-n}}{(k-n)!}\frac{z_2^n}{n!} \\ &= \frac{1}{k!}\sum_{n=0}^{k} \frac{k!}{(k-n)!\,n!} z_1^{k-n} z_2^n \\ &= \frac{1}{k!}(z_1 + z_2)^k. \end{aligned}$$

Thus formally we have the product of the series defining e^{z_1} and e^{z_2} is the series defining $e^{z_1+z_2}$, and these steps can be justified to give a proof of the equality (5.7). A consequence of (5.7) is that

$$(e^z)^n = e^{nz}$$

for every integer n. In particular $1/e^z = e^{-z}$.

Another property of the exponential function is that for all real θ,

$$e^{i\theta} = \cos\theta + i\sin\theta, \tag{5.8}$$

and the proof results from adding the series involved. Indeed, $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, etc., and thus

$$\begin{aligned} \cos\theta &= 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots \\ &= 1 + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^4}{4!} + \cdots, \end{aligned}$$

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots,$$

$$i\sin\theta = i\theta + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^5}{5!} + \cdots.$$

and is shorthand for the number Δ given by

$$\Delta = \sum(\pm)a_{1i_1}a_{2i_2}\cdots a_{ni_n},$$

where the sum is over all indices $i_1, \cdots, i_n$ such that $i_1, \cdots, i_n$ is a permutation of $1, \cdots, n$ and each term occurs with a $+$ or $-$ sign according as $i_1, \cdots, i_n$ is an even or odd permutation of $1, \cdots, n$. Thus

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

and

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

The principal results we require concerning determinants are contained in the following theorems. They are usually proved in elementary texts on linear algebra.

Theorem 2. *If the determinant Δ of the coefficients in* (6.1) *is not zero there is a unique solution of the system for $z_1, \cdots, z_n$. It is given by*

$$z_k = \frac{\Delta_k}{\Delta}, \quad (k = 1, \cdots, n),$$

where Δ_k is the determinant obtained from Δ by replacing its k*th column* $a_{1k}, \cdots, a_{nk}$ by $c_1, \cdots, c_n$.

Proof for the case $n = 2$. In this case suppose z_1, z_2 satisfy

$$\begin{aligned} a_{11}z_1 + a_{12}z_2 &= c_1 \\ a_{21}z_1 + a_{22}z_2 &= c_2. \end{aligned} \tag{6.2}$$

Multiply the first equation by a_{22}, the second equation by $-a_{12}$, and add. There results

$$z_1\Delta = a_{22}c_1 - a_{12}c_2 = \begin{vmatrix} c_1 & a_{12} \\ c_2 & a_{22} \end{vmatrix} = \Delta_1.$$

Multiply the first equation by $-a_{21}$, and the second by a_{11}, and add, obtaining

$$z_2\Delta = -a_{21}c_1 + a_{11}c_2 = \begin{vmatrix} a_{11} & c_1 \\ a_{21} & c_2 \end{vmatrix} = \Delta_2.$$

Thus if $\Delta \neq 0$, z_k must be Δ_k/Δ $(k = 1, 2)$, and it is readily verified that these values satisfy (6.2).

We note that for a homogeneous system ($c_1 = c_2 = \cdots = c_n = 0$ in (6.1)) there is always the solution

$$z_1 = z_2 = \cdots = z_n = 0.$$

This solution is called the *trivial solution.*

Theorem 3. *If* $c_1 = c_2 = \cdots = c_n = 0$ *in* (6.1), *and the determinant of the coefficients* $\Delta = 0$, *there is a solution of* (6.1) *such that not all the z_k are* 0.

Proof for the case $n = 2$. We are dealing with the case

$$\begin{aligned} a_{11}z_1 + a_{12}z_2 &= 0 \\ a_{21}z_1 + a_{22}z_2 &= 0, \end{aligned}$$

where

$$a_{11}a_{22} - a_{21}a_{12} = 0.$$

If $a_{11} \neq 0$,

$$z_1 = \frac{-a_{12}}{a_{11}}, \qquad z_2 = 1,$$

is a solution. If $a_{11} = 0$, and $a_{21} \neq 0$,

$$z_1 = \frac{-a_{22}}{a_{21}}, \qquad z_2 = 1,$$

is a solution. If $a_{11} = 0$, and $a_{21} = 0$,

$$z_1 = 1, \qquad z_2 = 0,$$

is a solution.

Combining Theorem 3 with Theorem 2 we obtain

Theorem 4. *The system of equations* (6.1) *has a unique solution if, and only if, the determinant Δ of the coefficients is not zero.*

Proof. If $\Delta \neq 0$ Theorem 2 says that there is a unique solution. Conversely, suppose there is a unique solution $z_1, \cdots, z_n$ of (6.1). If $\Delta = 0$, by and is shorthand for the number Δ given by

$$\Delta = \sum(\pm)a_{1i_1}a_{2i_2}\cdots a_{ni_n},$$

where the sum is over all indices $i_1, \cdots, i_n$ such that $i_1, \cdots, i_n$ is a permutation of $1, \cdots, n$ and each term occurs with a $+$ or $-$ sign according as $i_1, \cdots, i_n$ is an even or odd permutation of $1, \cdots, n$. Thus

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

and

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

The principal results we require concerning determinants are contained in the following theorems. They are usually proved in elementary texts on linear algebra.

Theorem 2. *If the determinant Δ of the coefficients in* (6.1) *is not zero there is a unique solution of the system for $z_1, \cdots, z_n$. It is given by*

$$z_k = \frac{\Delta_k}{\Delta}, \qquad (k = 1, \cdots, n),$$

where Δ_k is the determinant obtained from Δ by replacing its kth column $a_{1k}, \cdots, a_{nk}$ by $c_1, \cdots, c_n$.

Proof for the case $n = 2$. In this case suppose z_1, z_2 satisfy

$$\begin{aligned} a_{11}z_1 + a_{12}z_2 &= c_1 \\ a_{21}z_1 + a_{22}z_2 &= c_2. \end{aligned} \tag{6.2}$$

Multiply the first equation by a_{22}, the second equation by $-a_{12}$, and add. There results

$$z_1\Delta = a_{22}c_1 - a_{12}c_2 = \begin{vmatrix} c_1 & a_{12} \\ c_2 & a_{22} \end{vmatrix} = \Delta_1.$$

Multiply the first equation by $-a_{21}$, and the second by a_{11}, and add, obtaining

$$z_2\Delta = -a_{21}c_1 + a_{11}c_2 = \begin{vmatrix} a_{11} & c_1 \\ a_{21} & c_2 \end{vmatrix} = \Delta_2.$$

Thus if $\Delta \neq 0$, z_k must be Δ_k/Δ $(k = 1, 2)$, and it is readily verified that these values satisfy (6.2).

We note that for a homogeneous system ($c_1 = c_2 = \cdots = c_n = 0$ in (6.1)) there is always the solution

$$z_1 = z_2 = \cdots = z_n = 0.$$

This solution is called the *trivial solution.*

Theorem 3. *If* $c_1 = c_2 = \cdots = c_n = 0$ *in* (6.1), *and the determinant of the coefficients* $\Delta = 0$, *there is a solution of* (6.1) *such that not all the z_k are* 0.

Proof for the case $n = 2$. We are dealing with the case

$$\begin{aligned} a_{11}z_1 + a_{12}z_2 &= 0 \\ a_{21}z_1 + a_{22}z_2 &= 0, \end{aligned}$$

where

$$a_{11}a_{22} - a_{21}a_{12} = 0.$$

If $a_{11} \neq 0$,

$$z_1 = \frac{-a_{12}}{a_{11}}, \qquad z_2 = 1,$$

is a solution. If $a_{11} = 0$, and $a_{21} \neq 0$,

$$z_1 = \frac{-a_{22}}{a_{21}}, \qquad z_2 = 1,$$

is a solution. If $a_{11} = 0$, and $a_{21} = 0$,

$$z_1 = 1, \qquad z_2 = 0,$$

is a solution.

Combining Theorem 3 with Theorem 2 we obtain

Theorem 4. *The system of equations* (6.1) *has a unique solution if, and only if, the determinant Δ of the coefficients is not zero.*

Proof. If $\Delta \neq 0$ Theorem 2 says that there is a unique solution. Conversely, suppose there is a unique solution $z_1, \cdots, z_n$ of (6.1). If $\Delta = 0$, by

Theorem 3 there is a solution $\zeta_1, \cdots, \zeta_n$ of the corresponding homogeneous system, which is not the trivial solution. Then it is easy to check that $z_1 + \zeta_1, \cdots, z_n + \zeta_n$ is a solution of (6.1) distinct from $z_1, \cdots, z_n$, and forces us to conclude that $\Delta \neq 0$.

EXERCISES

1. Consider the system of equations

$$iz_1 + z_2 = 1 + i$$
$$2z_1 + (2 - i)z_2 = 1.$$

(a) Compute the determinant of the coefficients.
(b) Solve the system for z_1 and z_2.

2. Solve the following system for z_1, z_2 and z_3:

$$\begin{aligned} 3z_1 + z_2 - z_3 &= 0 \\ 2z_1 \qquad - z_3 &= 1 \\ z_2 + 2z_3 &= 2 \end{aligned}$$

3. Does the following system of equations have any solution other than $z_1 = z_2 = z_3 = 0$? If so find one.

$$\begin{aligned} 4z_1 + 2z_2 + 2z_3 &= 0 \\ 3z_1 + 7z_2 + 2z_3 &= 0 \\ 2z_1 + z_2 + z_3 &= 0 \end{aligned}$$

4. Consider the homogeneous system corresponding to (6.1) (the case $c_1 = c_2 = \cdots = c_n = 0$). Show that if the determinant of the coefficients $\Delta = 0$, there are an infinite number of solutions. (*Hint*: If $z_1, \cdots, z_n$ is a non-trivial solution, show that $\alpha z_1, \cdots, \alpha z_n$ is also a solution for any complex number α.)

5. Prove that if the determinant Δ of the coefficients in (6.1) is zero then either there is no solution of (6.1), or there are an infinite number of solutions. (*Hint*: Use Ex. 4.)

7. Remarks on methods of discovery and proof

Often a student studying mathematics has difficulty in understanding why or how a particular result, or method of proof, was ever conceived in the first place. Sometimes ideas seem to appear from nowhere. Now it is true that mathematical geniuses do invent radically new results, and methods for proving old results, which often appear quite strange. The most that ordinary people can do is to accept these brilliant ideas for what

they are, try to understand their consequences, and build on them to obtain further information. However, there are a few general principles which, if followed, can lead to a better understanding of mathematical discovery and proof.

Concerning discovery, we mention two principles:

(a) *use simple examples as a basis for conjecturing general results,*
(b) *argue in reverse.*

Both of these principles are illustrated in the proof we gave of Theorem 2 for the case $n = 2$. We were faced with trying to find out whether the system (6.1) of n linear equations has a solution or not, and what condition, or conditions, would guarantee a unique solution. We looked at the simplest example, which occurs for $n = 2$ (using (*a*)). Then we *assumed* that we had a solution (principle (*b*)), and found out what must be true for a solution, namely, that

$$z_1\Delta = \Delta_1, \qquad z_2\Delta = \Delta_2.$$

We immediately saw that if $\Delta \neq 0$, then

$$z_1 = \frac{\Delta_1}{\Delta}, \qquad z_2 = \frac{\Delta_2}{\Delta}. \tag{7.1}$$

Note that at this point we have not yet shown that there *is* a solution. All we have shown is that *if* z_1, z_2 is a solution, and $\Delta \neq 0$, it must be given by (7.1). We can now guess that if $\Delta \neq 0$, then z_1, z_2 given by (7.1) is a solution. This can be readily verified by substituting (7.1) into the given equations. An alternate procedure is to check that the steps leading to (7.1) can be reversed, if $\Delta \neq 0$. Once we have discovered the right condition for the case $n = 2$, it is natural to conjecture that a similar condition will work for a general n.

Three important methods of proving mathematical results are:

(i) *a constructive method,*
(ii) *method of contradiction,*
(iii) *method of induction.*

A typical example of a constructive method appears in the proof of Theorem 3 for the case $n = 2$. We wanted to show that nontrivial solutions of the two homogeneous equations exist if $\Delta = 0$. To do this we constructed solutions explicitly. An example of the method of contradiction appears in the proof of Theorem 4. We supposed that the system (6.1) had a unique solution. We assumed that $\Delta = 0$, and, using logical arguments, we arrived at the fact that (6.1) does not have a unique solution. This is a contradiction, and the only thing that can be wrong is our assumption that $\Delta = 0$.

The only other alternative is that $\Delta \neq 0$, which is the conclusion we desired.

The method of induction is concerned with proving an infinite number of statements $s_1, s_2, \cdots$, one for each positive integer n. If s_1 is true, and if for any positive integer k the statement s_k implies the statement s_{k+1}, then all the statements $s_1, s_2, \cdots$, are true. An example of a result which can be proved using induction is the formula

$$(fg)^{(k)} = \sum_{l=0}^{k} \binom{k}{l} f^{(k-l)} g^{(l)}, \qquad \binom{k}{l} = \frac{k!}{l!(k-l)!},$$

for the k-th derivative of the product of two complex-valued functions f, g which have k derivatives; see (4.3). The proof is the same as the induction used to prove the binomial formula

$$(a+b)^k = \sum_{l=0}^{k} \binom{k}{l} a^{k-l} b^l, \qquad (k = 1, 2, \cdots),$$

for the powers of the sum of two complex numbers a, b. The method of induction is equivalent to a property of the positive integers, and consequently we assume that this method is a valid method of proof.

The principles of discovery (*a*), (*b*), and the methods of proof (i), (ii), (iii), will be used many times throughout this book. It will be instructive for the student to identify which principles and methods are being used in any particular situation.

CHAPTER 1

Introduction—Linear Equations of the First Order

1. Introduction

In Sec. 2 we discuss what is meant by an ordinary differential equation and its solutions. Various problems which arise in connection with differential equations are considered in Sec. 3, notably initial value problems, boundary value problems, and the qualitative behavior of solutions. In a succession of easy steps we solve the linear equation of the first order in Secs. 4–7.

2. Differential equations

Suppose f is a complex-valued function defined for all real x in an interval I, and for complex y in some set S. The value of f at (x, y) is denoted by $f(x, y)$. An important problem associated with f is to find a (complex-valued) function ϕ on I, which is differentiable there, such that for all x on I,

(i) $\phi(x)$ is in S,

(ii) $\phi'(x) = f(x, \phi(x))$.

This problem is called an *ordinary differential equation of the first order*, and is denoted by

$$y' = f(x, y). \qquad (2.1)$$

The *ordinary* refers to the fact that only ordinary derivatives enter into the problem, and not partial derivatives. If such a function ϕ exists on I satisfying (i) and (ii) there, then ϕ is called a *solution* of (2.1) on I.