Complex Analysis & Differential Equations

Ulrik Egede

Adapted and enlarged by Martin McCall

Physics Department, Imperial College, London SW7 2AZ

First Term (2017-2018)

Contents

Chapter 1

Complex Numbers

Complex analysis is one of the most awe-inspiring areas of mathematics. Beginning with the notion of an imaginary unit, there follows an abundance of useful and unexpected results, methods, and concepts. The story begins in the mid 1550s, when the Italian mathematician Girolamo Cardano posed a problem that could not be solved with real numbers, namely, the existence of two numbers whose sum is 10 and whose product is 40. If we call these numbers x and y, then we are looking for a solution of

$$x + y = 10 \quad \text{and} \quad xy = 40\,.$$

By solving these simultaneous equations for, say, y in terms of x, we obtain a quadratic equation in x:

$$x^2 - 10x + 40 = 0\,.$$

The solutions to this equation yield

$$x = 5 \pm \sqrt{-15}\,, \qquad y = 5 \mp \sqrt{-15}\,,$$

from which we see directly that the sum of these numbers is 10 and their product is 40. Cardano did not pursue this, concluding that this result was 'as subtle as it is useless.'

Complex numbers did not arise from the above example, but in connection with the solution to quartic equations. Cardano presented formulae for the solutions of certain cubic and quartic equations, within which there are square roots of numbers that could be negative. Cardano had serious misgivings about expressions such as $2 + \sqrt{-2}$ and, in fact, referred to thinking about them as 'mental torture.'

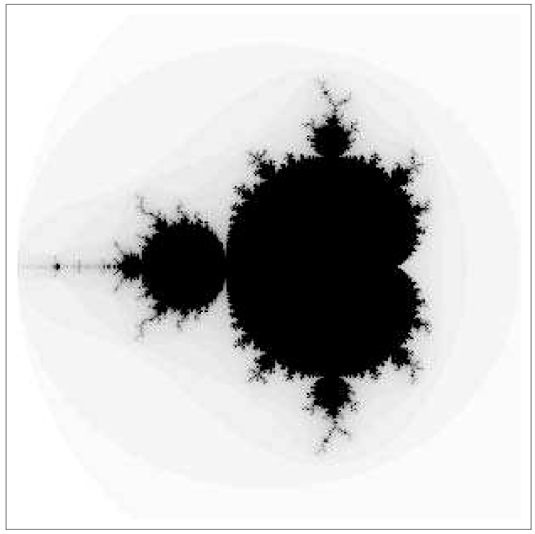

Figure 1.1: The Mandelbrot set, which is a set of complex numbers whose boundary is a fractal.

Example 1.1. *Cardano considered the equation*

$$x^3 = 15x + 4$$

where it is easy by inspection to see that $x = 4$ *is a solution. However, Cardano's general method gave*

$$x = \sqrt[3]{2+\sqrt{-121}} + \sqrt[3]{2-\sqrt{-121}}\,.$$

Now consider

$$\begin{aligned}(2+\sqrt{-1})^3 &= (3+4\sqrt{-1})(2+\sqrt{-1})\\ &= 2+11\sqrt{-1}\\ &= 2+\sqrt{-121}\,,\end{aligned}$$

so, if we are willing to pretend that the negative square root is not problematic, *we have*[1]

$$\begin{aligned} x &= \sqrt[3]{(2+\sqrt{-1})^3} + \sqrt[3]{(2-\sqrt{-1})^3} \\ &= 2+\sqrt{-1}+2-\sqrt{-1} \\ &= 4\,. \end{aligned}$$

Rafael Bombelli introduced the symbol i for $\sqrt{-1}$ in 1572 and René Descartes in 1637 called numbers such as $a+\sqrt{-b}$, where a is any real number and b is a *positive* number, *imaginary numbers*. The term *complex number*, which is the modern term for such numbers, seemed to have originated with Carl Friedrich Gauss in 1831, who also popularized the idea of endowing imaginary quantities with a 'real' existence as points in a plane.

The usage of complex numbers has developed tremendously in the intervening years, and now forms a natural part of coordinate systems, vectors, matrices, and quantum mechanics. As we discover more about advanced physics, complex numbers continue to become ever more significant. Examples include fractals, such as the Mandelbrot set (Fig. 1.1), quantum mechanics, Fourier transforms as used inside your phones and string theory. In fact throughout my career as an optical physicist (MWM), I can hardly remember a day when I haven't used a complex number at some point.

1.1 Imaginary and Complex Numbers

Boas 2.1, 2.2, 2.4

The solution of the quadratic equation

$$ax^2+bx+c=0\,, \tag{1.1}$$

in which a, b, and c are real constants, is the familiar formula

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}\,. \tag{1.2}$$

The key quantity here is the discriminant: b^2-4ac. If $b^2-4ac>0$, then there are two real solutions of Eq. (1.1) while, if $b^2-4ac=0$, there is only a single real solution. What happens if $b^2-4ac<0$? There is clearly no *real* solution of

[1] $\sqrt{3}-2$ and $-\sqrt{3}-2$ are also solutions but that is irrelevant for the example here.

Eq. (1.1) in this case. This is the situation faced by Cardano in his formula for the solution of cubic and quartic equations. If we introduce the symbol

$$i \equiv \sqrt{-1}\,,$$

such that $i^2 = -1$, then we have at least a symbolic solution to this impasse. The symbol i is called the **imaginary unit**.

Example 1.2. *Consider the quadratic equation*

$$x^2 + 4 = 0\,. \tag{1.3}$$

There are no real solutions to this equation, but by writing this equation as

$$x^2 = -4\,,$$

we have that

$$x = \sqrt{-4} = \sqrt{-1 \times 4} = \sqrt{-1} \times \sqrt{4} = \pm 2\,i\,. \tag{1.4}$$

Thus, there are two solutions to Eq. (1.3): $x = -2\,i$ and $x = 2\,i$.

Expressions of the form $a\,i$, in which a is any real number, are called **imaginary numbers**.

Example 1.3. *Consider the quadratic equation*

$$2x^2 - 2x + 1 = 0\,.$$

According to the quadratic formula (1.2), the solutions to this equation are

$$x = \frac{2 \pm \sqrt{4-8}}{4} = \frac{2 \pm \sqrt{-4}}{4} = \frac{1}{2} \pm \sqrt{-\frac{1}{4}}\,.$$

By proceeding as in (1.4), we find that the two solutions of this equation are

$$x = \frac{1}{2} - \frac{i}{2}\,, \qquad x = \frac{1}{2} + \frac{i}{2}\,.$$

Expressions of the form $a + b\,i$, in which a and b are real numbers, are called **complex numbers**. The number a is called the **real part** of the complex number and the number b, the coefficient of i, is called the **imaginary part**. When a complex number is a variable, the conventional notation is $z = x + i\,y$, where x, the real part of z is denoted as $\mathrm{Re}(z)$ and y, the imaginary part of z, as $\mathrm{Im}(z)$:

$$\boxed{z = x + i\,y\,; \quad \mathrm{Re}(z) = x\,, \quad \mathrm{Im}(z) = y} \tag{1.5}$$

1.2 Algebra of Complex Numbers

Boas 2.5

The algebra of complex numbers is similar to that for real numbers, with the proviso that the imaginary unit i has the property that $i^2 = -1$. The rules below show how to add, subtract, multiply, and divide complex numbers to obtain a result that is of the form $x+iy$. In deriving these rules, we will need two properties of complex numbers. Two complex numbers $x+iy$ and $x'+iy'$, in which x, y, x', and y' are real numbers, are equal if and only if their real and imaginary parts are separately equal:

$$x+iy = x'+iy' \quad \text{if and only if} \quad x = x' \quad \text{and} \quad y = y' \,. \tag{1.6}$$

As a special case of this statement, we have that a complex number $x+iy$ is equal to zero if only and only if the real and imaginary parts are each equal to zero:

$$x+iy = 0 \quad \text{if and only if} \quad x = 0 \quad \text{and} \quad y = 0 \,. \tag{1.7}$$

1.2.1 Binary Composition Operations

Consider the addition of two complex numbers $x+iy$ and $x'+iy'$, in which x, y, x', and y' are real numbers:

$$(x+iy)+(x'+iy') = (x+x')+i(y+y') \,, \tag{1.8}$$

that is, the real and imaginary parts are added separately. The rule for subtraction is similarly applied:

$$(x+iy)-(x'+iy') = (x-x')+i(y-y') \,, \tag{1.9}$$

The multiplication of two complex numbers proceeds with the usual rule for distributive property of multiplication over addition:

$$\begin{aligned}(x+iy)(x'+iy') &= x(x'+iy')+iy(x'+iy') \\ &= xx'+ixy'+iyx'+i^2yy' \\ &= (xx'-yy')+i(xy'+yx') \,.\end{aligned} \tag{1.10}$$

The division of complex numbers proceeds in two steps. Consider the quotient

$$\frac{x+iy}{x'+iy'} \,.$$

We first multiply the numerator and denominator by the quantity $x' - iy'$,

$$\left(\frac{x+iy}{x'+iy'}\right)\left(\frac{x'-iy'}{x'-iy'}\right) = \frac{(x+iy)(x'-iy')}{(x'+iy')(x'-iy')}\,.$$

and then carry out the multiplication in the numerator and denominator, using Eq. (1.10):

$$\begin{aligned}\frac{x+iy}{x'+iy'} &= \frac{(xx'+yy')+i(yx'-xy')}{x'^2+y'^2}\\ &= \frac{xx'+yy'}{x'^2+y'^2}+i\,\frac{yx'-xy'}{x'^2+y'^2}\,,\end{aligned} \tag{1.11}$$

where we must mandate that $x' + iy' \neq 0$ for this expression to be meaningful which, according to Eq. (1.7), means that $x' \neq 0$ or $y' \neq 0$. Notice that, in obtaining this quotient, we have used the fact that $i \times (-i) = 1$, which is a particular case of Eq. (1.10) with $x = x' = 0$, $y = 1$ and $y' = -1$. Note also the following special cases of Eq. (1.11). If $x = 1$ and $y = 0$, we obtain the reciprocal of a complex number as

$$\frac{1}{x'+iy'} = \frac{x'}{x'^2+y'^2} - i\,\frac{y'}{x'^2+y'^2}\,. \tag{1.12}$$

If in this equation we first set $x' = 0$, then setting $y' = 1$ and $y' = -1$ in turn, produces

$$\frac{1}{i} = -i \quad \text{and} \quad \frac{1}{-i} = i,$$

respectively.

To summarize, the algebraic rules for combining complex numbers are for real numbers x, y, x' and y'

$$\boxed{\begin{aligned}(x+iy)\pm(x'+iy') &= (x\pm x')+i(y\pm y')\,,\\ (x+iy)(x'+iy') &= (xx'-yy')+i(xy'+yx')\,,\\ \frac{x+iy}{x'+iy'} &= \frac{xx'+yy'}{x'^2+y'^2}+i\,\frac{yx'-xy'}{x'^2+y'^2}\,,\end{aligned}} \tag{1.13}$$

where, in the last operation, $x' \neq 0$ or $y' \neq 0$.

Example 1.4. *The sum of the complex numbers* $1+i$ *and* $2-3i$ *is carried out according to Eq. (1.8):*

$$(1+i)+(2-3i) = 3-2i\,.$$

The square of the complex number $1+i$ *is calculated according to the product rule in Eq. (1.10):*

$$\begin{aligned}(1+i)^2 &= (1+i)(1+i) = 1+2i+i^2 \\ &= 1+2i-1 = 2i\,.\end{aligned}$$

Finally, the quotient of the complex numbers $1-i$ *and* $1+i$ *is calculated according to Eq. (1.11):*

$$\begin{aligned}\frac{1-i}{1+i} &= \frac{(1-i)(1-i)}{(1+i)(1-i)} \\ &= \frac{1-2i+(-i)^2}{1+i-i+i(-i)} = -\frac{2i}{2} = -i\,.\end{aligned}$$

1.2.2 Complex Conjugation

Given a complex number $z = x+iy$, the complex conjugate of z, denoted by z^*, is the complex number

$$\boxed{z^* = x - iy\,.} \tag{1.14}$$

The addition rule for complex numbers in Eq. (1.10) can be used to obtain the real part of a complex number as

$$z + z^* = (x+iy)+(x-iy) = 2x\,,$$

so

$$\boxed{\mathrm{Re}(z) = \frac{z+z^*}{2}\,.} \tag{1.15}$$

Similarly, the imaginary part is obtained as

$$z - z^* = (x+iy)-(x-iy) = 2iy\,,$$

so

$$\boxed{\mathrm{Im}(z) = \frac{z - z^*}{2i}\,.} \tag{1.16}$$

Another useful property of the conjugate is

$$zz^* = (x + iy)(x - iy) = x^2 + y^2\,. \tag{1.17}$$

In particular, the reciprocal of z can be written as

$$\frac{1}{z} = \frac{1}{z}\left(\frac{z^*}{z^*}\right) = \frac{z^*}{zz^*} = \frac{x}{x^2 + y^2} - i\,\frac{y}{x^2 + y^2}\,, \tag{1.18}$$

which agrees with Eq. (1.12).

1.3 The Complex Plane

Boas 2.3

The conventional notation $z = x + iy$ of a generic complex number suggests a graphical representation of complex numbers based on Cartesian coordinates where z is associated with the ordered pair (x, y). The real axis corresponds to the abscissa in this coordinate system and the imaginary part to the ordinate, as depicted in Fig. 1.2(a). Thus, every point corresponds to a complex number. When the x-y plane is used to represent complex numbers in this way, it is referred to as the **complex plane** or as an **Argand diagram**.

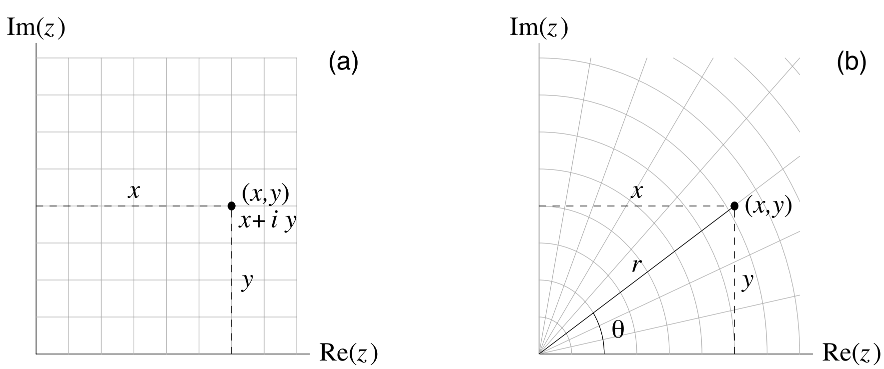

Figure 1.2: The representation of a complex number $z = x + iy$ in (a) rectangular and (b) polar coordinates.

An alternative representation of complex numbers that, for many purposes, is more convenient than Cartesian coordinates, is one based on *polar coordinates*. The basic construction is shown in Fig. 1.2(b). A straight line runs from the origin to the point (x, y). This line is characterized by two quantities: a length r that measures its length, and an angle θ that specifies the angle between this line and the x-axis. By convention, *positive* angles are taken in the *counterclockwise* direction from the x-axis. The relationship between the rectangular and the polar representations of complex numbers can be obtained from basic trigonometry. The rectangular coordinates corresponding to the point (r, θ) are

$$x = r\cos\theta\,, \qquad y = r\sin\theta\,, \tag{1.19}$$

and the polar coordinates of a point (x, y) are

$$r = \sqrt{x^2+y^2} \tag{1.20}$$

$$\theta = \begin{cases} \tan^{-1}\left(\frac{y}{x}\right) & \text{for } x > 0\,, \\ \tan^{-1}\left(\frac{y}{x}\right) + \pi & \text{for } x < 0\,. \end{cases} \tag{1.21}$$

Hence, we can write the polar form of a complex number $z = x + iy$ as

$$\boxed{z = x + iy = r(\cos\theta + i\sin\theta)\,.} \tag{1.22}$$

The quantity r is called the **modulus** or **magnitude** of z, and θ is called the **argument** or **phase**. In the next section we will find a more compact way of writing the right-hand side of this expression that will have many far-reaching consequences.

Example 1.5. *Consider the complex number $z = 4 + i3$. The polar representation of this number is specified by the modulus r and argument θ, which are obtained from the relations in (1.21) as*

$$r = \sqrt{4^2+3^2} = 5\,,$$

$$\theta = \tan^{-1}\left(\tfrac{3}{4}\right).$$

The magnitude of the complex conjugate $z^* = 4 - i3$ of z is the same as that for z, but the phase is now $-\theta$. This is immediately apparent from Eq. (1.22):

$$z^* = x - iy = r(\cos\theta - i\sin\theta) = r\big[\cos(-\theta) + i\sin(-\theta)\big]\,.$$

Notice that the complex conjugate of a complex number is obtained by reflection across the real axis.

1.4 Euler's Formula

Boas 2.9

The polar representation of a complex number in Eq. (1.22) contains the factor $\cos\theta + i\sin\theta$. We will show in this section that this factor has special properties. Suppose that we differentiate with respect to θ:

$$\frac{d}{d\theta}(\cos\theta + i\sin\theta) = -\sin\theta + i\cos\theta\,. \tag{1.23}$$

By writing the negative sign in front of the sine term as i^2, we obtain

$$\frac{d}{d\theta}(\cos\theta + i\sin\theta) = i^2\sin\theta + i\cos\theta = i(\cos\theta + i\sin\theta)\,, \tag{1.24}$$

so the effect of the differentiation is simply to multiply the original expression by a factor of i. Each successive derivative of this expression yields another multiplicative factor of i. Thus, the n^{th} derivative is

$$\frac{d^n}{d\theta^n}(\cos\theta + i\sin\theta) = i^n(\cos\theta + i\sin\theta)\,. \tag{1.25}$$

We now have the required mathematical input to perform a Taylor series expansion of the function $\cos\theta + i\sin\theta$ about $\theta = 0$. Recall that, for a function $f(x)$, the Taylor series about $x = 0$ (the Maclaurin series) has the form

$$\begin{aligned} f(x) &= f(0) + f^{(1)}(0)x + \frac{1}{2!}f^{(2)}(0)x^2 + \frac{1}{3!}f^{(3)}(0)x^3 + \cdots \\ &= \sum_{n=0}^{\infty} f^{(n)}(0)\frac{x^n}{n!}\,, \end{aligned}$$

in which the 'zeroth derivative' is the function itself and $0! = 1$. The notation $f^{(n)}(0)$ means that we take the n^{th} derivative of f and then set $x = 0$. For the function at hand, we have that $\cos 0 = 1$ and $\sin 0 = 0$. Hence, for $n = 1, 2, \ldots$,

$$\left.\frac{d^n}{dx^n}\right|_{\theta=0}(\cos\theta + i\sin\theta) = i^n\,, \tag{1.26}$$

from which we obtain the following Taylor series:

$$\begin{aligned} f(\theta) = \cos\theta + i\sin\theta &= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \cdots \\ &= \sum_{n=0}^{\infty}\frac{(i\theta)^n}{n!}\,. \end{aligned} \tag{1.27}$$

To appreciate the significance of this result, consider the Taylor series for e^x. Using the fact that

$$\frac{d(e^x)}{dx} = e^x\,,$$

and $e^0 = 1$, we obtain

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}\,. \tag{1.28}$$

Letting $x = i\theta$, the comparison of Eqs. (1.27) and (1.28) suggests the identification

$$\boxed{\cos\theta + i\sin\theta \equiv e^{i\theta}\,.} \tag{1.29}$$

This is known as **Euler's formula**. The elegant nature of this formula becomes evident when we evaluate both sides for $\theta = \pi$. With $\cos\pi = -1$ and $\sin\pi = 0$, we obtain

$$\boxed{e^{i\pi} = -1\,.} \tag{1.30}$$

The exponential function e^x for real x is *never* negative. But, by allowing for an imaginary variable in the argument of the exponential, negative values arise quite naturally. Indeed, the familiar rules for the manipulation of the exponential function, which can be derived from the Taylor series (1.28)[2], combined with imaginary arguments, leads to some new and useful results. Equations (1.22) and the polar representation in (1.29) imply an especially compact representation of $z = x + iy$:

$$\boxed{z = r\,e^{i\theta}\,.} \tag{1.31}$$

Example 1.6. *Consider the multiplication of two complex numbers with unit modulus* $(r = 1)$ *and with arguments* θ *and* ϕ*:*

$$\cos\theta + i\sin\theta \;=\; e^{i\theta}$$

$$\cos\phi + i\sin\phi \;=\; e^{i\phi}\,.$$

[2]We have, for the moment, side-stepped the question of convergence of the series (1.27). This will addressed in the next chapter.

The product of these equations is

$$(\cos\theta + i\sin\theta)(\cos\phi + i\sin\phi) = e^{i\theta}e^{i\phi}\,.$$

Expanding the left-hand side according to Eq. (1.10) yields

$$\cos\theta\cos\phi - \sin\theta\sin\phi + i(\sin\theta\cos\phi + \cos\theta\sin\phi)\,.$$

By using the standard rule for the product of exponential functions, i.e. $e^a e^b = e^{a+b}$*, together with Eq. (1.29), we obtain*

$$e^{i\theta}e^{i\phi} = e^{i(\theta+\phi)} = \cos(\theta+\phi) + i\sin(\theta+\phi)\,.$$

Equating the real and imaginary parts in the last two equations produces

$$\cos(\theta+\phi) = \cos\theta\cos\phi - \sin\theta\sin\phi\,,$$

$$\sin(\theta+\phi) = \sin\theta\cos\phi + \cos\theta\sin\phi\,,$$

which are the trigonometric identities for the sine and cosine of the sum of two angles! Other identities can be derived by following analogous steps. Quite apart from providing an additional level of confidence in our usage of Eq. (1.29), this procedure is far easier to apply than the conventional method of deriving such identities.

Finally, we note that the representation (1.31) is especially convenient for taking products and quotients of complex numbers. For any two complex numbers $z = re^{i\theta}$ and $z' = r'e^{i\theta'}$, we have

$$\boxed{\begin{aligned} zz' &= rr'e^{i(\theta+\theta')}\,,\\ \frac{z}{z'} &= \frac{r}{r'}e^{i(\theta-\theta')}\,. \end{aligned}} \qquad (1.32)$$

Chapter 2

Functions of Complex Variables

The binary operations of addition, subtraction, multiplication, and division of complex numbers are the basis for the assembly of composite expressions of complex quantities, such as polynomials and power series. This opens the way to extending functions of real variables to functions of variables that extend over the complex plane. Functions f of independent variables x and y that depend only on the combination $z = x + iy$ are called **functions of a complex variable** and are denoted by $f(z)$. We will be concerned in this course with what are called the elementary functions: powers and roots, trigonometric functions and their inverses, exponential and logarithmic functions, as well combinations of such functions. These functions can be defined by writing, for example, $\sin z$, e^z, and $\log z$, so that they become complex-valued quantities, having real and imaginary parts. Complex-valued functions can exhibit some quite unexpected behavior compared to their real counterparts, though most of the standard functions are real when their arguments are real. An obvious exception is the square root function, which becomes imaginary for negative arguments.

We will begin this chapter be examining the powers and roots of complex numbers. Since we can multiply z by itself and by any other complex number, we can form any polynomial in z and, by extension, any power series. This will enable us to define the power series of functions such as the exponential and trigonometric functions by their Taylor series expansions. By adapting the discussion of the convergence of real series to complex-valued series, the Taylor series of elementary functions of complex variables will be shown to retain the convergence properties of their real-valued counterparts. Since the properties of these functions can be derived from their Taylor series, they retain their familiar (and useful) properties for real arguments, while exhibiting a richer analytic structure with complex arguments.

2.1 Powers of Complex Numbers

Boas 2.11

As we noted in Sec. 1.4, products of complex numbers are most easily carried out in the polar form $z = r e^{i\theta}$. The n-fold product of z is

$$z^n = \left(r e^{i\theta}\right)^n = r^n e^{in\theta} .$$

Note that, for the particular case that $r = 1$, we have that

$$\begin{aligned}\left(e^{i\theta}\right)^n &= (\cos\theta + i\sin\theta)^n \\ &= e^{in\theta} = \cos n\theta + i\sin n\theta ,\end{aligned}$$

which leads to **De Moivre's theorem**

$$\boxed{(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta .} \tag{2.1}$$

Example 2.1. *Consider De Moivre's theorem for $n = 2$:*

$$\begin{aligned}(\cos\theta + i\sin\theta)^2 &= \cos^2\theta - \sin^2\theta + 2i\sin\theta\cos\theta \\ &= \cos 2\theta + i\sin 2\theta .\end{aligned}$$

By equating real and imaginary parts in this equation, we obtain

$$\begin{aligned}\cos 2\theta &= \cos^2\theta - \sin^2\theta , \\ \sin 2\theta &= 2\sin\theta\cos\theta ,\end{aligned}$$

which are standard double-angle trigonometric identities. Higher multiple-angle identities are derived with an analogous procedure.

Powers of complex numbers are used in polynomials and powers series. We consider first an example of successive powers of complex numbers.

Example 2.2. *Consider the powers of $z = 1 + i$. The successive powers z^n are straightforward to calculate:*

$$z^2 = 2i, \quad z^3 = -2 + 2i, \quad z^4 = -4 , \quad z^5 = -4 - 4i, \quad z^6 = -8i.$$

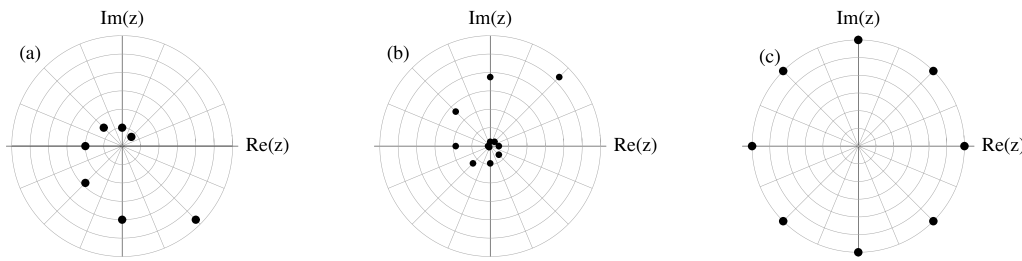

Figure 2.1: The powers z^n of (a) $z = 1+i$, (b) $z = \frac{1}{2}(1+i)$, and (c) $z = (1+i)/\sqrt{2}$. In each case, successive points are rotated by $\frac{\pi}{4}$ in the counterclockwise direction.

These points are plotted in the complex plane in Fig. 2.1(a). The expanding spiral of the sequence of these powers is evident. The reason for this type of structure can be seen from the polar representation

$$1+i = \sqrt{2}\, e^{i\frac{\pi}{4}} ,$$

in terms of which the nth power is

$$(1+i)^n = (\sqrt{2})^n\, e^{i\frac{n\pi}{4}} = 2^{n/2}\, e^{i\frac{n\pi}{4}} .$$

Thus, each successive power results in a rotation by $\frac{\pi}{4}$ in the counterclockwise direction and, since $\sqrt{2} > 1$, a larger radius, producing the expanding spiral.

Alternatively, if we now consider the sequence of powers of $\frac{1+i}{2}$, the corresponding polar representation yields

$$\left[\frac{1+i}{2}\right]^n = \left(\frac{\sqrt{2}}{2}\right)^n e^{i\frac{n\pi}{4}} = 2^{-n/2}\, e^{i\frac{n\pi}{4}} ,.$$

Each successive power still results in a rotation by $\frac{\pi}{4}$ in the counterclockwise direction but, because the magnitude $\frac{\sqrt{2}}{2} < 1$, the corresponding points produce a spiral that converges toward the origin, as shown in Fig. 2.1(b).

The marginal case is obtained for

$$z = \frac{1+i}{\sqrt{2}} ,$$

which has unit magnitude. The polar representations of the successive powers of this complex number are

$$\left[\frac{\sqrt{2}}{2}(1+i)\right]^n = e^{i\frac{n\pi}{4}} ,$$

which produces a cycle of eight points that are again separated by $\frac{\pi}{4}$ as shown in Fig. 2.1(c).

From these examples, the general expression for the powers of a complex number $x + iy = r\,e^{i\theta}$ is

$$(x+iy)^n = r^n\,e^{in\theta}\,.$$

We can use De Moivre's theorem (2.1) to express these products in rectangular form as

$$\boxed{(x+iy)^n = r^n\,e^{in\theta} = r^n\big(\cos n\theta + i\sin n\theta\big)\,.} \tag{2.2}$$

By invoking the usual rules for exponents, this expression is valid for *all* integers.

2.2 Roots of Complex Numbers

Boas 2.10

We now consider the inverse process of taking powers of complex numbers, taking their roots. The general problem is to find a solution of the equation $z^n = x' + iy'$, where $z = x + iy$ and x' and y' are real numbers. We will focus on the case of integer n in this section, and leave the more general case for a later section. We again proceed by working through an example.

Example 2.3. *Consider the solution of $z^4 = 16$, i.e. the fourth root of 16. Two solutions, $z = 2$ and $z = -2$, can be obtained without reference to complex numbers. But, to obtain all the roots of this equation, we must use complex numbers. We begin by writing the equation in terms of the polar representation of complex numbers, $z = r\,e^{i\theta}$*

$$(r\,e^{i\theta})^4 = r^4\,e^{i4\theta} = 16\,. \tag{2.3}$$

The equation for the magnitude, $r^4 = 16$, yields $r = 2$, since $r \geq 0$ always. To obtain the solutions for the phase, we first write

$$16 = 16\,e^{i2n\pi}\,, \tag{2.4}$$

for $n = 0, 1, 2, \ldots$, which clearly leaves the value of the magnitude unaffected since $e^{i2n\pi} = 1$ for any integer n. Thus, our solution for r is therefore also unaffected. However, the solutions for the phases of our roots now read

$$4\theta = 0,\, 2\pi,\, 4\pi,\, 6\pi,\, 8\pi, \ldots\,. \tag{2.5}$$

The reason for considering the additional angles in (2.4) becomes apparent when we divide both sides of this equation by 4 to obtain the solutions

$$\theta = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}, 2\pi, \ldots . \tag{2.6}$$

Since $\theta = 0$ and $\theta = 2\pi$ correspond to the same angle, and since each successive rotation by 2π on the right-hand side of (2.5) yields an additive factor of $\frac{\pi}{2}$ in (2.6), we find that there are four distinct values of θ. These yield four solutions of (2.3):

$$z_1 = 2\,e^{i0} = 2\,, \tag{2.7}$$

$$z_2 = 2\,e^{\frac{1}{2}i\pi} = 2\,i\,, \tag{2.8}$$

$$z_3 = 2\,e^{i\pi} = -2, \tag{2.9}$$

$$z_4 = 2\,e^{\frac{3}{2}i\pi} = -2\,i\,. \tag{2.10}$$

Thus, in addition to two real roots, we have also found two imaginary roots, which are complex conjugates.

Example 2.4. *Consider the solution of $z^3 = 1+i$, i.e. the cube roots of $1+i$. The most expedient method of solution is again based on the polar representation of complex numbers:*

$$z = r\,e^{i\theta}\,, \qquad 1+i = \sqrt{2}\,e^{i\frac{\pi}{4}}\,. \tag{2.11}$$

The equation to be solved is

$$\begin{aligned}(r\,e^{i\theta})^3 &= r^3\,e^{i3\theta} \\ &= \sqrt{2}\,e^{i\frac{\pi}{4}} = \sqrt{2}\,e^{i\left(\frac{\pi}{4}+2n\pi\right)}\,,\end{aligned} \tag{2.12}$$

where we have again added multiples of 2π to the phase of $1+i$. The equation for the magnitude, $r^3 = \sqrt{2}$, yields $2^{1/6}$. For the phase, we have

$$3\theta = \frac{\pi}{4}, \frac{\pi}{4}+2\pi, \frac{\pi}{4}+4\pi, \frac{\pi}{4}+6\pi, \ldots .$$

Dividing both sides of this equation by 3, we obtain the solutions

$$\begin{aligned}\theta &= \frac{\pi}{12}, \frac{\pi}{12}+\frac{2\pi}{3}, \frac{\pi}{12}+\frac{4\pi}{3}, \frac{\pi}{12}+\frac{6\pi}{3}, \ldots \\ &= \frac{\pi}{12}, \frac{9\pi}{12}, \frac{17\pi}{12}, \frac{25\pi}{12}, \ldots .\end{aligned}$$

Since

$$\frac{25\pi}{12} = \frac{\pi}{12} + 2\pi,$$

there are three distinct values of θ*:*

$$\theta = \frac{\pi}{12}, \frac{9\pi}{12}, \frac{17\pi}{12} = 15^\circ, 135^\circ, 255^\circ\,.$$

The three solutions of Eq. (2.12) are therefore given by

$$\begin{aligned} z_1 &= 2^{1/6}\, e^{i\frac{\pi}{12}} \\ z_2 &= 2^{1/6}\, e^{i\frac{9\pi}{12}} \\ z_3 &= 2^{1/6}\, e^{i\frac{17\pi}{12}}\,. \end{aligned}$$

There are several points about the calculations in Examples 2.3 and 2.4 that are general characteristics of roots of complex numbers. The nth root of a complex number has n distinct solutions, whose points in the complex plane are separated by an angle of $\frac{2\pi}{n}$. When connected by straight lines, these points form the vertices of a regular n-sided polygon. Figure 2.2 shows the roots calculated in Examples 2.3 and 2.4. In Example 2.3, the quartic roots of 16 lie on a circle of radius 2, separated by $\frac{\pi}{2} = 90^\circ$, forming a square, while the cube roots of $1+i$ in Example 2.4 lie on a circle of radius $2^{1/6}$ and are separated by $\frac{2\pi}{3} = 120^\circ$, forming an equilateral triangle.

We can now deduce the general expression for the roots of a complex number. The solutions of $z^n = x' + i\,y'$, where x' and y' are real numbers, are obtained by first writing the equation in polar form: $r^n\, e^{in\theta} = \rho\, e^{i\phi}$, where ρ and ϕ are the

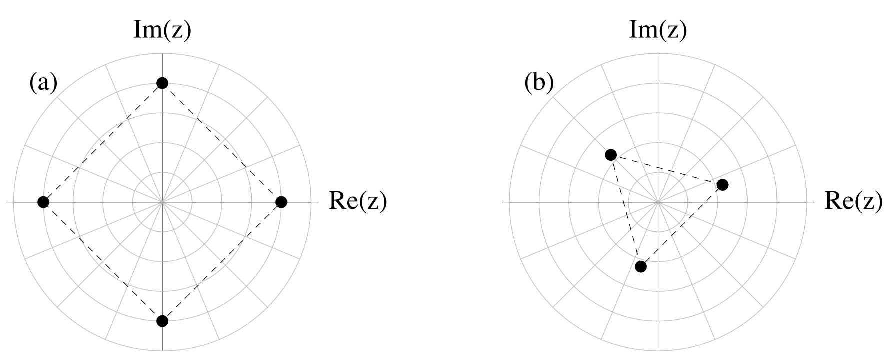

Figure 2.2: The solutions of (a) $z = 16^{1/4}$ and (b) $z = (1+i)^{1/3}$ plotted in the complex plane.

magnitude and phase, respectively, of $x' + iy'$. The nth roots of this number each have magnitude $\rho^{1/n}$ and phases given by the n values

$$\theta = \frac{\phi}{n}, \frac{\phi + 2\pi}{n}, \frac{\phi + 4\pi}{n}, \ldots, \frac{\phi + 2\pi(n-1)}{n}.$$

By invoking De Moivre's theorem (2.1) we can write this result as follows. Given a complex number $x' + iy'$ whose polar form is $\rho\, e^{i\phi}$, the solution of $z^n = \rho\, e^{i\phi}$ is given by the n complex numbers z_k, for $k = 0, \ldots, n-1$:

$$\begin{aligned} z_k &= \rho^{1/n} \exp\left[i \left(\frac{\phi}{n} + \frac{2k\pi}{n} \right) \right] \\ &= \rho^{1/n} \left[\cos\left(\frac{\phi}{n} + \frac{2k\pi}{n} \right) + i \sin\left(\frac{\phi}{n} + \frac{2k\pi}{n} \right) \right]. \end{aligned} \tag{2.13}$$

We can now understand how the plot of the nth roots of a complex number $\rho\, e^{i\phi}$ appear on the complex plane. There are n equally spaced points on a circle of radius $\rho^{1/n}$, with adjacent points separated by an angle of $\frac{2\pi}{n}$. The line that connects the points is a regular n-sided polygon, so the roots are the vertices of this polygon. The polygon is tilted by $\frac{\phi}{n}$, so that if $\phi = 0$, i.e. the number whose root is taken is real, at least one vertex (root) lies on the real axis.

2.3 Complex Power Series

Boas 2.7

A polynomial of order n in the real variable x has the form

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

in which the coefficients a_k are real numbers. A polynomial is a continuous function of x and is finite for any finite value of x. Only for $x \to \pm\infty$ does the polynomial become infinite. The derivatives of all orders exist and are continuous, although all but the first n are identically zero.

For any real function $f(x)$ we can write the infinite series

$$g(x) = a_0 + a_1 x + a_2 x^2 + \cdots, \tag{2.14}$$

where

$$a_n = \frac{f^{(n)}(0)}{n!} = \frac{1}{n!} \frac{d^n f}{dx^n}\bigg|_{x=0}. \tag{2.15}$$

If the derivatives are bounded, then the series converges to $f(x)$ everywhere. Remarkably, the values of the function everywhere are determined by the values of the function and its derivatives at the origin. By the convergence of an infinite series, we mean that the sum of enough terms is as close to a fixed value S as required, where S is the sum of the series. In other words, the sum of the series approaches the value S as a limit. We have based our discussion on real variables x, but polynomials and infinite series can be constructed where x is replaced by the complex variable $z = x + iy$.

There are several tests that determine if an infinite series converges. One of the simplest to apply is the **ratio test**. For an infinite series of complex numbers $\sum_{n=0}^{\infty} A_n$, consider the ratio ρ, defined by

$$\rho = \lim_{n\to\infty} \left| \frac{A_{n+1}}{A_n} \right| . \tag{2.16}$$

The complex series converges if $\rho < 1$ and diverges if $\rho > 1$. If $\rho = 1$, the test is inconclusive and another test for convergence must be used.

Example 2.5. *Consider the complex geometric series*

$$\sum_{n=0}^{\infty} z^n = 1 + z + z^2 + \cdots . \tag{2.17}$$

In the notation used above, $A_n = z^n$. Then, the quantity ρ in the ratio test is

$$\rho = \lim_{n\to\infty} \left| \frac{z^{n+1}}{z^n} \right| = |z| .$$

Thus, the geometric series (2.17) converges if $|z| < 1$. Since $|z| = 1$ can be written in polar form as $\sqrt{x^2 + y^2} = 1$, which is equivalent to $x^2 + y^2 = 1$, the geometric series converges for complex numbers z within a circle of unit radius centered at the origin. For this reason, we refer to the **radius of convergence** *of a series (even when the series resides on the real line).*

Figure 2.3 shows the sequence of partial sums

$$S_N = \sum_{n=0}^{N} z^n$$

for $z = \frac{1+i}{2}$, which lies within the radius of convergence of the geometric series (2.17), and for $z = 1 + i$, which lies outside of this radius. The convergence in the first case is seen by the spiral that converges toward $1 + i$, which is the sum of the series. In the second case, however, the partial sums lie on a diverging spiral, which is indicative of the divergence of the series.

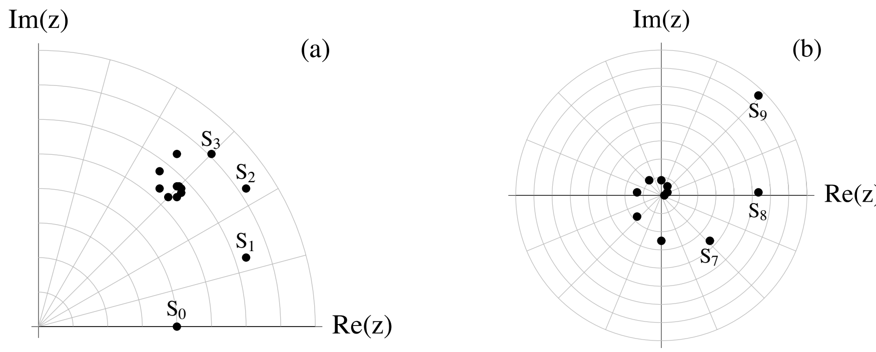

Figure 2.3: The partial sums of the geometric series (2.17) for (a) $z = \frac{1}{2}(1+i)$ and (b) $z = 1+i$. In (a) the partial sums are seen to converge toward $1+i$, which is the sum of the series, while in (b) the partial sums form an expanding spiral because $1+i$ lies outside the radius of convergence of this series.

For our purposes, the most important infinite series is the Maclaurin series for e^z, which is defined by

$$\sum_{n=0}^{\infty} \frac{z^n}{n!}\,. \tag{2.18}$$

This is a generalization of the function e^x for real x and a generalization of the function $e^{i\theta}$ that we introduced in Sec. 1.4. In the notation of the ratio test, we have $A_n = z^n/n!$, so

$$\rho = \lim_{n\to\infty}\left|\frac{z^{n+1}}{(n+1)!}\frac{n!}{z^n}\right| = \lim_{n\to\infty}\frac{|z|}{n+1} = 0\,.$$

Thus, for any complex number with a finite magnitude, we obtain $\rho < 1$, so the series (2.18) has an infinite radius of convergence, and we can now formally define e^z as

$$\boxed{e^z \equiv \sum_{n=0}^{\infty}\frac{z^n}{n!}\,,} \tag{2.19}$$

for any complex number $z = x+iy$. This representation is valid over the entire complex plane and subsumes the special cases just mentioned, namely, e^x and $e^{i\theta}$. As an immediate consequence, we have

$$e^z = e^{x+iy} = e^x e^{iy} = e^x(\cos y + i\sin y)\,.$$

We will explore the complex exponential in more detail in the next section.

2.4 The Complex Exponential

Boas 2.11

We have seen in the preceding section that the complex function e^z converges everywhere in the complex plane. In this section, we will show that this function has the all the properties of the real function e^x and, moreover, that these follow from the series representation (2.19).

2.4.1 The Cauchy Product

We first derive the product of power series. Suppose that we have two series,

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + \cdots .$$

$$\sum_{n=0}^{\infty} b_n = b_0 + b_1 + b_2 + \cdots .$$

Their product is formed as follows

$$\begin{aligned}\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) &= \left(a_0 + a_1 + a_2 + \cdots\right)\left(b_0 + b_1 + b_2 + \cdots\right) \\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + (a_0 b_2 + a_1 b_1 + a_2 b_0) + \cdots .\end{aligned} \tag{2.20}$$

Note that this product has been written as an ascending series of linear combinations of the $a_i b_j$ such that $i + j = n$. Hence, we can write this product as a sum

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) . \tag{2.21}$$

This is known as the **Cauchy product** of two series.

2.4.2 Products of Complex Exponentials

Consider the familiar rule: $e^{z_1+z_2} = e^{z_1} e^{z_2}$ for complex numbers z_1 and z_2. That this follows from the power series (2.19) can be seen by applying the Cauchy product

$$e^{z_1}\,e^{z_2} = \left(\sum_{n=0}^{\infty}\frac{z_1^n}{n!}\right)\left(\sum_{n=0}^{\infty}\frac{z_2^n}{n!}\right) = \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{z_1^k}{k!}\frac{z_2^{n-k}}{(n-k)!}\,. \tag{2.22}$$

The right-hand side has been obtained from (2.21) by making the identifications

$$a_n = \frac{z_1^n}{n!}\,, \qquad b_n = \frac{z_2^n}{n!}\,.$$

The interior sum on the right-hand side of (2.22) can be simplified by multiplying and dividing by $n!$

$$\sum_{k=0}^{n}\frac{z_1^k}{k!}\frac{z_2^{n-k}}{(n-k)!} = \frac{n!}{n!}\sum_{k=0}^{n}\frac{z_1^k}{k!}\frac{z_2^{n-k}}{(n-k)!} = \frac{1}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}z_1^k z_2^{n-k}\,.$$

By recalling the binomial theorem,

$$(a+b)^n = \sum_{k=0}^{n}\frac{n!}{k!(n-k)!}a^k b^{n-k}\,,$$

we see that

$$\frac{1}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}z_1^k z_2^{n-k} = \frac{(z_1+z_2)^n}{n!}\,,$$

from which we conclude that

$$e^{z_1}\,e^{z_2} = \sum_{n=0}^{\infty}\frac{(z_1+z_2)^n}{n!} = e^{z_1+z_2}\,. \tag{2.23}$$

As special cases, we have

$$e^z = e^{x+iy} = e^x\,e^{iy} = e^x(\cos y + i\sin y)\,,$$

and

$$\frac{1}{e^z} = \frac{1}{e^x(\cos y + i\sin y)} = \frac{e^{-x}}{\cos y + i\sin y}\left(\frac{\cos y - i\sin y}{\cos y - i\sin y}\right)$$

$$= e^{-x}(\cos y - i\sin y) = e^{-x}\,e^{-iy} = e^{-(x+iy)} = e^{-z}\,.$$

Thus, combining this result with that in (2.23), we find that, for any integer n, $(e^z)^n = e^{nz}$.

The main properties of the complex exponential function are summarized below

$$e^z = e^x(\cos y + i\sin y)\,, \tag{2.24}$$

$$|e^z| = e^x > 0\,, \tag{2.25}$$

$$e^{z_1+z_2} = e^{z_1}\,e^{z_2}\,, \tag{2.26}$$

$$e^{-z} = \frac{1}{e^z}\,, \tag{2.27}$$

$$(e^z)^n = e^{nz}\,, \quad \text{for any integer } n\,. \tag{2.28}$$

2.5 Complex Trigonometric Functions

Boas 2.11, 2.12

The properties of the complex exponential function can be used to define trigonometric functions with complex arguments. We begin with Euler's formula (1.29)

$$e^{i\theta} = \cos\theta + i\sin\theta\,, \tag{2.29}$$

$$e^{-i\theta} = \cos\theta - i\sin\theta\,. \tag{2.30}$$

Taking the sum of these equations causes the imaginary part to vanish, leaving

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

or,

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}\,. \tag{2.31}$$

Similarly, subtracting (2.30) from (2.29) yields

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

or,

$$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}\,. \tag{2.32}$$

With the trigonometric functions expressed in terms of the exponential function, we can now examine the properties of these functions with complex arguments. Consider first the complex cosine function:

$$\begin{aligned}
\cos z &= \cos(x+iy) = \tfrac{1}{2}\left[e^{i(x+iy)} + e^{-i(x+iy)}\right] \\
&= \tfrac{1}{2}\left(e^{ix}e^{-y} + e^{-ix}e^{y}\right) \\
&= \tfrac{1}{2}\left[(\cos x + i\sin x)\,e^{-y} + (\cos x - i\sin x)\,e^{y}\right] \\
&= \cos x\left(\frac{e^{y}+e^{-y}}{2}\right) - i\sin x\left(\frac{e^{y}-e^{-y}}{2}\right) \\
&= \cos x\cosh y - i\sin x\sinh y\,, \qquad (2.33)
\end{aligned}$$

where we have used the hyperbolic functions

$$\cosh y = \frac{e^{y}+e^{-y}}{2} \qquad (2.34)$$

$$\sinh y = \frac{e^{y}-e^{-y}}{2}\,. \qquad (2.35)$$

Since the trigonometric and hyperbolic functions are real-valued functions, we have that

$$\mathrm{Re}(\cos z) = \cos x\cosh y\,, \qquad \mathrm{Im}(\cos z) = -\sin x\sinh y\,.$$

Note, in particular, the special case where $x = 0$:

$$\boxed{\cos iy = \cosh y\,,}$$

so the hyperbolic cosine corresponds to the cosine of an imaginary angle. We could also have deduced this from the exponential representations of these functions. A similar calculation to that leading to (2.33) for the complex sine function yields

$$\sin(x+iy) = \sin x\cosh y + i\cos x\sinh y\,, \qquad (2.36)$$

and

$$\boxed{\sin iy = i\sinh y\,.}$$

Equations (2.33) and (2.36) can be used to determine the moduli of $\cos z$ and $\sin z$. For $\cos z$, we have

$$\begin{aligned}
|\cos z|^2 &= \cos^2 x \cosh^2 y + \sin^2 x \sinh^2 y \\
&= \cos^2 x (1 + \sinh^2 y) + \sin^2 x \sinh^2 y \\
&= \cos^2 x + \sinh^2 y (\cos^2 x + \sin^2 x) \\
&= \cos^2 x + \sinh^2 y ,
\end{aligned}$$

which implies that

$$|\cos z| = \sqrt{\cos^2 x + \sinh^2 y} .$$

A similar calculation produces

$$|\sin z| = \sqrt{\sin^2 x + \sinh^2 y} .$$

These relations show that $\cos z$ and $\sin z$ are unbounded. For example,

$$|\cos z| = \sqrt{\cos^2 x + \sinh^2 y} \geq \sqrt{\sinh^2 y} = |\sinh y| .$$

Referring to (2.35), we see that, since

$$\sinh y = \frac{e^y - e^{-y}}{2} ,$$

then, as $y \to \infty$, $\sinh y \to \infty$ and, as $y \to -\infty$, $\sinh y \to -\infty$. Hence, $|\cos z| \to \infty$ as $|\mathrm{Im}(z)| \to \infty$.

All of the properties of the complex trigonometric functions follow from the basic relations in (2.33) and (2.36):

$$\cos z = \cos x \cosh y - i \sin x \sinh y , \tag{2.37}$$

$$\sin z = \sin x \cosh y + i \cos x \sinh y . \tag{2.38}$$

The other standard trigonometric functions can be obtained from these relations. For example, the complex tangent $\tan z$ and the complex secant $\sec z$ are defined as

$$\tan z = \frac{\sin z}{\cos z} , \qquad \sec z = \frac{1}{\cos z} , \ (\cos z \neq 0) .$$

2.6 The Complex Logarithm

Boas 2.13

For real variables, the logarithm is defined as the inverse of the exponential function. Accordingly, if $y = e^x$, then $x = \ln y$, where 'ln' signifies the natural logarithm, i.e. the logarithm to base e. The same relationship exists between the complex exponential and the complex logarithm. For complex numbers z and w,

$$z = e^w \quad \longrightarrow \quad w = \ln z\,. \tag{2.39}$$

We see immediately from this definition that, for complex numbers z_1, z_2, w_1, and w_2, where $z_1 = e^{w_1}$ and $z_2 = e^{w_2}$,

$$z_1 z_2 = e^{w_1} e^{w_2} = e^{w_1 + w_2}\,,$$

which, according to (2.39), implies that

$$\ln z_1 z_2 = w_1 + w_2 = \ln z_1 + \ln z_2\,,$$

so the familiar rule for the logarithm of a product is recovered. The evaluation of the complex logarithm is most naturally carried out in terms of the polar representation $z = r e^{i\theta}$:

$$\begin{aligned} w = \ln z &= \ln(r e^{i\theta}) \\ &= \ln r + \ln e^{i\theta} \\ &= \ln r + i\theta\,. \end{aligned} \tag{2.40}$$

The first term on the right-hand side of this equation is the usual logarithm of the real positive number r. The second term has an inherent ambiguity, which can be seen from the polar representation. Since a rotation of θ by any integer multiple of 2π leaves the polar representation unaltered, i.e.

$$z = r e^{i\theta} = r e^{i(\theta + 2n\pi)}\,,$$

for any integer n, the logarithm is multi-valued, since for each value of n

$$w = \ln r + i(\theta + 2n\pi)\,,$$

corresponds to the same point. To define a *unique* logarithm associated with a complex number, we must restrict the range of θ to an interval of length 2π.

There are two such intervals that are commonly used: (i) $0 \leq \theta < 2\pi$, and (ii) $-\pi < \theta \leq \pi$. We will use (ii) here. Thus,

$$\boxed{\ln z = \ln r + i\theta\,, \quad -\pi < \theta \leq \pi\,.} \tag{2.41}$$

This definition means that the complex logarithmic function has a discontinuity along a line from the origin along the real axis in the negative direction.

Example 2.6. *Unlike the logarithm for real arguments, the complex logarithm can also be defined for negative real numbers. Suppose we have a negative number x, which we represent as $-|x|$. Since the negative real numbers correspond to $\theta = \pi$ in the complex plane,*

$$\ln(-|x|) = \ln|x| - i\pi.$$

Consider the logarithm of $z = 1 + i$. In polar form, $z = \sqrt{2}\,e^{\frac{1}{4}i\pi}$, so

$$\ln(1+i) = \ln\sqrt{2} + \tfrac{1}{4}\,i\pi.$$

2.7 Complex Powers

Boas 2.14

In analogy with real powers, we can define a complex power of a non-zero complex number by utilizing the inverse relationship between the exponential and the logarithm:

$$z^a = e^{\ln z^a} = e^{a\ln z}\,, \tag{2.42}$$

in which we interpret the logarithm as in the preceding section in terms of its principal value.

Example 2.7. *Consider the complex number $(1+i)^2$. We will evaluate this quantity in two ways: by direct expansion and by applying (2.42). We first calculate*

$$(1+i)^2 = (1+i)(1+i) = 1 + 2i - 1 = 2i.$$

Alternatively, with $1 + i = \sqrt{2}\,e^{\frac{1}{4}i\pi}$, we have

$$(1+i)^2 = e^{2\ln(1+i)} = e^{2(\ln\sqrt{2}+\frac{1}{4}i\pi)} = e^{\ln 2}\,e^{\frac{1}{2}i\pi} = 2i,$$

which agrees with the result of the direct expansion.

Consider now the evaluation of i^i. *With* $i = e^{\frac{1}{2}i\pi}$, *we obtain*

$$i^i = e^{i\ln i} = e^{i(\frac{1}{2}i\pi)} = e^{-\frac{1}{2}\pi},$$

which is a real number! Finally, we calculate $(1+i)^{1+i}$*:*

$$\begin{aligned}(1+i)^{1+i} &= e^{(1+i)\ln(1+i)} \\ &= e^{(1+i)(\ln\sqrt{2}+\frac{1}{4}i\pi)} \\ &= e^{(1+i)\ln\sqrt{2}+\frac{1}{4}(1+i)i\pi} \\ &= e^{\ln\sqrt{2}-\frac{1}{4}\pi}e^{i(\ln\sqrt{2}+\frac{1}{4}\pi)} \\ &= \sqrt{2}e^{-\frac{1}{4}\pi}\left[\cos\left(\ln\sqrt{2}+\tfrac{1}{4}\pi\right)+i\sin\left(\ln\sqrt{2}+\tfrac{1}{4}\pi\right)\right].\end{aligned}$$

2.8 Analytic functions

Boas 14.1, 14.2

Having extended first the basic operations and afterwards a set of functions from the real numbers to the complex numbers, it becomes natural to ask if there are other areas where one can do a similar extension. Can we differentiate complex numbers, can we integrate? This leads to the field of **complex algebra** where we will only touch on the very basic aspects. Already in the aspects of continuity and differrentiability we will see that complex algebra is a very rich field.

Is a function really a *function of z* rather than a function of the Cartesian components, i.e. can we express it as a direct function of $x+iy$. Let us consider a few examples

Example 2.8. *The function*

$$f(z) \equiv x^2 - y^2 + 2ixy$$

is a direct function of $z = x + iy$ *as it can be written as*

$$f(z) = (x+iy)^2.$$

Now consider the slightly different looking function

$$g(z) \equiv x^2 + y^2 + 2ixy.$$

If we try to express this function as a polynomial

$$g(z) = \alpha_0 + \alpha_1(x+iy) + \alpha_2(x+iy)^2$$

where $\alpha_i \in \mathbb{R}$. *Evaluating for* $y = 0$, *it is clear that* $\alpha_0 = 0$, $\alpha_1 = 0$ *and* $\alpha_2 = 1$. *But that implies*

$$g(z) = 0 + 0(x+iy) + 1(x+iy)^2 = (x+iy)^2 = x^2 - y^2 + 2ixy$$

which is manifestly not true. Hence g *is* not *a direct function of* $x+iy$. *Later we will see that* f *is an analytic function while* g *is not. Trying other functions of similar type to* f *and* g *will show that analytic functions are in fact quite rare.*

2.8.1 Continuity

For a real function, continuity at a point x_0 means that

$$\lim_{x \to x_0} f(x) = f(x_0),$$

where we have to consider the limit for the two cases $x < x_0$ and $x > x_0$. When trying to transfer this definition to complex numbers, the first thing we realise is that there is no **ordering** of complex numbers. In other words, it makes no sense to ask if $z < z'$ for $z, z' \in \mathbb{C}$. On the other hand, we can define a **metric** which measures the distance between two complex numbers. We have

$$d = |z - z'| \tag{2.43}$$

which leads to the fact that $d = 0$ if and only if $z = z'$.

We say that a complex function f is continuous at a point z_0 if and only if

$$\lim_{z \to z_0} f(z) = \lim_{|z - z_0| \to 0} f(z) = f(z_0). \tag{2.44}$$

For a real function to prove that it is continuous we just have to calculate the limit approaching from the left and from the right. However, for a complex function we can approach z_0 in an infinite number of different ways and the limit always has to be the same. To prove that a function is continuous at a point z_0 can thus be very hard. To prove that it is not continuous is easier; we simply have to approach z_0 in two different ways and show that we get different results (or that the limit doesn't exist at all).

Example 2.9. *Let $f(z) = z^2$ and test if f is continuous at z_0. We write $z = z_0 + \delta$ giving*

$$\lim_{z\to z_0} f(z) = \lim_{|\delta|\to 0} (z_0+\delta)^2 = z_0^2$$

as we have

$$|(z_0+\delta)^2 - z_0^2| = |\delta^2 + 2z\delta||\delta| \to 0 \text{ as } |\delta| \to 0.$$

We have thus shown that f is continuous at every point of the complex plane.

Being able to isolate the '$|\delta|$' term as in the above example is unusual. Usually we can't do this and can only establish when the definition of Eq. (2.44) fails as in the next example.

Example 2.10. *Consider*

$$f(z) = \frac{z^*}{z} - \frac{z}{z^*}$$

and let us see if $f(z)$ is continuous at $z_0 = 0$. [Note that here the definition of Eq. (2.44) *yields the indeterminate form $(\delta^*/\delta - \delta/\delta^*)$.] Using $z = x + iy$ we have*

$$f(z) = -\frac{4xyi}{x^2+y^2}\,.$$

Now consider approaching z_0 in two different ways. First approach along the line $y = x$ which gives

$$f(z)|_{y=x} = -\frac{4x^2 i}{2x^2} = -2i$$

and then along the the line $y = -x$ giving

$$f(z)|_{y=-x} = +\frac{4x^2 i}{2x^2} = 2i\,.$$

As the function is just constant along these lines, we have

$$\lim_{z\to 0} f(z)|_{y=x} = -2i \qquad \lim_{z\to 0} f(z)|_{y=-x} = 2i\,,$$

and as these are diffrent we have shown that f is not continuous at $z = 0$.

2.8.2 Differentiation

Just as for continuity, we can transfer the concept of differentiation from real to complex functions. Thus, if the limit

$$\lim_{\delta z\to 0} \frac{f(z_0+\delta z) - f(z_0)}{\delta z}\ ,\ \ \delta z \in \mathbb{C} \tag{2.45}$$

exists, we say that f is differentiable in z_0 and we define

$$\lim_{\delta z \to 0} \frac{f(z_0+\delta z)-f(z_0)}{\delta z} = \frac{\mathrm{d}f}{\mathrm{d}z} = f'(z_0)\,, \delta z \in \mathbb{C}\,. \tag{2.46}$$

Note that we divide by the complex value δz here and not by $|\delta z|$ as might naively be expected.

We say that a complex function $f(z)$ is **analytic** or **holomorphic** inside a region of the complex plane if the function has a derivative at every point inside the region. The study of holomorphic functions is what is called **complex analysis**.

Simply using the definition above, it is easy to show that all the usual rules of differentiation apply. If f and g are complex functions and $a \in \mathbb{C}$, then

$$(af)' = af' \tag{2.47}$$

$$(f+g)' = f'+g' \tag{2.48}$$

$$(fg)' = fg'+f'g \tag{2.49}$$

$$\left(\frac{f}{g}\right)' = \frac{gf'-fg'}{g^2} \tag{2.50}$$

$$(f(g))' = \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\mathrm{d}g}{\mathrm{d}z} \tag{2.51}$$

Just as for continuity, it is very hard to use the definition in Eq. (2.46) to determine if a function is differentiable in a given point; we have to show that the limit exists and is unique no matter how we approach z_0. However, as we will see below there is help at hand in terms of the Cauchy-Riemann equations.

Let us view f as two real functions mapping from $\mathbb{R}^2$ to $\mathbb{R}$,

$$f(z) = u(x,y)+iv(x,y) \quad , \quad z = x+iy \qquad x,y,u,v \in \mathbb{R}\,. \tag{2.52}$$

This is simply splitting up f into its real and imaginary parts.

Assume that f is differentiable in z_0. Approach z_0 going along the real axis, thus $x = x_0+\delta x$ with $\delta x \in \mathbb{R}$, $y = y_0$ and $\delta z = \delta x$. As f is assumed differentiable in z_0 we have

$$\begin{aligned}
&\lim_{\delta z \to 0} \frac{f(z_0+\delta z)-f(z_0)}{\delta z}\bigg|_{y=y_0} \\
&= \lim_{\delta x \to 0} \frac{u(x_0+\delta x,y_0)+iv(x_0+\delta x,y_0)-(u(x_0,y_0)+iv(x_0,y_0))}{\delta x} \\
&= \frac{\partial u}{\partial x}(x_0,y_0)+i\frac{\partial v}{\partial x}(x_0,y_0)\,.
\end{aligned} \tag{2.53}$$

Now approach z_0 going along the imaginary axis instead, giving $x = x_0$, $y = y_0 + \delta y$ with $\delta y \in \mathbb{R}$ and $\delta z = +i\delta y$. We then have

$$\begin{aligned}
&\lim_{\delta z \to 0} \frac{f(z_0+\delta z) - f(z_0)}{\delta z}\bigg|_{x=x_0} \\
&= \lim_{\delta y \to 0} \frac{u(x_0, y_0+\delta y) + iv(x_0, y_0+\delta y) - (u(x_0,y_0) + iv(x_0,y_0))}{+i\delta y} \\
&= \frac{1}{i}\left(\frac{\partial u}{\partial y}(x_0,y_0) + i\frac{\partial v}{\partial y}(x_0,y_0)\right) \\
&= \frac{\partial v}{\partial y}(x_0,y_0) - i\frac{\partial u}{\partial y}(x_0,y_0)\,.
\end{aligned} \tag{2.54}$$

Now as f is differentiable, the two preceding results are identical, and from equating the real and imaginary parts we have

$$\frac{\partial u}{\partial x}(x_0,y_0) = \frac{\partial v}{\partial y}(x_0,y_0) \tag{2.55}$$

$$\frac{\partial u}{\partial y}(x_0,y_0) = -\frac{\partial v}{\partial x}(x_0,y_0) \tag{2.56}$$

which are called the **Cauchy-Riemann equations**. Thus if a function is differentiable at z_0, then these equations will hold, i.e. it is a necessary condition of being differentiable. Much more interesting is that it is also a sufficient condition; if the Cauchy-Riemann equations are satisfied at z_0, and the partial derivatives continuous at z_0, then f is differentiable at z_0 with

$$f'(z) = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}\,. \tag{2.57}$$

We shall not prove this statement.

Example 2.11. *Take $f(z) = \cos z$. We have*

$$u(z) = \mathrm{Re}(\cos z) = \cos x \cosh y \tag{2.58}$$

$$v(z) = \mathrm{Im}(\cos z) = -\sin x \sinh y \tag{2.59}$$

which then gives the partial derivatives

$$\frac{\partial u}{\partial x} = -\sin x \cosh y \tag{2.60}$$

$$\frac{\partial u}{\partial y} = \quad \cos x \sinh y \tag{2.61}$$

$$\frac{\partial v}{\partial x} = -\cos x \sinh y \tag{2.62}$$

$$\frac{\partial v}{\partial y} = -\sin x \cosh y . \tag{2.63}$$

It can be seen that the Cauchy-Riemann equations are satisfied, and thus

$$f'(z) = -\sin x \cosh y - i \cos x \sinh y = -\sin z .$$

2.8.3 Interpretation of differentiation

It is well known that the differential $f'(x)$ of a real function mapping from $\mathbb{R}$ to $\mathbb{R}$ can be viewed as the slope of f at the point x. Here we will consider a different interpretation that allows for a direct analogy to complex differentiation.

The Taylor series of f evaluated at x is

$$f(x+\delta x) = f(x) + \delta x f'(x) + \mathscr{O}(\delta x^2) \tag{2.64}$$

or rearranging and ignoring terms of $\mathscr{O}(\delta x^2)$

$$\Delta f = \delta x f'(x) \;, \tag{2.65}$$

with $\Delta f = f(x+\delta x) - f(x)$. It can be seen from this that changing the point where we evaluate the function by δx leads to a change $\delta x f'(x)$ in the function. Or in other words the function at the point x *stretches* the real line by a factor $f'(x)$.

For a complex differentiable function g viewed as a map from $\mathbb{R}^2$ to $\mathbb{R}^2$, the Taylor expansion of g evaluated at z is of exactly the same form as above so

$$g(z+\delta z) = g(z) + \delta z \cdot g'(z) + \mathscr{O}(\delta z^2) \tag{2.66}$$

and

$$\Delta g = \delta z \cdot g'(z) \;, \tag{2.67}$$

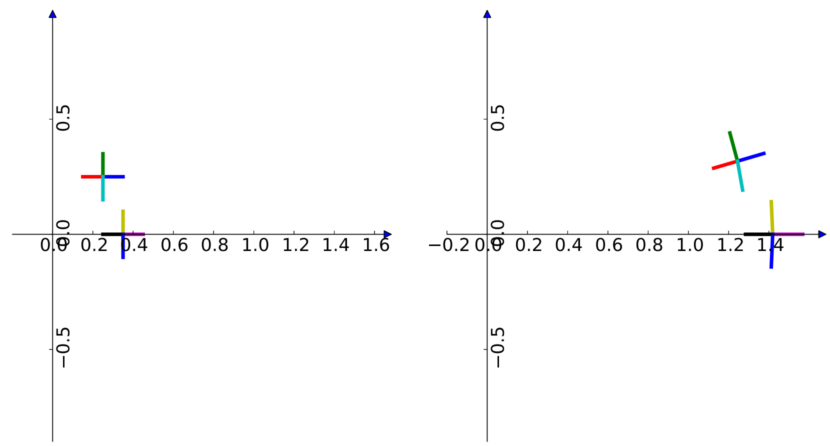

Figure 2.4: An illustration of how $\exp(z)$ when viewed as a mapping from the complex plane to the left to the complex plane to the right stretches and rotates the plane locally around a given point.

with $\Delta g = g(z+\delta z) - g(z)$. It should be remembered here that δz, $g'(z)$ and Δg are all complex numbers. If $g'(z)$ and δz are written in the polar notation,

$$g'(z) = r'\, e^{i\theta'}$$

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

the expression becomes

$$\Delta g = |\delta z|\, e^{i\theta} r'\, e^{i\theta'} = |\delta z| r'\, e^{i(\theta+\theta')}. \tag{2.68}$$

This expression states that if g is evaluated at a point at distance $|\delta z|$ away from z in the direction θ then it is mapped onto a point that is *stretched* by a factor r' and *rotated* by the angle θ' relative to $g(z)$. So, at the point z, the function $g(z)$ is stretching the complex plane in the same way as a real function but also rotating it. This is illustrated for the complex exponential function in Fig. 2.4. In Fig. 2.5 it can be seen how this leads to the idea that whilst the map preserves small shapes it will also stretch and rotate them. The figure also shows how a function that is not complex differentiable, such as complex conjugation, cannot be represented by such a stretch and rotation.

A mapping that preserves angles at any given point in the complex plane is called a **conformal** map. As an example, Google Maps uses the conformal

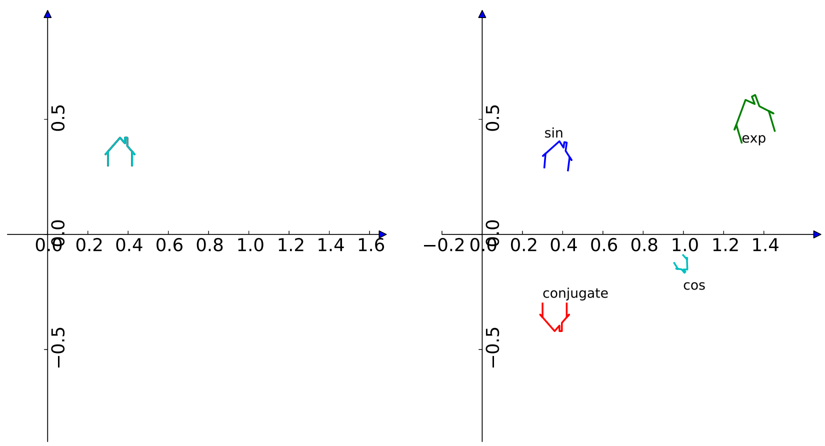

Figure 2.5: A complex differentiable function preserves small shapes when mapping from one complex plane to the other (though the result may be stretched and rotated). This is illustrated here with the complex function exp, cos and sin. For the complex conjugate it can be seen that the function is not complex differentiable as the function cannot be viewed as a stretch and a rotation (notice the location of the chimney).

mapping called a Mercator projection to map longitude and lattitude onto the flat screen. The transformation from one conformal projection to another will be an analytic function as we have seen that angles indeed are preserved in the local rotation.

Chapter 3

First-Order Differential Equations

Many physical phenomena are described in terms of a function whose value at a given point depends on its values at neighbouring points. Thus, the equation determining this function contains derivatives of the function, such as a first derivative to indicate the slope or a velocity, a second derivative to indicate the curvature or an acceleration, and so on. Such an equation, which establishes a relation between the function and its derivatives, is called a **differential equation**.

There are two main types of differential equation. A differential equation for a function of a single independent variable contains only *ordinary* derivatives of that function and is called an **ordinary differential equation**. A differential equation for a function of two or more independent variables contains *partial* derivatives of the function and therefore is called a **partial differential equation**. In this course, we will be concerned with ordinary differential equations.

Ordinary differential equations were introduced to describe the motion of *discrete* particles under the action of known applied forces. The groundwork for such applications was provided by Newton's work on mechanics, particularly the second law of motion, and the development of calculus by Newton and Leibniz. Such differential equations are expressed with time as the independent variable and the coordinates of the particles as the dependent variables. The study of phenomena associated with *continuous* media, such as the motion of fluids and the transmission of sound and other disturbances established the need for partial differential equations. In such cases, the independent variables are the position and time coordinates of points within the medium and the dependent variables are the quantities associated with the medium, such as the velocity of a fluid and its density.

The fundamental equations at the heart of almost all areas of science and engineering are expressed as differential equations. Among the best known of these are Newton's second law of motion in mechanics, Maxwell's equations in electromagnetism, Schrödinger's equation and Dirac's equations in quantum mechanics, the Navier–Stokes equation in fluid mechanics and aerodynamics, Einstein's equations in general relativity, the Fokker–Planck equation in non-equilibrium statistical mechanics, the Hodgkin–Huxley equation in cellular biology, and the Black–Scholes equation in quantitative finance. The widespread use of differential equations is evident in many aspects of modern life, including weather prediction, transportation, communication, and macroeconomic forecasting, to name just a few. In all of these cases, the differential equations embody the characteristics of specific natural or social phenomena, often manifesting unexpected complexity, which are most clearly revealed by examining their solutions in particular cases.

3.1 Notation and Nomenclature

Boas 8.1

An ordinary differential equation for a function y of a single independent variable x is a functional relationship between x, y and the derivatives of y. The **order** of a differential equation is the order of the highest derivative appearing in the differential equation. For example, the most general form of a first-order ordinary differential equation is

$$F(x,y,y') = 0\,,$$

which, for the equations we will study, is written as

$$\frac{dy}{dx} = f(x,y)\,.$$

The general form of an nth-order ordinary differential equation is given by the expression

$$F\left[x,y,y',\ldots,y^{(n)}\right] = 0\,. \tag{3.1}$$

If the function F in these equations is a polynomial in the highest-order derivative of y appearing in its argument list, then the **degree** of the differential equation is the power to which this highest derivative is raised, i.e. the degree of that polynomial. An equation is said to be **linear** if F is of first degree in y and in each of the derivatives appearing as arguments of F. Thus, the general form of a linear nth-order ordinary differential equation is

$$a_n\frac{d^ny}{dx^n} + \cdots + a_1\frac{dy}{dx} + a_0 y = f(x)\,, \tag{3.2}$$

where $f(x)$ and the coefficients $a_1, \dots, a_n$ are known quantities. If $f = 0$ this differential equation is said to be **homogeneous**; otherwise, it is **inhomogeneous**. In the next chapter, we will examine the solutions of second-order equations, both with and without the homogeneous term

$$a_2 \frac{d^2y}{dx^2} + a_1 \frac{dy}{dx} + a_0 y = f(x) .$$

If a set of differential equations share some of their parameters, they are said to form a set of **coupled differential equations**. For a set of n first order coupled differential equations, we can write this as

$$\begin{aligned} F_1\left[x, y_1, y_2, \dots, y_n, y_1', y_2', \dots, y_n'\right] &= 0 \\ F_2\left[x, y_1, y_2, \dots, y_n, y_1', y_2', \dots, y_n'\right] &= 0 \\ &\vdots \\ F_n\left[x, y_1, y_2, \dots, y_n, y_1', y_2', \dots, y_n'\right] &= 0 . \end{aligned}$$

If, for an nth order differential equation, we define variables $y_0 = y$, $y_1 = y'$, $y_2 = y''$, $\dots$, $y_n = y^{''\dots'} \equiv y^{(n)}$, then the nth order differential equation (3.1) can be rewritten as an algebraic relation between the variables, and a set of n coupled first order differential equations

$$\begin{aligned} F\left[x, y_0, y_1, y_2, \dots, y_n\right] &= 0 \\ y_1 &= y' \\ y_2 &= y'' = y_1' \\ &\vdots \\ y_n &= y^{''\dots'} = y^{(n)} = y_{n-1}' , \end{aligned}$$

and that solving an nth order differential equation is equivalent to solving n coupled first order differential equations, subject to the algebraic condition $F = 0$.

3.2 Methods for Solving 1st Order ODE's

Boas 8.2, 8.3

We begin our discussion with first-order equations. Our first example is based on the phenomenon of radioactive decay. We denote by $Q(t)$ the amount of material present at time t. This material decays at a rate r proportional to the amount of material present. The differential equation that describes this process is

$$\frac{dQ}{dt} = -rQ\,, \tag{3.3}$$

where the minus sign indicates that the amount of material decreases with time. We will solve this equation by using two standard methods.

3.2.1 Method 1: Trial solution

We attempt to solve this equation with a solution of the form $Q(t) = e^{\alpha t}$, where α is a constant that is to be determined. The method of trial solution with exponential functions is based on the property

$$\frac{d^n e^{\alpha x}}{dx^n} = \alpha^n e^{\alpha x}\,,$$

for any positive integer n. Substituting our trial solution into Eq. (3.3), we find

$$\frac{dQ}{dt} = \alpha e^{\alpha t} = -re^{\alpha t}\,,$$

or,

$$(\alpha + r)e^{\alpha t} = 0\,.$$

Since $e^{\alpha t} \neq 0$ we can satisfy this equation and obtain a solution if we set $\alpha = -r$. The most general solution we can write for Eq. (3.3) is, therefore,

$$Q(t) = A\,e^{-rt}\,,$$

where A is *any* constant. We can determine A by appealing to the physical situation described by our differential equation. If we set $t = 0$, then $Q(0)$ corresponds to the amount of material initially present, which we denote by Q_0. Accordingly, $Q(0) = A = Q_0$. Thus, the solution of Eq. (3.3) for the amount of material at time t is

$$\boxed{Q(t) = Q_0 e^{-rt}\,.} \tag{3.4}$$

This shows that a unique solution is obtained not just by solving the differential equation, but by also imposing **initial conditions** that are appropriate for circumstances of the physical problem at hand. The solution (3.4) is plotted in Fig. 3.1. The characteristic exponential decay is clearly evident. With increasing r, the rate of decay is considerably faster because this factor appears in the argument of an exponential function.

3.2.2 Method 2: Separation of Variables

The separation of variables is straightforward to set up and carry out, but some aspects of this implementation require justification. Suppose that we have Q at some time t. We can estimate Q at some later time $t+\Delta t$ by performing a Taylor series to first order in Δt

$$Q(t+\Delta t) \approx Q(t) + \frac{dQ}{dt}\Delta t\,. \tag{3.5}$$

We can now use the equation (3.3) governing radioactive decay to substitute for dQ/dt whereupon, after a simple rearrangement, we obtain

$$\frac{Q(t+\Delta t) - Q(t)}{Q(t)} \approx -r\Delta t\,. \tag{3.6}$$

This equation is valid at any time t and becomes more accurate as Δt becomes smaller. Suppose that we require the solution Q to (3.3) from an initial time 0 to some later time t. We divide this time interval into N equal segments $\Delta t_N = t/N$, so that the nth time increment $t_n = n\Delta t_N$, with $t_0 = 0$ and $t_N = t$. For each n, Eq. (3.6) can be written as

$$\frac{Q(t_{n+1}) - Q(t_n)}{Q(t_n)} \approx -r\Delta t\,.$$

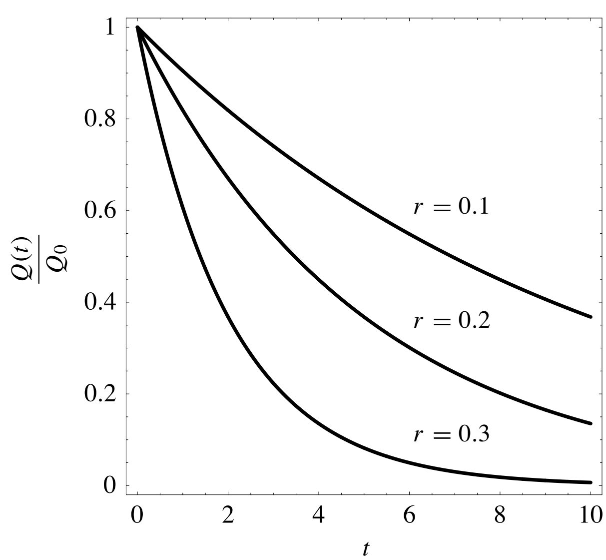

Figure 3.1: The solution in Eq. (3.4) plotted as $Q(t)/Q_0$ against t for three values of the rate constant r. With increasing r, the amount of material at time t decreases substantially.

We now sum this equation over n,

$$\sum_{n=0}^{N-1} \frac{Q(t_{n+1}) - Q(t_n)}{Q(t_n)} \approx -r \sum_{n=0}^{N-1} \Delta t \,,$$

where we have taken the sum only to $N-1$ because Q is evaluated at t_{n+1} which, at this upper limit, corresponds to $t_N = t$. Both sides of this equation are discrete approximations to integrals (called Riemann sums). By decreasing Δt toward zero (i.e. making N larger) these sums provide correspondingly better approximations to the integrals and, in the limit that $N \to \infty$, we have

$$\int_{Q(0)}^{Q(t)} \frac{dQ'}{Q'} = -r \int_0^t dt' \,. \tag{3.7}$$

The primes on the integration variables have been introduced to avoid confusion with the limits of integration. Because the dependent variable (Q) appears only on the left-hand side of the equation, and the independent variable (t) appears only on the right-hand side, these variables are said to have been *separated* and the resulting equation can be integrated directly. Thus, with $Q(0) = Q_0$, we can integrate (3.7) to obtain

$$\ln Q' \Big|_{Q_0}^{Q(t)} = \ln \left[\frac{Q(t)}{Q_0} \right] = -rt \,,$$

or, after exponentiating and solving for $Q(t)$,

$$Q(t) = Q_0 e^{-rt} \,,$$

which is the same as the solution in (3.4).

Although our development of the separation of variables method looks somewhat cumbersome, there is a short-cut that considerably simplifies the procedure. We begin by rearranging the equation (3.3) for radioactive decay as

$$\frac{dQ}{Q} = -r\, dt \,. \tag{3.8}$$

Equation (3.6) is the discrete analogue of this equation, which has been obtained by interpreting the derivative dQ/dt as a fraction, rather than as an operation on a function. This can be justified only as an intermediate step toward integrating this equation from $t = 0$, where $Q = Q(0) = Q_0$, to some later time t, where $Q = Q(t)$:

$$\int_{Q(0)}^{Q(t)} \frac{dQ'}{Q'} = -r \int_0^t dt' \,.$$

which is (3.7). The virtue of this somewhat loose interpretation of the mathematical formulation of the separation of variables method is that it is easy to apply and one can usually identify differential equations that are separable directly by inspection.

In summary, the advantage of the trial solution method is that it can be applied to higher-order equations, as we will show in the next chapter for second-order equations, but only to *linear* equations. The separation of variables method can be applied to certain types of nonlinear equations, as we will show in the next section, but only to *first-order* equations.

3.2.3 Method 3: Using an integrating factor

If a first order ODE has the form

$$\frac{dy}{dx} + P(x)y = Q(x) \tag{3.9}$$

it can be solved through by what is called an **integrating factor**. This is just a special implementation of separation of variables that applies to many problems. Consider first the homogeneous equation

$$\frac{dy}{dx} + P(x)y = 0 \tag{3.10}$$

which we can solve as

$$\frac{dy}{y} = -P(x)dx \tag{3.11}$$

$$\ln y = -\int P(x)dx + C \tag{3.12}$$

or

$$\begin{aligned} y &= \exp\left(-\int P(x)dx + C\right) \\ &= A\exp\left(-\int P(x)dx\right) \\ &= Ae^{-I}, \end{aligned} \tag{3.13}$$

where I is the **integrating factor**

$$I = \int P(x)dx\,. \tag{3.14}$$

To solve the inhomogeneous ODE in Eq. 3.9 we calculate

$$\begin{aligned}\frac{d}{dx}\left(ye^I\right) &= \frac{dy}{dx}e^I + ye^I\frac{dI}{dx} \\ &= \frac{dy}{dx}e^I + ye^I P(x) \\ &= e^I Q(x)\,.\end{aligned} \tag{3.15}$$

Integrating both sides and rearranging gives

$$\boxed{y = e^{-I}\int e^I Q(x)dx + Ce^{-I}\,.} \tag{3.16}$$

Example 3.1. *Solve*

$$x^2y' + 3xy = 1\,. \tag{3.17}$$

Rearranging to the form of Eq. 3.9 gives

$$P(x) = \frac{3}{x} \qquad Q(x) = \frac{1}{x^2} \tag{3.18}$$

and

$$I = \int \frac{3}{x}dx = \ln x^3 \tag{3.19}$$

and thus

$$e^I = x^3 \qquad e^{-I} = x^{-3}\,. \tag{3.20}$$

Inserting this into Eq. 3.16 gives

$$\begin{aligned}y &= x^{-3}\int x^3\frac{1}{x^2}dx + Cx^{-3} \\ &= \frac{1}{2x} + \frac{C}{x^3}\,.\end{aligned} \tag{3.21}$$

3.3 Numerical solutions

To solve a differential equation analytically has large advantages as it allows for calculations to arbitrary accuracy, is fast to use and is better suited for understanding the specific behaviour of the system. However, the reality is that the collection of differential equations that can be solved analytically is very small. In

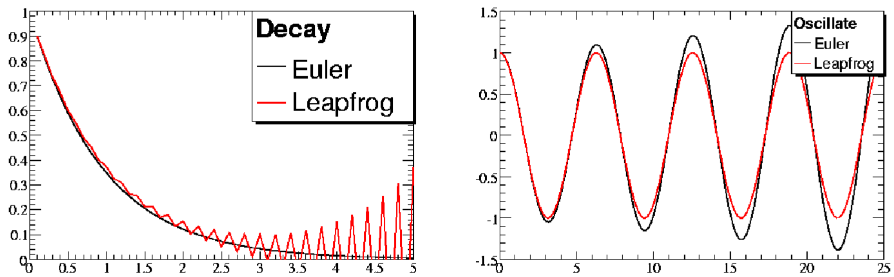

Figure 3.2: An illustration of two numerical methods used to solve a differential equation representing radioactive decay (left) and a swinging pendulum (right). The Euler method is a first order method while the leapfrog is a second order method. Both of these methods are not suitable for solving general problems.

modern science, nearly all differential equations are solved numerically. A good example is weather forecasting. The European Centre for Medium-Range Weather Forecasting (ECMWF) at Reading is where one of the largest supercomputers in Europe is situated. Every six hours it solves numerically the partial differential equations that describe the weather of the entire planet for the coming two weeks. The initial conditions are the latest observations of temperature, pressure, wind, cloud cover etc. as measured by Earth observation satellites and ground based weather stations. Another example is lattice QCD calculations where the nature of bound states of quarks is studied by running numerical models of how quantum chromodynamics behaves.

The principle of numerically solving a differential equation was already outlined in Sec. 3.2.2, where in Eq. (3.5) we saw how we could use the first element of the Taylor expansion to get an estimate for a *step* forward in time. This method of solving is called the *Euler* method and can be written as

$$y_{n+1} = y_n + hf(x_n, y_n) \,, \tag{3.22}$$

where n refers to the nth time step in the evaluation. The method is easy to implement but it takes a very small step size h to get good accuracy and completely fails for oscillating methods as illustrated in Fig. 3.2. Another well known method is the *leapfrog* method which is also illustrated in Fig. 3.2.

Any collection of numerical computational methods, such as the GNU Scientific library, contains methods for solving a set of coupled first order differential equations (and therefore equations of any order). The methods are in general, as the Euler method, based on the Taylor expansion but keeps higher order terms in the expansion. A popular choice is the Runge-Kutta-Feldberg method which combines a 5th and 6th order expansion and makes use of a step size that adapts to the requirements given.

3.4 Spread of Epidemics

An example of the use of first-order differential equations is to the spread of epidemics, first used by Daniel Bernoulli in 1760 to model the spread of smallpox. We will construct a simple model of an epidemic and then solve the resulting differential equation. This is essentially the way that models in epidemiology are constructed: assumptions and known characteristics are used to build a model, which is then solved and various scenarios are tested (e.g. inoculation or isolation) to develop strategies on how to respond to an epidemic.

Consider a model for the spread of a disease in which a population that is divided into two groups: a fraction x that has no disease, but is susceptible to the disease, and a fraction y that has the disease and can infect others. We suppose that everyone belongs to only one of these groups, so $x+y=1$. We now make three assumptions about how the disease is spread:

1. The disease spreads only by direct contact between infected and uninfected individuals. Direct can be taken to mean ‘close proximity,’ as in crowds, where colds and influenza viruses can easily spread.

2. The fraction of infected individuals increases at a rate ρ proportional to such contacts.

3. Both groups move freely among one another, so the number of direct contacts is xy. This is another way of saying that the x and y populations are uncorrelated.

The differential equation that embodies these assumptions is

$$\frac{dy}{dt} = \rho xy = \rho(1-y)y\,, \tag{3.23}$$

where ρ is a constant that specifies the ‘efficiency’ of the spreading at the point of contact, i.e. the likelihood of disease transmission once direct contact has occurred, and we have used the fact that $x+y=1$ to eliminate x in favour of y. In accordance with our experience in finding the solution for radioactive decay, we must supplement this equation by specifying the fraction of infected individuals initially: $y(0)=y_0$.

Equation (3.23) is a first-order *nonlinear* differential equation. Thus, we cannot use the trial solution method as formulated above. However, this equation can be arranged as

$$\frac{dy}{y(1-y)} = \rho\, dt\,,$$

so we can use the separation of variables method. Integrating both sides of the equation with respect to the indicated variables from $t = 0$, where $y = y(0) = y_0$ to a later time t, where $y = y(t)$, we obtain

$$\int_{y_0}^{y(t)} \frac{dy'}{y'(1-y')} = \rho \int_0^t dt' = \rho\, t\,.$$

The left-hand side of this equation can be integrated by the method of partial fractions. We first write

$$\frac{1}{y(1-y)} = \frac{A}{y} + \frac{B}{1-y}\,,$$

which implies that

$$A(1-y) + By = 1\,.$$

Choosing $y = 0$ yields $A = 1$, and choosing $y = 1$ yields $B = 1$, so

$$\frac{1}{y(1-y)} = \frac{1}{y} + \frac{1}{1-y}\,,$$

with which we obtain

$$\begin{aligned} \rho\, t &= \int_{y_0}^{y(t)} \frac{dy'}{y'} + \int_{y_0}^{y(t)} \frac{dy'}{1-y'} \\ &= \ln y' \Big|_{y_0}^{y(t)} - \ln(1-y') \Big|_{y_0}^{y(t)} \\ &= \ln\left[\frac{y(t)}{1-y(t)} \frac{1-y_0}{y_0}\right]. \end{aligned}$$

Solving for $y(t)$ yields,

$$\boxed{y(t) = \frac{y_0\, e^{\rho t}}{1 - y_0(1 - e^{\rho t})}\,.} \tag{3.24}$$

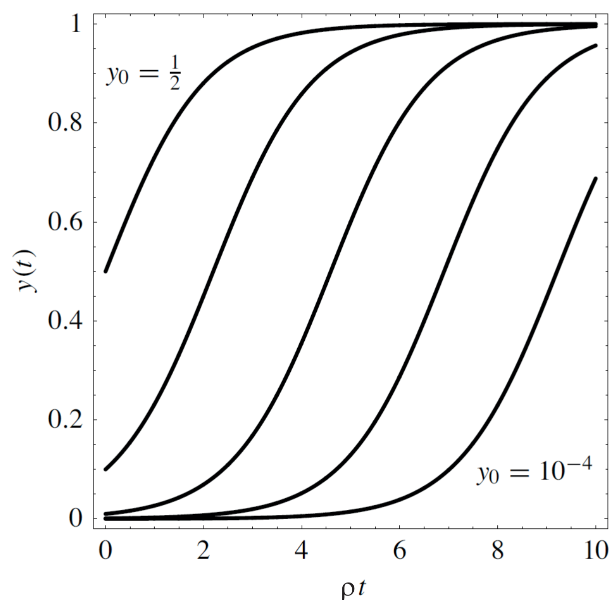

Figure 3.3: The solution in Eq. (3.24) shown as a function of ρt for values of y_0 in the range $10^{-4} \leq y_0 \leq \frac{1}{2}$. As y_0 decreases toward zero, the solution remains near $y = 0$ for longer times, while as y_0 increases toward unity, the solution approaches $y = 1$ for shorter times.

As $t \to \infty$, $y(t) \to 1$, provided that $y_0 \neq 0$ (Fig. 3.3). In other words, all of the population eventually becomes infected unless there is no infection initially. As long as $y_0 \neq 0$, no matter how small, the entire population becomes infected. Accordingly, the point $y = 1$ is said to be **stable** and the point $y = 0$ is said to be **unstable**. It can be seen from the figure that the time when the epidemic has spread to half the population depends strongly on y_0. But arbitrarily small y_0 makes no sense as we are dealing with an integer number of people.

A numerical simulation of the epidemic using an *agent based* or *individual based* approach is the only way that fluctuations due to integer numbers at the beginning of the outbreak can be correctly accounted for.

Chapter 4

Second-Order Ordinary Differential Equations

Among the simplest higher-order ordinary differential equations are linear homogeneous equations with constant coefficients. The second-order versions of these equations occur in many applications in science and engineering. Among the most prevalent of these involve Newton's second law of motion in mechanics and the flow of charge in electrical circuits. Such equations also occur as special cases of certain partial differential equations, for example, in quantum mechanics and problems of heat conduction. The solutions of second-order differential equations with constant coefficients will be shown to have the important feature of being expressible in terms of exponential functions, and their method of solution has evident extensions to higher-order equations with constant coefficients.

In this chapter, we illustrate the solution of equations with constant coefficients by focussing on second-order equations. The general form of a second-order linear homogeneous ordinary differential equation with constant coefficients is

$$a\,\frac{d^2y}{dx^2}+b\,\frac{dy}{dx}+c\,y=0\,, \tag{4.1}$$

where a, b and c are known constants that are usually real. The most important property of this equation is *linearity*. This means that if we have two solutions y_1 and y_2, then any linear combination of y_1 and y_2 is also a solution of this equation, i.e.

$$y(x)=Ay_1(x)+By_2(x)$$

is a solution for *any* choice of constants A and B. This is a direct consequence of the linearity of derivatives, which allows us to write

$$\begin{aligned}
& a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy \\
& = a\frac{d^2}{dx^2}(Ay_1+By_2)+b\frac{d}{dx}(Ay_1+By_2)+c(Ay_1+By_2) \\
& = A\left(a\frac{d^2y_1}{dx^2}+b\frac{dy_1}{dx}+cy_1\right)+B\left(a\frac{d^2y_2}{dx^2}+b\frac{dy_2}{dx}+cy_2\right) \\
& = 0\,.
\end{aligned} \tag{4.2}$$

The last equality follows from the fact that y_1 and y_2 are solutions of (4.1), so the coefficients of A and B are equal to zero. The same steps can be used to verify this statement for the more general case of a homogeneous linear differential equation with variable coefficients. This is the **superposition principle** for homogeneous linear differential equations. It lies at the heart of both the theory of these equations and the methodologies that have been developed for solving them.

4.1 The Characteristic Equation

Boas 8.5

The solution of (4.1) will be obtained by the method of trial solution. The recursive property of derivatives of the exponential function,

$$\frac{d}{dx}(e^{\alpha x})=\alpha\, e^{\alpha x}\,, \qquad \frac{d^2}{dx^2}(e^{\alpha x})=\alpha^2 e^{\alpha x}\,, \tag{4.3}$$

suggests that the trial solution method used for solving first-order equations in Sec. 3.2.1 can be applied to higher-order equations with constant coefficients. Suppose we try this for the differential equation (4.1). We substitute our trial solution $e^{\alpha x}$ into this equation and choose α by requiring the resulting expression to equal zero, i.e. that this function solves the equation. Substituting the derivatives in (4.3) into (4.1) yields

$$a\,\frac{d^2}{dx^2}(e^{\alpha x})+b\,\frac{d}{dx}(e^{\alpha x})+c\,(e^{\alpha x})=(a\alpha^2+b\alpha+c)\,e^{\alpha x}\,. \tag{4.4}$$

For the function $e^{\alpha x}$ to be a solution of (4.1), the coefficient of $e^{\alpha x}$ on the right-hand side of this equation must vanish (since the exponential is nonzero for finite

x). Thus, α must be chosen to be a root of the quadratic equation

$$\boxed{a\alpha^2 + b\alpha + c = 0\,.} \tag{4.5}$$

This is the **characteristic equation** of the differential equation (4.1) and the left-hand side of this equation is called the **characteristic polynomial**. The roots of the characteristic equation, which are given by the quadratic formula,

$$\boxed{\alpha = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}} \tag{4.6}$$

yield solutions of (4.1). This is the power of the trial solution method – the solution of a differential equation has been reduced to finding the roots of a quadratic equation.

By their appearance in the discriminant in this equation, the coefficients a, b and c are seen to be the central quantities for determining the number and type of roots of the characteristic polynomial and, through these roots, the behavior of the exponential solutions. In direct analogy to the discussion in Sec. 1.1, there are three cases to consider.

4.1.1 Case I: Real Distinct Roots

If $b^2 - 4ac > 0$, there are two distinct real roots of the characteristic equation, which we denote by α_1 and α_2. Two distinct solutions of (4.1) result:

$$y_1(x) = e^{\alpha_1 x} \qquad \text{and} \qquad y_2(x) = e^{\alpha_2 x}\,. \tag{4.7}$$

According to the procedure in (4.2), we can use these solutions to form the general solution of (4.1) by forming the linear combination

$$\boxed{y(x) = Ae^{\alpha_1 x} + Be^{\alpha_2 x}\,,} \tag{4.8}$$

where A and B are arbitrary constants. This is called the **general solution** of the differential equation. The constants A and B are determined by specifying initial conditions. Because there are two arbitrary constants in the solution, two initial conditions are needed to obtain a unique solution. These are often taken as the function y and its derivative evaluated at the origin,

$$y(0) = y_0\,, \qquad \left.\frac{dy}{dx}\right|_{x=0} \equiv y'(0) = y'_0\,, \tag{4.9}$$

although they can be specified at any point. It is also possible to specify an initial and a final state condition for y,

$$y(0) = y_0\,, \qquad y(x_{\text{end}}) = y_{\text{end}}\,. \tag{4.10}$$

Depending on the signs of α_1 and α_2, the solutions exhibit either exponential growth or exponential decay as functions of x as illustrated in Fig. 4.1.

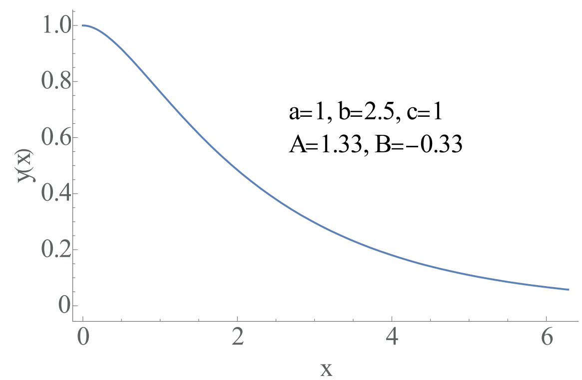

Figure 4.1: Plot of $y(x)$ for Case I: Real Distinct Roots. In this case $\alpha_1 = -0.5$ and $\alpha_2 = -2.0$.

4.1.2 Case II: Degenerate Roots

If $b^2 - 4ac = 0$, there is only a single real root, $\alpha_1 = -b/(2a)$, of the characteristic equation. Thus, this method produces only one solution of (4.1),

$$y_1(x) = e^{\alpha_1 x}\,. \tag{4.11}$$

We seem to have arrived at an impasse. The case of two real roots of the characteristic equation provided two distinct solutions with which we can obtain a unique solution from the general solution for a particular set of two initial conditions. A similar situation will arise in the case where the characteristic equation yields complex conjugate roots. However, if the discriminant vanishes, we appear to have only a single solution, which cannot be reconciled with two initial conditions. Thus, the method of trial solutions has failed to provide two solutions. But we can extend this method by making a few simple observations that will yield a second solution in a form that enables us to deal with the case of a vanishing discriminant in an analogous manner to the other two cases.

We begin by returning to (4.4), the left-hand side of which is

$$a\frac{d^2(e^{\alpha x})}{dx^2}+b\frac{d(e^{\alpha x})}{dx}+c(e^{\alpha x})\,. \tag{4.12}$$

This equation equals zero only if we set $\alpha=\alpha_1$, which shows that $e^{\alpha_1 x}$ is a solution

$$\left[a\frac{d^2(e^{\alpha x})}{dx^2}+b\frac{d(e^{\alpha x})}{dx}+c(e^{\alpha x})\right]\bigg|_{\alpha=\alpha_1}$$

$$=a\frac{d^2(e^{\alpha_1 x})}{dx^2}+b\frac{d(e^{\alpha_1 x})}{dx}+c(e^{\alpha_1 x})=0$$

Since α is a continuous variable, we can differentiate (4.12) with respect to α *before* setting α equal to α_1.

$$\left\{\frac{d}{d\alpha}\left[a\frac{d^2(e^{\alpha x})}{dx^2}+b\frac{d(e^{\alpha x})}{dx}+c(e^{\alpha x})\right]\right\}\bigg|_{\alpha=\alpha_1}$$

The order of the derivatives with respect to x and to α is immaterial, so we can take the α derivatives before the x derivatives, in which case we obtain,

$$a\,\frac{d^2}{dx^2}\left[\frac{d(e^{\alpha x})}{d\alpha}\right]+b\,\frac{d}{dx}\left[\frac{d(e^{\alpha x})}{d\alpha}\right]+c\left[\frac{d(e^{\alpha x})}{d\alpha}\right]$$

$$=a\,\frac{d^2}{dx^2}(xe^{\alpha x})+b\,\frac{d}{dx}(xe^{\alpha x})+c\,(xe^{\alpha x})\,.$$

The remaining derivatives are straightforward to calculate:

$$\frac{d}{dx}(xe^{\alpha x}) = e^{\alpha x}+\alpha xe^{\alpha x}\,,$$

$$\frac{d^2}{dx^2}(xe^{\alpha x}) = 2\alpha e^{\alpha x}+\alpha^2 xe^{\alpha x}\,,$$

whereupon we obtain

$$a\,\frac{d^2}{dx^2}\left[\frac{d(e^{\alpha x})}{d\alpha}\right]+b\,\frac{d}{dx}\left[\frac{d(e^{\alpha x})}{d\alpha}\right]+c\left[\frac{d(e^{\alpha x})}{d\alpha}\right]$$

$$=(2a\alpha+b)\,e^{\alpha x}+(a\alpha^2+b\alpha+c)xe^{\alpha x}\,. \tag{4.13}$$

By setting $\alpha = \alpha_1$ and using the fact that $\alpha_1 = -b/(2a)$, we find

$$2a\alpha_1 + b = -b + b = 0\,,$$
$$a\alpha_1^2 + b\alpha_1 + c = \frac{b^2}{4a} - \frac{b^2}{2a} + c = -\frac{b^2 - 4ac}{4a} = 0\,.$$

so the coefficient of each term on the right-hand side of (4.13) vanishes, leaving

$$a\,\frac{d^2}{dx^2}(x\,e^{\alpha_1 x}) + b\,\frac{d}{dx}(x\,e^{\alpha_1 x}) + c\,(x\,e^{\alpha_1 x}) = 0\,,$$

which shows that our second solution is, in this case,

$$y_2(x) = x\,e^{\alpha_1 x}\,.$$

The general solution is

$$\boxed{y(x) = (A + Bx)e^{\alpha_1 x}\,.} \tag{4.14}$$

Similar to those in (4.7), the solutions y_1 and y_2 exhibit either exponential growth or decay, depending on the sign of α_1. An example is shown in Fig. 4.2.

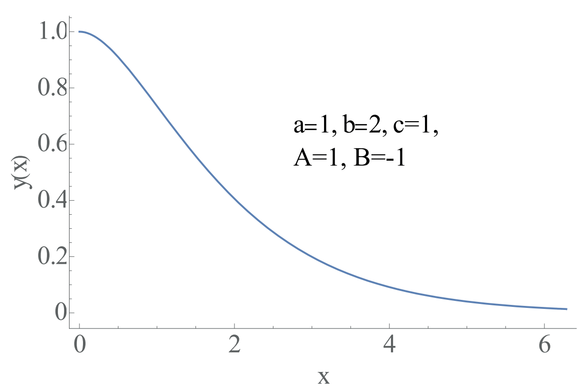

Figure 4.2: Plot of $y(x)$ for Case II: Degenerate Roots. In this case $\alpha_1 = \alpha_2 = -1.0$.

4.1.3 Case III: Complex Conjugate Roots

If $b^2 - 4ac < 0$, there are two complex roots, α_1 and α_2, which are complex conjugates: $\alpha_2 = \alpha_1^*$. The two solutions of (4.1) are thus given by

$$y_1(x) = e^{\alpha_1 x}, \qquad y_2(x) = e^{\alpha_1^* x}, \tag{4.15}$$

so the general solution is

$$\boxed{y(x) = A\, e^{\alpha_1 x} + B\, e^{\alpha_1^* x}.} \tag{4.16}$$

Since α_1 is a complex number, y_1 and y_2 are complex-valued functions. However, we can express the solutions to (4.1) solely in terms of real functions by utilizing Euler's formula (1.29). With α_1 and α_2 expressed in terms of their real and imaginary parts as

$$\alpha_1 = \mu + i\kappa, \qquad \alpha_1^* = \mu - i\kappa,$$

where μ and κ are real, we first write the solutions in (4.15) as

$$\begin{aligned} y_1(x) &= e^{(\mu+i\kappa)x} = e^{\mu x}(\cos\kappa x + i\sin\kappa x)\,, \\ y_2(x) &= e^{(\mu-i\kappa)x} = e^{\mu x}(\cos\kappa x - i\sin\kappa x)\,. \end{aligned}$$

Thus, linear combinations of y_1 and y_2 can be written as

$$\begin{aligned} y(x) &= Ay_1(x) + By_2(x) \\ &= A\Big[e^{\mu x}(\cos\kappa x + i\sin\kappa x)\Big] + B\Big[e^{\mu x}(\cos\kappa x - i\sin\kappa x)\Big] \\ &= (A+B)\,e^{\mu x}\cos\kappa x + i(A-B)\,e^{\mu x}\sin\kappa x\,. \end{aligned}$$

Since A and B are arbitrary quantities, then $A+B$ and $i(A-B)$ are as well, so we can write the general solution in an alternative form based on the real solutions

$$\tilde{y}_1(x) = e^{\mu x}\cos\kappa x \qquad \tilde{y}_2(x) = e^{\mu x}\sin\kappa x\,, \tag{4.17}$$

as

$$\boxed{y(x) = e^{\mu x}\,(C\cos\kappa x + D\sin\kappa x)\,,} \tag{4.18}$$

in which C and D are arbitrary constants whose values are obtained from the initial conditions (4.9). A useful equivalent form of Eq. (4.18) is given by

$$\boxed{y(x) = A_0 e^{\mu x} \cos(\kappa x + \phi)\,,} \tag{4.19}$$

where A_0 and ϕ are yet another pair of constants. As an exercise, see if you can show that $A_0 = (C^2 + D^2)^{1/2}$ and $\tan\phi = -D/C$. These solutions show that, as illustrated in Fig. 4.3, the imaginary parts of α_1 and α_2 produce oscillatory behavior, and their real parts, if nonzero, modulate this with either exponential growth or decay. The choice of whether to use the complex solutions of 4.15) or their real counterparts in (4.17) is largely a matter of taste and convenience. In the next section we will show how the three types of solution to the characteristic equation arise in a physical setting.

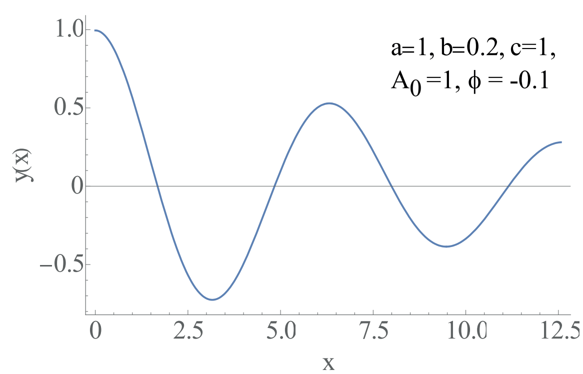

Figure 4.3: Plot of $y(x)$ for Case II: Complex Conjugate Roots. In this case $\alpha_1 = \alpha_2^* = -0.1 + i$.

4.2 The Damped Harmonic Oscillator

Boas 8.5

Consider the harmonic oscillator in Fig. 4.4, which consists of a mass m attached to a spring with stiffness k and damping b. Once displaced from equilibrium, there are two forces acting on the mass: the restoring spring force and friction. Newton's second law of motion for the position x of the oscillator is thus given by

$$m\ddot{x} = -kx - b\dot{x}\,, \tag{4.20}$$

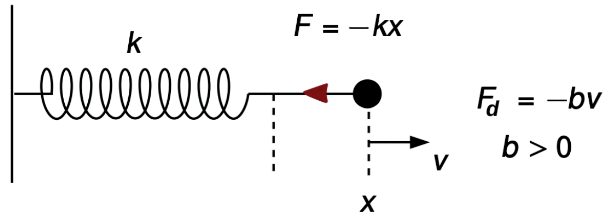

Figure 4.4: A mass m attached to a harmonic spring with stiffness k and damping b.

where a dot has been used as a shorthand for time derivative. Dividing through by m yields

$$\boxed{\ddot{x} + \gamma\dot{x} + \omega_0^2 x = 0\,,} \tag{4.21}$$

in which

$$\gamma = \frac{b}{m}\,, \qquad \omega_0^2 = \frac{k}{m}\,. \tag{4.22}$$

and ω_0 is called the *natural frequency* of the oscillator.

To obtain a specific solution for the position of the oscillator, we must supplement this equation with two initial conditions. We take

$$x(t=0) = x_0\,, \qquad \dot{x}(t=0) = 0\,, \tag{4.23}$$

which corresponds to the mass starting from rest at position x_0. Equation (4.21) has the form of Eq. (4.1), with

$$a = 1\,, \qquad b = \gamma, \qquad c = \omega_0^2\,,$$

and where the independent variable is now t and the dependant variable is x. The results of the previous section can immediately be used to determine $x(t)$ for the three cases.

4.2.1 Case I: $\gamma^2 - 4\omega_0^2 > 0$

This case is referred to as **overdamped**. The general solution is given by

$$x(t) = e^{-\gamma t/2}\left[Ae^{\rho t} + Be^{-\rho t}\right],$$

[cf. Eq. (4.8) with $\alpha_1 = -\gamma/2 + \rho$ and $\alpha_2 = -\gamma/2 - \rho$, where $\rho = (\gamma^2/4 - \omega_0^2)^{1/2}$.] The initial conditions of Eq. (4.23) require

$$x(0) = A + B = x_0\,, \qquad \dot{x} = \left(-\frac{\gamma}{2} + \rho\right)A + \left(-\frac{\gamma}{2} - \rho\right)B = 0\,,$$

yielding

$$A = \frac{1}{2}\left(1 + \frac{\gamma}{2\rho}\right)x_0\,, \qquad B = \frac{1}{2}\left(1 - \frac{\gamma}{2\rho}\right)x_0\,. \tag{4.24}$$

The solution for the position of the oscillator is therefore obtained as (exercise!)

$$\boxed{x(t) = x_0 e^{-\gamma t/2}\left(\cosh\rho t + \frac{\gamma}{2\rho}\sinh\rho t\right).} \tag{4.25}$$

Note that in Fig. 4.1, A and B were chosen to satisfy Eq. (4.24).

4.2.2 Case II: $\gamma^2 - 4\omega_0^2 = 0$

In this case the damping and oscillations are balanced. This is called the **critically damped** case and the general solution is given by (cf. Eq. (4.14) with $\alpha_1 = -b/2a = \gamma/2$):

$$x(t) = (A + Bt)e^{-\frac{1}{2}\gamma t}\,.$$

The initial conditions in Eq. (4.23) require,

$$x(0) = A = x_0\,, \qquad \dot{x}(0) = B - \tfrac{1}{2}\gamma A = 0\,,$$

yielding

$$A = x_0\,, \qquad B = \tfrac{1}{2}x_0\gamma. \tag{4.26}$$

The solution for the position of the oscillator is therefore obtained as

$$\boxed{x(t) = x_0\left(1 + \tfrac{1}{2}\gamma t\right)e^{-\frac{1}{2}\gamma t}\,.} \tag{4.27}$$

In Fig. 4.2, A and B were chosen to satisfy Eq. (4.26). Critical damping reflects the quickest way to restore to the position of the initial state. For this reason systems such as door closers and the suspension in a car are critically damped.

4.2.3 Case III: $\gamma^2 - 4\omega_0^2 < 0$

Here the oscillations dominate over the damping, so this is called the **underdamped** case. The general solution can be written as

$$x(t) = e^{-\gamma t/2}\left(Ae^{i\omega_d t} + Be^{-i\omega_d t}\right), \tag{4.28}$$

[cf. Eq. (4.16) with $\alpha_1 = -\gamma/2 + i\omega_d$ and $\alpha_2 = -\gamma/2 - i\omega_d$, where $\omega_d = (\omega_0^2 - \gamma^2/4)^{1/2}$]. The initial conditions in Eq. (4.23) require

$$x(0) = A + B = x_0\,, \qquad \dot{x}(0) = \left(-\frac{\gamma}{2} + i\omega_d\right)A + \left(-\frac{\gamma}{2} - i\omega_d\right)B = 0\,,$$

yielding

$$A = \frac{1}{2}\left(1 + \frac{\gamma}{2i\omega_d}\right), \qquad B = \frac{1}{2}\left(1 - i\frac{\gamma}{2i\omega_d}\right). \tag{4.29}$$

The solution for the position of the oscillator is therefore obtained as (exercise!)

$$\boxed{x(t) = x_0 e^{-\gamma t/2}\left(\cos\omega_d t + \frac{\gamma}{2\omega_d}\sin\omega_d t\right).} \tag{4.30}$$

In Fig. 4.3, A and B were chosen to satisfy Eq. (4.29).

The solution represented in Eq. (4.30) can be written equivalently (exercise!) as (cf. Eq. (4.19))

$$x(t) = x_0\left(\frac{\omega_0}{\omega_d}\right)e^{-\gamma t/2}\cos\left(\omega_d t + \phi\right),$$

where $\tan\phi = -\gamma/2\omega_d$. Note that the frequency of the oscillation is ω_d and not ω_0. Note also that the solutions illustrated in Figs. 4.1, 4.2 and 4.3 all start with unit initial displacement and zero slope. The overdamped case decays to zero monotonically, as does the critically damped solution, although the latter decays faster. The underdamped solution shows several periods of oscillation that have a decaying envelope. For all three cases, the equilibrium position is reached as $t \to \infty$ because of the presence of damping.

4.3 Inhomogeneous Equations

Boas 8.6

We can now try to consider systems which have the form

$$a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)\,.$$

These are called **inhomogeneous** because the function $f(x)$ on the right-hand side of this equation is specified independently of the solution. In physics these equations will pop up in many places where an external force is applied to a system. Resonance phenomena would not exist without it.

Inhomogeneous linear equations are solved by first supposing that there are two independent solutions $y^{(1)}(x)$ and $y^{(2)}(x)$ of this equation:

$$a\frac{d^2y^{(1)}}{dx^2}+b\frac{dy^{(1)}}{dx}+cy^{(1)} \;=\; f(x)\,, \tag{4.31}$$

$$a\frac{d^2y^{(2)}}{dx^2}+b\frac{dy^{(2)}}{dx}+cy^{(2)} \;=\; f(x)\,. \tag{4.32}$$

If we subtract one equation from the other, say Eq. (4.31) from (4.32), we obtain

$$a\frac{d^2[y^{(2)}-y^{(1)}]}{dt^2}+b\frac{d[y^{(2)}-y^{(1)}]}{dt}+c[y^{(2)}-y^{(1)}]=0\,,$$

i.e. the difference $y^{(2)}-y^{(1)}$ is a solution of the *homogeneous* equation! If we denote the general solution of the homogeneous equation by $Ay_1(x)+By_2(x)$, we conclude from the above that

$$y^{(2)}(x)=Ay_1(x)+By_2(x)+y^{(1)}(x)\,.$$

This suggests the following method of solution. Find just one solution $y_p(x)$ of the inhomogeneous equation, called a **particular solution**, by any means. The general solution $y(x)$ of the inhomogeneous equation is then given by

$$\boxed{y(x)=Ay_1(x)+By_2(x)+y_p(x)\,,} \tag{4.33}$$

in which y_1 and y_2 are solutions of the corresponding homogeneous equation.

Example 4.1. *Consider the equation*

$$\frac{d^2y}{dx^2}+\frac{dy}{dx}-6y=5$$

As the derivatives are zero for a constant, it is easy to see that $y_P = -\frac{5}{6}$ *is a particular solution. As the roots of the characteristic polynomial for the corresponding homogeneous equation are 2 and -3, the general solution is*

$$y(x) = Ae^{2x} + Be^{-3x} - \frac{5}{6}\,.$$

Example 4.2. *Now move on to having a polynomial on the right hand side. We can consider*

$$\frac{d^2y}{dx^2} + \frac{dy}{dx} - 6y = x^2 - 1\,.$$

For finding a particular solution, let us consider a polynomial of the form

$$y_P = \mu x^2 + \kappa x + \gamma\,.$$

Differentiating and equating terms in the ODE gives $\mu = -\frac{1}{6}$, $\kappa = -\frac{1}{18}$, *and* $\gamma = \frac{11}{108}$. *As the homogeneous equation is the same as in the previous example, we get the general solution*

$$y(x) = Ae^{2x} + Be^{-3x} - \frac{1}{6}x^2 - \frac{1}{18}x + \frac{11}{108}\,.$$

The most important case is where the function $f(x)$ is a (complex) exponential. Let us again illustrate this by an example.

Example 4.3. *To find the general solution to*

$$\frac{d^2y}{dx^2} + \frac{dy}{dx} - 6y = e^{2ix}\,,$$

we start by making a guess at a particular solutuon to be of exponential form, $y = Ce^{ax}$. *Inserting into the ODE gives*

$$Ce^{ax}(a^2 + a - 6) = e^{2ix}$$

which is satisfied when

$$a = 2i \qquad \text{and} \qquad C = \frac{-10-2i}{104}$$

thus giving the general solution

$$y(x) = Ae^{2x} + Be^{-3x} + \frac{-10-2i}{104}e^{2ix}\,.$$

What to do in the case where the right hand side has more than one term? We can use the principle of superpositon. So to solve

$$\frac{d^2y}{dx^2} + \frac{dy}{dx} - 6y = x^2 - 1 + e^{2ix}$$

we can simply take the particular solutions from the examples above and add them together to form a particualr solution of the combined equation. Then adding all the solutions to the homogeneous equation again gives the general solution.

As will be seen in next year's course on Fourier analysis, any periodic function can be written as an infinite sum of complex exponentials. The superposition princicple will then allow us to write the general solution for a system with any cyclic function applied as an external force.

4.4 Summary

We can both summarize and generalize the main results of this chapter as follows. The solution of *any* nth-order ordinary differential equation depends, in general, on n arbitrary constants $c_1, c_2, \ldots, c_n$:

$$y = \varphi(x; c_1, c_2, \ldots, c_n) \tag{4.34}$$

Thus, to obtain a unique solution for a particular problem, it is necessary to supplement the differential equation with auxiliary conditions. A common choice is for these constants to be determined from the initial values of the solution y and its first $n-1$ derivatives at some initial point x_0:

$$y(x_0) = A_0, \quad y'(x_0) = A_1, \quad \ldots \quad y^{(n)}(x_0) = A_n$$

The expression in (4.34) is a general solution if it is possible to satisfy these initial conditions for arbitrary values of the A_i with an appropriate choice of the c_j. This usually requires the solution of a system of algebraic equations.

For homogeneous linear nth-order equations,

$$a_n \frac{d^n y}{dx^n} + \cdots + a_1 \frac{dy}{dx} + a_0 y = 0$$

the general solution can be formed from any n linearly independent solutions $y_1, y_2, \ldots, y_n$ of this equation:

$$y = c_1 y_1 + c_2 y_2 + \cdots + c_n y_n$$

The determination of the c_j from the initial conditions now reduces to the solution of a system of n linear algebraic equations.

Chapter 5

Applications

In this final chapter we are going to discuss how complex numbers can be used in physical applications.

5.1 Using Complex Numbers in Differential Equations

5.1.1 Auxiliary Complex Variable - Application to Forced Oscillator

Boas 8.6

McCall 3.10

We have already seen in Section 4.1.3 how a real differential equation can give rise to complex numbers. However, the evolution of real physical systems are described using real numbers. Although the constants C and D in Eq. (4.17) can in general be complex, we found that when applied to a physical oscillator system, imposing the real initial conditions of Eq. (4.23), resulted in the solution represented in Eq. (4.25), being purely real.

We can take this as a clue that some simplifications might occur by representing physical quantities as the real parts of complex numbers. Say we have a physical quantity x that satisfies some differential equation in time t. An example is the damped oscillator we examined earlier, where we solved for the displacement of the mass as a function of time, $x(t)$. Now consider the lightly damped forced oscillator in Fig. 5.1,where everything is the same as before, except we now have a periodic external force that prevents the mass from decaying to its equilibrium position. The equation of motion is now:

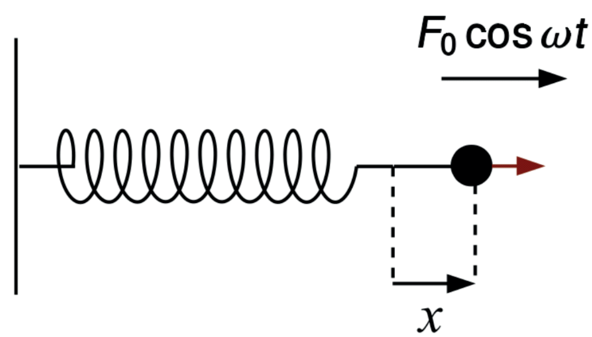

Figure 5.1: Forced oscillator.

$$\ddot{x} + \gamma \dot{x} + \omega_0^2 x = (F_0/m) \cos \omega t \,. \tag{5.1}$$

The general solution to this inhomogeneous equation is the sum of the complementary function and the particular integral. We already know the complementary function (cf. Eq. (4.28)), and are going to find the particular integral by defining an **auxiliary complex variable** $\tilde{z} = x + iy$ that satisfies the auxiliary differential equation

$$\ddot{\tilde{z}} + \gamma \dot{\tilde{z}} + \omega_0^2 \tilde{z} = (F_0/m) e^{i\omega t} \,. \tag{5.2}$$

Evidently Eq. (5.1) is found by taking the real part of Eq. (5.2). Inserting the trial solution $\tilde{z} = \tilde{A} e^{i\omega t}$ leads to

$$\left(-\omega^2 + i\gamma\omega + \omega_0^2\right) \tilde{A} e^{i\omega t} = (F_0/m) e^{i\omega t} \,,$$

implying that the *complex amplitude* $\tilde{A}$ is given by

$$\tilde{A} = \frac{F_0/m}{-\omega^2 + i\gamma\omega + \omega_0^2} \,,$$

Using the polar form $\tilde{A} = A e^{i\phi}$ we readily find that

$$A = \frac{F_0/m}{\left[(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2\right]^{1/2}} \,, \quad \tan \phi = -\frac{\omega \gamma}{\omega_0^2 - \omega^2} \,.$$

The particular integral $x(t) = \mathrm{Re}(\tilde{z})$ is given by

$$x(t) = \frac{F_0/m}{\left[(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2\right]^{1/2}} \cos(\omega t + \phi) \,. \tag{5.3}$$

Since the complementary function decays away, Eq. (5.3) represents the *steady state* of the forced oscillator. The physics of this solution will be reviewed in the Vibrations and Waves course. Here we will just note that the amplitude of the steady-state oscillations increases dramatically around the *resonance frequency*

$$\omega_r = \left(\omega_0^2 - \gamma^2/2\right)^{1/2} . \tag{5.4}$$

From the mathematical standpoint, the solution of a real differential equation was facilitated through solving an auxiliary complex differential equation, and then taking the real part. In fact, solving the forced oscillator problem without using complex variables is quite tricky[1].

Electrical Analogue

Boas 2.16

McCall 3.11

The mechanical forced oscillator has an electrical analogue. The circuit in Fig. 5.2 consists of an inductance, L, a resistance, R, and a capacitor, C, all in series with an alternating voltage source $V\cos(\omega t)$. In terms of the charge, $q(t)$, and the current, $I = \dot{q}$, passing through each element, the voltages dropped over the inductor, resistor and the capacitor are respectively $L(dI/dt) = L\ddot{q}$, $RI = R\dot{q}$ and q/C. According to Kirchhoff's law, the sum of these voltage drops equals the applied alternating voltage so that

$$L\ddot{q} + R\dot{q} + \frac{q}{C} = V\cos(\omega t). \tag{5.5}$$

Identifying $F_0 \leftrightarrow V$, $m \leftrightarrow L$, $b \leftrightarrow R$, $k \leftrightarrow C^{-1}$ and $x \leftrightarrow q$, we can, by comparison with Eqs. (4.20), (4.22) and (5.3), immediately write the steady state of Eq. (5.5) as

$$q(t) = \frac{V/L}{\left[(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2\right]^{1/2}} \cos(\omega t + \phi) \,, \tag{5.6}$$

[1]See pp. 51-57 of *Classical Mechanics, 2nd Ed.*, by M. W. McCall, Wiley (2011).

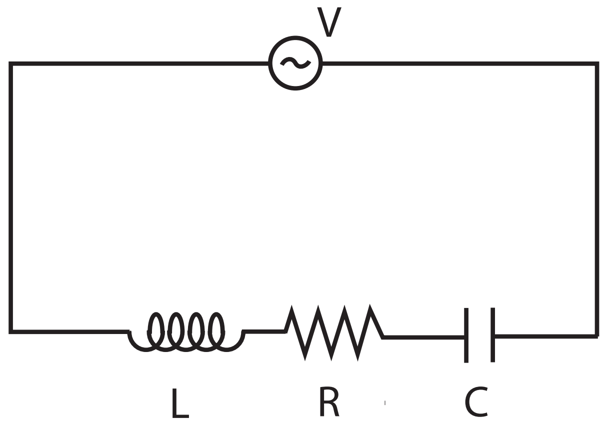

Figure 5.2: Forced oscillations in electronics.

where now

$$\omega_0 = (LC)^{-1/2} \ , \ \ \gamma = RL^{-1} \ , \tag{5.7}$$

and

$$\tan\phi = \frac{-\omega\gamma}{\omega_0^2 - \omega^2} \,. \tag{5.8}$$

Amplification at resonance of electrical signals in an LCR circuit is the basis of how a radio set is tuned.

5.1.2 Motion in the Argand Plane

In the Electricity and Magnetism course we will encounter the equation of motion for a particle of charge q and mass m moving in a uniform magnetic field $\mathbf{B}$:

$$m\frac{d\mathbf{v}}{dt} = q\mathbf{v} \times \mathbf{B} \,.$$

If the field is in the $\hat{\mathbf{z}}$ direction and we assume motion in the x-y plane, then the component equations of motion are

$$m\frac{dv_x}{dt} = qv_yB \,, \quad m\frac{dv_y}{dt} = -qv_xB \,. \tag{5.9}$$

This is a pair of coupled first order equations. Since there are two equations, there is a possibility that they can be combined into a single complex equation in the variable $\tilde{v} = v_x + iv_y$. Indeed it is easy to show that when we resolve the equation

$$\frac{d\tilde{v}}{dt} = -i\omega_c \tilde{v} \,, \tag{5.10}$$

into its real and imaginary parts, we recover Eqs. (5.9) provided $\omega_c = qB/m$. Now we can solve Eq. (5.10) easily using the methods developed in Chapter 3, but now applied to a complex first order differential equation. We find

$$\tilde{v} = \tilde{v}(0)e^{-i\omega_c t}\,, \tag{5.11}$$

where $\tilde{v}(0) = v_x(0) + iv_y(0)$ is a complex constant given by the initial value of $\tilde{v}$. Taking the real and imaginary parts of Eq. (5.11) yields the solution

$$v_x(t) = v_x(0)\cos\omega_c t + v_y(0)\sin\omega_c t\,,$$
$$v_y(t) = -v_x(0)\sin\omega_c t + v_y(0)\cos\omega_c t\,.$$

This represents the solution of Eqs. (5.9) for a given initial velocity. Noting that $v_x = dx/dt$ etc., a further integration of the above two equations yields

$$\int_{x(0)}^{x(t)} dx = x(t) - x(0) = \int_0^t v_x(t)dt = \omega_c^{-1}\left[v_x(0)\sin\omega_c t + v_y(0)\left(1-\cos\omega_c t\right)\right]\,,$$
$$\int_{y(0)}^{y(t)} dy = y(t) - y(0) = \int_0^t v_y(t)dt = \omega_c^{-1}\left[-\left(1-\cos\omega_c t\right)v_x(0) + v_y(0)\sin\omega_c t\right]\,,$$

which gives the full particle trajectory $[x(t), y(t)]$. We will see in the Electricity and Magnetism course that the particle trajectory is in fact a circle of radius $\omega_c^{-1}[v_x^2(0) + v_y^2(0)]^{1/2}$ whose centre lies at the point with coordinates $[x(0) + v_y(0)/\omega_c, y(0) - v_x(0)/\omega_c]$.

Another example that uses the same principle is the solution to the Foucault pendulum; see pp. 207-212 of *Classical Mechanics, 2nd Ed.*, by M. W. McCall, Wiley (2011).

5.2 Interference

The field of a spherical light wave of unit amplitude at radius r and time t may be written as[2]

$$E = \cos\left(\frac{2\pi}{\lambda}r - \omega t\right)\,,$$

where λ is the wavelength and ω is the (angular) frequency. Now consider the

[2]Strictly, there should be a factor of r^{-1} on the rhs, so that on squaring we get the inverse square law for the intensity fall-off of a spherical wave. But omitting this factor makes no difference to the final result.

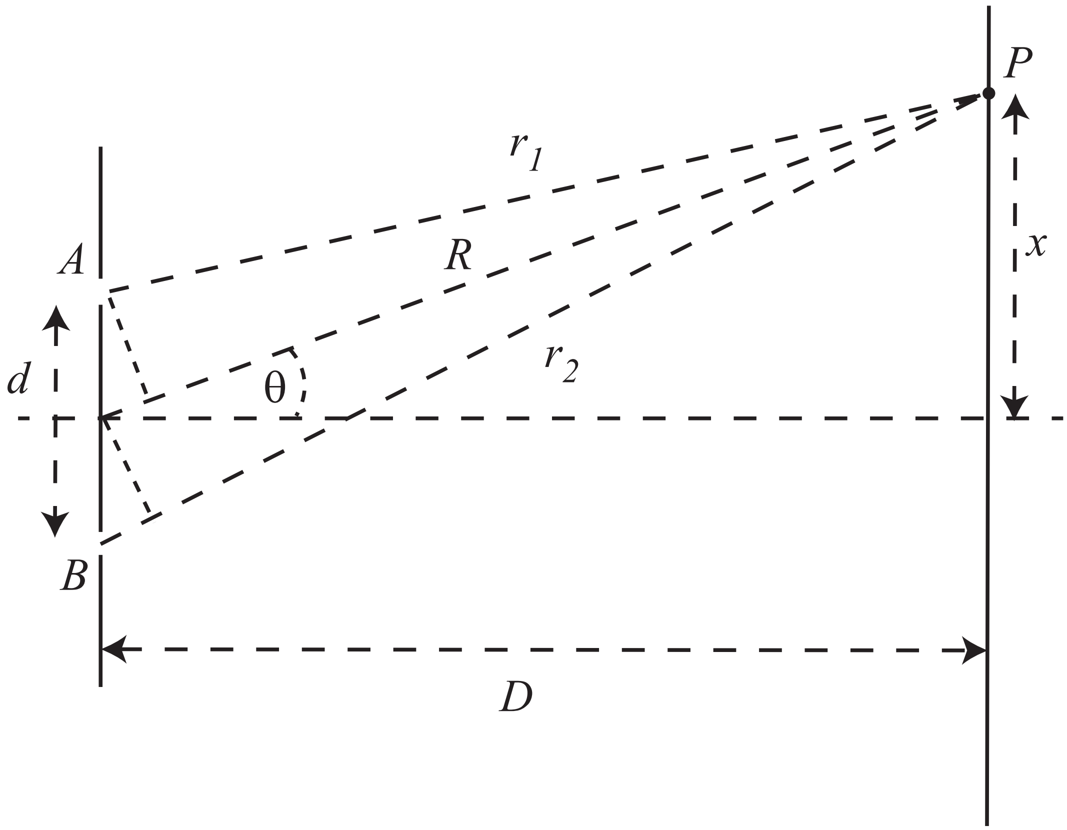

Figure 5.3: Two-slit interference.

spherical waves issuing from the slits labelled *A* and *B* in Fig. 5.3. The total field reaching the screen at point *P* is given by

$$E_1 + E_2 = \cos\left(\frac{2\pi}{\lambda} r_1 - \omega t\right) + \cos\left(\frac{2\pi}{\lambda} r_2 - \omega t\right).$$

The total intensity is given by the square of the total field

$$I(P,t) = |E_1 + E_2|^2 = \left|\cos\left(\frac{2\pi}{\lambda} r_1 - \omega t\right) + \cos\left(\frac{2\pi}{\lambda} r_2 - \omega t\right)\right|^2.$$

The result is an intensity variation in both position and time. Now most optical detectors are not fast enough to pick up the rapid temporal variation, and instead record a time-averaged intensity:

$$I_{av}(P) = \frac{1}{\tau}\int_0^{\tau} I(P,t)\,dt\,,$$

where the observation time τ is much longer than the short period ($= 2\pi/\omega$) of the wave's oscillation. You can see that inserting the expression for $I(P,t)$ will result in a complicated averaging process to obtain observed intensity, $I_{av}(P)$.

It turns out that we can get the correct answer much more slickly by using complex numbers. Define a complex field for a wave as

$$\tilde{E} = e^{i\left(\frac{2\pi}{\lambda}r - \omega t\right)}.$$

The total complex field at P is then given by

$$\tilde{E}_1 + \tilde{E}_2 = e^{i\left(\frac{2\pi}{\lambda}r_1 - \omega t\right)} + e^{i\left(\frac{2\pi}{\lambda}r_2 - \omega t\right)}.$$

In this complex representation the observed intensity at P is given by the squared amplitude of the total field:

$$I_{av}(P) = \left|\tilde{E}_1 + \tilde{E}_2\right|^2 = \left|e^{i\left(\frac{2\pi}{\lambda}r_1 - \omega t\right)} + e^{i\left(\frac{2\pi}{\lambda}r_2 - \omega t\right)}\right|^2 = \left|e^{i\frac{2\pi}{\lambda}r_1} + e^{i\frac{2\pi}{\lambda}r_2}\right|^2. \quad (5.12)$$

Notice that the time dependence $e^{-i\omega t}$ is common to both terms and is eliminated on calculating the squared modulus. The complicated time-averaging process has been finessed. Further simplification occurs if we note that provided the screen is placed a long way from the slits (i.e. $D \gg d$) then the radii r_1, r_2 and R are approximately parallel, and

$$r_{1,2} = R \mp \frac{d}{2}\sin\theta \approx R \mp \frac{dx}{2D}\,,$$

where we have further assumed that $x \ll D$ (so that $\sin\theta \approx \tan\theta = x/D$). Using these expressions for $r_{1,2}$ in Eq. (5.12) results in

$$I_{av}(P) = \left|e^{i\frac{2\pi}{\lambda}R}\left(e^{-i\frac{2\pi}{\lambda}\frac{xd}{2D}} + e^{i\frac{2\pi}{\lambda}\frac{xd}{2D}}\right)\right|^2 = 4\cos^2\left(\frac{\pi x d}{\lambda D}\right),$$

where now it's the common factor of $e^{i\frac{2\pi R}{\lambda}}$ that drops out on taking the squared modulus. The result for $I_{av}(P)$ is the familiar "cos squared" fringe pattern on the screen formed in this Young's slits experiment. It is not hard to show that the fringe spacing δx is given by

$$\delta x = \lambda\frac{D}{d}.$$

These results will also be very helpful in analysing interference of particle waves (i.e. de Broglie waves) in quantum mechanics.