2 A brief review of calculus

2.1 Introduction

While calculus is the mathematical key to an understanding of Nature, its roots lie really in problems of geometry.

Thus the **derivative** of a function $y = f(x)$, defined by

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}, \tag{2.1}$$

arises from the problem of finding the tangent to a given curve, and dy/dx represents the *slope* of that tangent, and hence of the curve itself at the point in question (Fig. 2.1). The notation $f'(x)$, rather than dy/dx, is sometimes used.

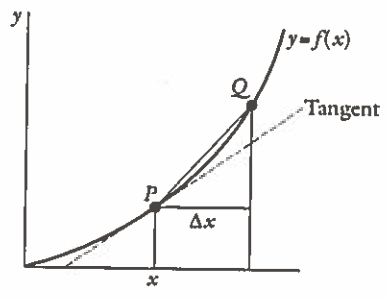


Fig. 2.1 Finding the tangent, by taking Q closer and closer to P.

From the point of view of the present book, **integration** is essentially the opposite process. Thus, given a function $f(x)$, we seek a function $I(x)$ such that

$$\frac{dI}{dx} = f(x), \tag{2.2}$$

and denote the outcome by

$$I = \int f(x)\,dx, \tag{2.3}$$

this **integral** being determined only to within an arbitrary additive constant. Yet there is, again, a geometrical interpretation of integral as the *area under a curve* (Fig. 2.2).

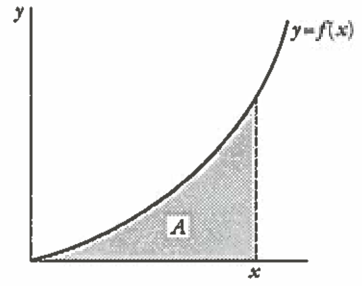


Fig. 2.2 Area under a curve: A has the property $dA/dx = f(x)$.

Now, it is not our purpose here to give a systematic account of the calculus; we aim only to collect together, for easy reference, some of the main results, and to recall some of the ways in which these results are related to one another.

2.2 Some elementary results

We assume first that the reader will be familiar with various special cases of differentiation and integration, including

$$\frac{d}{dx}(x^n) = nx^{n-1} \tag{2.4}$$

and its counterpart

$$\int x^n\,dx = \frac{x^{n+1}}{n+1}, \quad n \neq -1. \tag{2.5}$$

Here n need not be an integer; it is not difficult to show that the results hold equally well for *rational* n, i.e. $n = p/q$, where p and q are integers.

Two particularly useful results are

$$\begin{aligned} \frac{d}{dx}(\sin x) &= \cos x, \\ \frac{d}{dx}(\cos x) &= -\sin x, \end{aligned} \tag{2.6a, b}$$

where radian measure is used.

Fig. 2.3 G. W. Leibniz (1646–1716).

The calculus of more complicated functions is helped by the rules for **differentiating a product or quotient:**

$$\frac{\mathrm{d}}{\mathrm{d}x}(uv) = \frac{\mathrm{d}u}{\mathrm{d}x}v + u\frac{\mathrm{d}v}{\mathrm{d}x}, \tag{2.7a}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{u}{v}\right) = \frac{1}{v^2}\left(v\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}v}{\mathrm{d}x}\right), \tag{2.7b}$$

which Leibniz discovered in 1675. A consequence of the first of these is the rule for **integration by parts:**

$$\int \frac{\mathrm{d}u}{\mathrm{d}x}v\,\mathrm{d}x = uv - \int u\frac{\mathrm{d}v}{\mathrm{d}x}\,\mathrm{d}x. \tag{2.8}$$

It often happens that we have some variable y which is a function of x, while x itself is some function of another variable, say t. In this way y can be regarded as a function of t, and Leibniz's **chain rule** tells us that

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}. \tag{2.9}$$

This has, again, a natural counterpart in the formula for **integration by substitution:**

$$\int z\,\mathrm{d}x = \int z\frac{\mathrm{d}x}{\mathrm{d}t}\,\mathrm{d}t. \tag{2.10}$$

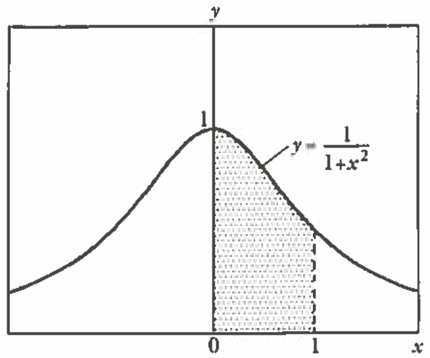


Fig. 2.4 The area corresponding to (2.11).

As an illustration of this last result, consider the area shown in Fig. 2.4, i.e.

$$A = \int_0^1 \frac{1}{1+x^2}\,\mathrm{d}x. \tag{2.11}$$

On introducing the change of variable $x = \tan\theta$ this can be rewritten

$$A = \int_0^{\pi/4} \frac{1}{\sec^2\theta}\frac{\mathrm{d}x}{\mathrm{d}\theta}\,\mathrm{d}\theta = \int_0^{\pi/4}\frac{1}{\sec^2\theta}\cdot\sec^2\theta\,\mathrm{d}\theta = \int_0^{\pi/4}\mathrm{d}\theta,$$

so that we obtain the elegant formula

$$\int_0^1 \frac{1}{1+x^2}\,\mathrm{d}x = \frac{\pi}{4}. \tag{2.12}$$

2.3 Taylor series

We often need an approximation to a function $y = f(x)$ for values of x near to some particular value $x = a$, say, and a crude but obvious way to do this is to draw a straight line through the point in question with the correct slope $f'(a)$, so that

$$y \doteq f(a) + (x-a)f'(a) \tag{2.13}$$

(Fig. 2.5).

This takes no account of the local curvature, of course, and we should be able to do rather better by taking a quadratic function of $(x-a)$, i.e. $y = c_0 + c_1(x-a) + c_2(x-a)^2$. We then choose the constants c_0, c_1, c_2 so that the values of y, y' and y'' are all correct at $x = a$; this requires taking $c_0 = f(a)$, $c_1 = f'(a)$, $c_2 = \frac{1}{2}f''(a)$. Continuing in this way we are led to the idea of a **Taylor series** representation for $y = f(x)$ about $x = a$:

$$y = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \frac{(x-a)^3}{3!}f'''(a) + \cdots. \tag{2.14}$$

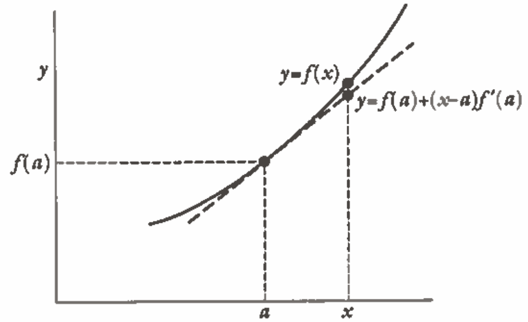


Fig. 2.5 Illustrating (2.13).

Taking $a = 0$, two particular examples are

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \cdots$$
$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \cdots. \qquad (2.15\text{a, b})$$

In the case of $\sin x$, Fig. 2.6 shows how the first few terms provide a very good approximation, provided $|x|$ is reasonably small. For larger $|x|$ we need to take more terms to get a good approximation, but the series (2.15a, b) may be used successfully in this way no matter how large $|x|$ is.

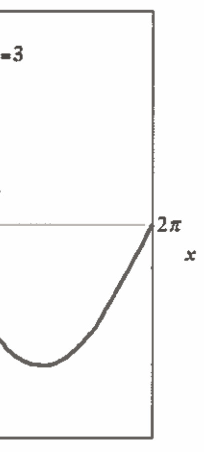


Fig. 2.6 Taylor series approximations to $y = \sin x$, about $x = 0$, using different numbers of terms. Over the interval shown, $n = 8$ gives an extremely good approximation to the actual curve.

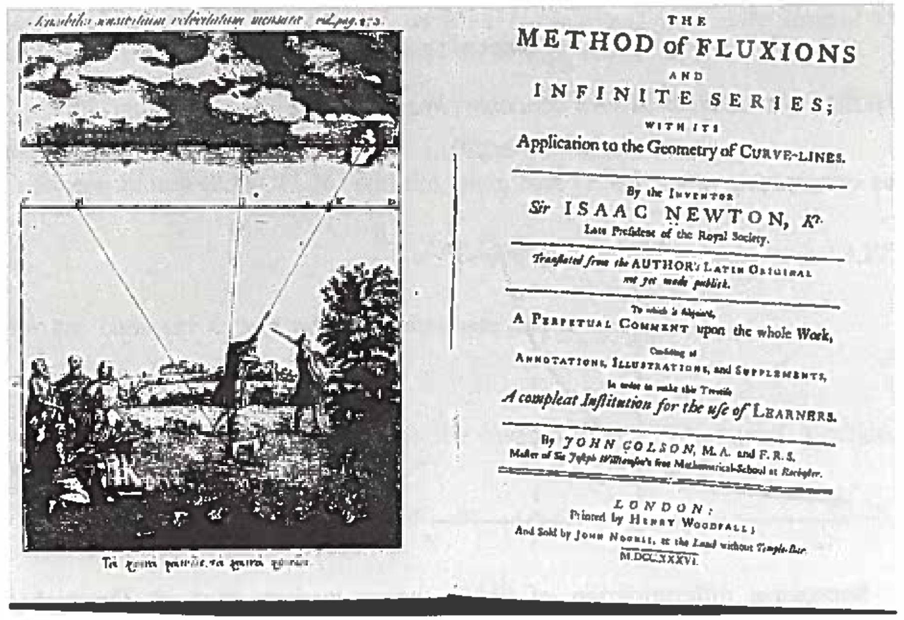
More typically, a Taylor series (2.14) converges only for $|x - a|$ less than some definite number R, called the radius of convergence. An important example of this is the **binomial series**

$$(1+x)^{\alpha} = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \cdots, \qquad (2.16)$$

which converges only for $|x| < 1$. Here α may be any real number, positive or negative, but the importance of the condition $|x| < 1$ can be seen very easily in the particular case $\alpha = -1$. The sum of the first n terms is then

$$1 - x + x^2 \cdots + (-1)^{n-1}x^{n-1} = \frac{1 + (-1)^{n-1}x^n}{1+x}, \qquad (2.17)$$

as we may confirm by multiplying both sides by $1 + x$ and observing all the cancellation. On taking the limit $n \to \infty$ we obtain $(1+x)^{-1}$, as desired, *provided* $|x| < 1$. This condition is vital, for only then does the term $(-1)^{n-1}x^n$ in the numerator of the right-hand side of (2.17) tend to zero, rather than oscillate ever more wildly as $n \to \infty$.



THE
METHOD of FLUXIONS
AND
INFINITE SERIES;
WITH ITS
Application to the Geometry of CURVE-LINES.

By the INVENTOR
Sir ISAAC NEWTON, Kt.
Late President of the Royal Society.

Translated from the AUTHOR's LATIN ORIGINAL
not yet made publick.

To which is subjoin'd,
A PERPETUAL COMMENT upon the whole Work,
Consisting of
ANNOTATIONS, ILLUSTRATIONS, and SUPPLEMENTS,
In order to make this Treatise
A compleat Institution for the use of LEARNERS.

By JOHN COLSON, M.A. and F.R.S.
Master of Sir Joseph Williamson's free Mathematical-School at Rochester.

LONDON:
Printed by HENRY WOODFALL;
And Sold by JOHN NOURSE, at the Lamb without Temple-Bar.
M.DCC.XXXVI.

Fig. 2.7 Newton's 1671 treatise on the calculus, eventually published in 1736.

While the main result, (2.14), was published by Brook Taylor in 1715, it was effectively known and used by Newton and others much earlier. Infinite series were, in fact, central to much of Newton's calculus, in the sense that he would often effect an integration by expanding the integrand in an infinite series and then integrate each term separately. If we borrow (2.12), an example of this type of argument is

$$\frac{\pi}{4} = \int_0^1 \frac{1}{1+x^2}\,dx = \int_0^1 1 - x^2 + x^4 - x^6 \cdots dx$$
$$= \left[x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \cdots\right]_0^1, \qquad (2.18)$$

giving the beautiful result

$$\frac{\pi}{4} = 1 - \tfrac{1}{3} + \tfrac{1}{5} - \tfrac{1}{7} + \cdots. \tag{2.19}$$

While this is often credited to Gregory and Leibniz, it was apparently first discovered by Indian mathematicians some 150 years earlier.

2.4 The functions e^x and $\log x$

From the point of view of the present book, the key property of the **exponential** function $y = \exp(x)$ is that

$$\frac{d}{dx}[\exp(x)] = \exp(x), \tag{2.20}$$

so that it is *equal to its own derivative*. We shall take this, together with

$$\exp(0) = 1, \tag{2.21}$$

as our starting point.

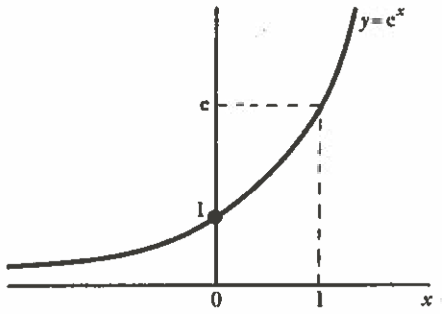


Fig. 2.8 The exponential function.

Successive differentiation of (2.20) shows, in fact, that all the higher derivatives of $\exp(x)$ are also equal to $\exp(x)$, and (2.21) then implies that they are all equal to 1 when $x = 0$. The Taylor series (2.14) for $y = \exp(x)$ about $x = 0$ is therefore

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \tag{2.22}$$

and it can be shown that this converges for all x.

If we *define* $e = \exp(1)$, then

$$\begin{aligned} e &= 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots \\ &= 2.71828\ 18284\ 59045\ldots. \end{aligned} \tag{2.23}$$

Using (2.20) and (2.21) we may show that

$$\exp(x + y) = \exp(x)\exp(y) \tag{2.24}$$

(Ex. 2.2). It follows that

$$\exp(nx) = [\exp(x)]^n,$$

and, in particular, that

$$\exp(n) = e^n \tag{2.25}$$

for positive integer values of n. The result can in fact be extended without too much difficulty to any *rational* value of n, and as elementary mathematics ascribes no clear meaning to e^x when x is irrational it is only natural to *define* e^x, when x is irrational, as $\exp(x)$. In this way

$$\exp(x) = e^x \tag{2.26}$$

for all real x.

By combining (2.20), (2.26) and the chain rule (2.9) we may show that

$$\frac{d}{dx}(e^{kx}) = k\,e^{kx} \tag{2.27}$$

for any constant k, and we shall often use this result.

The function $\log x$

We may define the function log as the inverse of the exponential function, so that

$$z = \log x \quad \Leftrightarrow \quad x = e^z. \tag{2.28}$$

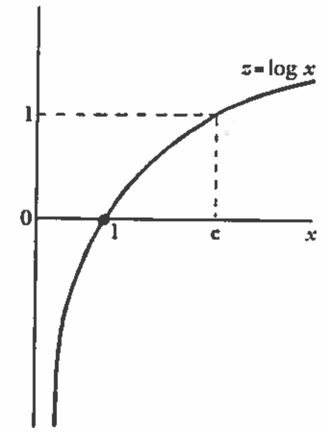


Fig. 2.9 The function $\log x$.

Some important consequences are

$$\log 1 = 0, \tag{2.29}$$

$$\log(ab) = \log a + \log b, \tag{2.30}$$

and it also follows that

$$\frac{\mathrm{d}}{\mathrm{d}x}(\log x) = \frac{1}{x}, \tag{2.31}$$

so that the 'missing' integral in the set (2.5) has now been identified: the integral of x^{-1} is $\log x$ (plus a constant).

Euler's formula for $e^{i\theta}$

The Taylor series (2.22) can be shown to be valid for all *real* x. If we are reckless enough to set $x = i\theta$, where $i = \sqrt{-1}$ and θ is real, we obtain

$$\begin{aligned} e^{i\theta} &= 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{i\theta^5}{5!} \cdots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} \cdots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} \cdots\right), \end{aligned} \tag{2.32}$$

and on using (2.15) this becomes

$$e^{i\theta} = \cos\theta + i\sin\theta. \tag{2.33}$$

This extraordinary formula was originally derived by Euler in a quite different way (Fig. 2.10). We must really view it as a *definition* of $e^{i\theta}$, for we have not yet established any meaning to a number raised to an imaginary power. Nonetheless, we soon find that $e^{i\theta}$ behaves according to all the 'usual' rules. In particular,

$$\frac{\mathrm{d}}{\mathrm{d}\theta}(e^{i\theta}) = ie^{i\theta} \tag{2.34}$$

and products behave according to the usual index law

$$e^{i\theta_1} e^{i\theta_2} = e^{i(\theta_1 + \theta_2)} \tag{2.35}$$

(Ex. 2.5). In consequence,

$$(e^{i\theta})^n = e^{in\theta},$$

where n is any positive integer, so

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta, \tag{2.36}$$

a result known as De Moivre's theorem.

Finally, on setting $\theta = \pi$ in Euler's formula (2.33) we obtain

$$e^{i\pi} = -1, \tag{2.37}$$

138. Ponatur denuo in formulis § 133 arcus z infinite parvus et sit n numerus infinite magnus i, ut iz obtineat valorem finitum v. Erit ergo $nz = v$ et $z = \frac{v}{i}$, unde $\sin. z = \frac{v}{i}$ et $\cos. z = 1$; his substitutis fit

$$\cos. v = \frac{\left(1 + \frac{v\sqrt{-1}}{i}\right)^i + \left(1 - \frac{v\sqrt{-1}}{i}\right)^i}{2}$$

atque

$$\sin. v = \frac{\left(1 + \frac{v\sqrt{-1}}{i}\right)^i - \left(1 - \frac{v\sqrt{-1}}{i}\right)^i}{2\sqrt{-1}}$$

In capite autem praecedente vidimus esse

$$\left(1 + \frac{z}{i}\right)^i = e^z$$

denotante e basin logarithmorum hyperbolicorum; scripto ergo pro z partim $+v\sqrt{-1}$ partim $-v\sqrt{-1}$ erit

$$\cos. v = \frac{e^{+v\sqrt{-1}} + e^{-v\sqrt{-1}}}{2}$$

et

$$\sin. v = \frac{e^{+v\sqrt{-1}} - e^{-v\sqrt{-1}}}{2\sqrt{-1}}$$

Ex quibus intelligitur, quomodo quantitates exponentiales imaginariae ad sinus et cosinus arcuum realium reducantur. Erit vero

$$e^{+v\sqrt{-1}} = \cos. v + \sqrt{-1} \cdot \sin. v$$

et

$$e^{-v\sqrt{-1}} = \cos. v - \sqrt{-1} \cdot \sin. v.$$

Fig. 2.10 Euler's formula (2.33), as first stated in his Introductio in analysin infinitorum *(1748). Compare the third displayed expression with (2.38); at this time the concept of a limit had still not been clearly formulated, and* i *is being used here to denote infinity (∞). Only later did Euler introduce the notation* $i = \sqrt{-1}$.

which relates the three fundamental quantities e, π and i and is widely regarded as one of the most beautiful equations in the whole of mathematics.

Exercises

2.1 *Stationary points*. A function $y = f(x)$ is said to have a stationary point wherever $\mathrm{d}y/\mathrm{d}x = 0$. Find all such points in the case

$$y = x^3 - ax + 1,$$

where a is a constant, and examine the sign of $\mathrm{d}^2y/\mathrm{d}x^2$ at each one to determine whether y has a local maximum or minimum there.

2.2 Show directly from (2.20) and (2.21) that

$$\exp(x+y)=\exp(x)\exp(y),$$

by considering

$$\frac{\exp(x+y)}{\exp(x)},$$

with y held constant.

2.3 *Another view of* e^x. The interesting result

$$\lim_{n\to\infty}\left(1+\frac{\alpha}{n}\right)^n = e^{\alpha}, \tag{2.38}$$

with α fixed, independent of the integer n, provides a quite different method from (2.23) for the calculation of e, though convergence is extremely slow (Table 2.1).

Table 2.1 Convergence to e using (2.38)

n	$(1+(1/n))^n$
1	2.0000
10	2.5937
100	2.7048
1000	2.7169
10000	2.7181

Prove (2.38) by first observing that (2.31) implies that

$$\lim_{\Delta x\to 0}\frac{\log(x+\Delta x)-\log x}{\Delta x}=\frac{1}{x},$$

and then setting $\Delta x = 1/n$.

2.4 *The functions* $\cosh x$ *and* $\sinh x$. These are defined as follows:

$$\cosh x = \tfrac{1}{2}(e^x+e^{-x}), \qquad \sinh x = \tfrac{1}{2}(e^x-e^{-x}).$$

Show that

$$\frac{d}{dx}\cosh x = \sinh x, \qquad \frac{d}{dx}\sinh x = \cosh x$$

and

$$\cosh^2 x - \sinh^2 x = 1.$$

2.5 Prove that

$$e^{i\theta_1}e^{i\theta_2}=e^{i(\theta_1+\theta_2)}.$$

3 Ordinary differential equations

3.1 Introduction

In this chapter we consider some methods for solving first and second order differential equations.

We begin with first order equations, i.e. equations of the form

$$\frac{dx}{dt}=f(x,t), \tag{3.1}$$

where $f(x,t)$ is some given function of x and t. There will also be an initial condition

$$x=x_0 \quad \text{at} \quad t=0, \tag{3.2}$$

x_0 being a given constant.

While we focus attention here on trying to find an exact expression for x in terms of t, we shall bear in mind that the question of most practical interest is often: 'what *eventually* happens to x?'. In particular,

(i) does $|x|\to\infty$?
(ii) does x tend to some finite limit?
(iii) does x settle down into some kind of regular oscillation?
(iv) might it perhaps do none of these things, but instead fluctuate erratically (while remaining bounded), even as $t\to\infty$?

(see Fig. 3.1).

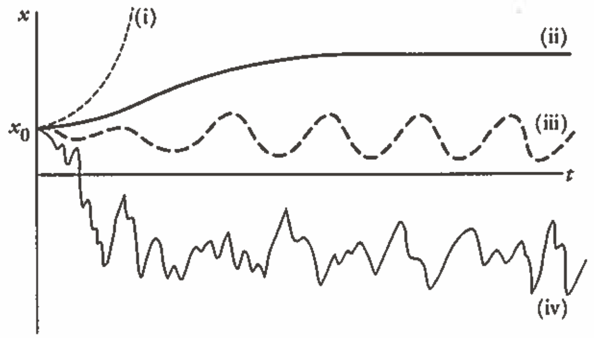


Fig. 3.1 Four conceivable possibilities for the eventual behaviour of the solution x(t) to a differential equation and given initial condition(s).