The Open University

MST210

Mathematical methods, models and modelling

Handbook

This publication forms part of an Open University module. Details of this and other Open University modules can be obtained from the Student Registration and Enquiry Service, The Open University, PO Box 197, Milton Keynes MK7 6BJ, United Kingdom (tel. +44 (0)845 300 6090, email general-enquiries@open.ac.uk).

Alternatively, you may visit the Open University website at www.open.ac.uk where you can learn more about the wide range of modules and packs offered at all levels by The Open University.

To purchase a selection of Open University materials visit www.ouw.co.uk, or contact Open University Worldwide, Walton Hall, Milton Keynes MK7 6AA, United Kingdom for a brochure (tel. +44 (0)1908 858779, fax +44 (0)1908 858787, email ouw-customer-services@open.ac.uk).

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published 2015.

Edited, designed and typeset by The Open University, using the Open University TEX System.

Printed in the United Kingdom by Bell & Bain Ltd, Glasgow.

SUP 03848 6

1.1

Contents

Introduction

This Handbook is a reference that you can take into the MST210 exam, and it may be valuable when you start to apply your knowledge in other modules. It will be more effective if you are already familiar with it before you sit the exam, and we suggest that you consult it when you attempt the assignments in the module. This Handbook is not designed as a teaching document, and reading it is not a substitute for studying the module units.

The first few sections consist of general mathematical reference material. It is not intended to be a self-contained or logically complete account of basic mathematics; it is just a set of definitions and results. Some of these results are used repeatedly in the module. Do not be alarmed if you see some things in this Handbook that are not discussed in the module, as the exam will be based solely on topics that are discussed in the units.

The later sections of the Handbook are brief summaries of the units, emphasising the most important results.

1 Notation

1.1 Greek alphabet

α	A	alpha	ι	I	iota	ρ	P	rho
β	B	beta	κ	K	kappa	σ	Σ	sigma
γ	Γ	gamma	λ	Λ	lambda	τ	T	tau
δ	Δ	delta	μ	M	mu	υ	Υ	upsilon
ε	E	epsilon	ν	N	nu	ϕ	Φ	phi
ζ	Z	zeta	ξ	Ξ	xi	χ	X	chi
η	H	eta	o	O	omicron	ψ	Ψ	psi
θ	Θ	theta	π	Π	pi	ω	Ω	omega

1.2 Symbols

$=$	is equal to
$\neq$	is not equal to
$\simeq$	is approximately equal to
$\pm$	plus or minus
$\mp$	minus or plus
$<$	less than
$\leq$	less than or equal to
$>$	greater than
$\geq$	greater than or equal to
$\sqrt{\ }$	positive square root
e	the number $2.718\,28\ldots$
π	the number $3.141\,59\ldots$
∞	infinity
$\mathbb{Z}$	the integers
$\mathbb{R}$	the real numbers
$\mathbb{C}$	the complex numbers

1.3 SI units (Système International d'Unités)

The SI system of units of measurement is based on seven *base units*, of which the following four are used in this module:

length:	the metre	(abbreviated to m);
mass:	the kilogram	(abbreviated to kg);
time:	the second	(abbreviated to s);
temperature:	the kelvin	(abbreviated to K).

Note that it is *not* 'degrees kelvin'.

Of these, only the kelvin may be unfamiliar. To express a temperature in kelvins, add 273.2 to the temperature expressed in degrees Celsius or 'degrees centigrade' (so the temperature of melting ice is 273.2 K, and that of boiling water is 373.2 K). The **absolute zero** of temperature is 0 K.

Units for other quantities are given in terms of the base units, as in the following examples:

volume:	metres cubed (or cubic metres)	(m^3);
velocity:	metres per second	($m\,s^{-1}$);
acceleration:	metres per second per second	($m\,s^{-2}$).

Notice that 'per' is represented by an index -1, and that the usual laws of indices apply, so 'metres per second per second' is written as $m\,s^{-2}$ rather than as $m\,s^{-1}\,s^{-1}$.

The following combinations of base units occur commonly enough to have special names:

force:	the newton (N)	= one kilogram metre per second per second	($kg\,m\,s^{-2}$);
energy:	the joule (J)	= one newton metre	($kg\,m^2\,s^{-2}$);
power:	the watt (W)	= one joule per second	($kg\,m^2\,s^{-3}$);
frequency:	the hertz (Hz)	= one cycle per second	(s^{-1}).

Other units are often expressed in terms of these subsidiary units, e.g.

stiffness of a spring (see Unit 9): newtons per metre ($N\,m^{-1}$).

To avoid very large or very small numbers, we also use multiple and fractional units, e.g. the kilometre (km), which is 1000 metres. The most important prefixes for forming these are given below.

giga	=	$1\,000\,000\,000 = 10^9$	(abbreviated to G)
mega	=	$1\,000\,000 = 10^6$	(abbreviated to M)
kilo	=	$1000 = 10^3$	(abbreviated to k)
milli	=	$1/1000 = 10^{-3}$	(abbreviated to m)
micro	=	$1/1\,000\,000 = 10^{-6}$	(abbreviated to μ)
nano	=	$1/1\,000\,000\,000 = 10^{-9}$	(abbreviated to n)

So, for example, the pressure of the atmosphere, which is about $10^5\,N\,m^{-2}$, is sometimes written as $100\,kN\,m^{-2}$, i.e. 100 kilonewtons per square metre.

1.4 Limits and sums

An ordered list of numbers $x_0, x_1, x_2, \ldots, x_n, \ldots$ is said to **converge** to the **limit** x if successive terms in the ordered list are better and better approximations to x. We write '$x_n \to x$ as $n \to \infty$' or '$\lim_{n\to\infty} x_n = x$'.

Given numbers $a_1, a_2, \ldots, a_n$, we define

$$\sum_{i=1}^{n} a_i \quad \text{to mean} \quad a_1 + a_2 + \cdots + a_{n-1} + a_n.$$

2 Numbers

2.1 Real numbers

The **integers** are the positive and negative whole numbers, together with zero. Non-integer numbers that can be expressed exactly as fractions are called **rational numbers**; those that cannot be so expressed, such as $\sqrt{2}$, e and π, are called **irrational**. The collection of the rational numbers (including the integers) and the irrational numbers is called the set of **real numbers**.

When a real number is expressed in decimal notation, if it is approximated then the approximation can be given to so many **decimal places**, or to so many **significant figures**. For example, 1.4142 is the approximation to $\sqrt{2}$ to four decimal places and five significant figures, while 0.000 000 342 is given to nine decimal places but three significant figures, and 342 000 has no decimal places but at least three significant figures. The process of reducing the number of decimal places or significant figures to which a number is expressed is referred to as **rounding**. To round a given number to n decimal places or n significant figures, take the number expressed to n decimal places or n significant figures that is closest to the given number, where it is conventional to round the digit 5 up. (Be aware that other conventions for rounding exist, and that computer programs and calculators do not always use this convention.)

A number given in the form $\pm b \times 10^c$, where $1 \leq b < 10$ and c is an integer, is said to be in **scientific notation**. For example, the number 342 000 can be expressed as 3.42×10^5 in scientific notation.

For any two real numbers a and b with $a < b$, we write:

$[a, b]$ for the set of all real numbers x such that $a \leq x \leq b$;
$[a, b)$ for the set of all real numbers x such that $a \leq x < b$;
$(a, b]$ for the set of all real numbers x such that $a < x \leq b$;
(a, b) for the set of all real numbers x such that $a < x < b$.

These sets of numbers are called **intervals**. The interval $[a, b]$ is a **closed** interval, (a, b) is an **open** interval, and the two others are **half-open** intervals. For physical problems the distinction between open and closed intervals is rarely significant. If we know that a real number x is 1.274 to three decimal places, then x lies in the interval $[1.2735, 1.2745)$.

2.2 Complex numbers

A **complex number** z is written in **Cartesian form** as $z = a + bi$, where a and b are real numbers, and $i^2 = -1$. We refer to a as the **real part** of z, written $\mathrm{Re}(z)$, and to b as the **imaginary part** of z, written $\mathrm{Im}(z)$.

Complex numbers can be added, e.g.

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$

or multiplied, e.g.

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

These formulas make use of the ordinary rules of algebra, together with the relation $i^2 = -1$.

The notation z^* is also used in some texts for the complex conjugate of z.

The **complex conjugate** of $z = a + bi$ is $\overline{z} = a - bi$. Note that

$$z\overline{z} = (a + bi)(a - bi) = a^2 + b^2$$

$\overline{z}$ and z^* are read as 'z bar' and 'z star'.

is a positive real number (unless $a = b = 0$). The **modulus** of z is the number $|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$.

To calculate a quotient of complex numbers, multiply top and bottom by the complex conjugate of the bottom, e.g.

$$\frac{a + bi}{c + di} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \left(\frac{ac + bd}{c^2 + d^2}\right) + \left(\frac{bc - ad}{c^2 + d^2}\right) i,$$

which expresses the quotient in the form $p + qi$ with p and q real.

The **Argand diagram** is a representation of complex numbers as points in a plane, where the complex number $a + bi$ is represented by the point with Cartesian coordinates (a, b). A point can also be represented in polar coordinates as (r, θ) and related to its Cartesian coordinates by

Polar coordinates are discussed in Subsection 5.2.

$$a = r\cos\theta, \quad b = r\sin\theta.$$

The angle coordinate θ is referred to as an **argument** of z, written $\arg(z)$, and the unique value of θ in the range $-\pi < \theta \leq \pi$ is referred to as the **principal value of the argument**, written $\mathrm{Arg}(z)$.

The multiplication of complex numbers in polar form is given by the rule $(r, \theta) \times (s, \phi) = (rs, \theta + \phi)$. So powers of a complex number can be expressed as $(r, \theta)^n = (r^n, n\theta)$, for n a positive integer. The special case when $r = 1$ is known as **de Moivre's theorem**:

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

Euler's formula is

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

This extends to any complex number $a + bi$ as

$$e^{a+bi} = e^a e^{bi} = e^a(\cos b + i\sin b).$$

The **exponential form** of a complex number $z = r(\cos\theta + i\sin\theta)$ is

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

This form is useful for multiplying, dividing and taking powers of complex numbers.

3 Functions and graphs

3.1 Functions

A **variable** is a quantity, represented by a symbol, that can vary over a set of values. If its value does not vary, then it is a **constant**.

Any expression or formula that involves a variable x, and whose value is uniquely determined by the value of x, is called a **function** of x.

If a variable y is a function of x (i.e. if y is equal to a function of x), then we call x the **independent variable** and y the **dependent variable**, and we may write $y = y(x)$. Here $y(x)$ stands for the function of x (i.e. for the formula involving x).

If f and g are two functions, then their **sum** is a function $f + g$ defined by

$$(f+g)(x) = f(x) + g(x) \quad \text{(for all } x\text{)}.$$

Moreover, if A and B are any two numbers, then the function $Af + Bg$ is defined by

$$(Af + Bg)(x) = A\,f(x) + B\,g(x) \quad \text{(for all } x\text{)}.$$

The function $f(g(x))$ is called the **composite function** or **composition** of the functions f and g.

The **graph** of a function $f(x)$ is the curve in the (x, y)-plane whose equation is $y = f(x)$.

A **continuous** function is one whose graph has no breaks or jumps in it, i.e. it can be drawn without lifting your pen from the paper.

A **constant function** $f(x)$ is one that assigns the same value to any input. Its graph is a straight line parallel to the x-axis. A special case is the **zero function**, which assigns the value 0 to any input.

A **linear function** is one having the form $a_1x + a_0$ (with $a_1 \neq 0$), where a_1 and a_0 are constants. Its graph is a straight line with **slope** a_1 and ***y*-intercept** a_0.

A **quadratic function** is one having the form $a_2x^2 + a_1x + a_0$ (with $a_2 \neq 0$), where a_2, a_1, a_0 are constants. Its graph is a parabola, similar in shape to the one shown in Subsection 5.3 if $a_2 > 0$, but the other way up if $a_2 < 0$.

A **cubic function** is one having the form $a_3x^3 + a_2x^2 + a_1x + a_0$ (with $a_3 \neq 0$), where a_3, a_2, a_1, a_0 are constants.

3.2 Polynomials

Linear, quadratic and cubic functions are all particular examples of **polynomial** functions, or simply polynomials. An ***n*th-order** polynomial, or polynomial of **degree *n***, is a function of the form

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0,$$

where n is a positive integer, x is a variable, and $a_0, a_1, \ldots, a_n$ are constants with $a_n \neq 0$. A linear polynomial has $n = 1$ (or $n = 0$), a quadratic polynomial has $n = 2$, and a cubic polynomial has $n = 3$.

The **roots** of a polynomial $p(x)$ are the solutions of the equation $p(x) = 0$. Every polynomial of degree n can be written as a product of a_n and n factors of the form $x - c_k$ $(k = 1, 2, \ldots, n)$, with each c_k a complex number (which may be real). Each of these factors corresponds to a root $x = c_k$ of the polynomial. If a factor $x - c$ occurs more than once, then the root $x = c$ is a **repeated root**; repeated roots are also sometimes referred to as **equal roots** or **coincident roots**.

The roots of a quadratic equation $ax^2 + bx + c = 0$, $a \neq 0$, are given by the **formula method** as

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

The quantity $b^2 - 4ac$ is referred to as the **discriminant** of the quadratic equation.

To **factorise** a polynomial is to express it as a product of two or more polynomials of lower degree. For example, the **difference of two squares** $x^2 - a^2$ factorises as

$$x^2 - a^2 = (x - a)(x + a),$$

and the **perfect square** $x^2 + 2ax + a^2$ factorises as

$$x^2 + 2ax + a^2 = (x + a)(x + a) = (x + a)^2.$$

3.3 Exponentials and logarithms

A function of the form $y = ba^x$, where a and b are non-zero constants (with $a > 0$ and $a \neq 1$), and x is real, is said to exhibit **exponential behaviour**. In a^x, a is referred to as the **base** and x as the **exponent** (or **index** or **power**). Properties of such a function include

$$a^0 = 1,$$
$$a^{-x} = 1/a^x,$$
$$a^x \times a^y = a^{x+y},$$
$$a^x/a^y = a^{x-y},$$
$$(a^x)^y = a^{xy} = (a^y)^x.$$

The function e^x, where $e = 2.718\,28\ldots$, is referred to as the **exponential function**. It is also written as $\exp x$.

Some texts refer to $y = ba^x$ as *an* exponential function; this is not to be confused with *the* exponential function, $\exp x$.

The **natural logarithm function** $\ln x$ is defined to be the **inverse function** of the exponential function $\exp x$, i.e. each reverses the effect of the other, so that

$$\ln(\exp x) = x \quad \text{for all real } x,$$
$$\exp(\ln x) = x \quad \text{for all real } x > 0.$$

In other words, if $e^y = x$, then $y = \ln x$, and vice versa.

The natural logarithm function has the properties

$$\ln 1 = 0,$$
$$\ln(1/x) = -\ln x, \quad x > 0,$$
$$\ln(xy) = \ln x + \ln y, \quad x, y > 0,$$
$$\ln(x/y) = \ln x - \ln y, \quad x, y > 0,$$
$$\ln x^y = y \ln x, \quad x > 0.$$

Any function $y = ba^x$ can be written in the form $y = be^{kx}$, where $k = \ln a$.

Another logarithm function is $\log_{10} x$, for $x > 0$, where $y = \log_{10} x$ if $10^y = x$ (and vice versa). The properties given above for ln also hold for $\log_{10}$.

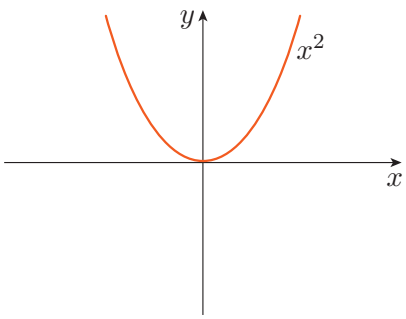
Although they are not used in this module, we note for general reference purposes that the **hyperbolic functions** $\sinh x$, $\cosh x$ and $\tanh x$ are defined as combinations of exponential functions:

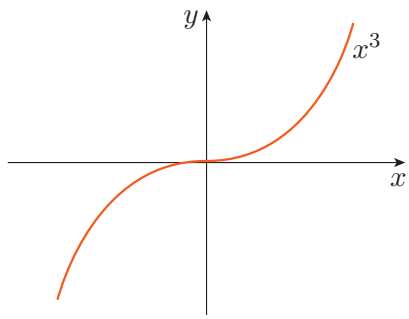
$$\sinh x = \tfrac{1}{2}(e^x - e^{-x}),$$
$$\cosh x = \tfrac{1}{2}(e^x + e^{-x}),$$
$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

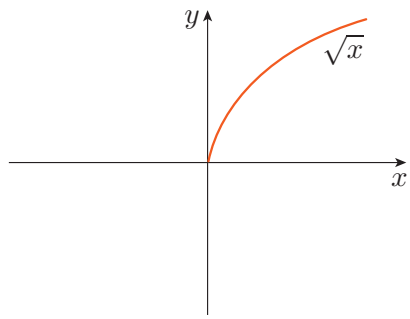
The inverses of these functions are

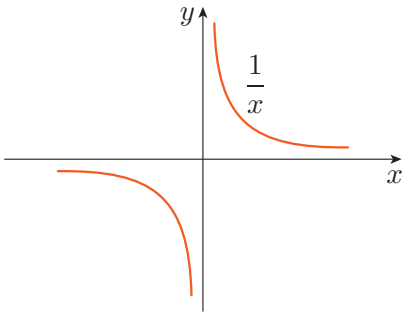
$$\operatorname{arcsinh} x = \ln(x + \sqrt{x^2 + 1}),$$
$$\operatorname{arccosh} x = \ln(x + \sqrt{x^2 - 1}), \quad x \geq 1,$$
$$\operatorname{arctanh} x = \tfrac{1}{2} \ln\left(\frac{1+x}{1-x}\right), \quad |x| < 1.$$

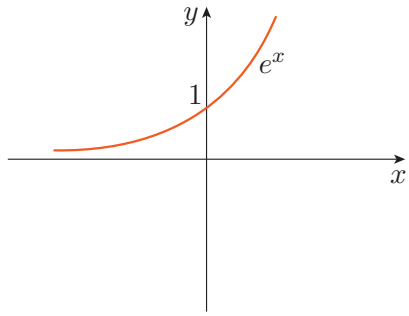
3.4 Graphs of some common functions

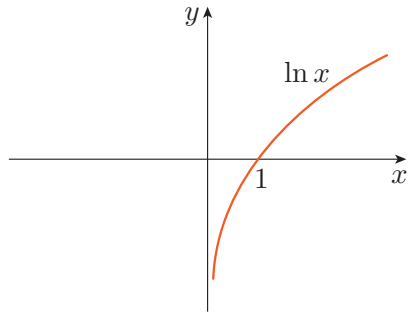


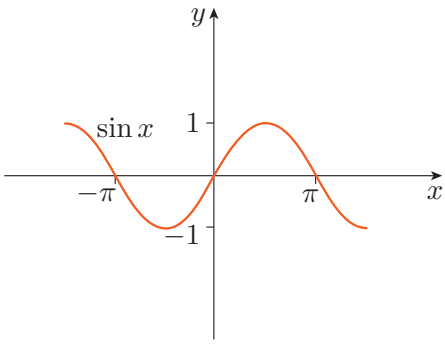


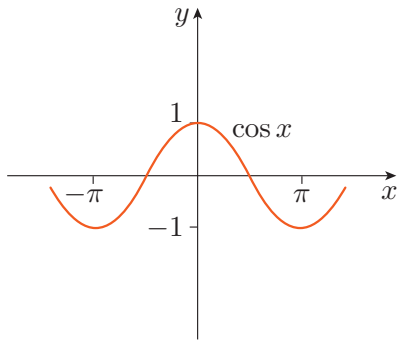


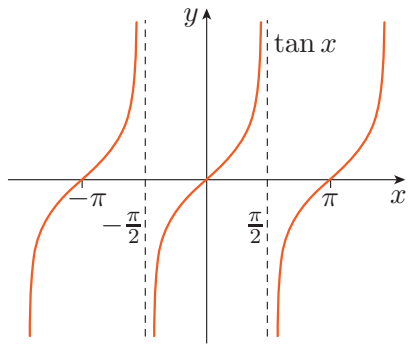


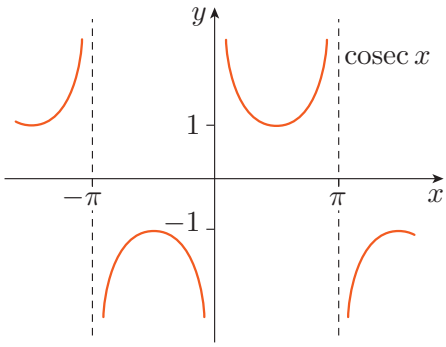


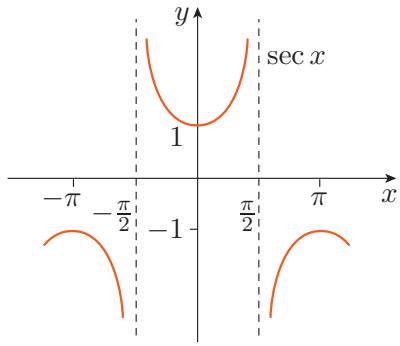


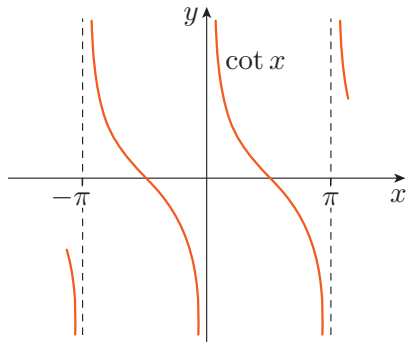


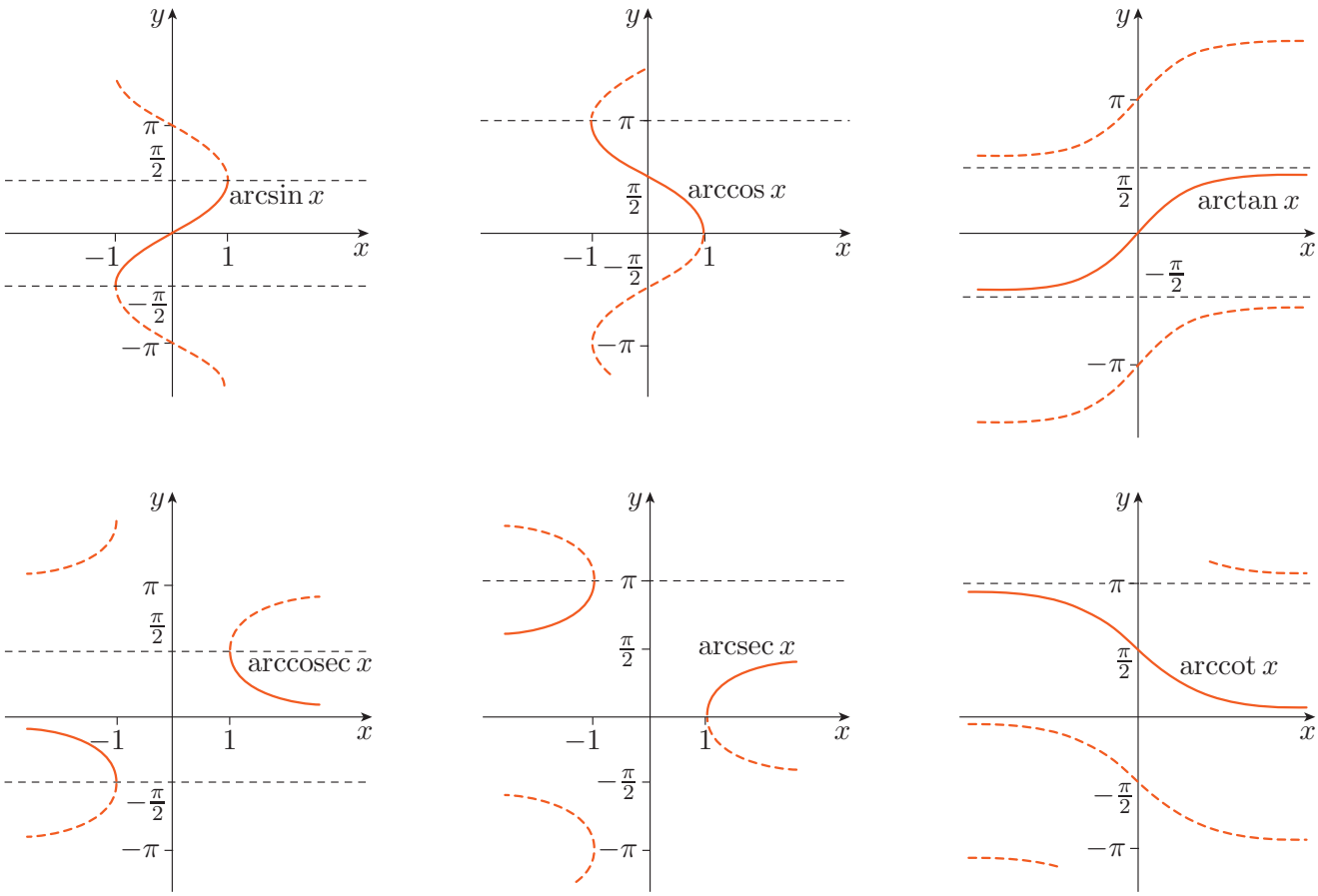












4 Trigonometry

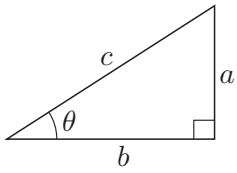
4.1 Radians and degrees

In this module we usually measure angles in **radians** rather than degrees. There are 2π radians in a full circle, corresponding to $360°$, so 1 radian is $(180/\pi)° \simeq 57°$. An advantage of working in radians is the simplicity of the formula for the arc length subtended by an angle in a circle of radius r: the length of the arc subtended by an angle of θ radians is simply $r\theta$. The following radian measures of standard angles are worth knowing:

- a right angle is $\frac{\pi}{2}$ radians
- the angles of an equilateral triangle are $\frac{\pi}{3}$ radians.

An angle is **acute** if its radian measure lies between 0 and $\frac{\pi}{2}$, and **obtuse** if its radian measure lies between $\frac{\pi}{2}$ and π.

4.2 Trigonometric functions and their inverses

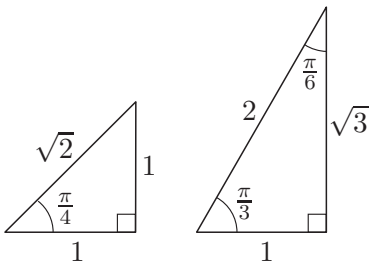


For an acute angle θ, the values of trigonometric functions are related to the ratios of lengths of the sides of a triangle as follows.

Function	Definition for acute angles in terms of triangle shown	Definition in terms of sin and cos
$\sin\theta$	$\frac{a}{c}$	
$\cos\theta$	$\frac{b}{c}$	
$\tan\theta$	$\frac{a}{b}$	$\frac{\sin\theta}{\cos\theta}$
$\cot\theta$	$\frac{b}{a}$	$\frac{\cos\theta}{\sin\theta}$
$\sec\theta$	$\frac{c}{b}$	$\frac{1}{\cos\theta}$
$\operatorname{cosec}\theta$	$\frac{c}{a}$	$\frac{1}{\sin\theta}$

Inverse function	Definition
$\arcsin x$	$= \theta$ where $\sin\theta = x$
$\arccos x$	$= \theta$ where $\cos\theta = x$
$\arctan x$	$= \theta$ where $\tan\theta = x$
$\operatorname{arccot} x$	$= \theta$ where $\cot\theta = x$
$\operatorname{arcsec} x$	$= \theta$ where $\sec\theta = x$
$\operatorname{arccosec} x$	$= \theta$ where $\operatorname{cosec}\theta = x$

4.3 Two useful triangles



From the two triangles in the margin it can be seen that:

$$\sin\tfrac{\pi}{6} = \tfrac{1}{2}, \quad \cos\tfrac{\pi}{6} = \tfrac{\sqrt{3}}{2}, \quad \tan\tfrac{\pi}{6} = \tfrac{1}{\sqrt{3}},$$
$$\operatorname{cosec}\tfrac{\pi}{6} = 2, \quad \sec\tfrac{\pi}{6} = \tfrac{2}{\sqrt{3}}, \quad \cot\tfrac{\pi}{6} = \sqrt{3};$$
$$\sin\tfrac{\pi}{4} = \tfrac{1}{\sqrt{2}}, \quad \cos\tfrac{\pi}{4} = \tfrac{1}{\sqrt{2}}, \quad \tan\tfrac{\pi}{4} = 1,$$
$$\operatorname{cosec}\tfrac{\pi}{4} = \sqrt{2}, \quad \sec\tfrac{\pi}{4} = \sqrt{2}, \quad \cot\tfrac{\pi}{4} = 1;$$
$$\sin\tfrac{\pi}{3} = \tfrac{\sqrt{3}}{2}, \quad \cos\tfrac{\pi}{3} = \tfrac{1}{2}, \quad \tan\tfrac{\pi}{3} = \sqrt{3},$$
$$\operatorname{cosec}\tfrac{\pi}{3} = \tfrac{2}{\sqrt{3}}, \quad \sec\tfrac{\pi}{3} = 2, \quad \cot\tfrac{\pi}{3} = \tfrac{1}{\sqrt{3}}.$$

Other values of the trigonometric functions worth remembering are:

$$\sin 0 = 0, \qquad \cos 0 = 1, \qquad \tan 0 = 0;$$
$$\sin \tfrac{\pi}{2} = 1, \qquad \cos \tfrac{\pi}{2} = 0;$$
$$\sin \pi = 0, \qquad \cos \pi = -1, \qquad \tan \pi = 0;$$
$$\sin \tfrac{3\pi}{2} = -1, \qquad \cos \tfrac{3\pi}{2} = 0.$$

4.4 Trigonometric identities

Pythagoras' theorem states that for any right-angled triangle, if c is the length of the hypotenuse (the side opposite the right angle) and a and b are the lengths of the other two sides, then

$$c^2 = a^2 + b^2.$$

This leads to the following trigonometric identities:

$$\sin^2\theta + \cos^2\theta = 1,$$
$$\tan^2\theta + 1 = \sec^2\theta,$$
$$1 + \cot^2\theta = \operatorname{cosec}^2\theta.$$

Addition formulas

$$\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta,$$
$$\sin(\alpha-\beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta,$$
$$\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta,$$
$$\cos(\alpha-\beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta,$$
$$\tan(\alpha+\beta) = \frac{\tan\alpha + \tan\beta}{1 - \tan\alpha\tan\beta},$$
$$\tan(\alpha-\beta) = \frac{\tan\alpha - \tan\beta}{1 + \tan\alpha\tan\beta};$$

$$\sin\alpha\cos\beta = \tfrac{1}{2}\sin(\alpha+\beta) + \tfrac{1}{2}\sin(\alpha-\beta),$$
$$\cos\alpha\sin\beta = \tfrac{1}{2}\sin(\alpha+\beta) - \tfrac{1}{2}\sin(\alpha-\beta),$$
$$\cos\alpha\cos\beta = \tfrac{1}{2}\cos(\alpha+\beta) + \tfrac{1}{2}\cos(\alpha-\beta),$$
$$\sin\alpha\sin\beta = \tfrac{1}{2}\cos(\alpha-\beta) - \tfrac{1}{2}\cos(\alpha+\beta).$$

In particular, these formulas give

$$\sin(\alpha+2\pi) = \sin\alpha, \quad \cos(\alpha+2\pi) = \cos\alpha, \quad \tan(\alpha+\pi) = \tan\alpha;$$
$$\sin(-\alpha) = -\sin\alpha, \quad \cos(-\alpha) = \cos\alpha, \quad \tan(-\alpha) = -\tan\alpha.$$

Double-angle formulas

$$\sin 2\alpha = 2\sin\alpha\cos\alpha,$$
$$\cos 2\alpha = \cos^2\alpha - \sin^2\alpha = 1 - 2\sin^2\alpha = 2\cos^2\alpha - 1,$$
$$\tan 2\alpha = \frac{2\tan\alpha}{1-\tan^2\alpha},$$
$$\sin^2\alpha = \tfrac{1}{2}(1-\cos 2\alpha),$$
$$\cos^2\alpha = \tfrac{1}{2}(1+\cos 2\alpha).$$

Cosines of related angles

$$\cos(\tfrac{\pi}{2}-\alpha) = \sin\alpha, \qquad \cos(\tfrac{\pi}{2}+\alpha) = -\sin\alpha,$$
$$\cos(\pi-\alpha) = -\cos\alpha, \quad \cos(\pi+\alpha) = -\cos\alpha.$$

4.5 General sinusoidal functions

A **sinusoidal function** or **sinusoid** is a function $x(t)$ of time t of the form

$$x = x_0 + A\cos(\omega t + \phi) = x_0 + A\sin(\omega t + \psi),$$

where x_0 is a constant, A is a positive constant called the **amplitude**, ω is a positive constant called the **angular frequency**, and ϕ and ψ are constants called **phase constants**.

A sinusoidal function oscillates between $x_0 - A$ and $x_0 + A$, repeating the same pattern of oscillations through each time interval of length $2\pi/\omega$, known as the **period** of the function. For these reasons, sinusoidal functions are examples of **oscillatory** functions and of **periodic** functions.

The phase constants in the two forms of the sinusoidal function are related according to $\phi = \frac{3\pi}{2} + \psi$ (or equivalently $\phi = \psi - \frac{\pi}{2}$). Alternative forms of the sinusoidal functions are given by

$$x = x_0 + A\cos(\omega t + \phi) = x_0 + B\cos(\omega t) + C\sin(\omega t),$$
$$x = x_0 + A\sin(\omega t + \psi) = x_0 + D\sin(\omega t) + E\cos(\omega t),$$

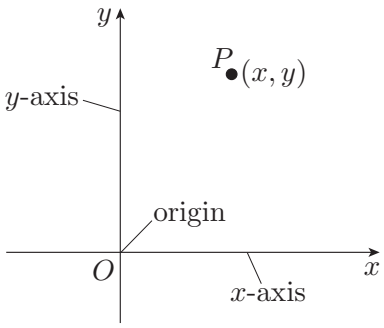
where $B = A\cos\phi$, $C = -A\sin\phi$, $D = A\cos\psi$, $E = A\sin\psi$.

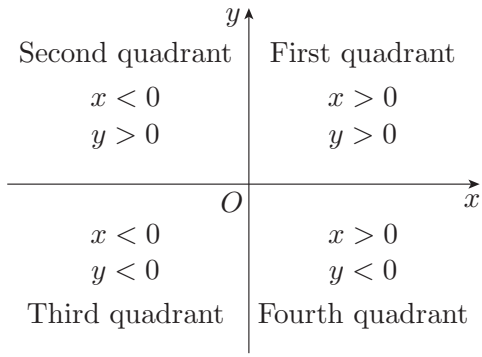
5 Geometry

5.1 Cartesian coordinates

The **Cartesian coordinates** (x, y) of a point P in a plane specify the position of that point relative to two perpendicular axes, the ***x*-axis** (or **horizontal axis**) and ***y*-axis** (or **vertical axis**), which meet at a point O called the **origin**, with Cartesian coordinates $(0, 0)$. The directions of the axes indicate increasing numerical values for the x- and y-coordinates.

Values of x to the right of the y-axis are positive, and those to the left are negative; similarly, values of y above the x-axis are positive, and those below are negative. The four parts into which a plane is divided by Cartesian coordinate axes are known as **quadrants** of the plane. A plane on which Cartesian coordinate axes have been specified is often referred to as the **(x, y)-plane**.



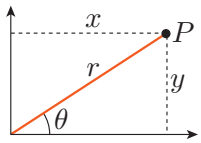


5.2 Polar coordinates

The point P whose polar coordinates are (r, θ) has Cartesian coordinates (x, y) where

$$x = r\cos\theta, \quad y = r\sin\theta.$$

The value of r is always positive (except at the origin, where it is zero). For a given point P, the value of θ is not unique: we can add or subtract any integer multiple of 2π and obtain another value for θ that describes the same point. The value of θ satisfying $-\pi < \theta \leq \pi$ is called the **principal value** of θ.



The relationship between polar and Cartesian coordinates

5.3 Plane figures and curves

A closed plane figure with straight sides is called a **polygon**. A polygon with 3 sides is a **triangle**, one with 4 sides is a **quadrilateral**, one with 5 sides is a **pentagon**, one with 6 sides is a **hexagon**, and in general one with n sides is called an ***n*-gon**.

A polygon is said to be **regular** if all its sides have equal length and all its angles are equal. A regular triangle is referred to as an **equilateral triangle**, and a regular quadrilateral is a **square**.

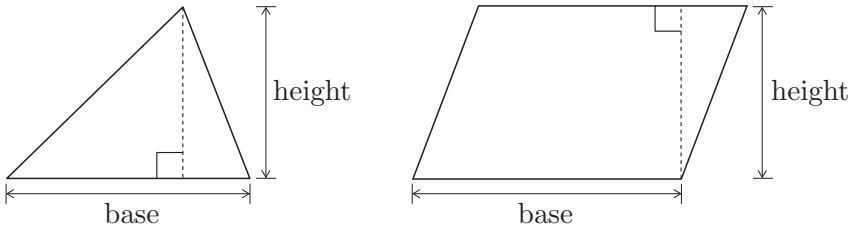
An **isosceles triangle** is one with two sides of equal length (or equivalently with two equal angles). A **right-angled triangle** is one in which one angle is a right angle. The **angle sum of a triangle** is π radians (180°).

A **parallelogram** is a quadrilateral with opposite sides parallel.
A **rectangle** is a parallelogram all of whose angles are right angles.
A **square** is a rectangle all of whose sides have equal length.

The **angle sum of a quadrilateral** is 2π radians (360°).

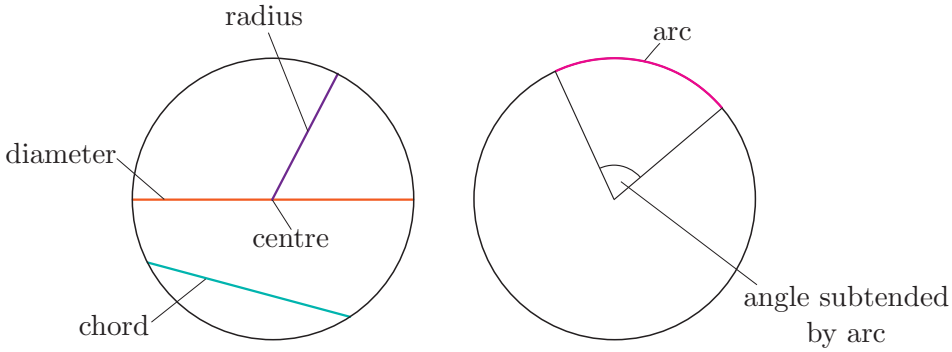
The **area of a triangle** $= \frac{1}{2} \times$ base $\times$ height.

The **area of a parallelogram** $=$ base $\times$ height.



The areas of a triangle and parallelogram

A **circle** is a set of points in a plane that are a constant distance from a fixed point in the plane. The fixed point is the **centre** of the circle, and the constant distance is its **radius**. If a straight line cuts a circle at two points, then the segment of that straight line within the circle is known as a **chord** of the circle. The length of a chord that passes through the centre of a circle is the **diameter** of the circle. The terms *diameter* and *radius* are also used to refer to a chord through the centre of a circle and to a straight line from a point on the circle to its centre, respectively.



Each continuous segment of a circle is known as an **arc** of the circle; the angle made at the centre of a circle by two radii drawn from the ends of the arc is known as the **angle subtended by the arc**. (In such circumstances the arc itself is sometimes called the arc subtended by the angle.)

The distance around a circle is known as its **circumference**, and for a circle of radius r this is given by $2\pi r$. The **area of a circle** of radius r is πr^2.

The **equation of a circle** in the (x, y)-plane with centre (a, b) and radius r is $(x - a)^2 + (y - b)^2 = r^2$.

Other curves in the (x, y)-plane that can be represented by quadratic formulas are the **ellipse**, the **parabola** and the **hyperbola**, examples of which are shown below.

Ellipse Parabola Hyperbola

A straight line between two distinct points on a curve is known as a **chord** of the curve. A straight line that just touches a curve is known as a **tangent** to that curve at the point where it touches.

6 Differentiation

6.1 Notation and terminology

If $f(x)$ is a function, then its **derived function** or **derivative** $f'(x)$ is defined by

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

The process of calculating $f'(x)$ from $f(x)$ is called **differentiation** of $f(x)$ with respect to x. Differentiation with respect to x can also be denoted by the symbol $\dfrac{d}{dx}$ written to the left of the expression or variable being differentiated, so that, for example,

$$\frac{d}{dx}(f(x)) \quad \text{and} \quad \frac{df(x)}{dx} \quad \text{both mean} \quad f'(x).$$

If $y = f(x)$, then $f'(x)$ can also be written as y' or $\dfrac{dy}{dx}$, where to save space we often print dy/dx in place of $\dfrac{dy}{dx}$.

When the independent variable is t (time), we often use a dot to indicate a derivative, so that $\dot{u}$ means the same thing as $u'(t)$ or du/dt.

The notation $f'(x)$ is referred to as **function** notation, dy/dx as **Leibniz** notation, and $\dot{u}$ as **Newton's** or **Newtonian** notation.

The derivative of a derivative is called a **second derivative**. For example, the second derivative of the function $f(x)$, denoted by $f''(x)$, is the derivative of $f'(x)$. If $y = f(x)$, then the second derivative is also written as y'' or d^2y/dx^2. If u is a function of t, then its second derivative can be written as $\ddot{u}$.

Third and **higher derivatives** are defined analogously. The nth derivative of f is denoted by $f^{(n)}$ or, if $y = f(x)$, by $d^n y/dx^n$; n is referred to as the **order** of the derivative. The prime and dot notations are not used for higher derivatives, except that f''' is sometimes used in place of $f^{(3)}$.

A **complex-valued** function $f(x) = g(x) + i\,h(x)$, where g and h are real-valued functions, can be differentiated in a natural way as

$$f'(x) = g'(x) + i\,h'(x).$$

6.2 Rules of differentiation

Constant multiple rule: If k is a constant and u is a function of x, then

$$(ku)' = ku', \quad \text{or equivalently} \quad \frac{d}{dx}(ku) = k\frac{du}{dx}.$$

Sum rule: If u and v are functions of x, then

$$(u+v)' = u' + v', \quad \text{or equivalently} \quad \frac{d}{dx}(u+v) = \frac{du}{dx} + \frac{dv}{dx}.$$

Product and **quotient rules**: If u and v are functions of x, then

$$(uv)' = u'v + uv', \quad \text{or equivalently} \quad \frac{d}{dx}(uv) = \frac{du}{dx}v + u\frac{dv}{dx},$$

and

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}, \quad \text{or equivalently} \quad \frac{d}{dx}\left(\frac{u}{v}\right) = \left(\frac{du}{dx}v - u\frac{dv}{dx}\right)\Big/v^2.$$

This is sometimes called the **'function of a function' rule**.

Composite rule or **chain rule**: If g and u are two functions, and $h(x) = g(u(x))$, then

$$h'(x) = g'(u(x))\,u'(x).$$

Another way to write this is as

$$\frac{dh}{dx} = \frac{dh}{du}\,\frac{du}{dx}.$$

Implicit differentiation: Given an equation connecting two variables x and y, we can use implicit differentiation to calculate dy/dx by differentiating both sides with respect to x and then solving the resulting equation algebraically for dy/dx (instead of solving for y before differentiating).

6.3 Standard derivatives

In each case, a is a constant.

Function	Derivative
a	0
x^a	ax^{a-1}
e^{ax}	ae^{ax}
$\ln(ax)$	$\dfrac{1}{x}$
$\sin(ax)$	$a\cos(ax)$
$\cos(ax)$	$-a\sin(ax)$
$\tan(ax)$	$a\sec^2(ax)$
$\cot(ax)$	$-a\operatorname{cosec}^2(ax)$
$\sec(ax)$	$a\sec(ax)\tan(ax)$
$\operatorname{cosec}(ax)$	$-a\operatorname{cosec}(ax)\cot(ax)$
$\arcsin(ax)$	$\dfrac{a}{\sqrt{1-a^2x^2}}$
$\arccos(ax)$	$-\dfrac{a}{\sqrt{1-a^2x^2}}$
$\arctan(ax)$	$\dfrac{a}{1+a^2x^2}$
$\operatorname{arccot}(ax)$	$-\dfrac{a}{1+a^2x^2}$
$\operatorname{arcsec}(ax)$	$\dfrac{a}{\lvert ax\rvert\sqrt{a^2x^2-1}}$
$\operatorname{arccosec}(ax)$	$-\dfrac{a}{\lvert ax\rvert\sqrt{a^2x^2-1}}$

6.4 Stationary points

The **gradient** of a function f at a point x_0 is the slope of the tangent to the graph of f at that point, and is given by the derivative of f at that point, i.e. $f'(x_0)$. A function is **increasing** on an interval if its gradient is positive throughout that interval; it is **decreasing** if its gradient is negative throughout that interval. A **stationary point** of f is a point x_0 where the gradient is zero, i.e. $f'(x_0) = 0$.

A function is **smooth** if it is continuous and has a continuous derivative.

Consider a stationary point x_0 of a smooth function f.

- x_0 is a **local maximum** if, for all x in the immediate vicinity of x_0, $f'(x) > 0$ if $x < x_0$ and $f'(x) < 0$ if $x > x_0$. An alternative condition is $f''(x_0) < 0$.
- x_0 is a **local minimum** if, for all x in the immediate vicinity of x_0, $f'(x) < 0$ if $x < x_0$ and $f'(x) > 0$ if $x > x_0$. An alternative condition is $f''(x_0) > 0$.
- x_0 is a **point of inflection** if $f''(x_0) = 0$ and $f''(x)$ changes sign as x increases through x_0.

A **global maximum** of a function f is a point x_0 such that $f(x_0) \geq f(x)$ for all x where f is defined. A **global minimum** of a function f is a point x_0 such that $f(x_0) \leq f(x)$ for all x where f is defined. A function f is **bounded above** by an **upper bound** A if $f(x) \leq A$ for all x where f is defined. A function f is **bounded below** by a **lower bound** B if $f(x) \geq B$ for all x where f is defined.

6.5 Curve sketching

The following is a possible procedure for sketching the graph of $y = f(x)$, where $f(x)$ is some given function.

1. Check whether $f(x)$ is a standard function whose graph you already know, or is a simple modification of such a function. If not, proceed to Step 2.
2. Determine how y behaves when x is very large and positive, and when x is very large and negative.
3. Look for any obvious symmetries or repetitions in the behaviour of f.
4. Find where the curve crosses the x- and y-axes, if at all.
5. Look for any values of x at which $f(x)$ is undefined, and examine the behaviour of $f(x)$ near these values of x.
6. Find the locations of any local maxima, local minima or points of inflection.
7. Try to determine whether there are any intervals over which the function is increasing or decreasing.
8. Transfer the information found in Steps 4 and 6 to a sketch graph, then use this information together with any information found in Steps 2, 3, 5 and 7 to try to sketch a smooth curve. If you are still unsure about any parts of the curve, choose suitable values of x and plot the corresponding points $(x, f(x))$ before completing the curve.

Lines $y = c$ where $f(x) \to c$ as $x \to \pm\infty$, and lines $x = c$ where $f(x) \to \pm\infty$ as $x \to c$, where c is a constant, are known as **asymptotes** of the graph of the function $f(x)$.

6.6 Taylor polynomials and series

Factorials

For any positive integer n, we define ***n* factorial**, written $n!$, by

$$n! = 1 \times 2 \times 3 \times \cdots \times (n-1) \times n.$$

The first few factorials are $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$. We also define $0! = 1$.

Taylor polynomials

For a function $f(x)$ with n continuous derivatives near $x = a$, the **Taylor polynomial of degree *n*** about $x = a$ or the ***n*th-order Taylor polynomial** about $x = a$ is

$$p_n(x) = f(a) + (x-a)f'(a) + \frac{1}{2!}(x-a)^2 f''(a) + \cdots + \frac{1}{n!}(x-a)^n f^{(n)}(a).$$

When used to approximate $f(x)$ near $x = a$, we refer to this polynomial as the ***n*th-order Taylor approximation** to $f(x)$ near $x = a$, and write

$$f(x) \simeq f(a) + (x-a)f'(a) + \frac{1}{2!}(x-a)^2 f''(a) + \cdots + \frac{1}{n!}(x-a)^n f^{(n)}(a).$$

In particular, $n = 1$ gives the **tangent approximation**

$$f(x) \simeq f(a) + (x-a)f'(a),$$

and $n = 2$ gives the **quadratic approximation**

$$f(x) \simeq f(a) + (x-a)f'(a) + \tfrac{1}{2}(x-a)^2 f''(a).$$

These approximations are good when x is close to a.

Taylor series

The **Taylor series about** $x = a$ for a function $f(x)$ with infinitely many continuous derivatives near $x = a$ is

$$\begin{aligned} f(x) &= f(a) + (x-a)f'(a) + \frac{1}{2!}(x-a)^2 f''(a) + \cdots \\ &\quad + \frac{1}{n!}(x-a)^n f^{(n)}(a) + \cdots \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}(x-a)^n f^{(n)}(a). \end{aligned}$$

Some standard Taylor series about $x = 0$

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots + (-1)^{n-1}\frac{1}{(2n-1)!}x^{2n-1} + \cdots,$$

$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots + (-1)^{n}\frac{1}{(2n)!}x^{2n} + \cdots,$$

$$e^x = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots,$$

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots + (-1)^{n-1}\frac{1}{n}x^n + \cdots \quad (-1 < x < 1).$$

7 Integration

7.1 Notation and terminology

The **indefinite integral** of a continuous function $f(x)$ is

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

where F is a function such that $F'(x) = f(x)$, known as an **integral** or **antiderivative** of f, and C is a constant, often referred to as an **arbitrary constant** or **constant of integration**.

The **definite integral** of a continuous function $f(x)$ from a to b is

$$\int_a^b f(x)\,dx = \big[F(x)\big]_a^b = F(b) - F(a),$$

where F is any integral of f. The numbers a and b are called the **lower limit of integration** and **upper limit of integration**, respectively. If the areas bounded by the graph of $f(x)$ above and below the x-axis between a and b are A_1 and A_2, respectively, then $\int_a^b f(x)\,dx = A_1 - A_2$.

The process of finding an indefinite or definite integral is known as **integration**, and the function f being integrated is known as the **integrand**. If F_1 and F_2 are two integrals of f, then they differ by a constant, i.e. $F_1(x) = F_2(x) + C$, where C is a constant.

7.2 Rules of integration

Constant multiple rule:

$$\int k\,f(x)\,dx = k\int f(x)\,dx \quad \text{(where } k \text{ is a constant)}.$$

Sum rule:

$$\int \big(f(x)+g(x)\big)\,dx = \int f(x)\,dx + \int g(x)\,dx.$$

Integration by substitution:

$$\int f\big(g(x)\big)\,g'(x)\,dx = \int f(u)\,du \quad \text{(where } u = g(x)\text{)},$$

or in Leibniz notation

$$\int f(u)\,\frac{du}{dx}\,dx = \int f(u)\,du.$$

The following formula, which can be derived by integration by substitution, is also useful:

$$\int \frac{g'(x)}{g(x)}\,dx = \ln|g(x)| + C \quad \text{(where } g(x) \neq 0\text{)}.$$

Integration by parts:

$$\int f(x)\,g'(x)\,dx = f(x)\,g(x) - \int f'(x)\,g(x)\,dx.$$

For definite integrals,

$$\int_a^b f(x)\,g'(x)\,dx = \big[f(x)\,g(x)\big]_a^b - \int_a^b f'(x)\,g(x)\,dx.$$

7.3 Standard integrals

In each case, a is a non-zero constant, b is any constant, and n is any integer. When using the table to obtain indefinite integrals, add an arbitrary constant.

Function	Integral
a	ax
$x^a \ (a \neq -1)$	$\dfrac{x^{a+1}}{a+1}$
$\dfrac{1}{ax+b}$	$\dfrac{1}{a}\ln\lvert ax+b\rvert$
e^{ax}	$\dfrac{1}{a}e^{ax}$
$\ln(ax)$	$x(\ln(ax)-1)$
$\sin(ax)$	$-\dfrac{1}{a}\cos(ax)$
$\cos(ax)$	$\dfrac{1}{a}\sin(ax)$
$\tan(ax)$	$-\dfrac{1}{a}\ln\lvert\cos(ax)\rvert$
$\cot(ax)$	$\dfrac{1}{a}\ln\lvert\sin(ax)\rvert$
$\sec(ax)$	$\dfrac{1}{a}\ln\lvert\sec(ax)+\tan(ax)\rvert$
$\operatorname{cosec}(ax)$	$\dfrac{1}{a}\ln\lvert\operatorname{cosec}(ax)-\cot(ax)\rvert$
$\sec^2(ax)$	$\dfrac{1}{a}\tan(ax)$
$\operatorname{cosec}^2(ax)$	$-\dfrac{1}{a}\cot(ax)$
$\dfrac{1}{x^2+a^2}$	$\dfrac{1}{a}\arctan\left(\dfrac{x}{a}\right)$
$\dfrac{1}{(x-a)(x-b)}$	$\dfrac{1}{a-b}\ln\left\lvert\dfrac{a-x}{x-b}\right\rvert$
$\dfrac{1}{\sqrt{x^2+a^2}}$	$\ln(x+\sqrt{x^2+a^2}\,)$
$\dfrac{1}{\sqrt{x^2-a^2}}$	$\ln\lvert x+\sqrt{x^2-a^2}\,\rvert$
$\dfrac{1}{\sqrt{a^2-x^2}}$	$\arcsin\left(\dfrac{x}{a}\right)$

Unit summaries

Unit 1 First- and second-order differential equations

1. A **differential equation** is an equation that relates an independent variable x, a dependent variable y, and one or more derivatives of y. Its **order** is the order of the highest derivative that appears.
2. A **solution** of a differential equation is a function $y = y(x)$ that satisfies it. A solution that is written in the form '$y =$ function of x' is an **explicit solution**; otherwise it is an **implicit solution**, i.e. an equation of the form $F(x, y) = 0$ for some function F.
3. The **general solution** of a differential equation is the collection of all of its solutions. It is usually possible to give the general solution of a first-order differential equation as a formula containing one **arbitrary constant**. A **particular solution** of a differential equation is a single solution containing no arbitrary constant.
4. An **initial condition** for a first-order differential equation $dy/dx = f(x, y)$ is an assignment of a value y_0 that the dependent variable y must take when the independent variable x takes some given value x_0. An initial condition may be specified in the form '$y = y_0$ when $x = x_0$' or '$y(x_0) = y_0$'; x_0 and y_0 are referred to as **initial values**.

   An **initial-value problem** is to find the particular solution of a differential equation that satisfies a given initial condition. Given the general solution of a differential equation involving an arbitrary constant C, we can determine the solution satisfying an initial condition by substituting the initial values into the general solution; this gives an equation from which the required value of C can, in principle, be found.
5. Some differential equations have an **analytic solution**, i.e. an explicit general solution derived using calculus. To decide whether an equation of the form $dy/dx = f(x, y)$ may be solved analytically by one of the methods described in the module, proceed as follows.
   (a) If $f(x, y)$ is independent of y, so that $f(x, y) = f(x)$, then the equation may be solved by **direct integration**: its general solution is $y = F(x) + C$, where $F(x)$ is an integral of $f(x)$, and C is an arbitrary constant.
   (b) If $f(x, y)$ has the form $f(x, y) = g(x)\,h(y)$, then use the method of separation of variables described below.
   (c) If $f(x, y)$ has the form $f(x, y) = h(x) - g(x)\,y$ (so that the equation is of the form $dy/dx + g(x)\,y = h(x)$), then the equation is linear, and may be solved by the integrating factor method described below.

6. To solve the differential equation

$$\frac{dy}{dx} = g(x)\,h(y), \quad \text{where } h(y) \neq 0,$$

using the method of **separation of variables**, proceed as follows.

(a) Divide both sides of the equation by $h(y)$, and integrate with respect to x, to obtain

$$\int \frac{1}{h(y)}\,dy = \int g(x)\,dx.$$

(b) If possible, perform the two integrations, obtaining an implicit form of the general solution, which should include one arbitrary constant.

(c) If possible, rearrange the formula found in Step (b) to give y in terms of x; this is the explicit general solution of the differential equation.

7. A first-order differential equation is **linear** if it can be expressed in the form

$$\frac{dy}{dx} + g(x)\,y = h(x).$$

It is **homogeneous** if $h(x) = 0$ for all x, **inhomogeneous** otherwise.

If the functions $g(x)$ and $h(x)$ are continuous throughout an interval (a, b) and x_0 belongs to this interval, then the initial-value problem

$$\frac{dy}{dx} + g(x)\,y = h(x), \quad y(x_0) = y_0,$$

has a unique solution on the interval.

This includes the possibilities that either $a = -\infty$ or $b = \infty$ (or both), so the interval might be all of the real line.

8. The **integrating factor method** applies to differential equations of the form

$$\frac{dy}{dx} + g(x)\,y = h(x),$$

and has the following steps.

(a) Determine the **integrating factor**

$$p(x) = \exp\left(\int g(x)\,dx\right).$$

A constant of integration is not needed here.

You can, if you wish, check that you have found p correctly by checking that

$$p(x)\,\frac{dy}{dx} + p(x)\,g(x)\,y = \frac{d}{dx}(p(x)\,y),$$

i.e. by checking that $dp/dx = p(x)\,g(x)$.

(b) Multiply the differential equation by $p(x)$ to recast it as

$$p(x)\,\frac{dy}{dx} + p(x)\,g(x)\,y = p(x)\,h(x).$$

(c) Rewrite the differential equation as

$$\frac{d}{dx}(p(x)\,y) = p(x)\,h(x).$$

(d) Integrate this last equation, to obtain

$$p(x)\,y = \int p(x)\,h(x)\,dx.$$

The integral in Step (d) will involve an arbitrary constant C.

(e) Divide through by $p(x)$, to obtain the general solution in explicit form.

It is a good idea to check, by substitution into the original equation, that the function obtained is indeed a solution.

9. A **direction field** associates a unique direction to each point within a specified region of the (x, y)-plane. The direction corresponding to the point (x, y) may be thought of as the slope of a short line segment through the point.

   In particular, the direction field for the differential equation

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

   associates the direction $f(x, y)$ with the point (x, y).

10. To apply **Euler's method** to the initial-value problem

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0,$$

    proceed as follows.

    (a) Take x_0 and $Y_0 = y_0$ as starting values, choose a step size h, and set $i = 0$.

    (b) Calculate the x-coordinate x_{i+1} using the formula

$$x_{i+1} = x_i + h.$$

    (c) Calculate a corresponding approximation Y_{i+1} to $y(x_{i+1})$ using the formula

$$Y_{i+1} = Y_i + h\,f(x_i, Y_i).$$

    (d) If further approximate values are required, increase i by 1 and return to Step (b).

11. A **linear** second-order differential equation is one that can be written in the form

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

    where $a(x)$, $b(x)$, $c(x)$ and $f(x)$ are given functions. If $f(x)$ is identically zero, then the equation is **homogeneous**; otherwise it is **inhomogeneous**.

12. A **linear constant-coefficient** second-order differential equation is one in which the functions $a(x)$, $b(x)$ and $c(x)$ are constant (with $a \neq 0$), i.e.

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

    and $f(x)$ is a given function.

13. The **principle of superposition** states that if $y_1(x)$ is a solution of $ay'' + by' + cy = f_1(x)$, and $y_2(x)$ is a solution of $ay'' + by' + cy = f_2(x)$, then for any constants k_1 and k_2,

$$y(x) = k_1\, y_1(x) + k_2\, y_2(x)$$

is a solution of the equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = k_1\, f_1(x) + k_2\, f_2(x).$$

14. The **general solution of the homogeneous equation**

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

is obtained by following the procedure below.

(a) Solve for λ the **auxiliary equation**

$$a\lambda^2 + b\lambda + c = 0.$$

(b) (i) If the auxiliary equation has distinct real roots λ_1 and λ_2, then the general solution of the differential equation is

$$y(x) = Ce^{\lambda_1 x} + De^{\lambda_2 x}.$$

(ii) If the auxiliary equation has equal real roots $\lambda_1 = \lambda_2$ (i.e. $a\lambda^2 + b\lambda + c$ is a perfect square), then the general solution of the differential equation is

$$y(x) = (C + Dx)e^{\lambda_1 x}.$$

(iii) If the auxiliary equation has complex conjugate roots $\lambda_1 = \alpha + i\beta$ and $\lambda_2 = \alpha - i\beta$, then the general solution of the differential equation is

$$y(x) = e^{\alpha x}(C\cos\beta x + D\sin\beta x).$$

In each case, C and D are arbitrary constants.

15. The **general solution of the inhomogeneous equation**

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

is given by $y = y_c + y_p$, where:

- y_c, the **complementary function**, is the general solution of the **associated homogeneous equation**

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

- y_p, a **particular integral**, is any particular solution of the original inhomogeneous equation.

This leads to the following procedure for finding the general solution of an inhomogeneous equation.

(a) Find the complementary function, by solving the auxiliary equation of the associated homogeneous equation.

(b) Find a particular integral, as described below.

(c) Add the particular integral to the complementary function.

16. For certain inhomogeneous equations, particular solutions can be found by the **method of undetermined coefficients**. This method works when the function $f(x)$ is a polynomial, an exponential function or a sinusoidal function, i.e. a linear combination of a sine and a cosine.

The method involves using a **trial solution** $y(x)$ of a form similar to that of $f(x)$, but with coefficients that initially are undetermined. The coefficients are determined by substituting the trial solution into the differential equation and choosing the coefficients so that the equation is satisfied.

Suitable trial solutions $y(x)$ for specified right-hand-side functions $f(x)$ are shown in the following table.

$f(x)$	Trial solution $y(x)$
$m_n x^n + m_{n-1}x^{n-1} + \cdots + m_1 x + m_0$	$p_n x^n + p_{n-1}x^{n-1} + \cdots + p_1 x + p_0$
me^{kx}	pe^{kx}
$m\cos kx + n\sin kx$	$p\cos kx + q\sin kx$

The coefficients in the left-hand column, $m, k, m_0, m_1, \ldots$, are given constants. Those in the right-hand column, $p, p_0, p_1, \ldots$, are constants to be determined, by differentiating the expression for $y(x)$ twice, substituting into the left-hand side of the differential equation, and equating coefficients of corresponding terms on the left- and right-hand sides.

Note that even if some of the given coefficients in $f(x)$ are zero – for example $f(x) = 3x$ or $f(x) = 2\cos x$ – it is still necessary in general to use the full expression from the right-hand column of the table.

Exceptionally, if a term in the complementary function is part of the trial solution, then the differential equation may have no solution of the suggested form. In that case, multiply the suggested function by x (one or more times) and try again.

17. An **initial-value problem** for a second-order differential equation is a problem in which one has to find the particular solution $y = y(x)$ of a given equation such that y and its derivative y' take specified values y_0 and z_0, respectively, when the independent variable x takes the value x_0. The numbers x_0, y_0 and z_0 are called **initial values**. The relationships between initial values are called **initial conditions**. These may be specified either as '$y = y_0$ and $y' = z_0$ when $x = x_0$', or as '$y(x_0) = y_0$, $y'(x_0) = z_0$'.

The initial-value problem

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x), \quad y(x_0) = y_0,\ y'(x_0) = z_0,$$

where a, b, c are real constants with $a \neq 0$, and $f(x)$ is a given continuous real-valued function on an interval (r, s), with $x_0 \in (r, s)$, has a unique solution on that interval.

18. In a **boundary-value problem**, a condition is placed on the value of either y or its derivative, or some combination of the two, at each of two different values of x. The conditions are referred to as **boundary conditions**, and values of x and y in these conditions are **boundary values**. Such a problem may have a unique solution, or no solution, or an infinite number of solutions.

Unit 2 Vector algebra and statics

1. A **vector** is a mathematical object consisting of a non-negative real number called its **magnitude**, and a **direction**.

   In printed text, vectors appear in **bold**; in handwritten work they are written underlined.

   - Two vectors are **equal** if and only if they have the same magnitude and direction.
   - The magnitude of a vector $\mathbf{v}$ is denoted by $|\mathbf{v}|$ or sometimes by v.
   - The **zero vector 0** has magnitude zero; no direction is defined for it.
   - The **displacement** from a point P to a point Q is represented by the **displacement vector** $\overrightarrow{PQ}$.

2. In discussions involving vectors, the word **scalar** is used to denote a real number (positive, negative or zero).

3. **Scaling a vector** or **scalar multiplication** of a vector is the process of multiplying a vector by a scalar.

   For a vector $\mathbf{v}$ and (non-zero) scalar m, the **scalar multiple** $m\mathbf{v}$ is the vector whose magnitude is $|m|\,|\mathbf{v}|$, and whose direction is

   - in the same direction as $\mathbf{v}$ if $m > 0$
   - in the opposite direction to $\mathbf{v}$ if $m < 0$.

   Note that $-\mathbf{v} = (-1)\mathbf{v}$ is the vector with the same magnitude as $\mathbf{v}$ but pointing in the opposite direction.

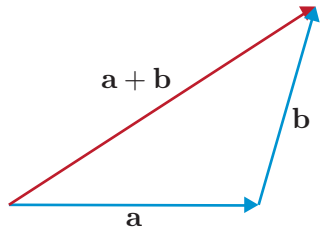
   Also note that $0\mathbf{v} = m\mathbf{0} = \mathbf{0}$ for any vector $\mathbf{v}$ and scalar m.

4. A **unit vector** is a vector whose magnitude is 1. For $\mathbf{v} \neq \mathbf{0}$, the vector $\widehat{\mathbf{v}} = (1/|\mathbf{v}|)\mathbf{v}$ is a unit vector in the same direction as $\mathbf{v}$. Unit vectors along the Cartesian axes are called **Cartesian unit vectors**, and are denoted by $\mathbf{i}$, $\mathbf{j}$ and (in three dimensions) $\mathbf{k}$.

5. **Vector addition** is defined geometrically by the **triangle rule** (see the diagram in the margin).

   The vector $\mathbf{a}+\mathbf{b}$ is called the **sum** or **resultant** of $\mathbf{a}$ and $\mathbf{b}$. **Vector subtraction** is defined by $\mathbf{a}-\mathbf{b}=\mathbf{a}+(-\mathbf{b})$.

   Note that $\mathbf{a}-\mathbf{a}=\mathbf{0}$ and $\mathbf{0}+\mathbf{a}=\mathbf{a}$ for any vector $\mathbf{a}$.



Triangle rule

6. A three-dimensional Cartesian coordinate system consists of three mutually perpendicular axes, usually labelled x, y and z, that meet at a point called the origin.

   A **right-handed coordinate system** is one satisfying the **right-hand grip rule**:

   - Point the straightened fingers of your right hand in the direction of the positive x-axis, and rotate your wrist until you find that you can bend your fingers in the direction of the positive y-axis.
   - Extend the thumb of your right hand. This is the direction of the positive z-axis for a right-handed coordinate system.

7. Given a Cartesian coordinate system, any vector $\mathbf{a}$ can be written uniquely in **component form** as

   $$\mathbf{a}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}, \quad \text{or equivalently} \quad \mathbf{a}=(a_1 \quad a_2 \quad a_3)^T,$$

   where $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are unit vectors in the directions of the positive x-, y- and z-axes, respectively. The numbers a_1, a_2 and a_3 are called the (**Cartesian**) **components** of $\mathbf{a}$ in the directions of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. The process of finding the Cartesian components of a vector is called **resolving a vector into its components**.

   Here and below, we give the formulas for the three-dimensional case; the corresponding two-dimensional versions are obtained by omitting the third component.

   Let $\mathbf{a}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$ and $\mathbf{b}=b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k}$. Then:

   - $|\mathbf{a}|=\sqrt{a_1^2+a_2^2+a_3^2}$
   - $\mathbf{a}+\mathbf{b}=(a_1+b_1)\mathbf{i}+(a_2+b_2)\mathbf{j}+(a_3+b_3)\mathbf{k}$
   - $m\mathbf{a}=(ma_1)\mathbf{i}+(ma_2)\mathbf{j}+(ma_3)\mathbf{k}$.

8. The **position vector** of a point A is the displacement vector $\overrightarrow{OA}$, where O is the origin. So if A has coordinates (x,y,z), then its position vector is $\overrightarrow{OA}=x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$.

   If P and Q are two points with coordinates (p_1,p_2,p_3) and (q_1,q_2,q_3), respectively, then the displacement vector $\overrightarrow{PQ}$ is

   $$\overrightarrow{PQ}=\overrightarrow{OQ}-\overrightarrow{OP}=(q_1-p_1)\mathbf{i}+(q_2-p_2)\mathbf{j}+(q_3-p_3)\mathbf{k}.$$

9. The **vector equation of the straight line** joining points P and Q whose position vectors are $\mathbf{p}$ and $\mathbf{q}$, respectively, is

   $$\mathbf{r}(s)=(1-s)\mathbf{p}+s\mathbf{q}.$$

   If s is allowed to vary over all real numbers, then $\mathbf{r}(s)$ traces out the entire straight line through the points P and Q.

10. The **dot product** (or **scalar product**) of two vectors $\mathbf{a}$ and $\mathbf{b}$ is the scalar defined by

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \cos\theta,$$

where θ (for $0 \leq \theta \leq \pi$) is the angle between the directions of $\mathbf{a}$ and $\mathbf{b}$. In particular, non-zero vectors $\mathbf{a}$ and $\mathbf{b}$ are perpendicular if and only if $\mathbf{a} \cdot \mathbf{b} = 0$. Also, $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$.

- The dot products of the Cartesian unit vectors are

$$\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1, \quad \mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = 0.$$

- In component form, if $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, then

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3.$$

  The formula $|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 + a_3^2$ is a particular case.

- The angle θ between two (non-zero) vectors $\mathbf{a}$ and $\mathbf{b}$ is given by

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} = \frac{a_1b_1 + a_2b_2 + a_3b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2}\sqrt{b_1^2 + b_2^2 + b_3^2}}.$$

11. The component of a vector $\mathbf{a}$ in the direction of a unit vector $\widehat{\mathbf{u}}$ is

$$\mathbf{a} \cdot \widehat{\mathbf{u}} = |\mathbf{a}| \cos\theta,$$

where θ is the angle between $\mathbf{a}$ and $\widehat{\mathbf{u}}$.

The Cartesian components of $\mathbf{a}$ are $\mathbf{a} \cdot \mathbf{i}$, $\mathbf{a} \cdot \mathbf{j}$ and $\mathbf{a} \cdot \mathbf{k}$; in other words,

$$\mathbf{a} = (\mathbf{a} \cdot \mathbf{i})\,\mathbf{i} + (\mathbf{a} \cdot \mathbf{j})\,\mathbf{j} + (\mathbf{a} \cdot \mathbf{k})\,\mathbf{k}.$$

12. The **cross product** (or **vector product**) of two vectors $\mathbf{a}$ and $\mathbf{b}$ is the vector defined by

$$\mathbf{a} \times \mathbf{b} = (|\mathbf{a}|\,|\mathbf{b}| \sin\theta)\,\widehat{\mathbf{c}},$$

where θ (for $0 \leq \theta \leq \pi$) is the angle between the directions of $\mathbf{a}$ and $\mathbf{b}$, and $\widehat{\mathbf{c}}$ is a unit vector at right angles to both $\mathbf{a}$ and $\mathbf{b}$, whose sense is given by the right-hand rule for cross products: this is like the right-hand rule for a coordinate system given above, but with the x-, y-, z-axes replaced by $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, respectively.

The cross product has the following algebraic properties:

- $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$.
- $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) + (\mathbf{a} \times \mathbf{c})$ and $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \times \mathbf{c}) + (\mathbf{b} \times \mathbf{c})$.
- For any scalar λ, $(\lambda\mathbf{a}) \times \mathbf{b} = \lambda(\mathbf{a} \times \mathbf{b}) = \mathbf{a} \times (\lambda\mathbf{b})$.
- $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ if and only if one of $\mathbf{a}$ and $\mathbf{b}$ is a scalar multiple of the other – in other words, one of $\mathbf{a}$ and $\mathbf{b}$ is zero, or they are parallel or antiparallel. (In particular, $\mathbf{a} \times \mathbf{a} = \mathbf{0}$.)
- In general, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$.

- The cross products of the Cartesian unit vectors are

  $$\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0},$$
  $$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{i} = \mathbf{j}.$$

- In component form, the cross product of the vectors $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ is

  $$\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2)\,\mathbf{i} + (a_3b_1 - a_1b_3)\,\mathbf{j} + (a_1b_2 - a_2b_1)\,\mathbf{k}.$$

There is an easy way to remember this formula in terms of determinants: see items 39 and 42 of Unit 4.

13. The **area of a parallelogram** whose sides are defined by vectors $\mathbf{a}$ and $\mathbf{b}$ is $|\mathbf{a} \times \mathbf{b}|$.

    The **area of a triangle** two of whose sides are $\mathbf{a}$ and $\mathbf{b}$ is $\frac{1}{2}|\mathbf{a} \times \mathbf{b}|$.

    The **volume of a parallelepiped** whose sides through one vertex are $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, is $|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$.

14. **Statics** is the analysis of the conditions under which objects remain stationary when subjected to **forces**. **Dynamics** is the study of the motion of objects subject to forces. A force is represented mathematically by a vector. Each force acting on an object can be modelled as acting at a particular point of the object, called the **point of action** of the force. Whenever several forces act at the same point, their overall effect is the same as their **resultant** force, i.e. the single force corresponding to the vector that is the sum of the vectors representing the individual forces.

15. A **particle** is a material object whose size and internal structure may be neglected. It has mass but no size and so occupies a single point in space. Where a particle model is appropriate, all the forces acting on the object are modelled as acting through the point in space occupied by the particle.

16. A **force diagram** is a diagram where a particle is represented by a black dot and each force acting on it is represented by a vector arrow whose tail coincides with the particle and whose direction corresponds to the direction in which the force acts.

17. An object that does not move is said to be in **equilibrium**. If the total force (i.e. the vector sum of all the individual forces) acting on a particle is not zero (i.e. not the zero vector), then the particle will move; so when a particle is in equilibrium, the total force acting on it must be zero. Thus for a particle subjected to forces $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$ to be (and remain) in equilibrium, the forces acting on it must sum to the zero vector:

    $$\mathbf{F}_1 + \mathbf{F}_2 + \cdots + \mathbf{F}_n = \mathbf{0}.$$

    This is the **equilibrium condition for particles**.

18. The force of attraction of objects to the Earth is called the **force of gravity** or the **gravitational force**. The gravitational force acting on a particular object is called the **weight** of the object. It can be assumed that the weight of a particular object is constant near the Earth's surface; it is a *vector* quantity, whose magnitude is measured in newtons (N) in the SI system, and whose direction is downwards towards the centre of the Earth.

An object's weight is to be distinguished from its mass. The **mass** of an object is the amount of matter in the object, and is independent of the object's position in the Universe; it is a *scalar* quantity, measured in kilograms (kg) in the SI system.

Mass and weight are related as follows: an object of mass m has weight $\mathbf{W}$ of magnitude $|\mathbf{W}| = mg$, where g is a constant known as the **magnitude of the acceleration due to gravity**; near the Earth's surface, g has the value $9.81\,\mathrm{m\,s^{-2}}$ (approximately). If the Cartesian unit vector $\mathbf{k}$ points vertically upwards from the surface of the Earth, then $\mathbf{W} = -mg\mathbf{k}$.

19. The force exerted by a surface on an object in contact with it is called the **normal reaction force**, or simply the **normal reaction**, of the surface on the object. This force acts in the direction at right angles to the common tangent at the point of contact between the objects. The magnitude of the normal reaction exerted by a given surface varies according to the mass of the object placed on it.

    It is usually assumed that the magnitude of the normal reaction is potentially unlimited, although actual objects, such as tables on whose surfaces objects rest, are capable of supporting only limited weights without breaking.

20. The force exerted by a taut string, rope or cable on an object attached to it is called a **tension force**. A **model string** is an object that has a length but no area, volume or mass, and that does not stretch. A model string may be described as **light** (it has no mass) and **inextensible** (it does not stretch). The magnitude of the tension force exerted by a given string is referred to as the **tension in the string**.

    It is usually assumed that the magnitude of the tension force is potentially unlimited, although actual strings are capable of supporting only limited weights without breaking. The tension force on a particle attached to a string always acts along the length of the string and away from the point of its attachment to the particle.

21. The force that inhibits the movement of one surface over another, due to roughness, is called a **friction force**, and has the following properties.

    - The friction force $\mathbf{F}$ acts in a direction perpendicular to the normal reaction $\mathbf{N}$ between the surfaces, and opposite to any possible motion along the common tangent to the surfaces.
    - $|\mathbf{F}| \leq \mu|\mathbf{N}|$, where μ is a constant called the **coefficient of static friction** for the two surfaces involved.
    - $|\mathbf{F}| = \mu|\mathbf{N}|$ when the two surfaces are on the verge of slipping. This equality is sometimes referred to as describing a situation of **limiting friction**.
    - If one of the surfaces is designated as being **smooth**, then it may be assumed that there is no friction present when this surface is in contact with another, regardless of the roughness of the other surface.

22. A **model pulley** is an object with no mass and no size, over which a model string may pass. It offers no resistance to rotation. The tension in a model string passing over a model pulley is the same either side of the pulley.

23. The following procedure for **solving statics problems** may be used.

    (a) Draw a sketch of the physical situation, and annotate it with any relevant information. ◀ Draw picture ▶

    (b) Choose coordinate axes, and mark them on your sketch. ◀ Choose axes ▶

    (c) Draw a force diagram or diagrams. ◀ Draw force diagram(s) ▶

    (d) Use the equilibrium condition and any other appropriate law(s) to obtain equation(s). ◀ Apply law(s) ▶

    (e) Solve the equation(s). ◀ Solve equation(s) ▶

    (f) Interpret the solution in terms of the original problem. ◀ Interpret solution ▶

    The equation obtained by using the equilibrium condition is generally a vector equation. To solve this (to find the magnitudes of forces acting, or the directions in which they act), it is often necessary to express the force vectors in component form, i.e. to resolve the vectors into their components with respect to the Cartesian unit vectors in the directions of the axes of the chosen coordinate system. To determine the component of a force in a given direction, find the angle between the force vector and a unit vector in that direction; the resolved component is the magnitude of the force multiplied by the cosine of this angle.

    You can use the following formulas to simplify the cosine of an angle:

    $$\cos\left(\tfrac{\pi}{2} - \alpha\right) = \sin\alpha, \quad \cos\left(\tfrac{\pi}{2} + \alpha\right) = -\sin\alpha,$$
    $$\cos(\pi - \alpha) = -\cos\alpha, \quad \cos(\pi + \alpha) = -\cos\alpha.$$

    Care in the choice of axes often simplifies the solution of a statics problem. For example, for a problem involving a particle on an *inclined plane*, it is often advantageous to use axes parallel and perpendicular to the plane.

24. **Newton's third law** of mechanics states that for each force exerted by one object on another, there is a force of equal magnitude acting in the opposite direction, exerted by the second object on the first.

25. An **extended body** is a material object that has one or more of length, breadth and depth, but whose internal structure may be neglected. An extended body, like a particle, has mass; but unlike a particle, it has a size of some sort and occupies more than a single point in space. An extended body is said to be **rigid** if it does not change its shape when forces act on it. A rigid body in equilibrium can change that state only by slipping or turning.

    When considering the equilibrium of a rigid body – or any extended body – it is important to take into account whereabouts on the body the forces to which it is subject act. The weight of an extended body acts through its **centre of mass**; for a symmetric body made of uniform material, the centre of mass coincides with the centre of symmetry, or geometric centre.

26. A **model rod** is a rigid body that has length but no breadth or depth.

    A **model pivot** has a single point of contact with any object balancing on it.

27. The **line of action** of a force acting on a rigid body is a straight line through the point on the body at which the force acts, in the direction in which the force acts.

    The **torque** $\boldsymbol{\Gamma}$ of a force $\mathbf{F}$ about a fixed point O is given by

    $$\boldsymbol{\Gamma} = \mathbf{r} \times \mathbf{F},$$

    where $\mathbf{r}$ is the position vector, relative to O, of any point on the line of action of the force. In particular, if O lies on the line of action of $\mathbf{F}$, then $\boldsymbol{\Gamma} = \mathbf{0}$. The torque of a force about a point O gives its turning effect about O.

    In SI units, the units of torque are newton metres, written as N m (or $\mathrm{kg\,m^2\,s^{-2}}$).

28. For a rigid body subject to forces $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$ to be (and remain) in equilibrium, the forces must sum to zero, and the torques $\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \ldots, \boldsymbol{\Gamma}_n$ of the forces relative to any fixed point O must also sum to zero:

    $$\sum_{i=1}^{n} \mathbf{F}_i = \mathbf{0} \quad \text{and} \quad \sum_{i=1}^{n} \boldsymbol{\Gamma}_i = \sum_{i=1}^{n} \mathbf{r}_i \times \mathbf{F}_i = \mathbf{0},$$

    where $\mathbf{r}_i$ is the position vector, relative to O, of a point on the line of action of $\mathbf{F}_i$. This is the **equilibrium condition for rigid bodies**.

    A different choice of the point O about which to calculate the torques leads to an equivalent condition. When solving a statics problem for a rigid body, therefore, you may choose the point about which to calculate torques to suit your convenience; a point through which a number of the forces acts is often a good choice, because the torques of those forces will then be zero.

Unit 3 Dynamics

1. The description of the motion of a particle in space begins with the representation of its **position** by a vector $\mathbf{r}(t)$ whose components are functions of the independent variable t (the time); $\mathbf{r}(t)$ is an example of a **vector function**. As t varies, $\mathbf{r}(t)$ traces out the path of the particle as a curve in space. The explicit representation of the path by a vector function is called a **parametrisation** of the path, and the independent variable is sometimes called a **parameter**.

2. The **derivative** of a vector function

    $$\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k},$$

    where $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are (fixed-direction) Cartesian unit vectors, is the vector function

    $$\frac{d\mathbf{r}}{dt} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j} + \frac{dz}{dt}\mathbf{k}.$$

The derivative is often written in mechanics as $\dot{\mathbf{r}}(t)$, and the argument (t) can be omitted when the sense is clear.

3. The **velocity**, or for emphasis velocity vector, of a moving particle with position $\mathbf{r} = \mathbf{r}(t)$ is

$$\mathbf{v} = \mathbf{v}(t) = \frac{d\mathbf{r}}{dt} = \dot{\mathbf{r}}.$$

A particle's **speed** is the magnitude of its velocity, $|\mathbf{v}|$.

4. The **acceleration** (vector) of a particle with velocity $\mathbf{v}$ is defined as

$$\mathbf{a} = \mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2} = \ddot{\mathbf{r}}.$$

Note that acceleration is defined as the rate of *change* of *velocity*, not (as in everyday usage) the rate of increase of speed.

5. An **equation of motion** is any equation relating two or more of acceleration, velocity, position and time. In mechanics problems it is usually required to find a particle's position as a function of time from an equation of motion, which involves solving a differential equation.

6. When a particle moves in such a way that its path is a straight line, its motion is said to be **one-dimensional**. For one-dimensional motion, when the x-axis is taken along the particle's path, the position, velocity and acceleration of the particle can be expressed as

$$\mathbf{r}(t) = x(t)\,\mathbf{i}, \quad \mathbf{v}(t) = \dot{x}(t)\,\mathbf{i}, \quad \mathbf{a}(t) = \ddot{x}(t)\,\mathbf{i},$$

so the position of the particle is specified by the function $x(t)$, the velocity by the function $v(t) = \dot{x}(t)$, and the acceleration by the function $a(t) = \dot{v}(t) = \ddot{x}(t)$. It is often convenient to write x, v and a instead of $x(t)$, $v(t)$ and $a(t)$.

The acceleration in one-dimensional motion can also be expressed as

$$a = \frac{dv}{dt} = \frac{d^2x}{dt^2} = v\,\frac{dv}{dx}.$$

7. If a particle is moving in a straight line along the x-axis with **constant acceleration** $\mathbf{a}(t) = a_0\mathbf{i}$ (where the subscript 0 on a symbol indicates that that symbol represents a constant), and if at time $t = 0$ the particle is at position $x_0\mathbf{i}$ and its velocity is $v_0\mathbf{i}$, then its acceleration, velocity and position at time t are given by

$$a = a_0, \quad v = v_0 + a_0 t, \quad x = x_0 + v_0 t + \tfrac{1}{2}a_0 t^2.$$

Furthermore,

$$v^2 = v_0^2 + 2a_0(x - x_0).$$

These results can be used *only* if the acceleration is constant.

8. The **mass** of a particle is expressed by a single positive number. It is an inherent property of the particle and does not depend on time, position, force or any other variable.

9. The **fundamental laws of Newtonian mechanics**, which concern the motion of a particle, may be stated precisely as follows.

   - If an object modelled as a particle is composed of a number of parts, then the total mass of the particle is the sum of the masses of the parts. This is the **law of addition of mass**.
   - If several forces act simultaneously on a particle, then the resultant or total force is the vector sum of the individual forces. This is the **law of addition of forces**.
   - If a particle of mass m experiences a total force $\mathbf{F}$, then its acceleration $\mathbf{a}$ is given by

     $$\mathbf{F} = m\mathbf{a}.$$

     This is **Newton's second law**.
   - When $\mathbf{F}$ is zero, $\mathbf{a}$ is zero: in the absence of a force, a particle either stays permanently at rest or moves at constant velocity, i.e. at a constant speed in a straight line. This is **Newton's first law**.

   A force of magnitude one **newton** (1 N) is the force required to accelerate a particle whose mass is one kilogram at one metre per second per second (i.e. $1\,\mathrm{N} = 1\,\mathrm{kg\,m\,s^{-2}}$). The statement of Newton's second law given above applies to particles of constant mass, which are the only particles studied in this module.

10. The following procedure for **solving mechanics problems** involving one-dimensional motion may be used.

◀ Draw picture ▶ (a) Draw a sketch of the physical situation, and annotate it with any relevant information.

◀ Choose axes ▶ (b) Choose the x-axis to lie along the direction of motion, and select an origin. Mark the x-axis, its direction and the origin on your sketch.

◀ State assumptions ▶ (c) State any assumptions that you make about the object and the forces acting on it.

◀ Draw force diagram ▶ (d) Draw a force diagram.

◀ Apply Newton's 2nd law ▶ (e) Apply Newton's second law to obtain a vector equation. Resolve each force along the chosen axes in order to resolve the vector equation into a scalar equation.

◀ Solve differential equation ▶ (f) Substitute $v\,dv/dx$, dv/dt or d^2x/dt^2 for the acceleration a in the equation of motion, and solve the resulting differential equation(s) to obtain the velocity v in terms of the position x or time t, or the position x in terms of the time t, as appropriate, making use of any given initial conditions.

◀ Interpret solution ▶ (g) Interpret the solution in terms of the original problem.

11. A model for **sliding friction** is:
    - the friction force $\mathbf{F}$ acts in a direction perpendicular to the normal reaction $\mathbf{N}$ and opposite to the motion
    - $|\mathbf{F}| = \mu'|\mathbf{N}|$, where μ' is the **coefficient of sliding friction** for the two surfaces involved.

    The numerical value of the coefficient of sliding friction μ' is always smaller than the numerical value of the coefficient of static friction μ.
12. **Air resistance** is a force whose direction is opposite to that of the motion of an object, and whose magnitude depends on the object's speed, shape and size as follows:
    - it tends to slow the object down
    - it has little effect at low speeds, but at higher speeds it becomes more noticeable
    - its effects are reduced for an object that presents a smaller profile.

    The **air resistance force R** on a smooth spherical object of diameter D travelling with velocity $\mathbf{v}$ can be modelled as

$$\mathbf{R} = \begin{cases} -c_1 D\mathbf{v} & \text{for } D|\mathbf{v}| \lesssim 10^{-5} \quad \text{(linear model)}, \\ -c_2 D^2|\mathbf{v}|\mathbf{v} & \text{for } 10^{-2} \lesssim D|\mathbf{v}| \lesssim 1 \quad \text{(quadratic model)}, \end{cases}$$

    where $c_1 \simeq 1.7 \times 10^{-4}$ and $c_2 \simeq 0.20$ are the constants for air.

    For motion through **water** we have $c_1 \simeq 9.4 \times 10^{-3}$ for $D|\mathbf{v}| \lesssim 10^{-6}$, and $c_2 \simeq 156$ for $10^{-3} \lesssim D|\mathbf{v}| \lesssim 10^{-1}$.
13. These models for air resistance and water resistance apply only to smooth spherical objects. To adapt these models to other objects, we have to model the objects as smooth spheres. The diameter of the sphere used to model an object is referred to as the **effective diameter** of the object.
14. The **terminal speed** of an object falling under gravity and a resistive force depending on speed (such as air resistance or water resistance) is the constant speed that the object will acquire as time tends to infinity, and is the speed at which the resistive force just balances the object's weight. It can be found by setting the acceleration to zero in the equation of motion of the object.
15. A **projectile** is any object moving in three-dimensional space subject only to the force of gravity. The following terminology is used in the description of the motion of a projectile. While the projectile is off the ground and subject only to the force of gravity, it is said to be **in flight**. The start of the flight is the **launch**, and the projectile's initial velocity is its **launch velocity**. The angle to the horizontal at which the projectile is launched is the **launch angle**. The flight ends with an **impact**. The **time of flight** is the time between the moment of launch and the moment of impact. The path of the projectile while in fight is its **trajectory**. We define the **range** of the projectile to be the horizontal distance between the point of launch and the point of impact.

16. If a projectile is launched from the origin at time $t = 0$ with launch speed u at an upwards angle θ to the horizontal, then the initial conditions are $\mathbf{r}(0) = \mathbf{0}$ and $\dot{\mathbf{r}}(0) = (u\cos\theta)\mathbf{i} + (u\sin\theta)\mathbf{j}$, where $\mathbf{i}$ and $\mathbf{j}$ are unit vectors in the vertical plane containing the launch velocity, with $\mathbf{i}$ horizontal and $\mathbf{j}$ vertically upwards. The solution of the equation of motion is

$$\mathbf{r}(t) = (ut\cos\theta)\mathbf{i} + (ut\sin\theta - \tfrac{1}{2}gt^2)\mathbf{j}.$$

The projectile's **trajectory** is a parabola whose equation is

$$y = x\tan\theta - x^2\frac{g}{2u^2}(1 + \tan^2\theta).$$

Unit 4 Matrices and determinants

1. A **matrix** is a rectangular array of numbers (in this module, usually real numbers), called its elements, components or entries. If a matrix has m rows and n columns, then it is said to be an $m \times n$ matrix, or to be of **order** or **size** $m \times n$. A matrix with one column is often referred to as a **column vector**; a matrix with one row is a **row vector**. A vector in the sense of Unit 2, when expressed in component form, can be represented as a column vector.

2. The general $m \times n$ matrix $\mathbf{A}$ that has a_{ij} as its element in the ith row and the jth column is written as

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

or $\mathbf{A} = (a_{ij})$ for short.

3. Two matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ are **equal** if they have the same order ($m \times n$, say) and $a_{ij} = b_{ij}$ for all $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

4. The $m \times n$ **zero matrix 0** is the $m \times n$ matrix all of whose elements are zero. In particular, a $1 \times n$ zero matrix is called a **zero row vector**, and an $m \times 1$ zero matrix is called a **zero column vector**.

5. A **square matrix** is one that has the same number of rows as columns. The **leading** or **main diagonal** of an $n \times n$ square matrix is its diagonal from top left to bottom right, consisting of the entries a_{ii}, $i = 1, 2, \ldots, n$.

6. An **upper triangular matrix** is a square matrix in which each entry below the leading diagonal is 0. A **lower triangular matrix** is a square matrix in which each entry above the leading diagonal is 0. A matrix that is upper triangular, lower triangular or both (i.e. diagonal) is sometimes referred to simply as a **triangular matrix**. The transpose of an upper triangular matrix is a lower triangular matrix, and vice versa.

7. Consider a **system of linear equations**

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n, \end{aligned}$$

where there is the same number n of equations as of unknowns $x_1, x_2, \ldots, x_n$. The **coefficient matrix** of such a system of equations is the $n \times n$ square matrix $\mathbf{A} = (a_{ij})$ formed from the coefficients of the unknowns in the expressions on the left-hand sides of the equations.

The **augmented matrix** of the system is the $n \times (n+1)$ matrix that has the coefficient matrix for its first n columns and whose final column is the column vector $\mathbf{b}$ consisting of the right-hand sides of the equations. The augmented matrix is usually written $\mathbf{A}|\mathbf{b}$, with a vertical bar separating the coefficient matrix $\mathbf{A}$ from the **right-hand-side vector b**.

8. **Back substitution** is the process of solving a system of linear equations whose coefficient matrix is upper triangular, with none of the entries on the leading diagonal being zero. The final equation of such a system takes the form $a_{nn}x_n = b_n$, with $a_{nn} \neq 0$, so x_n is easily found. The penultimate equation is $a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = b_{n-1}$; since the value of x_n is already known, and $a_{n-1,n-1} \neq 0$, this equation is easily solved for x_{n-1}. We may continue in the same way, working back from the last equation to the first equation, at each stage substituting values already known in order to determine the value of the next unknown.

9. A **linear combination** of row vectors $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_m$ (for example, the rows of an $m \times n$ matrix) is a row vector $\mathbf{R}$ of the form

$$\mathbf{R} = q_1\mathbf{R}_1 + q_2\mathbf{R}_2 + \cdots + q_m\mathbf{R}_m,$$

where $q_1, q_2, \ldots, q_m$ are numbers. A **non-trivial** linear combination is one in which at least one of $q_1, q_2, \ldots, q_m$ is non-zero.

10. **Gaussian elimination** is an efficient and systematic method of obtaining the solution of a system of linear equations. It is based on two principles. The first is that the solution of a system of linear equations whose coefficient matrix is upper triangular, with none of the entries on the leading diagonal zero, can be found simply by back substitution. The second is that if one replaces any row $\mathbf{R}_i$ of the augmented matrix of a system of equations by the linear combination $\mathbf{R}_i + q_k\mathbf{R}_k$, where $\mathbf{R}_k$ is any other row, then the system of equations corresponding to the new matrix has the same solution as the original system.

11. To solve a system of n linear equations in n unknowns, with coefficient matrix $\mathbf{A}$ and right-hand-side vector $\mathbf{b}$, by Gaussian elimination, carry out the following steps (if possible).

    (a) Write down the augmented matrix $\mathbf{A}|\mathbf{b}$, denoting its rows by $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_n$.

    (b) Subtract appropriate scalar multiples of $\mathbf{R}_1$ from $\mathbf{R}_2, \mathbf{R}_3, \ldots, \mathbf{R}_n$ so as to reduce all the elements below the leading diagonal in the first column to zero.

    In the matrix obtained, subtract appropriate scalar multiples of $\mathbf{R}_2$ from $\mathbf{R}_3, \mathbf{R}_4, \ldots, \mathbf{R}_n$ so as to reduce the elements below the leading diagonal in the second column to zero.

    Continue this process until $\mathbf{A}|\mathbf{b}$ is reduced to $\mathbf{U}|\mathbf{c}$, where $\mathbf{U}$ is an upper triangular matrix.

    (c) Solve the system of equations with coefficient matrix $\mathbf{U}$ and right-hand-side vector $\mathbf{c}$ by back substitution.

    The process can be thought of as having two stages: in the first, the elimination stage, the coefficient matrix is reduced to upper triangular form; in the second, the equations are solved by back substitution.

    At the kth stage in the elimination, the kth row of the augmented matrix will take the form

    $$(0 \quad \cdots \quad 0 \quad a_{kk} \quad \cdots \quad a_{kn} \mid b_k),$$

    and any row below it will take the form

    $$(0 \quad \cdots \quad 0 \quad a_{lk} \quad \cdots \quad a_{ln} \mid b_l)$$

    with $l > k$. In order to reduce the coefficient a_{lk} to 0, we subtract a_{lk}/a_{kk} times the kth row from the lth row. (We assume that $a_{kk} \neq 0$.) The number a_{lk}/a_{kk} is a **multiplier**. The number that we divide by in forming the multiplier, namely a_{kk}, is the **pivot**, or **pivot element**, and the row in which it lies is the **pivot row**.

12. **Gaussian elimination with essential row interchanges** is a modification of simple Gaussian elimination designed to overcome the difficulty that arises if a pivot is zero. In elimination with essential row interchanges, whenever one of the pivots is zero, we interchange the pivot row with the first available row below it that would lead to a non-zero pivot, effectively reordering the original system of equations.

13. If Gaussian elimination (with essential row interchanges), applied to the augmented matrix of a system of linear equations, produces an upper triangular coefficient matrix $\mathbf{U}$ in which the final element on the leading diagonal is zero, then the back substitution process breaks down.

    - If the final element of the right-hand-side vector $\mathbf{c}$ obtained by the elimination is non-zero, then the equations have no solution, and are said to be **inconsistent**.

- If the final element of $\mathbf{c}$ is zero, and the remaining elements on the leading diagonal of $\mathbf{U}$ are all non-zero, then the equations have infinitely many solutions: the value of x_n may be chosen freely, but once a choice is made, the value of each of the other unknowns is completely determined.
- If any of the other elements on the leading diagonal of $\mathbf{U}$ is zero, then further investigation is required to discover whether the equations are inconsistent or whether they have infinitely many solutions.
- The equations have a unique solution if and only if none of the elements on the leading diagonal of $\mathbf{U}$ is zero.

14. For systems of three linear equations in three unknowns, the reasons for inconsistency or for the existence of infinitely many solutions can be illustrated geometrically. A linear equation $ax + by + cz = d$ in three unknowns x, y and z, where a, b, c and d are constants, defines a plane in three-dimensional space. (A point (x, y, z) lies on the plane if and only if its coordinates satisfy the equation.) A system of three equations in three unknowns determines three planes.

    If these planes intersect at a single point, then the system of equations has a unique solution, given by the coordinates of the point of intersection.

    But the planes may not have any point in common – they may intersect in pairs in three distinct parallel lines, for example – in which case the equations have no solution and are inconsistent.

    On the other hand, the planes may have many points in common – they may have a common line of intersection, or all of them may coincide – in which case the equations have an infinite number of solutions.

15. The row vectors $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_m$ are **linearly dependent** if there is a non-trivial linear combination of them that is equal to the zero row vector $\mathbf{0}$, i.e. if there are numbers $q_1, q_2, \ldots, q_m$, not all zero, such that

    $$q_1\mathbf{R}_1 + q_2\mathbf{R}_2 + \cdots + q_m\mathbf{R}_m = \mathbf{0}.$$

    Row vectors that are not linearly dependent are **linearly independent**.

    In order to test whether or not a given set of row vectors is linearly dependent, one can form them into a matrix and carry out the Gaussian elimination process (with essential row interchanges) on this matrix. If a row of zeros occurs in the elimination, then the row vectors are linearly dependent; if not, then they are linearly independent. A row of zeros may occur at any stage in the elimination.

16. If the rows of the augmented matrix $\mathbf{A}|\mathbf{b}$ of a system of equations are linearly dependent, then the equations have an infinite number of solutions.

    If the rows of $\mathbf{A}$ are linearly dependent and the rows of $\mathbf{A}|\mathbf{b}$ are linearly independent, then the equations are inconsistent and have no solution.

    If the rows of $\mathbf{A}$ are linearly independent, then the equations have a unique solution.

17. The **sum** $\mathbf{A}+\mathbf{B}$ of two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same order is the matrix of the same order as $\mathbf{A}$ and $\mathbf{B}$ given by $\mathbf{A}+\mathbf{B}=(a_{ij}+b_{ij})$, where $\mathbf{A}=(a_{ij})$ and $\mathbf{B}=(b_{ij})$. Thus the elements of $\mathbf{A}+\mathbf{B}$ are obtained by adding together the elements of $\mathbf{A}$ and $\mathbf{B}$ in the same positions.

18. Matrix addition is **commutative**: for any matrices $\mathbf{A}$ and $\mathbf{B}$ of the same order,

    $$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}.$$

    Matrix addition is **associative**: for any matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ of the same order,

    $$(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C}).$$

    For any matrix $\mathbf{A}$,

    $$\mathbf{A}+\mathbf{0}=\mathbf{A}$$

    (where $\mathbf{0}$ is the zero matrix of the same order as $\mathbf{A}$).

19. The **scalar multiple** of a matrix $\mathbf{A}=(a_{ij})$ by a number k is given by

    $$k\mathbf{A}=(ka_{ij}).$$

    Thus $k\mathbf{A}$ is obtained by multiplying each of the elements of $\mathbf{A}$ by k.

20. The **negative** of the matrix $\mathbf{A}=(a_{ij})$ is $-\mathbf{A}=(-a_{ij})$, so

    $$\mathbf{A}+(-\mathbf{A})=\mathbf{0}.$$

    For two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same order,

    $$\mathbf{A}-\mathbf{B}=\mathbf{A}+(-\mathbf{B}).$$

21. The **product** of an $m\times p$ matrix $\mathbf{A}$ and a $p\times n$ matrix $\mathbf{B}$ is the $m\times n$ matrix $\mathbf{C}=\mathbf{AB}$, where c_{ij} is formed using the ith row of $\mathbf{A}$ and the jth column of $\mathbf{B}$, to give

    $$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{ip}b_{pj}.$$

    The product of two matrices can be formed only if the number of columns of the first matrix is the same as the number of rows of the second.

22. The definitions of matrix multiplication and equality of matrices allow us to represent a system of linear equations as $\mathbf{Ax}=\mathbf{b}$, where $\mathbf{A}$ is the coefficient matrix, $\mathbf{x}$ is the column vector of unknowns, and $\mathbf{b}$ is the right-hand-side column vector.

23. In general, matrix multiplication is not commutative, so $\mathbf{AB}$ may not be equal to $\mathbf{BA}$ (even when $\mathbf{A}$ and $\mathbf{B}$ are square matrices of the same size, so that both products can be formed).

    For any matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, of the appropriate sizes so that all products can be formed, matrix multiplication is **associative**:

    $$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}).$$

24. The **powers** of a square matrix are defined in the obvious way: $\mathbf{A}^2 = \mathbf{AA}$, $\mathbf{A}^3 = \mathbf{AAA}$, and so on.

25. The **transpose** $\mathbf{A}^T$ of a matrix $\mathbf{A}$ is the matrix obtained by interchanging the rows and columns of $\mathbf{A}$. If we denote the matrix $\mathbf{A}$ by (a_{ij}) and $\mathbf{A}^T$ by (a_{ij}^T), then $a_{ij}^T = a_{ji}$. If $\mathbf{A}$ is an $m \times n$ matrix, then $\mathbf{A}^T$ is an $n \times m$ matrix.

26. For any matrix $\mathbf{A}$,

    $$(\mathbf{A}^T)^T = \mathbf{A}.$$

    For any matrices $\mathbf{A}$ and $\mathbf{B}$ of the same size,

    $$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T.$$

    If $\mathbf{A}$ is an $m \times p$ matrix and $\mathbf{B}$ is a $p \times n$ matrix, then

    $$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T;$$

    note the change in order.

27. A square matrix $\mathbf{A}$ is **symmetric** if $\mathbf{A} = \mathbf{A}^T$, i.e. if it is symmetric about its leading diagonal.

28. A **diagonal matrix** is a square matrix in which all the elements off the leading diagonal are 0. A diagonal matrix is symmetric.

29. The $n \times n$ **identity matrix** $\mathbf{I}$ is the $n \times n$ diagonal matrix each of whose diagonal entries is 1. For any $m \times n$ matrix $\mathbf{A}$, $\mathbf{AI} = \mathbf{A}$, and for any $n \times p$ matrix $\mathbf{B}$, $\mathbf{IB} = \mathbf{B}$. In particular, if $\mathbf{A}$ is an $n \times n$ matrix and $\mathbf{I}$ is the $n \times n$ identity matrix, then

    $$\mathbf{IA} = \mathbf{AI} = \mathbf{A}.$$

30. An **inverse** of a square matrix $\mathbf{A}$ is a matrix $\mathbf{B}$ of the same size such that $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, if such a matrix exists. Only a square matrix can have an inverse, but many square matrices do not have inverses; those that do not are called **non-invertible** (or **singular**), while those that do are called **invertible** (or **non-singular**).

31. To find the inverse of an invertible square matrix $\mathbf{A}$, do the following.

    (a) Form the augmented matrix $\mathbf{A}|\mathbf{I}$, where $\mathbf{I}$ is the identity matrix of the same size as $\mathbf{A}$.

    (b) Use row operations on the augmented matrix to reduce the left-hand side to the identity matrix $\mathbf{I}$.

    (c) Then the matrix on the right-hand side is the inverse of $\mathbf{A}$.

32. If $\mathbf{A}$ does have an inverse, then its inverse is unique, and is written $\mathbf{A}^{-1}$. Moreover, $\mathbf{A}^{-1}$ is invertible, and $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$. If $\mathbf{A}$ and $\mathbf{B}$ are invertible matrices of the same size, then $\mathbf{AB}$ is invertible and

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1};$$

note the change in order.

33. If a square matrix is invertible, then its rows are linearly independent, and so are its columns. If the rows (columns) of a square matrix are linearly independent, then so are its columns (rows), and the matrix is invertible.

34. A 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is invertible if and only if $ad - bc \neq 0$. When this condition holds,

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Thus to find the inverse of an invertible 2×2 matrix, interchange the diagonal entries, take the negatives of the other two entries, and divide each resulting entry by the non-zero quantity $ad - bc$. There is no comparable simple formula for the inverse of an invertible $n \times n$ matrix for $n > 2$.

35. The **determinant** of the 2×2 matrix $\mathbf{A}$ above, denoted by $\det \mathbf{A}$ or

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix},$$

is given by $\det \mathbf{A} = ad - bc$. Thus $\mathbf{A}$ is invertible or non-invertible according as $\det \mathbf{A} \neq 0$ or $\det \mathbf{A} = 0$, respectively.

36. A **linear transformation** of the plane is a function that maps each two-dimensional vector $(x \quad y)^T$ to the image vector $(ax + by \quad cx + dy)^T$, where a, b, c and d are real numbers. Such a linear transformation can be represented by the **transformation matrix**

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix};$$

the image of $\mathbf{x} = (x \quad y)^T$ is $\mathbf{Ax}$.

The images of the Cartesian unit vectors $\mathbf{i} = (1 \quad 0)^T$ and $\mathbf{j} = (0 \quad 1)^T$ are the columns of the matrix $\mathbf{A}$. The area of the parallelogram that is the image of the unit square whose sides are $\mathbf{i}$ and $\mathbf{j}$ is $|\det \mathbf{A}|$, the modulus of $\det \mathbf{A}$.

37. The system of linear equations

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

can be written in matrix form as $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A} = (a_{ij})$. Solving the equations is equivalent to finding a vector $\mathbf{x}$ that is mapped to the vector $\mathbf{b}$ by the linear transformation with matrix $\mathbf{A}$. If $\mathbf{A}$ is invertible, then the equations have a unique solution given by $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

This matrix approach to solving a system of linear equations can be used whenever $\mathbf{A}^{-1}$ exists. However, the calculation of inverses is not straightforward except for 2×2 matrices, so it is more efficient to use the Gaussian elimination method to solve the equations.

38. If $\mathbf{A}$ is a square matrix and $\mathbf{Ax} = \mathbf{0}$ for some $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{A}$ is non-invertible (and $\det \mathbf{A} = 0$). Conversely, if $\mathbf{A}$ is invertible, then the only solution of the equation $\mathbf{Ax} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$.

39. The **determinant** of the 3×3 matrix

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

is

$$\begin{aligned} \det \mathbf{A} &= a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \\ &= a_1b_2c_3 - a_1b_3c_2 - a_2b_1c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1. \end{aligned}$$

40. We do not give a formula for the determinant of an $n \times n$ matrix for $n > 3$. However, all determinants satisfy the following rules, which can be used to calculate determinants of any size.

    (a) If $\mathbf{A}$ is a diagonal, upper triangular or lower triangular matrix, then $\det \mathbf{A}$ is the product of the diagonal entries.

    (b) Interchanging any two rows or any two columns of $\mathbf{A}$ changes the sign of $\det \mathbf{A}$.

    (c) $\det(\mathbf{A}^T) = \det \mathbf{A}$.

    (d) Multiplying any row or any column of $\mathbf{A}$ by a scalar k multiplies $\det \mathbf{A}$ by k.

    (e) Adding a multiple of one row of $\mathbf{A}$ to another row does not change $\det \mathbf{A}$.

    (f) $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$.

    To calculate the determinant of a square matrix $\mathbf{A}$, use the Gaussian elimination process to reduce $\mathbf{A}$ to an upper triangular matrix, then apply (a). This works because of (e). However, if it is necessary to make an essential row interchange to avoid a zero pivot, then this will change the sign of the determinant, by (b).

41. If the rows or columns of $\mathbf{A}$ are linearly dependent, then $\det \mathbf{A} = 0$; otherwise $\det \mathbf{A} \neq 0$. The matrix $\mathbf{A}$ is invertible if and only if $\det \mathbf{A} \neq 0$. If $\det \mathbf{A} \neq 0$, then $\det(\mathbf{A}^{-1}) = (\det \mathbf{A})^{-1}$.

42. The cross product $\mathbf{b} \times \mathbf{c}$ of vectors $\mathbf{b} = (b_1 \quad b_2 \quad b_3)^T$ and $\mathbf{c} = (c_1 \quad c_2 \quad c_3)^T$ can be expressed as a 3×3 determinant:

$$\mathbf{b} \times \mathbf{c} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

43. Let $\mathbf{a} = (a_1 \quad a_2 \quad a_3)^T$, $\mathbf{b} = (b_1 \quad b_2 \quad b_3)^T$ and $\mathbf{c} = (c_1 \quad c_2 \quad c_3)^T$; then the **scalar triple product** $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ is given by

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

The **volume of the parallelepiped** with sides defined by the vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ is given by the modulus of the scalar triple product, and therefore by the modulus of the determinant.

44. The **area of the triangle** whose vertices have position vectors $\mathbf{a} = (a_1 \quad a_2 \quad 0)^T$, $\mathbf{b} = (b_1 \quad b_2 \quad 0)^T$ and $\mathbf{c} = (c_1 \quad c_2 \quad 0)^T$ is

$$\tfrac{1}{2} \left| \det \begin{pmatrix} 1 & 1 & 1 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} \right|.$$

45. To determine the **interpolating polynomial** of degree n

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

through $n+1$ data points (x_i, y_i), $i = 0, 1, \ldots, n$, solve the system of equations $\mathbf{X}\mathbf{a} = \mathbf{y}$, where

$$\mathbf{X} = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix},$$

for the coefficients $a_0, a_1, \ldots, a_n$.

46. The **error** in an estimate $\overline{x}$ of the exact value x of a real variable is $\overline{x} - x$. The **absolute error** is $|\overline{x} - x|$.

47. The **norm** $\|\mathbf{x}\|$ of a vector $\mathbf{x} = (x_1 \quad x_2 \quad \cdots \quad x_n)^T$ is the magnitude of the element of largest magnitude in $\mathbf{x}$, i.e.

$$\|\mathbf{x}\| = \max_{i=1,\ldots,n} |x_i|.$$

Suppose that $\overline{\mathbf{x}}$ is an estimate of the exact vector $\mathbf{x}$. The **change** in $\mathbf{x}$ is $\delta\mathbf{x} = \overline{\mathbf{x}} - \mathbf{x}$, and the **absolute change** is $\|\delta\mathbf{x}\| = \|\overline{\mathbf{x}} - \mathbf{x}\|$. The term 'change' is used here rather than 'error' since we anticipate making small changes to the data of a problem to see the effect on the solution.

48. The **absolute condition number** k_{a} for the problem of solving $\mathbf{Ax} = \mathbf{b}$, when $\mathbf{b}$ is subject to small changes of up to ε in size, is the largest possible value of the ratio of $\|\delta\mathbf{x}\|$ to $\|\delta\mathbf{b}\|$, i.e.

$$k_{\mathrm{a}} = \max_{\|\delta\mathbf{b}\| \leq \varepsilon} \frac{\|\delta\mathbf{x}\|}{\|\delta\mathbf{b}\|}.$$

The absolute condition number for small changes to the right-hand-side vector $\mathbf{b}$ in the solution of $\mathbf{Ax} = \mathbf{b}$ is given by the maximum row sum of the magnitudes of the elements of $\mathbf{A}^{-1}$, i.e.

$$k_{\mathrm{a}} = \max_i \{|c_{i1}| + |c_{i2}| + \cdots + |c_{in}|\},$$

where c_{ij} are the elements of the matrix $\mathbf{C} = \mathbf{A}^{-1}$.

49. Suppose that small changes are made in the data for a problem. The problem is **absolutely ill-conditioned** if it is possible for the absolute change in the solution to be significantly larger than the absolute change in the data; otherwise, it is **absolutely well-conditioned**.

The problem of solving $\mathbf{Ax} = \mathbf{b}$ is judged to be:

- absolutely well-conditioned if the absolute condition number k_{a} for the problem is less than 5
- neither absolutely well-conditioned nor absolutely ill-conditioned if k_{a} is greater than 5, but less than 10
- absolutely ill-conditioned if k_{a} is greater than 10.

Unit 5 Eigenvalues and eigenvectors

1. An **eigenvector** of a square matrix $\mathbf{A}$ is a non-zero vector $\mathbf{v}$ such that $\mathbf{Av} = \lambda\mathbf{v}$ for some scalar λ. The number λ is the **eigenvalue** corresponding to the eigenvector $\mathbf{v}$. Any non-zero scalar multiple of an eigenvector is also an eigenvector with the same eigenvalue.

2. An $n \times n$ matrix $\mathbf{A}$ has **characteristic equation**

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0,$$

where $\mathbf{I}$ is the $n \times n$ identity matrix. It is a polynomial of degree n in λ. Its roots are the eigenvalues of $\mathbf{A}$, because the equation $\mathbf{Av} = \lambda\mathbf{v}$ can have a non-zero solution $\mathbf{v}$ if and only if the matrix $\mathbf{A} - \lambda\mathbf{I}$ is non-invertible.

3. In principle, the eigenvalues and eigenvectors of a square matrix $\mathbf{A}$ can be found as follows. Find the roots of the characteristic equation of $\mathbf{A}$ to obtain the eigenvalues. For each eigenvalue λ, solve the **eigenvector equations** $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$ to obtain the corresponding eigenvector. The eigenvector equations must necessarily have infinitely many solutions in which the components of $\mathbf{v}$ are not all zero; to specify the eigenvectors corresponding to λ, it is usually sufficient to choose a suitable representative eigenvector, or a finite number of representative eigenvectors.

4. Vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are **linearly dependent** if some non-trivial linear combination of these vectors gives the zero vector, i.e. if there are numbers $q_1, q_2, \ldots, q_m$, not all zero, such that

$$q_1\mathbf{v}_1 + q_2\mathbf{v}_2 + \cdots + q_m\mathbf{v}_m = \mathbf{0}.$$

Vectors that are not linearly dependent are said to be **linearly independent**.

5. If $\mathbf{v}_i$ is an eigenvector of an $n \times n$ matrix $\mathbf{A}$ with eigenvalue λ_i, for $i = 1, 2, \ldots, m$, where $m \leq n$, and the eigenvalues λ_i are distinct ($\lambda_i \neq \lambda_j$ for $i \neq j$), then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are linearly independent.

6. If the characteristic equation has a **complex root** $\lambda = a + ib$, then the complex conjugate $\overline{\lambda} = a - ib$ is also a root (since the coefficients of the characteristic equation are real). The eigenvectors associated with λ will have complex components. Furthermore, if $\mathbf{v}$ is a complex eigenvector associated with λ, then $\overline{\mathbf{v}}$, the vector whose components are the complex conjugates of those of $\mathbf{v}$, is an eigenvector associated with $\overline{\lambda}$.

   The components of an eigenvector corresponding to a real eigenvalue can be taken to be real.

7. We call an eigenvalue λ_0 a p-fold **repeated eigenvalue** if $(\lambda - \lambda_0)^{p+1}$ is a factor of the characteristic equation. An eigenvalue is called a repeated eigenvalue if it is a p-fold repeated eigenvalue for some $p > 0$. A p-fold repeated eigenvalue, real or complex, may have associated with it any number of linearly independent eigenvectors between 1 and $p + 1$, depending on the nature of $\mathbf{A}$; so if $\mathbf{A}$ has repeated eigenvalues, then it may not be possible to find n linearly independent eigenvectors of $\mathbf{A}$.

8. The eigenvalues of a triangular matrix are the diagonal entries.

9. A matrix is invertible if and only if all of its eigenvalues are non-zero.

10. The **trace** of a square matrix is the sum of the elements on its leading diagonal. The trace of $\mathbf{A}$ is denoted by $\operatorname{tr}\mathbf{A}$.

11. The sum of the eigenvalues of $\mathbf{A}$ is $\operatorname{tr}\mathbf{A}$. The product of the eigenvalues of $\mathbf{A}$ is $\det\mathbf{A}$. (Here each eigenvalue must be counted according to its multiplicity, i.e. a repeated eigenvalue must be included in the sum or the product *each* time it is repeated.)

12. The **characteristic equation** of a 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} \\ &= \lambda^2 - (a + d)\lambda + (ad - bc) = 0. \end{aligned}$$

The eigenvalues of $\mathbf{A}$ are the roots of this quadratic equation for λ.

The **eigenvector equations** for an eigenvalue λ are

$$(a-\lambda)x + by = 0,$$
$$cx + (d-\lambda)y = 0.$$

This pair of equations typically reduces to a single equation of the form $py = qx$, with solution $x = p$, $y = q$, so that $(p \quad q)^T$ is an eigenvector. Any scalar multiple $(kp \quad kq)^T$ with $k \neq 0$ is also an eigenvector corresponding to the same eigenvalue.

13. Consider a 2×2 matrix with real entries.
    - The matrix may have two distinct real eigenvalues, in which case it is possible to find two linearly independent real eigenvectors, one corresponding to each eigenvalue.
    - The matrix may have a complex conjugate pair of eigenvalues, in which case one can find two linearly independent complex eigenvectors by taking an eigenvector corresponding to one of the eigenvalues and the complex conjugate vector.
    - If the matrix has a repeated eigenvalue with two associated linearly independent eigenvectors, then it is a scalar multiple of the 2×2 identity matrix (i.e. it is a diagonal matrix with entries on the diagonal both the same).
    - The only other possibility is that the matrix has a repeated eigenvalue but a single corresponding eigenvector (up to scalar multiples).
14. If $\mathbf{A}$ is an arbitrary matrix and λ is one of its eigenvalues, then:
    - λ^k is an eigenvalue of $\mathbf{A}^k$ for any positive integer k
    - if $\mathbf{A}$ is invertible, then λ^{-1} is an eigenvalue of $\mathbf{A}^{-1}$
    - $\lambda - p$ is an eigenvalue of $\mathbf{A} - p\mathbf{I}$ for any number p
    - $(\lambda - p)^{-1}$ is an eigenvalue of $(\mathbf{A} - p\mathbf{I})^{-1}$ for any number p that is not an eigenvalue of $\mathbf{A}$
    - $p\lambda$ is an eigenvalue of $p\mathbf{A}$ for any number p.

    In each case, an eigenvector of $\mathbf{A}$ is also an eigenvector of the associated matrix.
15. If two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$ satisfy the condition

    $$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n = 0,$$

    then they are said to be **orthogonal** to one another.
16. If $\mathbf{A}$ is a real symmetric $n \times n$ matrix, then:
    - the eigenvalues of $\mathbf{A}$ are real
    - the eigenvectors corresponding to distinct eigenvalues of $\mathbf{A}$ are orthogonal, and those corresponding to repeated eigenvalues are linearly independent and can be chosen to be orthogonal.

17. **Direct iteration** is a numerical method for calculating the eigenvalue of largest magnitude $\lambda_{\max}$ and corresponding eigenvector of a square matrix $\mathbf{A}$.

    Choose any non-zero vector $\mathbf{e}_0$. For $n = 0, 1, 2, \ldots$, do the following.

    (a) Calculate $\mathbf{z}_{n+1} = \mathbf{A}\mathbf{e}_n$.

    (b) Find α_{n+1}, the component of largest magnitude of $\mathbf{z}_{n+1}$.

    (c) Put $\mathbf{e}_{n+1} = \mathbf{z}_{n+1}/\alpha_{n+1}$.

    If this process converges (and it usually will for $\lambda_{\max}$ real and distinct in magnitude), then for sufficiently large n, $\mathbf{e}_n$ will be a good approximation to the eigenvector corresponding to $\lambda_{\max}$, and α_{n+1} will be a good approximation to $\lambda_{\max}$.

18. **Inverse iteration** is a numerical method for calculating the eigenvalue of smallest magnitude $\lambda_{\min}$ and corresponding eigenvector of an invertible square matrix $\mathbf{A}$ (where $\lambda_{\min}$ must be real and distinct in magnitude). (Requiring that $\mathbf{A}$ is invertible is equivalent to requiring that $\lambda_{\min} \neq 0$.)

    Choose any non-zero vector $\mathbf{e}_0$. For $n = 0, 1, 2, \ldots$, do the following.

    (a) Calculate $\mathbf{z}_{n+1} = \mathbf{A}^{-1}\mathbf{e}_n$.

    (b) Find α_{n+1}, the component of largest magnitude of $\mathbf{z}_{n+1}$.

    (c) Put $\mathbf{e}_{n+1} = \mathbf{z}_{n+1}/\alpha_{n+1}$.

    Generally, for sufficiently large n, $\mathbf{e}_n$ will be a good approximation to the eigenvector corresponding to $\lambda_{\min}$, and α_{n+1} will be a good approximation to $1/\lambda_{\min}$.

    Calculating $\mathbf{A}^{-1}$ is difficult for large matrices, and in such cases it is better to substitute the following for Step (a):

    (a) Calculate $\mathbf{z}_{n+1}$ by solving the equation $\mathbf{A}\mathbf{z}_{n+1} = \mathbf{e}_n$.

19. **Modified inverse iteration** is a numerical method for calculating an eigenvalue and corresponding eigenvector of a suitable square matrix. Let p be a real number, and let $\mathbf{A}$ be a square matrix for which there is a single non-zero real eigenvalue λ_1 such that $|\lambda_1 - p| \neq 0$ is least.

    Choose any non-zero vector $\mathbf{e}_0$. For $n = 0, 1, 2, \ldots$, do the following.

    (a) Calculate $\mathbf{z}_{n+1} = (\mathbf{A} - p\mathbf{I})^{-1}\mathbf{e}_n$.

    (b) Find α_{n+1}, the component of largest magnitude of $\mathbf{z}_{n+1}$.

    (c) Put $\mathbf{e}_{n+1} = \mathbf{z}_{n+1}/\alpha_{n+1}$.

    Generally, for sufficiently large n, $\mathbf{e}_n$ will be a good approximation to the eigenvector corresponding to λ_1, and α_{n+1} will be a good approximation to $1/(\lambda_1 - p)$.

    Calculating $(\mathbf{A} - p\mathbf{I})^{-1}$ is difficult for large matrices, and in such cases it is better to substitute the following for Step (a):

    (a) Calculate $\mathbf{z}_{n+1}$ by solving $(\mathbf{A} - p\mathbf{I})\mathbf{z}_{n+1} = \mathbf{e}_n$.

20. If the iterative process for direct, inverse or modified inverse iteration fails to converge, it may be because of an unsuitable choice of $\mathbf{e}_0$, and a different choice may rectify matters. It may, however, be because the chosen eigenvalue is complex, repeated or close to another eigenvalue, or it may be because there is another eigenvalue equally distant from (or nearly equally distant from) the required eigenvalue.

Unit 6 Systems of linear differential equations

1. A **system of linear constant-coefficient first-order differential equations** for unknowns $x_1, x_2, \ldots, x_n$ takes the form

$$\begin{aligned}
\dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + h_1(t),\\
\dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + h_2(t),\\
&\vdots\\
\dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + h_n(t),
\end{aligned}$$

where there is the same number n of equations as of unknowns. The independent variable t often represents time. The dependent variables may be written as above; but for systems with $n = 2$ or $n = 3$, x, y and z are often used instead of x_1, x_2 and x_3. The coefficients a_{ij} are constants.

The equations may be expressed in matrix form as

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{h},$$

where

$$\mathbf{x} = (x_1 \quad x_2 \quad \cdots \quad x_n)^T$$

and

$$\dot{\mathbf{x}} = (\dot{x}_1 \quad \dot{x}_2 \quad \cdots \quad \dot{x}_n)^T;$$

here $\mathbf{A} = (a_{ij})$ is a constant matrix, but the vector $\mathbf{h} = (h_1 \quad h_2 \quad \cdots \quad h_n)^T$ depends on t.

2. The **general solution** of a system of linear constant-coefficient first-order differential equations is a collection of all possible solutions of the system of equations. The general solution of a first-order system can usually be expressed as a solution containing the same number of arbitrary constants as there are unknowns.

3. A **particular solution** of a system of linear constant-coefficient first-order differential equations is a solution containing no arbitrary constants and satisfying given conditions.

4. The **principle of superposition** for first-order systems of differential equations states that if $\mathbf{x}_1$ is a solution of the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{h}_1$ and $\mathbf{x}_2$ is a solution of the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{h}_2$, then for any constants p and q, $p\mathbf{x}_1 + q\mathbf{x}_2$ is a solution of the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + p\mathbf{h}_1 + q\mathbf{h}_2$.

5. A **homogeneous** system is a system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{h}$ with $\mathbf{h} = \mathbf{0}$, or in other words, a system of the form $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. An **inhomogeneous** system is one that is not homogeneous.

6. If λ is a real eigenvalue of the matrix $\mathbf{A}$, and $\mathbf{v}$ is a corresponding eigenvector, then

$$\mathbf{x}(t) = \mathbf{v}e^{\lambda t}$$

is a solution of the system of differential equations $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. (Note that in the expression for $\mathbf{x}(t)$, $\mathbf{v}$ is a constant vector.)

7. If λ is a repeated eigenvalue of the matrix $\mathbf{A}$ with a linearly independent pair of corresponding eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$, then

$$\mathbf{x}(t) = \mathbf{v}_1 e^{\lambda t} \quad \text{and} \quad \mathbf{x}(t) = \mathbf{v}_2 e^{\lambda t}$$

are linearly independent solutions of the system of differential equations $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$.

8. If λ is a repeated eigenvalue of the matrix $\mathbf{A}$ with a corresponding eigenvector $\mathbf{v}$ that is unique (up to scalar multiples), then there is a vector $\mathbf{b}$ such that $(\mathbf{A} - \lambda\mathbf{I})\mathbf{b} = \mathbf{v}$, and

$$\mathbf{x}(t) = (\mathbf{v}t + \mathbf{b})e^{\lambda t}$$

is a solution of the system of differential equations $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, in addition to $\mathbf{x}(t) = \mathbf{v}e^{\lambda t}$.

9. If λ is a complex eigenvalue of the matrix $\mathbf{A}$ with a corresponding complex eigenvector $\mathbf{v}$, then

$$\mathbf{x}(t) = \operatorname{Re}\bigl(\mathbf{v}e^{\lambda t}\bigr) \quad \text{and} \quad \mathbf{x}(t) = \operatorname{Im}\bigl(\mathbf{v}e^{\lambda t}\bigr)$$

are (real) solutions of the system of differential equations $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$.

10. If $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions of the homogeneous system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, then by the principle of superposition so is $C_1\mathbf{x}_1 + C_2\mathbf{x}_2$ for any constants C_1 and C_2.

11. The **general solution of a homogeneous system** of differential equations $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ can be found, in favourable circumstances, by first finding the eigenvalues and a set of corresponding eigenvectors of $\mathbf{A}$, and then writing the solution as a sum of terms $\mathbf{x}_1 + \mathbf{x}_2 + \cdots$ where there is a summand of the form:

- $C\mathbf{v}e^{\lambda t}$ for every non-repeated real eigenvalue λ, where $\mathbf{v}$ is a corresponding eigenvector and C is an arbitrary constant
- $(C\mathbf{v}_1 + D\mathbf{v}_2)e^{\lambda t}$ for every once-repeated real eigenvalue λ, where $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent eigenvectors corresponding to λ, and C and D are arbitrary constants
- $C\mathbf{v}e^{\lambda t} + D(\mathbf{v}t + \mathbf{b})e^{\lambda t}$ for every once-repeated real eigenvalue λ with a corresponding eigenvector $\mathbf{v}$ that is unique (up to scalar multiples), where $\mathbf{b}$ satisfies $(\mathbf{A} - \lambda\mathbf{I})\mathbf{b} = \mathbf{v}$, and C and D are arbitrary constants
- $C\operatorname{Re}\bigl(\mathbf{v}e^{\lambda t}\bigr) + D\operatorname{Im}\bigl(\mathbf{v}e^{\lambda t}\bigr)$ for every non-repeated complex eigenvalue λ with a corresponding complex eigenvector $\mathbf{v}$, where C and D are arbitrary constants.

If, when all eigenvalues have been dealt with, the resulting expression contains sufficiently many arbitrary constants, it will be the general solution. In particular, if an $n \times n$ matrix $\mathbf{A}$ has n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, with $\mathbf{v}_i$ corresponding to an eigenvalue λ_i, $i = 1, 2, \ldots, n$, then the general solution of $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ is

$$\mathbf{x} = C_1\mathbf{v}_1 e^{\lambda_1 t} + C_2\mathbf{v}_2 e^{\lambda_2 t} + \cdots + C_n\mathbf{v}_n e^{\lambda_n t},$$

where $C_1, C_2, \ldots, C_n$ are arbitrary constants. This formula will apply when $\mathbf{A}$ has n distinct eigenvalues, real or complex, and may apply when some of the eigenvalues of $\mathbf{A}$ are repeated. Complex eigenvalues will appear in complex conjugate pairs. Replacing each complex conjugate pair of terms $\mathbf{v}e^{\lambda t}$ and $\overline{\mathbf{v}}e^{\overline{\lambda} t}$ with $\mathrm{Re}\big(\mathbf{v}e^{\lambda t}\big)$ and $\mathrm{Im}\big(\mathbf{v}e^{\lambda t}\big)$ gives the general real solution (restricting the constants to be real also).

The procedure will fail to give the general solution if the matrix $\mathbf{A}$ has a repeated real eigenvalue that is repeated more than once, or if it has a repeated complex eigenvalue.

12. The **complementary function** of an inhomogeneous system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{h}$ is the general solution of the **corresponding homogeneous system** $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. A **particular integral** of the inhomogeneous system is any solution of it.

13. The principle of superposition, applied to an inhomogeneous system, states that the general solution of the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{h}$ takes the form $\mathbf{x}_\mathrm{c} + \mathbf{x}_\mathrm{p}$, where $\mathbf{x}_\mathrm{c}$ is the complementary function and $\mathbf{x}_\mathrm{p}$ is a particular integral.

14. To find a **particular integral** $\mathbf{x}_\mathrm{p} = (x_\mathrm{p} \quad y_\mathrm{p})^T$ of $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{h}$, where $\mathbf{A}$ is a 2×2 matrix, try a solution constructed as follows.

    - When the elements of $\mathbf{h}$ are polynomials of degree less than or equal to k, choose x_p and y_p to be polynomials of degree k.
    - When the elements of $\mathbf{h}$ are multiples of the same exponential function, choose x_p and y_p to be multiples of this exponential function.

    In order for a trial solution to satisfy the system of differential equations, the undetermined coefficients in the trial solution must satisfy a system of linear algebraic equations, obtained by substituting the trial solution into the system of differential equations and equating coefficients of corresponding terms. The coefficients are obtained by solving these algebraic equations (if possible).

15. The trial solution method applied to $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{h}$ may fail because (part of) $\mathbf{h}$ may be a solution of the homogeneous system, which will lead to the algebraic equations for the undetermined coefficients being inconsistent. In this case it may help to multiply the suggested trial solution by t and try again.

For example, suppose that the 2×2 matrix $\mathbf{A}$ has distinct real eigenvalues λ_1 and λ_2 with corresponding linearly independent eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$. Further suppose that $\mathbf{h} = \mathbf{k}e^{\lambda_2 t}$, so that $\mathbf{h}$ is a solution of the homogeneous system. To find a particular integral, find p and q such that $\mathbf{k} = p\mathbf{v}_1 + q\mathbf{v}_2$. Then the particular integral is

$$\left(\frac{p\mathbf{v}_1}{\lambda_2 - \lambda_1} + q\mathbf{v}_2 t\right) e^{\lambda_2 t}.$$

16. A **system of homogeneous linear constant-coefficient second-order differential equations** for unknowns $x_1, x_2, \ldots, x_n$ takes the form

$$\begin{aligned}
\ddot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\
\ddot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\
&\vdots \\
\ddot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n,
\end{aligned}$$

where there is the same number n of equations as of unknowns.

The equations may be expressed in matrix form as

$$\ddot{\mathbf{x}} = \mathbf{A}\mathbf{x},$$

where $\mathbf{A}$ is a constant matrix.

17. If μ^2 is an eigenvalue of the matrix $\mathbf{A}$, and $\mathbf{v}$ is a corresponding eigenvector, then

$$\mathbf{x}(t) = \mathbf{v}e^{\mu t}$$

is a solution of the system of differential equations $\ddot{\mathbf{x}} = \mathbf{A}\mathbf{x}$.

18. Consider a system $\ddot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ where $\mathbf{A}$ is an $n \times n$ matrix with n distinct real eigenvalues. Then:

- each positive eigenvalue μ^2 of $\mathbf{A}$ gives rise to two linearly independent solutions of the system,

  $$\mathbf{v}e^{\mu t} \quad \text{and} \quad \mathbf{v}e^{-\mu t},$$

  where $\mathbf{v}$ is any eigenvector corresponding to μ^2
- each negative eigenvalue $-\omega^2$ of $\mathbf{A}$ gives rise to two linearly independent solutions of the system,

  $$\mathbf{v}\cos\omega t \quad \text{and} \quad \mathbf{v}\sin\omega t,$$

  where $\mathbf{v}$ is any eigenvector corresponding to $-\omega^2$
- a zero eigenvalue gives rise to two linearly independent solutions of the system,

  $$\mathbf{v} \quad \text{and} \quad \mathbf{v}t,$$

  where $\mathbf{v}$ is any eigenvector corresponding to the eigenvalue 0.

The **general solution** of the system of differential equations is a linear combination of the $2n$ linearly independent solutions given above. It contains $2n$ arbitrary constants.

Complex eigenvalues of $\mathbf{A}$ are not discussed here, but they can be dealt with in a fashion similar to that used for the first-order case.

Unit 7 Functions of several variables

1. A **function of two variables** is a function f whose domain is $\mathbb{R}^2$ (or a subset of $\mathbb{R}^2$) and whose codomain is $\mathbb{R}$. Thus for each point (x, y) in the domain of f, there is a unique value z defined by

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

   A **function of n variables** is a function whose domain is (a subset of) $\mathbb{R}^n$ and whose codomain is $\mathbb{R}$.

2. The **surface** with equation $z = f(x, y)$ is the set of all points with coordinates $(x, y, z) = (x, y, f(x, y))$, as (x, y) ranges over $\mathbb{R}^2$ (or that part of it on which $f(x, y)$ is defined), plotted in a three-dimensional Cartesian coordinate system. For example:

   - the surface corresponding to any linear function of two variables $z = f(x, y) = Ax + By + C$ (where A, B and C are constants) is a **plane**
   - the surface corresponding to $f(x, y) = x^2 + y^2$ is a **paraboloid** (which can be obtained by plotting the parabola $z = x^2$ in the (x, z)-plane and then rotating it about the z-axis)
   - the surface corresponding to $f(x, y) = y^2 - x^2$ is a **hyperboloid**.

3. The equation for a plane can be written in vector form as

   $$(\mathbf{r} - \mathbf{a}) \cdot \widehat{\mathbf{n}} = 0,$$

   where $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ is the position vector for any point on the plane, $\mathbf{a} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ with (a, b, c) being some point that the plane passes through, and $\widehat{\mathbf{n}}$ is a unit vector that is normal to the plane.

4. The **section function** with x fixed at a (a constant), of a function of two variables $f(x, y)$, is the function of a single variable y obtained by fixing the value of x to be a; in other words, it is the function $f(a, y)$ (considered as a function of y). Likewise, the section function with y fixed at b of $f(x, y)$ is the function $f(x, b)$ (considered as a function of x).

5. For a function of two variables $f(x, y)$, the curve in the (x, y)-plane defined by the equation $f(x, y) = c$, for some constant c, is called a **contour curve**. The family of contour curves $f(x, y) = c$, plotted on the (x, y)-plane for different values of c chosen from a discrete set of constants $c_1, c_2, \ldots$, is called a **contour map**.

6. Given a function f of two variables x and y, the (first-order) **partial derivative** of f with respect to x, denoted by $\partial f/\partial x$, is obtained by differentiating $f(x, y)$ with respect to x while treating y as a constant. Similarly, the partial derivative of f with respect to y, $\partial f/\partial y$, is obtained by differentiating $f(x, y)$ with respect to y while treating x as a constant.

The partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ represent the slopes of the surface $z = f(x, y)$ at the point $(x, y, f(x, y))$ in the x- and y-directions, respectively, i.e. the slopes of the graphs of the section functions.

We write

$$\frac{\partial f}{\partial x}(a, b)$$

for the value at the point (a, b) of the partial derivative of f with respect to x, and

$$\frac{\partial f}{\partial x}(x, y)$$

when we want to emphasise that the partial derivative is itself a function of two variables (and similarly for other partial derivatives).

The notations f_x and f_y are used as alternatives to $\partial f/\partial x$ and $\partial f/\partial y$, and $f_x(a, b)$ and $f_y(a, b)$ as alternatives to $(\partial f/\partial x)(a, b)$ and $(\partial f/\partial y)(a, b)$, respectively.

7. The partial derivatives of a function $f(x, y)$ with respect to x and y may be defined more formally as

$$\frac{\partial f}{\partial x} = \lim_{\delta x \to 0} \frac{f(x + \delta x, y) - f(x, y)}{\delta x},$$
$$\frac{\partial f}{\partial y} = \lim_{\delta y \to 0} \frac{f(x, y + \delta y) - f(x, y)}{\delta y},$$

where δx and δy denote (small) increments in the values of x and y, respectively.

8. Let $z = f(x, y)$, and let δz be the increment in the value of z (or $f(x, y)$) corresponding to small increments δx and δy in the values of x and y, from a and b to $a + \delta x$ and $b + \delta y$, respectively. Then

$$\delta z \simeq \frac{\partial f}{\partial x}(a, b)\, \delta x + \frac{\partial f}{\partial y}(a, b)\, \delta y,$$

which may be written in the form

$$\delta z \simeq \frac{\partial z}{\partial x}\, \delta x + \frac{\partial z}{\partial y}\, \delta y.$$

This approximation has an important application to error analysis: if the values of a and b are approximate, with errors δx and δy, then the formula above for δz may be used to estimate the possible corresponding error in the calculated value of z.

9. If $z = f(x, y)$, then the rate of change of z with respect to t along a parametrised curve $(x(t), y(t))$ is given by

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

This is a version of the **chain rule**.

10. The **gradient** of the function $f(x, y)$ at the point (a, b) is the vector $f_x(a, b)\,\mathbf{i} + f_y(a, b)\,\mathbf{j}$. It is denoted by $\boldsymbol{\nabla} f(a, b)$ or by $\mathbf{grad}\, f(a, b)$. The **gradient function** is the vector-valued function of two variables given by $\boldsymbol{\nabla} f(x, y) = f_x(x, y)\,\mathbf{i} + f_y(x, y)\,\mathbf{j}$.
11. The **slope** of the surface $z = f(x, y)$ at the point $(a, b, f(a, b))$ in the direction of the unit vector $\widehat{\mathbf{d}} = (\cos\alpha)\mathbf{i} + (\sin\alpha)\mathbf{j}$ in the (x, y)-plane (in other words, in the direction making an anticlockwise angle α with the positive x-axis) is

    $$f_x(a, b)\cos\alpha + f_y(a, b)\sin\alpha = (\boldsymbol{\nabla} f(a, b)) \cdot \widehat{\mathbf{d}}.$$

    This is also the rate of change of $f(x, y)$ in the given direction. The slope of the surface at the point $(a, b, f(a, b))$ (and the rate of change of f) is greatest in the direction determined by the gradient vector $\boldsymbol{\nabla} f(a, b)$. Moreover, $\boldsymbol{\nabla} f(a, b)$ points in a direction normal to the contour curve passing through (a, b).
12. The **second-order partial derivatives**, or second partial derivatives, of a function $f(x, y)$ are

    $$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right),$$
    $$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial^2 f}{\partial y\,\partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

    They are often abbreviated as f_{xx}, f_{yy}, f_{xy} and f_{yx}, respectively.
13. The **mixed derivative theorem** states that for any function $f(x, y)$ of two variables that is sufficiently smooth for the second-order partial derivatives to exist and to be continuous,

    $$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial^2 f}{\partial y\,\partial x}, \quad \text{or equivalently} \quad f_{xy} = f_{yx}.$$

14. The **first-order Taylor polynomial** about (a, b) of the function $f(x, y)$ is

    $$p_1(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

    This is also called the **Taylor polynomial of degree 1** about (a, b).

    The **tangent plane** to the surface $z = f(x, y)$ at $(a, b, f(a, b))$ is given by $z = p_1(x, y)$. As a consequence, the first-order Taylor polynomial is often called the **tangent approximation** to $f(x, y)$ near (a, b).
15. The **second-order Taylor polynomial** about (a, b) of the function $f(x, y)$ is

    $$\begin{aligned} p_2(x, y) = {} & f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) \\ & + \tfrac{1}{2!}(f_{xx}(a, b)(x - a)^2 + 2f_{xy}(a, b)(x - a)(y - b) \\ & \quad + f_{yy}(a, b)(y - b)^2). \end{aligned}$$

    This is also called the **Taylor polynomial of degree 2** about (a, b), or the **quadratic approximation** to $f(x, y)$ near (a, b).

16. A **stationary point** of a function $f(x, y)$ is a point (a, b) at which

$$f_x(a, b) = f_y(a, b) = 0.$$

The corresponding point $(a, b, f(a, b))$ on the surface $z = f(x, y)$ is called a stationary point on the surface. The tangent plane to a surface at a stationary point is horizontal. To find the stationary point(s), it is necessary to solve the simultaneous equations $f_x(x, y) = 0$, $f_y(x, y) = 0$; every solution (x, y) gives the coordinates of a stationary point. The equations will not, in general, be linear.

17. A function $f(x, y)$, defined on a domain D, has a **local minimum** at (a, b) in D if for all (x, y) in D sufficiently close to (a, b), we have $f(x, y) \geq f(a, b)$.

A function $f(x, y)$, defined on a domain D, has a **local maximum** at (a, b) in D if for all (x, y) in D sufficiently close to (a, b), we have $f(x, y) \leq f(a, b)$.

A point that is either a local maximum or a local minimum is an **extremum**.

18. Every extremum of $f(x, y)$ is a stationary point of $f(x, y)$. However, not every stationary point is necessarily an extremum. A stationary point that is not an extremum is a **saddle point**.

19. The **classification of stationary points** may be carried out using the **eigenvalue test for classifying stationary points** as follows.

Suppose that the function $f(x, y)$ has a stationary point at (a, b). Let

$$A = f_{xx}(a, b), \quad B = f_{xy}(a, b), \quad C = f_{yy}(a, b),$$
$$\mathbf{H} = \begin{pmatrix} A & B \\ B & C \end{pmatrix},$$

and let λ_1 and λ_2 be the eigenvalues of $\mathbf{H}$.

(a) If λ_1 and λ_2 are both positive, then there is a local minimum at (a, b).

(b) If λ_1 and λ_2 are both negative, then there is a local maximum at (a, b).

(c) If λ_1 and λ_2 are non-zero and opposite in sign, then there is a saddle point at (a, b).

(d) If either λ_1 or λ_2 is zero, then the test is inconclusive.

20. The **classification of stationary points** may also be carried out using the $\boldsymbol{AC - B^2}$ **test** as follows.

Suppose that the function $f(x, y)$ has a stationary point at (a, b). Let

$$A = f_{xx}(a, b), \quad B = f_{xy}(a, b), \quad C = f_{yy}(a, b).$$

(a) If $AC - B^2 > 0$, then there is:

- a local minimum at (a, b) if $A > 0$
- a local maximum at (a, b) if $A < 0$.

(b) If $AC - B^2 < 0$, then there is a saddle point at (a, b).

(c) If $AC - B^2 = 0$, then the test is unable to classify the stationary point.

21. If f is a function of n variables, then a **stationary point** of f is a point $(a_1, a_2, \ldots, a_n)$ at which all n first partial derivatives are zero. There will be n^2 second partial derivatives, and their values at $(a_1, a_2, \ldots, a_n)$ can be written in the form of an $n \times n$ matrix $\mathbf{H}$, the (i, j)-entry being

$$\frac{\partial^2}{\partial x_i \, \partial x_j} f(a_1, \ldots, a_n).$$

This matrix is called the **Hessian matrix** of f at the point $(a_1, \ldots, a_n)$; provided that f is sufficiently smooth, it will be a symmetric matrix, by the mixed derivative theorem. The eigenvalues of $\mathbf{H}$ will all be real.

The matrix $\mathbf{H}$ may be used to classify the stationary points of f as follows.

- If all the eigenvalues are positive, then there is a local minimum at $(a_1, \ldots, a_n)$.
- If all the eigenvalues are negative, then there is a local maximum at $(a_1, \ldots, a_n)$.
- If all the eigenvalues are non-zero but they are not all of the same sign, then there is a saddle point at $(a_1, \ldots, a_n)$.
- If one or more of the eigenvalues is zero, then the test is inconclusive.

Unit 8 Mathematical modelling

1. The term **mathematical modelling** refers to the processes by which mathematics is used to represent, understand and predict events in the real world. At the core of mathematical modelling is the construction of mathematical models: a **mathematical model** is a mathematical relationship between variables representing the possible states of a real-world system, which simulates the behaviour of the system.

2. The **mathematical modelling process**, which is sometimes called the **mathematical modelling cycle**, takes a problem in the real world and translates it into a mathematical problem whose solution should enable one to predict the behaviour of the real-world situation. The key stages in the mathematical modelling process are as follows.

   1 **Specify the purpose of the model**:
   define the problem;
   decide which aspects of the problem to investigate.

   2 **Create the model**:
   state assumptions;
   choose variables and parameters;
   formulate mathematical relationships.

3 **Do the mathematics**:
solve equations;
draw graphs;
derive results.

4 **Interpret the results**:
collect relevant data;
describe the mathematical solution in words;
decide what results to compare with reality.

5 **Evaluate the model**:
test the model by comparing its predictions with reality;
criticise the model.

3. The **input–output principle** states that

$$\boxed{\text{accumulation}} = \boxed{\text{input}} - \boxed{\text{output}},$$

It is frequently used in modelling, as a way of setting up a model of a dynamic process.

4. Mass, length and time are fundamental properties and are referred to as **base dimensions**. The base dimension mass is denoted by M, length by L, and time by T.

For example, the dimensions of speed are length divided by time; this is written as

$$[\text{speed}] = \mathrm{L\,T^{-1}},$$

where the square brackets are read as 'the dimensions of'.

5. A number, such as the coefficient $\frac{1}{2}$ in the formula $\frac{1}{2}mv^2$ for kinetic energy, has no dimension; it is a **dimensionless quantity**, and we denote this by writing $\left[\frac{1}{2}\right] = 1$.

6. Below are the dimensions of some of the most commonly used physical quantities in this module. For a vector quantity, we give the dimensions of the magnitude of the vector (and any one of its components will also have the same dimensions).

Physical quantity	Dimensions
Acceleration	$\mathrm{L\,T^{-2}}$
Angle	1
Area	$\mathrm{L^2}$
Density	$\mathrm{M\,L^{-3}}$
Energy	$\mathrm{M\,L^2\,T^{-2}}$
Force	$\mathrm{M\,L\,T^{-2}}$
Speed	$\mathrm{L\,T^{-1}}$
Torque	$\mathrm{M\,L^2\,T^{-2}}$
Volume	$\mathrm{L^3}$

The dimensions of a derivative are given by the ratio of the dimensions of the dependent and independent variables. For example, the dimensions of the rate of change of force with area are given by $\mathrm{M\,L\,T^{-2}/L^2 = M\,L^{-1}\,T^{-2}}$.

7. It is often possible to check for errors in the equations proposed for a mathematical model by using the principle of **dimensional consistency**. In any equation occurring in a mathematical model, the dimensions must be the same for each of the additive terms on either side of the equation. If they are not, then a mistake has been made in the derivation of the equation.
8. The **base units** of the SI system that are used in this module are the metre (abbreviated to m) for length, the kilogram (kg) for mass, the second (s) for time, and the kelvin (K) for temperature.

   Additionally, various **derived units** are commonly used. For example, the newton (N) and the joule (J) are defined in terms of the base units by $1\,\mathrm{N} = 1\,\mathrm{kg\,m\,s^{-2}}$ and $1\,\mathrm{J} = 1\,\mathrm{kg\,m^2\,s^{-2}}$.

Unit 9 Oscillations and energy

1. A **model spring** is an object with variable length and zero mass that can exert a force on any object attached to either of its ends. The magnitude of this force is called the **tension in the spring**. A model spring is characterised by two positive constants, its **natural length** l_0 and its **stiffness** k.
2. The **force law for a model spring** is known as **Hooke's law**, which may be expressed in vector form as follows. The force $\mathbf{H}$ exerted by a model spring of natural length l_0 and stiffness k on an object attached to one of its ends is given by

   $$\mathbf{H} = k(l - l_0)\,\widehat{\mathbf{s}},$$

   where l is the (current) length of the spring, and $\widehat{\mathbf{s}}$ is a unit vector in the direction along the axis from the end where the object is attached towards the centre of the spring.
3. To find the force on a particle due to a spring, do the following.
   (a) Determine the length l of the spring; this is best done by drawing a diagram to show the various lengths in the system. Note that l may be in terms of unknown displacements such as x.
   (b) Use the given natural length l_0 of the spring to find the extension $l - l_0$.
   (c) From the diagram, determine $\widehat{\mathbf{s}}$, the unit vector from the particle towards the centre of the spring.
   (d) The spring tension force is $\mathbf{H} = k(l - l_0)\,\widehat{\mathbf{s}}$, where k is the stiffness of the spring.

4. For a particle moving in a straight line subject only to the force exerted on it by a model spring, where the origin is taken at the particle's equilibrium position and the unit vector **i** points in the direction in which the spring extends, the force exerted by the spring on the particle is $-kx\mathbf{i}$, where k is the stiffness of the spring. The particle's **equation of motion** is then

$$m\ddot{x} + kx = 0,$$

or equivalently,

$$\ddot{x} + \omega^2 x = 0,$$

where $\omega^2 = k/m$. This equation is called the **simple harmonic motion equation** (or **SHM** equation). It is the simplest form of equation of motion for a particle subject to spring forces. Problems concerned with the motion of particles attached to springs generally lead to equations of motion closely related to the simple harmonic motion equation.

5. The general solution of the equation $\ddot{x} + \omega^2 x = \omega^2 x_{\text{eq}}$ can be written in the form

$$x(t) = B\cos(\omega t) + C\sin(\omega t) + x_{\text{eq}},$$

or alternatively as

$$x(t) = A\cos(\omega t + \phi) + x_{\text{eq}},$$

where $A > 0$ and $-\pi < \phi \leq \pi$. The pairs of constants B, C and A, ϕ are related as follows:

$$A = \sqrt{B^2 + C^2}, \quad \cos\phi = B/A, \quad \sin\phi = -C/A.$$

The following names are given to quantities that are characteristic of simple harmonic motion:

- the **angular frequency** ω (rad s^{-1}) of the oscillations is the number of cycles in 2π seconds
- the **period** τ (seconds) is the time for one complete cycle, $\tau = 2\pi/\omega$
- the **frequency** f (Hz (hertz), or s^{-1}) is the number of cycles per second, $f = 1/\tau = \omega/(2\pi)$
- the **amplitude** A of the oscillations is the maximum departure from the mean position
- the **phase angle** or **phase** ϕ of the oscillations is such that $A\cos\phi$ is the particle's position at $t = 0$
- the **equilibrium position** about which the particle oscillates is x_{eq}.

6. The **kinetic energy** $T(v)$ of a particle of mass m moving with speed v is given by

$$T(v) = \tfrac{1}{2}mv^2.$$

7. For a particle under the influence of a force $\mathbf{F} = F(x)\,\mathbf{i}$ that is either constant or depends only on the position $\mathbf{r} = x\mathbf{i}$, the **potential energy function** $U(x)$ is

$$U(x) = -\int F(x)\,dx,$$

or equivalently,

$$F(x) = -\frac{dU}{dx}.$$

8. The **total mechanical energy** of a particle is given by the sum of its kinetic and potential energies, i.e.

$$E = \tfrac{1}{2}mv^2 + U(x).$$

If there is more than one force, then the total potential energy $U(x)$ is the sum of the potential energy functions $U_1(x), U_2(x), \ldots, U_n(x)$ due to each of the n forces.

9. The **gravitational potential energy** of a particle of mass m is given by

$$U(x) = mgx,$$

where the height x is measured vertically from some chosen **datum**, and g is the magnitude of the acceleration due to gravity.

10. The **potential energy stored in a model spring** of stiffness k and natural length l_0 is

$$U(l) = \tfrac{1}{2}k(l - l_0)^2,$$

where l is the length of the spring, and the datum is chosen to be at the natural length of the spring.

11. The **law of conservation of mechanical energy** states that if the force on a particle depends only on the particle's position, then the total mechanical energy of the particle is constant. In other words,

$$E = \tfrac{1}{2}mv^2 + U(x)$$

is a constant.

12. The following points about conservation of mechanical energy should be noted.

- The SI unit for energy is the **joule** (J), where $1\,\mathrm{J} = 1\,\mathrm{kg\,m^2\,s^{-2}}$.
- The conservation of mechanical energy for one-dimensional motion is applicable only for forces that depend on position alone (and not on time, velocity or any other variable).
- The law of conservation of mechanical energy for one-dimensional motion says nothing more and nothing less than Newton's second law. It is just an alternative starting point for solving a problem.
- The (constant) value of the total mechanical energy depends on the datum chosen for the potential energy function.

13. A mechanics problem involving the one-dimensional motion of a particle in which the total force depends only on the position of the particle (or is constant), where the question to be answered refers to a relationship between the position and velocity of the particle, can be solved using the following procedure for **applying the law of conservation of mechanical energy**.

◀ Draw picture ▶
(a) Draw a picture.

◀ State assumptions ▶
(b) State any assumptions used, such as ignoring friction or air resistance.

◀ Choose datum ▶
(c) Choose a datum. For gravitational potential energy, it is usual to choose the origin. For a spring, the datum is usually chosen to be the point of zero deformation. (For other forces, this step could be deferred until the potential energy function is known; choose the datum to simplify the expression.)

◀ Find potential energy ▶
(d) If gravity or model springs are present, simply add together the corresponding potential energy functions. Otherwise, identify the x-component $F(x)$ of the total force acting on the particle, then apply the potential energy function definition

$$U(x) = -\int F(x)\,dx,$$

where x is measured in the direction of motion.

◀ Find E ▶
(e) Use the initial conditions to calculate the value of the constant E, the total mechanical energy of the system.

◀ Find x or v ▶
(f) Solve the resulting equation either for x (at a specified value of v) or for v (at a specified value of x).

◀ Interpret solution ▶
(g) Interpret the solution in terms of the original problem.

14. A point where a particle changes its direction of motion (where $v = 0$) is called a **turning point** and is given by solving the equation $E - U(x) = 0$. The **region of motion** of the particle satisfies the inequality $E - U(x) \geq 0$.

15. For a system with potential energy function $U(x)$, **positions of equilibrium** x_{eq} occur whenever $U'(x_{\text{eq}}) = 0$.

    If $U''(x_{\text{eq}}) > 0$, then for small oscillations the motion is simple harmonic with angular frequency $\omega = \sqrt{U''(x_{\text{eq}})/m}$.

    If $U''(x_{\text{eq}}) < 0$, then the equilibrium point is unstable.

Unit 10 Forcing, damping and resonance

1. **Damping** is the frictional effect that causes the vibrations of an oscillating system to die away in the absence of external forces. An oscillating system subjected to an additional time-dependent force is said to be **forced**. **Resonance** is the phenomenon that occurs when a system subjected to periodic forcing produces vibrations of large amplitude at certain frequencies of the forcing.

2. The **linear damping model** assumes that the resistance force $\mathbf{R}$ acting on a particle is proportional to its velocity, but in the opposite direction. If the velocity is $\dot{x}\mathbf{i}$, then the resistance force is modelled by

   $$\mathbf{R} = -r\dot{x}\mathbf{i},$$

   where r is a positive constant called the **damping constant**.

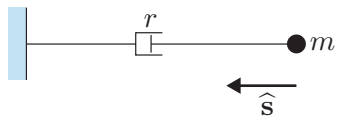
3. If the damping constant is sufficiently large in relation to the other parameters in the system, then there will be no oscillations. Such damping is **strong damping**. Damping where there are decaying oscillations is **weak damping**. For a system in which the level of damping can be varied continuously, there will be a crossover point between weak and strong damping, and the corresponding level of damping is **critical damping**.

4. A **model damper**, with a particle attached to one of its ends, provides a force on the particle that opposes its motion relative to the other end. The magnitude of this resistance force is proportional to the rate of change of the length of the model damper. The force provided by the model damper (when compressing or extending) is therefore

   $$\mathbf{R} = r\dot{l}\,\widehat{\mathbf{s}},$$

   where $\dot{l}$ is the rate of change of the length of the model damper, r is the damping constant, and $\widehat{\mathbf{s}}$ is a unit vector in the direction from the particle towards the centre of the model damper.

   A model damper has zero mass. It is represented diagrammatically as shown in the margin.



5. The equation of motion of a system consisting of a particle attached to a fixed point by a model damper and a model spring takes the form

   $$m\ddot{x} + r\dot{x} + kx = kx_{\text{eq}},$$

   where m is the particle's mass, r is the damping constant, k is the stiffness of the spring, and x_{eq} is the equilibrium position of the particle. A mechanical system whose equation of motion is of this form is a **damped linear harmonic oscillator**, or **damped harmonic oscillator** for short. (Note that the equation of motion is linear, and that in the absence of damping it would reduce to the simple harmonic motion equation.) If the origin of the x-coordinate is taken at the equilibrium position of the particle, then the right-hand side of the equation is zero, i.e. the equation is homogeneous.

6. The **damping ratio** for a damped harmonic oscillator is the dimensionless quantity

   $$\alpha = \frac{r}{2\sqrt{mk}}.$$

   The damping ratio gives a measure of how important damping is, relative to the mass m and the spring stiffness k of the system. The oscillator is strongly, weakly or critically damped according as $\alpha > 1$, $\alpha < 1$ or $\alpha = 1$, respectively.

7. The **natural angular frequency** of a damped harmonic oscillator is

$$\omega = \sqrt{k/m}.$$

It is the angular frequency of an undamped harmonic oscillator that has the same ratio of spring stiffness to mass as the damped one.

8. For **strongly damped motion**, i.e. when $\alpha > 1$ (or equivalently, when $r^2 - 4mk > 0$), the auxiliary equation of the damped harmonic motion equation has distinct real roots

$$\lambda_1 = \frac{-r + \sqrt{r^2 - 4mk}}{2m}, \quad \lambda_2 = \frac{-r - \sqrt{r^2 - 4mk}}{2m}.$$

The general solution of the homogeneous damped harmonic motion equation in this case is therefore

$$x(t) = Be^{\lambda_1 t} + Ce^{\lambda_2 t},$$

where B and C are arbitrary constants. The roots of the auxiliary equation can be expressed in terms of the damping ratio and the natural angular frequency as

$$\lambda_1 = \omega\big(-\alpha + \sqrt{\alpha^2 - 1}\,\big), \quad \lambda_2 = \omega\big(-\alpha - \sqrt{\alpha^2 - 1}\,\big).$$

Both λ_1 and λ_2 are negative since $\alpha > \sqrt{\alpha^2 - 1}$. Hence the solution of the damped harmonic motion equation is a sum of two decaying exponentials. Solutions of this type predict a return to the equilibrium position, without oscillation (although the graph of x against t may cross the t-axis once).

9. For **weakly damped motion**, i.e. when $\alpha < 1$ (or equivalently, when $r^2 - 4mk < 0$), the auxiliary equation of the damped harmonic motion equation has complex roots. The general solution of the homogeneous damped harmonic motion equation is the product of a decaying exponential and a sinusoidal function,

$$x(t) = Ae^{-\rho t}\cos(\nu t + \phi),$$

where $\rho = r/(2m)$, $\nu = \sqrt{4mk - r^2}/(2m)$, and A and ϕ are arbitrary constants (with A positive and ϕ in the range $-\pi < \phi \leq \pi$). The solution can also be expressed in terms of the damping ratio and the natural angular frequency as

$$x(t) = Ae^{-\omega\alpha t}\cos\big((\omega\sqrt{1 - \alpha^2})t + \phi\big).$$

The decaying exponential factor ensures that the oscillations represented by the sinusoidal factor reduce in amplitude. The **period** τ of the motion is the time between successive zeros of x, where the particle is moving in the same direction. The **amplitude** is the positive and continually changing quantity $Ae^{-\omega\alpha t}$. Over one cycle of period τ, where

$$\tau = \frac{2\pi}{\omega\sqrt{1 - \alpha^2}},$$

the amplitude is multiplied by the factor

$$e^{-\omega\alpha\tau} = \exp\left(-\frac{2\pi\alpha}{\sqrt{1-\alpha^2}}\right).$$

This is the **amplitude decay factor** per cycle. The angular frequency of the oscillations, $\nu = \omega\sqrt{1-\alpha^2}$, is less than the natural angular frequency ω.

10. For **critically damped motion**, i.e. when $\alpha = 1$ (or equivalently, when $r^2 - 4mk = 0$), the auxiliary equation of the damped harmonic oscillator equation has equal roots

$$\lambda_1 = \lambda_2 = -\frac{r}{2m} = -\omega\alpha = -\omega.$$

The corresponding general solution of the homogeneous damped harmonic oscillator equation is

$$x(t) = (Bt + C)e^{-\omega t},$$

where B and C are arbitrary constants. Solutions of this form do not represent oscillations and there is at most one time when $x(t) = 0$; the graph of a solution resembles that for strong damping, but the system returns more quickly towards the equilibrium position.

11. A typical equation of motion for a **forced damped harmonic oscillator** is

$$m\ddot{x} + r\dot{x} + kx = kx_{\text{eq}} + P\cos(\Omega t).$$

It differs from the damped harmonic oscillator equation by the addition of the sinusoidal term $P\cos(\Omega t)$ to the right-hand side: this is the forcing term. (Non-sinusoidal forcing is not considered in this unit.) There are two ways in which the forcing term may arise: **direct forcing** occurs when there is an additional sinusoidally varying force; **forcing by displacement** occurs when the system is subject to a sinusoidally varying constraint. It is important to note that the equation of motion is of the same form in both cases.

12. The general solution of the equation of motion of a forced damped harmonic oscillator is the sum of the complementary function and a particular integral. The complementary function is the general solution of a homogeneous equation, which corresponds to a damped system with no forcing. So the complementary function dies away with increasing time, whatever the initial conditions (this was shown in Unit 1), thus is called the **transient** part of the solution.

When the origin of the x-coordinate is the equilibrium position of the system, the particular integral is a sinusoidal function with the same angular frequency Ω as the forcing term but, in general, with a different amplitude and with a non-zero phase. The particular integral does not die away, and eventually becomes the main component of the motion. It is called the **steady-state solution** of the equation of motion.

13. The steady-state solution of the equation

$$m\ddot{x} + r\dot{x} + kx = P\cos(\Omega t)$$

(where the origin is the equilibrium position) is $x_\mathrm{p} = A\cos(\Omega t + \phi)$, where the amplitude A is given by

$$A = \frac{P}{\sqrt{\left(k - m\Omega^2\right)^2 + r^2\Omega^2}}$$

and the phase ϕ is given by

$$\phi = -\arccos\left(\frac{k - m\Omega^2}{\sqrt{\left(k - m\Omega^2\right)^2 + r^2\Omega^2}}\right).$$

14. The **magnification factor** produced by a sinusoidally varying forcing process is the ratio of the amplitude A of the (sinusoidal) steady-state solution to the amplitude of the input oscillations. (Note that the amplitude P of the forcing term need not be the same as the amplitude of the input oscillations. The precise relation between the two will depend on how the forcing is applied.)

15. For a given damping ratio α, the magnification factor is a function of β, the ratio of the forcing angular frequency to the natural angular frequency of the unforced system (i.e. $\beta = \Omega/\omega$). Considering only the case when the forcing is applied by moving the free end of the spring, the magnification factor function is

$$M(\beta) = \left((1-\beta^2)^2 + 4\alpha^2\beta^2\right)^{-1/2}.$$

If $\alpha < 1/\sqrt{2}$, then the magnification factor has a maximum at $\beta = \sqrt{1-2\alpha^2}$, and this phenomenon is called **resonance**.
If $\alpha \geq 1/\sqrt{2}$, then the magnification factor is a decreasing function of β, i.e. no resonance occurs.

Unit 11 Normal modes

1. The number of **degrees of freedom** of a system is the smallest number of coordinates needed to describe its configuration (i.e. the positions of the constituent parts of the system) at any instant in time.

2. A **normal mode** is a type of motion of a system of particles in which the coordinates of the particles all vary sinusoidally with the same frequency. The angular frequency of the sinusoidal motion is the **normal mode angular frequency**, with SI units $\mathrm{rad\,s^{-1}}$.

3. The general solution of the equation of motion for an oscillating mechanical system (without damping or forcing) whose dynamic matrix has distinct non-zero eigenvalues can be written as a linear combination of linearly independent normal modes of the system.

4. To analyse the motion of an oscillatory mechanical system (without damping or forcing) with n degrees of freedom, proceed as follows.

(a) Model the system using particles in conjunction with model springs, model rods, etc. Draw a sketch of the physical situation and annotate it with any relevant information. ◀ Draw picture ▶

(b) Choose n coordinates, denoted by $x_1, \ldots, x_n$, say, and corresponding origins. ◀ Choose coordinates ▶

(c) State any assumptions that you make about the objects and the forces acting on them. ◀ State assumption(s) ▶

(d) Draw a force diagram for each particle separately. ◀ Draw force diagram(s) ▶

(e) Apply Newton's second law to each particle separately, and resolve each force in the directions of appropriate axes to obtain n linear second-order differential equations of motion. ◀ Apply Newton's 2nd law ▶

(f) Write the equations of motion in matrix form as $\ddot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is an $n \times n$ matrix, called the **dynamic matrix** of the system. ◀ Solve equation(s) ▶

(g) Find the eigenvalues $\lambda_1, \ldots, \lambda_n$ and a corresponding set of eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of $\mathbf{A}$. The eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are often referred to as the **normal mode eigenvectors** of the system.

(h) Determine the normal mode angular frequencies $\omega_1, \ldots, \omega_n$ from the formula $\omega_i = \sqrt{-\lambda_i}$.

(i) Write down the general solution of the equations of motion in the form

$$x(t) = C_1\mathbf{v}_1 \cos(\omega_1 t + \phi_1) + \cdots + C_n\mathbf{v}_n \cos(\omega_n t + \phi_n),$$

where $C_1, \ldots, C_n$ and $\phi_1, \ldots, \phi_n$ are arbitrary constants.

(j) Interpret the solution in terms of the original problem. ◀ Interpret solution ▶

5. If one of the initial conditions of a mechanical system is that all the **particles start from rest**, then it follows that all the phase angles are zero.

6. If all the particles within a mechanical system start from rest at positions given by the elements of a normal mode eigenvector, then the system will oscillate in a normal mode.

7. For a normal mode motion, the relative motion of a pair of particles with non-zero coordinates is determined by the components of the corresponding normal mode eigenvector. The motion is:
   - **in-phase** if the components have the same sign
   - **phase-opposed** if the components have opposite signs.

8. If a particle of mass m, displaced from an equilibrium position, is acted on by forces $\mathbf{F}_i$ $(i = 1, \ldots, n)$, then its acceleration $\ddot{\mathbf{r}}$ is given by

$$m\ddot{\mathbf{r}} = \sum_{i=1}^{n} \Delta\mathbf{F}_i,$$

where $\mathbf{r}$ is the displacement of the particle from the equilibrium position, and $\Delta\mathbf{F}_i = \mathbf{F}_i - \mathbf{F}_{i,\text{eq}}$ $(i = 1, \ldots, n)$, where $\mathbf{F}_{i,\text{eq}}$ is the force corresponding to $\mathbf{F}_i$ when the particle is at its equilibrium position.

9. If a spring is displaced along its length, then the change in force $\Delta\mathbf{H}$ exerted by the spring is given by

$$\Delta\mathbf{H} = k\,\Delta l\,\widehat{\mathbf{s}},$$

where k is the stiffness of the spring, $\Delta l = l - l_{\text{eq}}$ is the change in the spring's length from its equilibrium value, and $\widehat{\mathbf{s}}$ is a (constant) unit vector directed from the particle towards the centre of the spring.

10. Near the Earth's surface, the weight of a particle of constant mass is constant, so $\Delta\mathbf{W} = \mathbf{0}$. If a normal reaction force balances only the weight of a particle (such as an object resting on a horizontal table), then $\mathbf{N}$ is constant, i.e. $\Delta\mathbf{N} = \mathbf{0}$. In these circumstances the weights and normal reaction forces do not contribute terms to the equation of motion and may be safely ignored.

11. **Rigid body motion** is the motion that occurs when all the particles of a system move with the distances between them remaining invariant; that is, the particles move as if part of a rigid body.

A mechanical system, whose equation of motion is $\ddot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, is capable of rigid body motion if the dynamic matrix $\mathbf{A}$ has zero as an eigenvalue. An eigenvector corresponding to a zero eigenvalue determines the velocities that the particles need to be given for the system to move in the corresponding rigid body motion. (If the particles start from rest with positions given by such an eigenvector, then the system remains at rest.)

12. Vibrations of a stretched string, such as a guitar string, that are aligned with the string are **longitudinal**, while vibrations at right angles to the string are **transverse**. In this context, the **fundamental** is the normal mode of a system that has the lowest non-zero normal mode angular frequency. The frequency of this mode is the **fundamental frequency**. **Harmonics** are the normal modes of a system that are not rigid body motions.

13. A model of a real-world system in which properties of the system are lumped together into discrete components (e.g. particles and model springs) is called a **lumped parameter model**.

14. The **equilibrium tension** in a model spring is $T_{\text{eq}} = k(l_{\text{eq}} - l_0)$, where k is the stiffness of the spring, l_0 is its natural length, and l_{eq} is its equilibrium length.

When a small displacement $y\mathbf{j}$ is made at right angles to the equilibrium alignment of a model spring, the approximate change in the spring force is given by

$$\Delta\mathbf{H} \simeq -\frac{T_{\text{eq}}}{l_{\text{eq}}}\,y\,\mathbf{j}.$$

Unit 12 Systems of differential equations

1. An **autonomous differential equation**, or system of differential equations, is one where the independent variable does not appear explicitly.

2. A **vector field** is a rule that assigns a vector

   $$\mathbf{u} = (u(x, y) \quad v(x, y))^T$$

   to each point (x, y) in a two-dimensional region. Since $\mathbf{u}$ is dependent on two values, x and y, we often write it in the form $\mathbf{u}(x, y)$, or $\mathbf{u}(\mathbf{x})$, in order to emphasise this dependence.

   A vector field determines a system of (in general) non-linear differential equations

   $$(\dot{x} \quad \dot{y})^T = \mathbf{u}(x, y),$$

   which can be written as

   $$\dot{\mathbf{x}} = \mathbf{u}(\mathbf{x}).$$

   A particular solution $(x(t) \quad y(t))^T$ of this system determines a path in the (x, y)-plane, parametrised by t, whose tangent vector at any point (x, y) on it is the vector $\mathbf{u}(x, y)$. Such a solution curve is called a **phase path**. The (x, y)-plane containing the solution curves is called the **phase plane**, and a diagram showing the phase paths is called a **phase portrait**.

3. The evolution of two interacting populations x and y can be modelled by the **Lotka–Volterra equations**

   $$\dot{x} = kx\left(1 - \frac{y}{Y}\right), \quad \dot{y} = -hy\left(1 - \frac{x}{X}\right) \quad (x \geq 0,\ y \geq 0),$$

   where x is the population of prey and y is the population of predators, and k, h, X and Y are positive constants.

4. A second-order differential equation $\ddot{x} = f(x, \dot{x})$ can be converted into a pair of simultaneous first-order equations by setting $y = \dot{x}$; the equivalent pair is

   $$\dot{x} = y, \quad \dot{y} = f(x, y).$$

5. An **equilibrium point** of a system of differential equations $\dot{\mathbf{x}} = \mathbf{u}(x, y)$ is a point (X, Y) such that $x(t) = X$, $y(t) = Y$ is a constant solution of the system; that is, (X, Y) is a point at which $\dot{x}(t) = 0$ and $\dot{y}(t) = 0$.

6. To find the equilibrium points of the system of differential equations $\dot{\mathbf{x}} = \mathbf{u}(x, y)$ for some vector field $\mathbf{u}$, solve the equation $\mathbf{u}(x, y) = \mathbf{0}$.

7. An equilibrium point is said to be:

   - **stable** when all points in the neighbourhood of the point remain in the neighbourhood of the point as time increases
   - **unstable** otherwise.

8. Suppose that the system of differential equations $\dot{\mathbf{x}} = \mathbf{u}(\mathbf{x})$ has an equilibrium point at $x = X$, $y = Y$. To linearise this system, carry out the following steps.

   (a) Find the **Jacobian matrix**

   $$\mathbf{J}(x, y) = \begin{pmatrix} u_x(x, y) & u_y(x, y) \\ v_x(x, y) & v_y(x, y) \end{pmatrix},$$

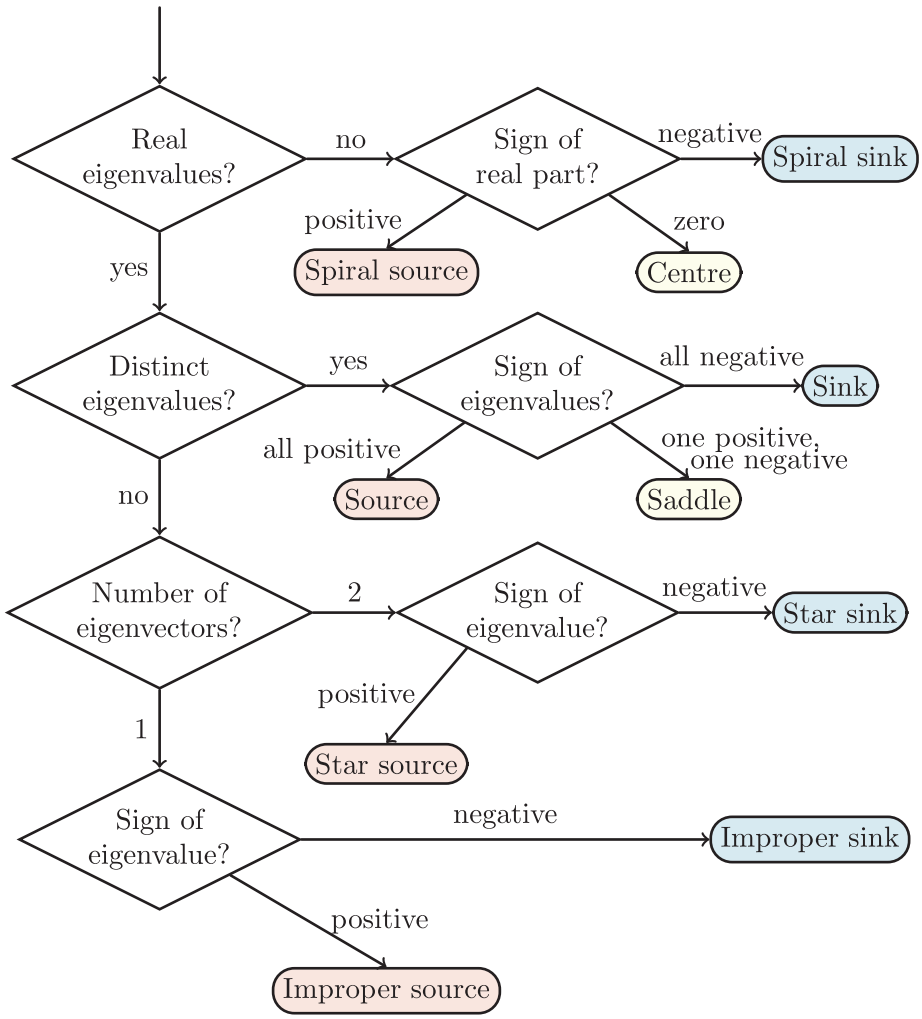
   where $u_x(x, y)$ is the x-derivative of u, i.e. $\partial u/\partial x$.

   (b) In the neighbourhood of equilibrium point (X, Y), the differential equations can be approximated by the **linearised form**

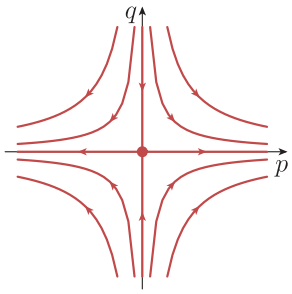
   $$\begin{pmatrix} \dot{p} \\ \dot{q} \end{pmatrix} = \mathbf{J} \begin{pmatrix} p \\ q \end{pmatrix},$$

   where $x(t) = X + p(t)$ and $y(t) = Y + q(t)$, and $\mathbf{J} = \mathbf{J}(x, y)$.

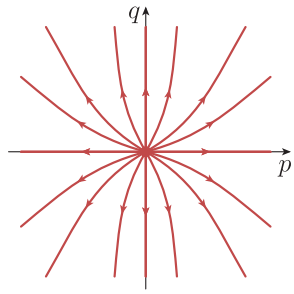
9. Consider a linear system $\dot{\mathbf{p}} = \mathbf{J}\mathbf{p}$ for a 2×2 matrix $\mathbf{J}$. The nature of the equilibrium point at $p = 0$, $q = 0$ is determined by the eigenvalues and eigenvectors of $\mathbf{J}$. A decision tree for classifying equilibrium points is shown below.



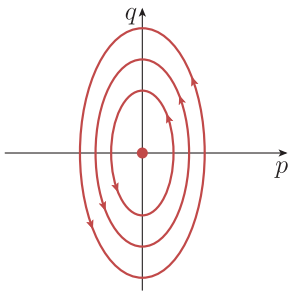
10. To **classify the equilibrium points of a system of non-linear differential equations**, carry out the following steps.
    (a) Find the equilibrium points by solving the equation $\mathbf{u}(x, y) = \mathbf{0}$.
    (b) Use the Jacobian matrix to find a linear system that approximates the non-linear system in the neighbourhood of each equilibrium point.
    (c) Use the above decision tree to classify the linear system.
    (d) For each equilibrium point, the behaviour of the original non-linear system is the same as that of the linear approximation, except when the linear system has a centre. If the linear system has a centre, then the equilibrium point of the original non-linear system may be a centre, a spiral sink or a spiral source.
11. Near an equilibrium point, the phase paths are similar to the following diagrams. (The diagrams for the various types of sink are the same as for the sources except with the directions of the arrows changed.)



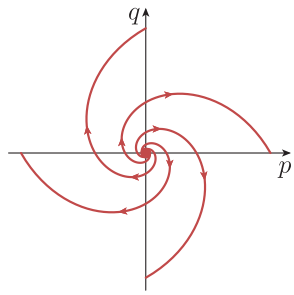
Saddle point



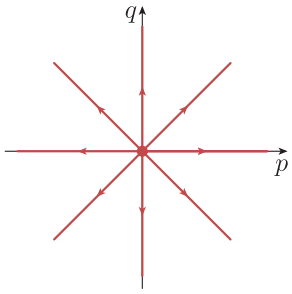
Source



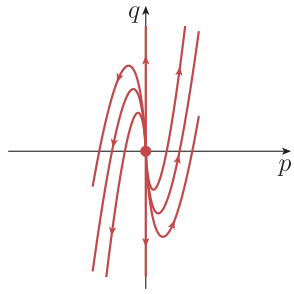
Centre



Spiral source



Star source



Improper source

12. The solutions of $\dot{x} = 0$ or $\dot{y} = 0$ are called **nullclines**. When we need to distinguish between the two sets of nullclines, the solutions of $\dot{x} = 0$ will be called the nullclines for $\dot{x}$, and the solutions of $\dot{y} = 0$ will be called the nullclines for $\dot{y}$.

13. In order to sketch the **phase portrait** of a system of differential equations in x and y, carry out the following steps.

◀ Equilibrium points ▶

(a) Find and classify the equilibrium points. Mark these points on a sketch, and draw small sketches of the local behaviour of the paths near the equilibrium points.

◀ Nullclines ▶

(b) Find the nullclines by separately solving $\dot{x} = 0$ and $\dot{y} = 0$. Draw these on your sketch, using two different colours for the two sets of nullclines. (Note that the equilibrium points occur where nullclines for x and y intersect – this is a useful check.)

◀ Nullcline crossings ▶

(c) For each equilibrium point on the nullclines for $\dot{x}$, evaluate the sign of $\dot{y}$ on either side. Mark this on your sketch by adding up or down arrows.

For each equilibrium point on the nullclines for $\dot{y}$, evaluate the sign of $\dot{x}$ on either side. Mark this on your sketch by adding left or right arrows.

◀ Nullcline regions ▶

(d) The nullclines divide the phase plane into several regions. Label each region with a NE, NW, SE or SW arrow to show the general direction of the arrows in that region.

◀ Complete paths ▶

(e) Extend the paths from the equilibrium points by curving the paths in the direction of the arrows in each region. Add any extra paths that do not start or end at equilibrium points, so that each region is crossed by at least one path.

14. A **limit cycle** of a system of differential equations is a closed solution curve to which nearby solution curves tend (either forwards or backwards in time).

Unit 13 Fourier series

1. A function $f(t)$ is **periodic** if there is some positive number τ such that $f(t + \tau) = f(t)$ for all t. The number τ is a **period** of the function f. If τ is a period of a function, so is $n\tau$ for any positive integer n. The **fundamental period** of a periodic function is the smallest (positive) value for the period. Unless it is specified otherwise, assume that 'period' means 'fundamental period'.
A **fundamental interval** for a periodic function is any interval whose length is the fundamental period.

2. A **Fourier series** is an infinite series of sinusoidal functions of the form

$$A_0 + A_1 \cos\left(\frac{2\pi t}{\tau}\right) + A_2 \cos\left(\frac{4\pi t}{\tau}\right) + \cdots$$
$$+ B_1 \sin\left(\frac{2\pi t}{\tau}\right) + B_2 \sin\left(\frac{4\pi t}{\tau}\right) + \cdots,$$

or equivalently,

$$A_0 + \sum_{n=1}^{\infty} A_n \cos\left(\frac{2n\pi t}{\tau}\right) + \sum_{n=1}^{\infty} B_n \sin\left(\frac{2n\pi t}{\tau}\right),$$

where $A_0, A_1, A_2, \ldots, B_1, B_2, \ldots$ and τ are constants, with $\tau > 0$.

3. A function $f(t)$ is an **even function** if

   $$f(-t) = f(t) \quad \text{for all values of } t;$$

   it is an **odd function** if

   $$f(-t) = -f(t) \quad \text{for all values of } t.$$

4. Even and odd functions combine under addition and multiplication as follows.
   - The sum of two even functions is even.
   - The sum of two odd functions is odd.
   - The sum of an even function and an odd function is neither even nor odd (unless one of the functions is the zero function).
   - The product of two even functions is even.
   - The product of two odd functions is even.
   - The product of an even and an odd function is odd.

5. Let $f(t)$ be periodic of period $2a$. Then $f(t)$ is even provided that it is even over the interval $[-a, a]$. Similarly, $f(t)$ is odd provided that it is odd over the interval $[-a, a]$.

6. If g is an even function of period $2a$, then

   $$\int_{-a}^{a} g(t)\,dt = 2\int_{0}^{a} g(t)\,dt.$$

   If f is an odd function of period $2a$, then

   $$\int_{-a}^{a} f(t)\,dt = 0.$$

7. For a a non-zero constant and C a constant,

   $$\int t\sin(at)\,dt = \frac{1}{a^2}\big(\sin(at) - at\cos(at)\big) + C,$$
   $$\int t\cos(at)\,dt = \frac{1}{a^2}\big(\cos(at) + at\sin(at)\big) + C.$$

8. The **pointwise convergence theorem** for Fourier series can be stated as follows. If, on the interval $[-\pi, \pi]$, the function f has a continuous derivative except at a finite number of points, then at each point $x_0 \in [-\pi, \pi]$, the Fourier series for f converges to

   $$\tfrac{1}{2}\left(f(x_0^+) + f(x_0^-)\right).$$

   Here $f(x_0^+)$ is the limit of $f(x)$ as x approaches x_0 from above, and $f(x_0^-)$ is the limit of $f(x)$ as x approaches x_0 from below.

9. The integrals of the sinusoidal functions

$$\cos\left(\frac{2n\pi t}{\tau}\right) \quad \text{and} \quad \sin\left(\frac{2n\pi t}{\tau}\right)$$

over the interval $\left[-\frac{\tau}{2}, \frac{\tau}{2}\right]$ have the following important properties.

For any positive integers m and n:

$$\int_{-\tau/2}^{\tau/2} \cos\left(\frac{2m\pi t}{\tau}\right) dt = 0,$$

$$\int_{-\tau/2}^{\tau/2} \cos\left(\frac{2m\pi t}{\tau}\right) \cos\left(\frac{2n\pi t}{\tau}\right) dt = 0 \quad (m \neq n),$$

$$\int_{-\tau/2}^{\tau/2} \cos^2\left(\frac{2m\pi t}{\tau}\right) dt = \tfrac{1}{2}\tau,$$

$$\int_{-\tau/2}^{\tau/2} \sin\left(\frac{2m\pi t}{\tau}\right) dt = 0,$$

$$\int_{-\tau/2}^{\tau/2} \sin\left(\frac{2m\pi t}{\tau}\right) \sin\left(\frac{2n\pi t}{\tau}\right) dt = 0 \quad (m \neq n),$$

$$\int_{-\tau/2}^{\tau/2} \sin^2\left(\frac{2m\pi t}{\tau}\right) dt = \tfrac{1}{2}\tau,$$

$$\int_{-\tau/2}^{\tau/2} \sin\left(\frac{2m\pi t}{\tau}\right) \cos\left(\frac{2n\pi t}{\tau}\right) dt = 0.$$

10. For a periodic function $f(t)$, with period τ and fundamental interval $\left[-\frac{\tau}{2}, \frac{\tau}{2}\right]$, the Fourier series

$$F(t) = A_0 + \sum_{n=1}^{\infty} A_n \cos\left(\frac{2n\pi t}{\tau}\right) + \sum_{n=1}^{\infty} B_n \sin\left(\frac{2n\pi t}{\tau}\right)$$

is found by using the formulas

$$A_0 = \frac{1}{\tau}\int_{-\tau/2}^{\tau/2} f(t)\, dt,$$

$$A_n = \frac{2}{\tau}\int_{-\tau/2}^{\tau/2} f(t) \cos\left(\frac{2n\pi t}{\tau}\right) dt \quad (n = 1, 2, \ldots),$$

$$B_n = \frac{2}{\tau}\int_{-\tau/2}^{\tau/2} f(t) \sin\left(\frac{2n\pi t}{\tau}\right) dt \quad (n = 1, 2, \ldots).$$

11. Consider a function $f(t)$ defined over a finite domain $0 \leq t \leq T$.

The **even periodic extension** of $f(t)$ is given by

$$f_{\text{even}}(t) = \begin{cases} f(t) & \text{for } 0 \leq t < T, \\ f(-t) & \text{for } -T \leq t < 0, \end{cases}$$

$$f_{\text{even}}(t + 2T) = f_{\text{even}}(t).$$

The **odd periodic extension** of $f(t)$ is given by

$$f_{\text{odd}}(t) = \begin{cases} f(t) & \text{for } 0 < t < T, \\ -f(-t) & \text{for } -T < t < 0, \\ 0 & \text{for } t = 0 \text{ or } t = T, \end{cases}$$

$$f_{\text{odd}}(t + 2T) = f_{\text{odd}}(t).$$

Unit 14 Partial differential equations

1. A **partial differential equation** relates a dependent variable and two or more independent variables through the partial derivatives of the dependent variable. (A differential equation involving only a single independent variable, and therefore only ordinary derivatives, may be called an ordinary differential equation if it is necessary to differentiate it from a partial differential equation.)

   A given function of the independent variables is a solution of a given partial differential equation if the equation is satisfied when the function and its appropriate partial derivatives are substituted into it. To find a particular solution of a partial differential equation, it is necessary to impose initial and boundary conditions that the solution must satisfy (in addition to satisfying the partial differential equation itself).

2. The **order** of a partial differential equation is the order of the highest derivative that occurs in it.

3. A **linear** partial differential equation is one that contains no products or non-linear functions of terms involving the dependent variable and its partial derivatives. For example, a linear first-order partial differential equation where the dependent variable is u and the independent variables are x and y can be written as

   $$A(x, y)\, u_x + B(x, y)\, u_y + C(x, y)\, u = D(x, y)$$

   for some functions A, B, C and D.

4. A partial differential equation (or boundary condition or initial condition) all of whose additive terms involve the dependent variable or its derivatives (so that the equation contains no constant terms or terms involving solely the independent variables) is **homogeneous**.

5. If $\Theta(x, t)$ satisfies a partial differential equation with an **inhomogeneous boundary condition**, then this can be recast as a problem with homogeneous boundary conditions if there is a function $f(x)$ such that $u(x, t) = \Theta(x, t) - f(x)$ satisfies the same partial differential equation with homogeneous boundary conditions. For example, if $\Theta(x, t)$ has boundary conditions $\Theta(0, t) = \Theta(1, t) = C$, where C is a non-zero constant, then $u(x, t) = \Theta(x, t) - C$ will satisfy the homogeneous boundary conditions $u(0, t) = u(1, t) = 0$.

6. The **principle of superposition** states that if u and v are solutions of a linear homogeneous partial differential equation, then $Au + Bv$ is also a solution of the same equation for any constants A and B. Furthermore, if u and v also satisfy homogeneous boundary conditions such as $u(0,t) = 0$ or $u_x(L,t) = 0$, then the linear combination $Au + Bv$ will also satisfy them.

7. The general solution of the equation $T'(t) - \mu\, T(t) = 0$ is

$$T(t) = C\exp(\mu t),$$

where C is a constant.

8. The general solution of the equation $X''(x) - \mu\, X(x) = 0$ is

$$X(x) = \begin{cases} Ae^{cx} + Be^{-cx} & \text{for } \mu > 0, \\ Ax + B & \text{for } \mu = 0, \\ A\cos kx + B\sin kx & \text{for } \mu < 0, \end{cases}$$

where A and B are constants, $c = \sqrt{\mu}$ and $k = \sqrt{-\mu}$.

9. The procedure called **separation of variables** is a procedure to find the solution of a homogeneous linear partial differential equation with dependent variable u and independent variables x and t, subject to boundary and initial conditions, by carrying out the following steps.

◀ Separate variables ▶

(a) Separate the variables by substituting the trial solution

$$u(x,t) = X(x)\,T(t)$$

into the partial differential equation and rearranging so that each side of the equation involves only one of the variables. Both sides must then be equal to a **separation constant** μ.

Rearranging then gives two separate ordinary differential equations for X and T. The boundary conditions for u will give boundary conditions for X.

◀ Solve ODEs ▶

(b) Find the general solution of the ordinary differential equations for X and T found above. Use the boundary conditions for X to find the **normal mode solutions** $u_n(x,t)$.

◀ Initial conditions ▶

(c) Write down the general solution as a linear combination of the normal mode solutions:

$$u(x,t) = \sum_{n=0}^{\infty} a_n\, u_n(x,t).$$

The initial conditions (and results about Fourier series) can be used to determine the constants a_n appearing in this solution.

10. The transverse vibrations of a **taut string** can be modelled as follows. It is assumed that the string is fixed at both ends, that it is uniform, that it remains taut, and that there is no damping. The mass of the string is M, its equilibrium length is L, and the equilibrium tension is T_{eq} (a constant). The string is released from rest, its initial shape

being specified by a function $f(x)$; the subsequent transverse displacement of the string is modelled by a function $u(x,t)$, where (when the system is in equilibrium) x is the distance from one fixed end, and t is the time from the moment of release.

The model consists of the homogeneous linear second-order partial differential equation

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}, \quad \text{where } c^2 = \frac{TL}{M},$$

which is called the **wave equation**, the boundary conditions that correspond to the two ends being fixed,

$$u(0,t) = u(L,t) = 0 \quad (t \geq 0),$$

and the initial conditions that correspond to the string being released from rest with initial shape $f(x)$,

$$u(x,0) = f(x), \quad u_t(x,0) = 0 \quad (0 \leq x \leq L).$$

11. If there is linear damping of the vibrations of a plucked string, then the partial differential equation in the model is modified into the **damped wave equation**

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\left(\frac{\partial^2 u}{\partial t^2} + 2\varepsilon\frac{\partial u}{\partial t}\right),$$

where ε is a positive constant.

12. **Newton's law of cooling** states that for a given object, the rate of decrease of temperature is proportional to the excess temperature over the environment.

13. The change in temperature distribution along a rod as it cools can be modelled by assuming that Newton's law of cooling applies and that the rod is uniform and loses heat only at its ends. The temperature distribution $\Theta(x,t)$ along the bar is a function of the distance x from one fixed end, and time t.

    The model consists of the homogeneous linear second-order partial differential equation called the **heat equation**,

$$\frac{\partial \Theta}{\partial t} = \alpha\frac{\partial^2 \Theta}{\partial x^2},$$

    together with appropriate boundary conditions (such as the end of the rod being insulated) and initial conditions (such as a given initial temperature distribution).

14. If, instead of the bar being insulated along its length, there is heat loss from the sides of the rod, then the partial differential equation can be modified to the **uninsulated rod equation**

$$\frac{\partial \Theta}{\partial t} = \alpha\frac{\partial^2 \Theta}{\partial x^2} - \gamma(\Theta - \Theta_0).$$

Unit 15 Vector calculus

1. An **orthogonal matrix A** has an inverse equal to its transpose, so $\mathbf{A}^T\mathbf{A} = \mathbf{I}$. A rotation of three-dimensional space is represented by an orthogonal matrix. Any transformation that takes one right-handed set of unit vectors into another such set is a rotation, and can therefore be represented by an orthogonal matrix.
2. A **scalar field** is a distribution of scalar values on a two- or three-dimensional region and is represented mathematically by a (scalar) function of two or three variables. Scalar physical quantities such as temperature and atmospheric pressure may be modelled as scalar fields. The terms 'scalar field' and 'scalar function' are used interchangeably.
3. The family of curves given by $f(x, y) = C$, for different values of the constant C, are the **contour curves** of the two-dimensional scalar field f.
4. The family of surfaces given by $f(x, y, z) = C$, for different values of the constant C, are the **contour surfaces** of the scalar field f.
5. A **vector field** is a distribution of vectors on a two- or three-dimensional region and is represented mathematically by a vector function of two or three variables. Physical quantities that are vectorial in nature may be modelled as vector fields: examples include forces (force fields), velocities (velocity fields) and magnetic fields. The terms 'vector field' and 'vector function' are used interchangeably.

   A vector field $\mathbf{F}$ may be expressed in the form

   $$\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k},$$

   where $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are Cartesian unit vectors. F_1, F_2 and F_3, which are functions of the (Cartesian) coordinates x, y and z, are the components or component fields of $\mathbf{F}$. We may write a vector field as $\mathbf{F}(x, y, z)$ or $\mathbf{F}(\mathbf{r})$ to emphasise its dependence on position.
6. The **vector field lines** of a vector field $\mathbf{F}$ are continuous curves in the domain of $\mathbf{F}$ such that at any point (x, y, z), the tangent to the curve at (x, y, z) is in the direction of the vector $\mathbf{F}(x, y, z)$.
7. The **gradient** of a three-dimensional scalar field f is the vector field **grad** f whose expression in terms of Cartesian coordinates (x, y, z) is

   $$\textbf{grad}\, f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}.$$

   The gradient of a two-dimensional scalar field is defined similarly except that it has no $\mathbf{k}$-component. The gradient of a scalar field f at a point is always normal (i.e. perpendicular) to the tangent (line or plane) to the contour (curve or surface) of f that passes through that point.

8. The value of the **maximum derivative** of the scalar field $f(x, y)$ at the point (a, b) is given by $|\mathbf{grad}\, f(a, b)|$, and the **direction** of this maximum derivative is given by the unit vector

$$\frac{\mathbf{grad}\, f(a, b)}{|\mathbf{grad}\, f(a, b)|}.$$

9. The **derivative** or **directional derivative** of a scalar field f in the direction specified by a unit vector $\widehat{\mathbf{d}}$ is given by

$$\mathbf{grad}\, f \cdot \widehat{\mathbf{d}}.$$

10. The **vector differential operator** $\boldsymbol{\nabla}$, read as 'del' or 'nabla', is the operator given in Cartesian coordinates by

$$\boldsymbol{\nabla} = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}.$$

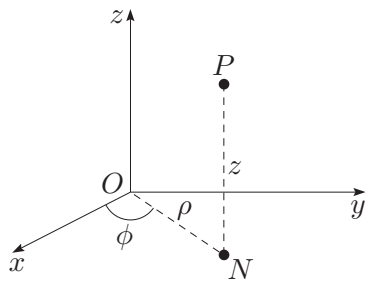
The gradient of a scalar field f can be written either as $\mathbf{grad}\, f$ or as $\boldsymbol{\nabla} f$.

11. The **gradient function in polar coordinates** of a scalar field f is

$$\mathbf{grad}\, f = \mathbf{e}_r \frac{\partial f}{\partial r} + \mathbf{e}_\theta \frac{1}{r}\frac{\partial f}{\partial \theta},$$

where $\mathbf{e}_r$ and $\mathbf{e}_\theta$ are the unit vectors in the r- and θ-directions, respectively.

12. Any point P can be represented in **cylindrical coordinates** by the triple (ρ, ϕ, z), where z is the distance of P from the (x, y)-plane and (ρ, ϕ) are the polar coordinates of the projection N of P onto the (x, y)-plane (see margin figure).



Cylindrical coordinates

Cylindrical coordinates are related to the Cartesian coordinates (x, y, z) by

$$x = \rho\cos\phi, \quad y = \rho\sin\phi, \quad z = z,$$
$$\rho = (x^2 + y^2)^{1/2}, \quad \cos\phi = x/\rho, \quad \sin\phi = y/\rho.$$

We require that

$$\rho \geq 0, \quad -\pi < \phi \leq \pi, \quad z \in \mathbb{R}.$$

The value of ϕ for points on the z-axis ($\rho = 0$) is undefined. By convention, we put $\phi = 0$ for such points.

13. The **cylindrical unit vectors** $\mathbf{e}_\rho$, $\mathbf{e}_\phi$ and $\mathbf{e}_z$ can be expressed in terms of the Cartesian unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ as

$$\mathbf{e}_\rho = \mathbf{i}\cos\phi + \mathbf{j}\sin\phi = \frac{x\mathbf{i} + y\mathbf{j}}{\sqrt{x^2 + y^2}},$$
$$\mathbf{e}_\phi = -\mathbf{i}\sin\phi + \mathbf{j}\cos\phi = \frac{-y\mathbf{i} + x\mathbf{j}}{\sqrt{x^2 + y^2}},$$
$$\mathbf{e}_z = \mathbf{k}.$$

14. A vector $\mathbf{F}$ can be expressed in terms of the cylindrical unit vectors $\mathbf{e}_\rho$, $\mathbf{e}_\phi$ and $\mathbf{e}_z$ as

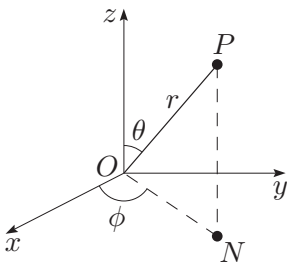
$$\mathbf{F} = F_\rho \mathbf{e}_\rho + F_\phi \mathbf{e}_\phi + F_z \mathbf{e}_z,$$

where $F_\rho = \mathbf{F} \cdot \mathbf{e}_\rho$, $F_\phi = \mathbf{F} \cdot \mathbf{e}_\phi$ and $F_z = \mathbf{F} \cdot \mathbf{e}_z$. The scalar quantities F_ρ, F_ϕ and F_z are the **cylindrical components** of $\mathbf{F}$.

15. The **gradient function in cylindrical coordinates** of a scalar field f is

$$\mathbf{grad}\, f = \mathbf{e}_\rho \frac{\partial f}{\partial \rho} + \mathbf{e}_\phi \frac{1}{\rho} \frac{\partial f}{\partial \phi} + \mathbf{e}_z \frac{\partial f}{\partial z},$$

where $\mathbf{e}_\rho$, $\mathbf{e}_\phi$ and $\mathbf{e}_z$ are unit vectors in the ρ-, ϕ- and z-directions, respectively.



Spherical coordinates

16. Any point P can be represented in **spherical coordinates** by the triple (r, θ, ϕ), where r is the distance of P from the origin, and θ and ϕ are the polar and azimuthal angles, respectively. In the margin figure, N is the projection of P onto the (x, y)-plane.

The spherical coordinates of P are related to the Cartesian coordinates (x, y, z) by

$$x = ON \cos\phi = r \sin\theta \cos\phi,$$
$$y = ON \sin\phi = r \sin\theta \sin\phi,$$
$$z = r \cos\theta,$$
$$r = \sqrt{x^2 + y^2 + z^2}.$$

We require that

$$r \geq 0, \quad -\pi < \phi \leq \pi, \quad 0 \leq \theta \leq \pi.$$

17. The **spherical unit vectors** $\mathbf{e}_r$, $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ can be expressed in terms of the Cartesian unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ as

$$\mathbf{e}_r = \mathbf{i} \sin\theta \cos\phi + \mathbf{j} \sin\theta \sin\phi + \mathbf{k} \cos\theta,$$
$$\mathbf{e}_\theta = \mathbf{i} \cos\theta \cos\phi + \mathbf{j} \cos\theta \sin\phi - \mathbf{k} \sin\theta,$$
$$\mathbf{e}_\phi = -\mathbf{i} \sin\phi + \mathbf{j} \cos\phi.$$

18. A vector $\mathbf{F}$ can be expressed in terms of the spherical unit vectors $\mathbf{e}_r$, $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ as

$$\mathbf{F} = F_r \mathbf{e}_r + F_\theta \mathbf{e}_\theta + F_\phi \mathbf{e}_\phi,$$

where $F_r = \mathbf{F} \cdot \mathbf{e}_r$, $F_\theta = \mathbf{F} \cdot \mathbf{e}_\theta$ and $F_\phi = \mathbf{F} \cdot \mathbf{e}_\phi$. The scalars F_r, F_θ and F_ϕ are the **spherical components** of $\mathbf{F}$.

19. The **gradient function in spherical coordinates** of a scalar field f is

$$\mathbf{grad}\, f = \mathbf{e}_r \frac{\partial f}{\partial r} + \mathbf{e}_\theta \frac{1}{r} \frac{\partial f}{\partial \theta} + \mathbf{e}_\phi \frac{1}{r \sin\theta} \frac{\partial f}{\partial \phi},$$

where $\mathbf{e}_r$, $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ are unit vectors in the r-, θ- and ϕ-directions, respectively.

Unit 16 Further vector calculus

1. The **divergence** of a vector field $\mathbf{F}$ is the scalar field given by

$$\operatorname{div}\mathbf{F} = \boldsymbol{\nabla}\cdot\mathbf{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z},$$

where x, y and z are Cartesian coordinates, and F_1, F_2 and F_3 are the Cartesian components of $\mathbf{F}$. The notations $\operatorname{div}\mathbf{F}$ and $\boldsymbol{\nabla}\cdot\mathbf{F}$ are interchangeable.

2. The **divergence** of a vector field $\mathbf{F}(\rho,\theta,z) = F_\rho\mathbf{e}_\rho + F_\phi\mathbf{e}_\phi + F_z\mathbf{e}_z$ in **cylindrical coordinates** is

$$\operatorname{div}\mathbf{F} = \boldsymbol{\nabla}\cdot\mathbf{F} = \frac{\partial F_\rho}{\partial\rho} + \frac{1}{\rho}F_\rho + \frac{1}{\rho}\frac{\partial F_\phi}{\partial\phi} + \frac{\partial F_z}{\partial z}.$$

The divergence of a vector field $\mathbf{F}(r,\theta,\phi) = F_r\mathbf{e}_r + F_\theta\mathbf{e}_\theta + F_\phi\mathbf{e}_\phi$ in **spherical coordinates** is

$$\operatorname{div}\mathbf{F} = \frac{\partial F_r}{\partial r} + \frac{1}{r}\left(\frac{\partial F_\theta}{\partial\theta} + 2F_r\right) + \frac{1}{r\sin\theta}\left(\frac{\partial F_\phi}{\partial\phi} + F_\theta\cos\theta\right).$$

3. The **curl** of a vector field $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$ in **Cartesian coordinates** is

$$\begin{aligned}\mathbf{curl}\,\mathbf{F} &= \boldsymbol{\nabla}\times\mathbf{F}\\ &= \begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ F_1 & F_2 & F_3\end{vmatrix}\\ &= \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\mathbf{k}.\end{aligned}$$

The notations $\mathbf{curl}\,\mathbf{F}$ and $\boldsymbol{\nabla}\times\mathbf{F}$ are interchangeable.

4. The **curl of a two-dimensional vector field** $\mathbf{F}(x,y)$ is

$$\mathbf{curl}\,\mathbf{F} = \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\mathbf{k}.$$

5. The **curl** of a vector field $\mathbf{F} = \mathbf{F}(\rho,\phi,z) = F_\rho\mathbf{e}_\rho + F_\phi\mathbf{e}_\phi + F_z\mathbf{e}_z$ in **cylindrical coordinates** is

$$\begin{aligned}\boldsymbol{\nabla}\times\mathbf{F} = &\left(\frac{1}{\rho}\frac{\partial F_z}{\partial\phi} - \frac{\partial F_\phi}{\partial z}\right)\mathbf{e}_\rho + \left(\frac{\partial F_\rho}{\partial z} - \frac{\partial F_z}{\partial\rho}\right)\mathbf{e}_\phi\\ &+ \left(\frac{\partial F_\phi}{\partial\rho} - \frac{1}{\rho}\frac{\partial F_\rho}{\partial\phi} + \frac{1}{\rho}F_\phi\right)\mathbf{e}_z.\end{aligned}$$

The curl of a vector field $\mathbf{F}(r,\theta,\phi) = F_r\mathbf{e}_r + F_\theta\mathbf{e}_\theta + F_\phi\mathbf{e}_\phi$ in **spherical coordinates** is

$$\begin{aligned}\boldsymbol{\nabla}\times\mathbf{F} = &\left(\frac{1}{r}\frac{\partial F_\phi}{\partial\theta} - \frac{1}{r\sin\theta}\frac{\partial F_\theta}{\partial\phi} + \frac{\cot\theta}{r}F_\phi\right)\mathbf{e}_r\\ &+ \left(-\frac{\partial F_\phi}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial F_r}{\partial\phi} - \frac{1}{r}F_\phi\right)\mathbf{e}_\theta\\ &+ \left(\frac{\partial F_\theta}{\partial r} - \frac{1}{r}\frac{\partial F_r}{\partial\theta} + \frac{1}{r}F_\theta\right)\mathbf{e}_\phi.\end{aligned}$$

6. The **curl** of a two-dimensional vector field $\mathbf{F} = \mathbf{F}(\rho, \phi) = F_\rho \mathbf{e}_\rho + F_\phi \mathbf{e}_\phi$ in **cylindrical coordinates** is

$$\boldsymbol{\nabla} \times \mathbf{F} = \left(\frac{\partial F_\phi}{\partial \rho} + \frac{F_\phi}{\rho} - \frac{1}{\rho} \frac{\partial F_\rho}{\partial \phi} \right) \mathbf{e}_z.$$

7. The **work done by a force** $\mathbf{F}(\mathbf{r})$, given in Cartesian coordinates, in moving a particle along a path C given by $\mathbf{r} = \mathbf{r}(t)$ from $\mathbf{r}(t_0)$ to $\mathbf{r}(t_1)$ is

$$\begin{aligned} W &= \int_C \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} \\ &= \int_{t_0}^{t_1} \left(F_1(t) \frac{dx}{dt} + F_2(t) \frac{dy}{dt} + F_3(t) \frac{dz}{dt} \right) dt. \end{aligned}$$

8. The **scalar line integral** of a vector field $\mathbf{F}(\mathbf{r})$ along a path C given by $\mathbf{r} = \mathbf{r}(t)$, from $\mathbf{r}(t_0)$ to $\mathbf{r}(t_1)$, is

$$\int_C \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} = \int_{t_0}^{t_1} \mathbf{F}(t) \cdot \frac{d\mathbf{r}}{dt} \, dt.$$

   The term $\mathbf{F}(t)$ on the right-hand side of this equation is defined to be $\mathbf{F}(\mathbf{r}(t))$, i.e. it is the value of the vector field along the path C as a function of the parameter t. To calculate the integrand, take the dot product of $\mathbf{F}(t)$ and the derivative of $\mathbf{r}(t)$ to obtain an expression involving t. The integral on the right-hand side is then an ordinary integral in t. The value of this integral is unchanged if the parametrisation of C is changed, provided that the path is traversed in the same direction.

9. When it is desirable to emphasise the dependence of a scalar line integral on the endpoints A, B of the path, write it as $\int_{AB} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r}$, where the order AB indicates the direction in which the path is traversed. However, it must be remembered that in general, the value of a scalar line integral depends not just on the points A and B, but also on the path joining them. The effect of traversing the path in the opposite sense, starting at B and ending at A, is to change the sign of the line integral:

$$\int_{BA} \mathbf{F}(r) \cdot d\mathbf{r} = -\int_{AB} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r}.$$

10. Line integrals can be evaluated along segments of **open curves**, where the starting point A and the endpoint B are distinct points, or around **closed curves**, for which A and B coincide. A scalar line integral around a closed curve C is written as

$$\oint_C \mathbf{F} \cdot d\mathbf{r}.$$

11. A scalar field U such that $\mathbf{F} = -\mathbf{grad}\, U$ is a **potential field** or a **potential** for $\mathbf{F}$. The value of U can be chosen to be zero at some convenient point, called the **datum**, and then the potential is uniquely determined.

The scalar line integral along a path from A to B can be evaluated by calculating the potential at the endpoints:

$$\int_{AB} \mathbf{F} \cdot d\mathbf{r} = U(\mathbf{r}_A) - U(\mathbf{r}_B),$$

where $\mathbf{r}_A$ and $\mathbf{r}_B$ are the position vectors of the endpoints.

12. A **simply-connected region** is a region where *any* closed curve contained in it can be continuously shrunk to a point without leaving the region.
13. If the scalar line integral of a vector field $\mathbf{F}$ along a path from A to B depends only on A and B and not on the path itself, then the scalar line integral is **path-independent**; otherwise it is **path-dependent**.
14. Suppose that we are given the force $\mathbf{F}(\mathbf{r})$ whose line integral is path-independent. To determine the potential function $U(\mathbf{r})$, with the datum set at the origin, carry out the following steps.
    (a) Take C to be the direct path from $(0, 0, 0)$ to the general point (a, b, c) parametrised by
    $$\mathbf{r} = at\mathbf{i} + bt\mathbf{j} + ct\mathbf{k} \quad (0 \leq t \leq 1).$$
    (b) With this choice of parametrisation, calculate the scalar line integral
    $$U(a, b, c) = -\int_C \mathbf{F} \cdot d\mathbf{r} = -\int_0^1 \mathbf{F} \cdot \frac{d\mathbf{r}}{dt}\, dt,$$
    and write $U(\mathbf{r}) = U(x, y, z)$.
15. A **conservative** vector field $\mathbf{F}$ is a vector field satisfying property (a) below.

    A conservative vector field has the following properties.

    (a) All line integrals of $\mathbf{F}$ between any two fixed points in the domain are path-independent.
    (b) The line integrals of $\mathbf{F}$ around all closed curves in the domain are zero.
    (c) The curl of $\mathbf{F}$ is zero everywhere in the domain.
    (d) All gradient fields are conservative, and a conservative vector field can be expressed as the gradient of a scalar field.
    (e) For a fixed origin O (the datum), there exists a unique potential field defined by
    $$U(\mathbf{r}) = -\int_{OP} \mathbf{F} \cdot d\mathbf{r}$$
    with $\mathbf{F} = -\mathbf{grad}\, U$, where $\overrightarrow{OP} = \mathbf{r}$.

    Properties (a), (b), (d) and (e) are equivalent. Property (c) is also equivalent if the domain is a simply-connected region.
16. The **curl test** states that if $\mathbf{curl}\, \mathbf{F} = \mathbf{0}$ everywhere, then $\mathbf{F}$ is a conservative field provided that the domain of $\mathbf{F}$ is simply-connected.

Unit 17 Multiple integrals

1. The **area integral** of $f(x, y)$ over a **region of integration** S in the (x, y)-plane, subdivided into N area elements where element i contains the point (x_i, y_i) and is of area δA_i, is

$$\int_S f(x, y)\, dA = \lim_{N\to\infty} \sum_{i=1}^{N} f(x_i, y_i)\, \delta A_i,$$

where $\delta A_i \to 0$ for each i as $N \to \infty$.

2. The **area integral in Cartesian coordinates** of a function $f(x, y)$ over a rectangular region S contained between the lines $x = a$, $x = b$ and $y = c$, $y = d$ is obtained as two successive integrals as follows:

$$\int_S f(x, y)\, dA = \int_{x=a}^{x=b} \left(\int_{y=c}^{y=d} f(x, y)\, dy \right) dx.$$

3. To evaluate an area integral

$$\int_S f(x, y)\, dA,$$

carry out the following steps.

   (a) Draw a diagram showing the region of integration S.

   (b) Draw on the diagram a strip parallel to the y-axis, and show the lower limit $y = \alpha(x)$ and the upper limit $y = \beta(x)$ of this strip. These are the limits of the y-integration, i.e. for the 'inner' integral.

   (c) Determine the lower limit a and upper limit b of x for points on the boundary of S. These are the limits of the x-integration, i.e. for the 'outer' integral.

   (d) Write the area integral as two single integrals, making sure that the outer limits are constants:

$$\int_S f(x, y)\, dA = \int_{x=a}^{x=b} \left(\int_{y=\alpha(x)}^{y=\beta(x)} f(x, y)\, dy \right) dx.$$

   (e) Evaluate the inner integral, holding x constant, to give

$$g(x) = \int_{y=\alpha(x)}^{y=\beta(x)} f(x, y)\, dy.$$

   (f) Evaluate the outer integral to give

$$\int_S f(x, y)\, dA = \int_{x=a}^{x=b} g(x)\, dx.$$

4. The **area integral in polar coordinates** of a function $f(r, \theta)$ over a disc D of radius a is

$$\int_D f\, dA = \int_{\theta=-\pi}^{\theta=\pi} \left(\int_{r=0}^{r=a} f(r, \theta)\, r\, dr \right) d\theta.$$

5. The quantity $I = md^2$, the product of a particle's mass m and the square of its perpendicular distance d from a fixed axis, is called the **moment of inertia** of the particle about the axis.

6. A **lamina** is a region of a plane used to model a sheet of material of uniform thickness. The distribution of mass in a lamina is given by its **surface density**, whose dimensions are $\mathrm{M\,L^{-2}}$; when the lamina is of uniform composition, its surface density is constant. Imagine the lamina placed in the (x, y)-plane; then its surface density is a function $f(x, y)$, known as the **surface density function**, defined on the region S occupied by the lamina. The total mass of the lamina is

$$M = \int_S f(x, y)\, dA.$$

7. The **moment of inertia of a lamina**, about the z-axis, with surface density function f, occupying a region S of the (x, y)-plane, is

$$I = \int_S f d^2\, dA,$$

where d is the perpendicular distance from the z-axis, so

$$f d^2 = \begin{cases} f(x, y)\,(x^2 + y^2) & \text{in Cartesian coordinates,} \\ f(r, \theta)\, r^2 & \text{in polar coordinates.} \end{cases}$$

8. The **volume integral** of $f(x, y, z)$ over a **region of integration** B, subdivided into N elements where element i contains the point (x_i, y_i, z_i) and is of volume δV_i, is

$$\int_B f(x, y, z)\, dV = \lim_{N \to \infty} \sum_{i=1}^{N} f(x_i, y_i, z_i)\, \delta V_i,$$

where $\delta V_i \to 0$ as $N \to \infty$.

9. The **volume integral in Cartesian coordinates** of a function $f(x, y, z)$ over a rectangular cuboid B whose faces lie in coordinate planes $x = a$, $x = b$, $y = c$, $y = d$, $z = p$, $z = q$ is obtained as three successive integrals as follows:

$$\int_B f(x, y, z)\, dV = \int_{x=a}^{x=b} \left(\int_{y=c}^{y=d} \left(\int_{z=p}^{z=q} f(x, y, z)\, dz \right) dy \right) dx.$$

10. The **volume integral in cylindrical coordinates** is given by

$$\int_B f\, dV = \int_B f(\rho, \phi, z)\, \rho\, dz\, d\phi\, d\rho.$$

For example, if the region of integration is a cylinder B of radius a and height h, with the base of the cylinder in the (x, y)-plane, then in cylindrical coordinates the limits of integration are $z = 0$ and $z = h$, $\phi = -\pi$ and $\phi = \pi$, $\rho = 0$ and $\rho = a$. So the volume integral of a scalar field f over a cylinder is given by

$$\int_B f\, dV = \int_{\rho=0}^{\rho=a} \left(\int_{\phi=-\pi}^{\phi=\pi} \left(\int_{z=0}^{z=h} f \rho\, dz \right) d\phi \right) d\rho.$$

If f is independent of ϕ, then

$$\int_B f\, dV = 2\pi \int_{\rho=0}^{\rho=a} \left(\int_{z=0}^{z=h} f \rho\, dz \right) d\rho.$$

11. The **volume integral in spherical coordinates** is given by

$$\int_B f\,dV = \int_B f(r,\theta,\phi)\,r^2\sin\theta\,d\phi\,d\theta\,dr.$$

For a region contained between spherical shells of radii R_1 and R_2, where $R_1 < R_2$, for example, we have

$$\int_B f\,dV = \int_{r=R_1}^{r=R_2}\left(\int_{\theta=0}^{\theta=\pi}\left(\int_{\phi=-\pi}^{\phi=\pi} f(r,\theta,\phi)\,r^2\sin\theta\,d\phi\right)d\theta\right)dr.$$

If the scalar field f is *spherically symmetric*, so that f depends on r only, and not on θ or ϕ, then

$$\int_B f\,dV = 4\pi\int_{R_1}^{R_2} f(r)\,r^2 dr.$$

12. The **moment of inertia of a rigid body**, about the z-axis, with density function f, occupying a region B, is

$$I = \int_B f d^2\,dV,$$

where d is the distance from the z-axis, so

$$fd^2 = \begin{cases} f(x,y,z)\,(x^2+y^2) & \text{in Cartesians,} \\ f(\rho,\phi,z)\,\rho^2 & \text{in cylindrical coordinates,} \\ f(r,\theta,\phi)\,(r\sin\theta)^2 & \text{in spherical coordinates.} \end{cases}$$

13. The **volume integral over a general volume** is calculated as

$$\int_B f(x,y,z)\,dV = \int_{x=a}^{x=b}\left(\int_{y=\alpha(x)}^{y=\beta(x)}\left(\int_{z=\gamma(x,y)}^{z=\psi(x,y)} f(x,y,z)\,dz\right)dy\right)dx.$$

14. To evaluate a volume integral

$$\int_B f(x,y,z)\,dV,$$

carry out the following steps.

(a) Draw two diagrams, showing the region of integration B with the equations of the upper and lower boundaries marked, and the projection S of this region onto the (x,y)-plane.

(b) Within the region B, draw a column perpendicular to the (x,y)-plane and determine the limits of the z-integration, $z=\gamma(x,y)$ and $z=\psi(x,y)$, say.

(c) Evaluate the single integral of $f(x,y,z)$ over z between $z=\gamma(x,y)$ and $z=\psi(x,y)$, keeping x and y constant, to find the function $h(x,y)$ defined by

$$h(x,y) = \int_{z=\gamma(x,y)}^{z=\psi(x,y)} f(x,y,z)\,dz.$$

(d) Evaluate the area integral of $h(x,y)$ over the region S using the procedure for evaluating area integrals.

15. Suppose that a surface is described in the form $z = f(x, y)$. To find the area Σ of a portion of the surface lying over the region S in the (x, y)-plane, carry out the following steps.

    (a) Form the function

    $$g(x, y) = \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}.$$

    (b) Calculate the area integral

    $$\Sigma = \int_S g(x, y)\, dA,$$

    as in the procedure for evaluating area integrals.

Unit 18 Reviewing the model

1. The **method of dimensional analysis** aims to find dimensionally consistent relationships by carrying out the following steps.

    (a) List the parameters $y, x_1, x_2, \ldots, x_n$ that represent the important features involved in the situation being modelled, and determine their dimensions.

    (b) Assume a relationship involving the powers of these parameters, namely

    $$y = k\, x_1^{\alpha}\, x_2^{\beta}\, x_3^{\gamma} \cdots x_n^{\nu},$$

    where k is a dimensionless constant.

    (c) Use the principle of dimensional consistency to write

    $$[y] = [x_1]^{\alpha}\, [x_2]^{\beta}\, [x_3]^{\gamma} \cdots [x_n]^{\nu},$$

    and equate the powers of M, L and T on both sides of this equation. (As usual, $[x]$ denotes the dimensions of the quantity x.)

    (d) Solve the four simultaneous equations obtained in Step (c) for the powers $\alpha, \beta, \ldots, \nu$. This solution will usually involve unknown parameters.

    (e) Substitute for the powers in the expression in Step (b), and rewrite it in terms of dimensionless groups. Hence write down a general expression for y, which will usually involve an unknown function of these dimensionless groups.

2. The **absolute sensitivity** of a solution value y to a parameter value x is defined as $\delta y/\delta x$, where δy is the change in y caused by a small change δx in the value of the parameter x.

    The **relative sensitivity** of a solution value y to a parameter value x is defined as

    $$\frac{\delta y/y}{\delta x/x} = \frac{x}{y}\frac{\delta y}{\delta x},$$

    where $\delta y/y$ is the relative (or proportionate) change in y for a small relative change $\delta x/x$ in the value of the parameter x.

3. For the analytical approach to calculating sensitivity, where $y = f(x_1, x_2, \ldots, x_n)$, the absolute sensitivity of y with respect to x_i for infinitesimally small changes in x_i is $\partial y/\partial x_i$; the relative sensitivity is
$$\frac{x_i}{y}\frac{\partial y}{\partial x_i}.$$
4. Suppose that small absolute or relative changes are made in the parameters of a model. The model is absolutely (relatively) sensitive if it is possible for the absolute (relative) change in the solution to be significantly larger than the absolute (relative) change in one or more of the parameters.

   Usually the interpretation of *significantly larger* is dependent on the context. However, for the sake of clarity and certainty, we adopt the module convention. A model is judged to be:
   - absolutely (relatively) insensitive with respect to a particular parameter if a small absolute (relative) change in the parameter is magnified by a factor less than or equal to 5
   - neither absolutely (relatively) sensitive nor absolutely (relatively) insensitive with respect to a particular parameter if a small absolute (relative) change in the parameter is magnified by a factor of between 5 and 10
   - absolutely (relatively) sensitive with respect to a particular parameter if a small absolute (relative) change in the parameter is magnified by a factor greater than or equal to 10.

Unit 19 Systems of particles

1. An n-**particle system** is a system modelled so that the total mass of the system is divided among n particles.
2. The **centre of mass** of an n-particle system, whose particles have masses $m_1, m_2, \ldots, m_n$, and position vectors $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n$ relative to an origin O, has position vector
$$\mathbf{r}_G = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_n\mathbf{r}_n}{m_1 + m_2 + \cdots + m_n} = \frac{\sum_{i=1}^{n} m_i\mathbf{r}_i}{M}$$
relative to O, where $M = \sum_{i=1}^{n} m_i$ is the total mass of the system.
3. In an n-particle system, two kinds of force act: **internal forces** arise from the interactions of the constituent particles with each other, and **external forces** are due to agencies outside the system.
4. **Newton's third law** of mechanics states that for each force exerted by one object on another, there is a force of equal magnitude acting in the opposite direction, exerted by the second object on the first. The internal forces in an n-particle system obey Newton's third law.
5. The motion of the centre of mass $\mathbf{r}_G$ of an n-particle system of total mass M subject to external forces with sum $\mathbf{F}^{\text{ext}}$ is given by a form of Newton's second law as
$$\mathbf{F}^{\text{ext}} = M\ddot{\mathbf{r}}_G.$$

6. A **homogeneous rigid body** is a rigid body of uniform density.
7. Consider an object with a square base resting on a flat horizontal surface. The object will topple over if and only if the centre of mass of the object is not vertically above the base.
8. The coordinates (x_G, y_G) of the **centre of mass of a lamina** S are

$$x_G = \frac{\int_S x f \, dA}{M} = \frac{\int_S x f \, dA}{\int_S f \, dA} \quad \text{and} \quad y_G = \frac{\int_S y f \, dA}{M} = \frac{\int_S y f \, dA}{\int_S f \, dA}.$$

   For a uniform lamina, the surface density f is a constant and can be taken outside the integral.
9. The total gravitational potential energy of an n-particle system is equal to the gravitational potential energy of a single particle, whose mass is the total mass of the system, located at the system's centre of mass.
10. The **linear momentum p** of a particle with position vector $\mathbf{r}$ is a vector quantity given as the product of its mass m and its velocity $\dot{\mathbf{r}}$, so

$$\mathbf{p} = m\dot{\mathbf{r}}.$$

    The **linear momentum P** of an n-particle system is the vector sum of the linear momenta of the individual particles, so

$$\mathbf{P} = \sum_{i=1}^{n} m_i \dot{\mathbf{r}}_i = M\dot{\mathbf{r}}_G,$$

    where m_i and $\mathbf{r}_i$ represent the mass and position of particle i, M is the total mass of the system, and $\mathbf{r}_G$ is its centre of mass.
11. The total linear momentum of an n-particle system is not affected by (instantaneous) collisions among the particles, which is known as the **principle of conservation of linear momentum**. Furthermore, in the absence of external forces, the total linear momentum of the system remains constant.
12. If the kinetic energy of a system is the same before and after a collision within the system, then the collision is said to be **elastic**; otherwise, it is said to be **inelastic**.
13. **Newton's law of restitution** states that in an (instantaneous) collision between two smooth non-rotating objects, where the area of contact at the moment of impact lies on a common tangent plane, the velocities parallel to the tangent plane remain unchanged before and after impact. Also, we have

$$\begin{pmatrix}\text{relative velocity component}\\ \text{parallel to the common}\\ \text{normal after impact}\end{pmatrix} = -e \times \begin{pmatrix}\text{relative velocity component}\\ \text{parallel to the common}\\ \text{normal before impact}\end{pmatrix},$$

    where e is the **coefficient of restitution** for a collision between the two objects.

Unit 20 Circular motion

1. For a particle moving in a circle of radius R with **constant angular speed** ω:
   - the speed (the magnitude of the velocity) is a constant, $|\mathbf{v}| = \omega R$
   - the direction of the velocity is tangential to the circle
   - the acceleration has constant magnitude, $|\mathbf{a}| = \omega^2 R = |\mathbf{v}|^2/R$
   - the direction of the acceleration is towards the centre of the circle (this inward component is commonly called the **centripetal acceleration**).
2. **Differentiation of vector functions** obeys the following rules, where the scalar h and the vectors $\mathbf{f}$ and $\mathbf{g}$ are functions of the scalar variable t:
$$\frac{d}{dt}(\mathbf{f}+\mathbf{g}) = \frac{d\mathbf{f}}{dt} + \frac{d\mathbf{g}}{dt},$$
$$\frac{d}{dt}(h\mathbf{f}) = \frac{dh}{dt}\,\mathbf{f} + h\,\frac{d\mathbf{f}}{dt},$$
$$\frac{d}{dt}(\mathbf{f}\cdot\mathbf{g}) = \frac{d\mathbf{f}}{dt}\cdot\mathbf{g} + \mathbf{f}\cdot\frac{d\mathbf{g}}{dt},$$
$$\frac{d}{dt}(\mathbf{f}\times\mathbf{g}) = \frac{d\mathbf{f}}{dt}\times\mathbf{g} + \mathbf{f}\times\frac{d\mathbf{g}}{dt}.$$
   The rules for differentiating products bear a strong resemblance to the familiar rule for differentiating the product of two scalar functions. However, the order of the terms in the cross products is important.
3. Given a polar coordinate system and a two-dimensional Cartesian coordinate system with the same origin O, where the positive x-axis corresponds to the axis from which the polar coordinate θ is measured, the **polar unit vectors** $\mathbf{e}_r$ (the **radial direction**) and $\mathbf{e}_\theta$ (the **tangential direction**) are related to the Cartesian unit vectors $\mathbf{i}$ and $\mathbf{j}$ by
$$\mathbf{e}_r = (\cos\theta)\mathbf{i} + (\sin\theta)\mathbf{j},$$
$$\mathbf{e}_\theta = (-\sin\theta)\mathbf{i} + (\cos\theta)\mathbf{j}.$$
   The derivatives with respect to time of the polar unit vectors are
$$\dot{\mathbf{e}}_r = \dot{\theta}\mathbf{e}_\theta,$$
$$\dot{\mathbf{e}}_\theta = -\dot{\theta}\mathbf{e}_r.$$
   The polar unit vectors are mutually **orthogonal** (i.e. perpendicular).

4. In polar coordinates, the position, velocity and acceleration of a particle moving in the circle $r = R$ are, respectively,

$$\mathbf{r} = R\mathbf{e}_r,$$
$$\dot{\mathbf{r}} = R\dot{\theta}\mathbf{e}_\theta,$$
$$\ddot{\mathbf{r}} = -R\dot{\theta}^2\mathbf{e}_r + R\ddot{\theta}\mathbf{e}_\theta.$$

The quantity $\dot{\theta}$ is the particle's **rate of rotation**, and the quantity $\omega = |\dot{\theta}|$ is the particle's **angular speed**. The SI units of angular speed are $\mathrm{rad\,s^{-1}}$.

5. **Uniform circular motion** is when a particle is moving in a circle of constant radius R, with centre at the origin O, having a constant rate of rotation $\dot{\theta}$ and hence a constant angular speed $\omega = |\dot{\theta}|$.

- The velocity of the particle is

$$\dot{\mathbf{r}} = R\dot{\theta}\mathbf{e}_\theta,$$

i.e. the velocity of the particle is purely tangential and has constant magnitude $v = |\dot{\mathbf{r}}| = R|\dot{\theta}| = R\omega$.

- The acceleration of the particle is

$$\ddot{\mathbf{r}} = -R\dot{\theta}^2\mathbf{e}_r = -R\omega^2\mathbf{e}_r = -\frac{v^2}{R}\,\mathbf{e}_r,$$

i.e. the acceleration is directed towards the centre of the circle and has constant magnitude $|\ddot{\mathbf{r}}| = R\dot{\theta}^2 = R\omega^2 = v^2/R$.

- The time taken for one complete revolution of the circle is

$$T = \frac{2\pi}{\omega}.$$

Circular motion in which $\dot{\theta}$ is not constant is called **non-uniform circular motion**.

6. **Newton's law of universal gravitation** states that the gravitational force of attraction exerted on a particle of mass m_1 by a particle of mass m_2 is

$$\mathbf{F} = -\frac{Gm_1m_2}{|\mathbf{r}|^3}\,\mathbf{r},$$

where $\mathbf{r}$ is the position vector of the particle of mass m_1, relative to the particle of mass m_2, and $G = 6.674 \times 10^{-11}$ (in $\mathrm{N\,m^2\,kg^{-2}}$) is the **gravitational constant**.

7. The **angular velocity** $\boldsymbol{\omega}$ of a particle moving in a circle, or of a rigid body rotating about a fixed axis, is a vector with magnitude equal to the angular speed and with direction along the axis of rotation in the sense given by the right-hand grip rule.

   The angular velocity $\boldsymbol{\omega}$ of a particle moving in a circle about an axis in the $\mathbf{k}$-direction, with angular rate of rotation given by $\dot{\theta}$, is

   $$\boldsymbol{\omega} = \dot{\theta}\mathbf{k}.$$

8. The velocity $\mathbf{v}$ of a particle moving in a circle with angular velocity $\boldsymbol{\omega}$ is

   $$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r},$$

   where $\mathbf{r}$ is the position vector of the particle relative to an origin on the axis of rotation.

9. For a particle that has linear momentum $\mathbf{p} = m\dot{\mathbf{r}}$ and position vector $\mathbf{r}$ relative to an origin O, its **angular momentum** $\mathbf{L}$ about O is

   $$\mathbf{L} = \mathbf{r} \times \mathbf{p} = \mathbf{r} \times m\dot{\mathbf{r}}.$$

10. The **torque law** for a particle states that the rate of change of a particle's angular momentum about a fixed point is equal to the applied torque about that point, i.e.

    $$\dot{\mathbf{L}} = \boldsymbol{\Gamma}.$$

    (The torque $\boldsymbol{\Gamma}$ of a force $\mathbf{F}$ acting at the point with position vector $\mathbf{r}$, relative to the given origin, is the vector $\mathbf{r} \times \mathbf{F}$.)

11. The angular momentum $\mathbf{L}$ about the origin of a particle of mass m with motion (not necessarily circular) confined to the (x, y)-plane is

    $$\mathbf{L} = mr^2\dot{\theta}\mathbf{k}.$$

12. The **law of conservation of angular momentum** states that if the total torque acting on a particle about a fixed point is zero, then the angular momentum of the particle about that point is constant. In other words, if $\boldsymbol{\Gamma} = \mathbf{0}$, then $\mathbf{L}$ is constant.

Unit 21 Rotating bodies and angular momentum

1. An **extended body** is an object that has size, i.e. one or more of length, breadth and depth, and whose size and shape are to be taken into consideration when its motion is being investigated. A **rigid body** is an extended body whose size and shape do not change. We model extended bodies as systems of interacting particles, and rigid bodies as systems of particles such that all the inter-particle distances remain constant.

2. **Newton's third law** may be re-stated as follows. The force $\mathbf{I}_{12}$ exerted on particle 1 by particle 2 is equal in magnitude but opposite in direction to the force $\mathbf{I}_{21}$ exerted on particle 2 by particle 1, with both forces acting along the line joining the two particles. An equivalent condition in symbols is

   $$\mathbf{I}_{12} + \mathbf{I}_{21} = \mathbf{0} \quad \text{and} \quad \boldsymbol{\Gamma}_{12} + \boldsymbol{\Gamma}_{21} = \mathbf{0},$$

   where $\boldsymbol{\Gamma}_{12} = \mathbf{r}_1 \times \mathbf{I}_{12}$ and $\boldsymbol{\Gamma}_{21} = \mathbf{r}_2 \times \mathbf{I}_{21}$, with $\mathbf{r}_1$ and $\mathbf{r}_2$ being the position vectors of particles 1 and 2, respectively, relative to the origin.

3. Consider a system of n particles. Let $\mathbf{r}_i$ be the position vector (relative to a fixed origin O) of particle i, and let $\mathbf{F}_i$ be the total external force on particle i, for $i = 1, 2, \ldots, n$. The total external torque on the system about O is

   $$\boldsymbol{\Gamma} = \sum_{i=1}^{n} \mathbf{r}_i \times \mathbf{F}_i.$$

   The **torque law** states that the rate of change of the total angular momentum $\mathbf{L}$ of the system about O equals the total external torque acting on the system, i.e.

   $$\dot{\mathbf{L}} = \boldsymbol{\Gamma}.$$

   In particular, when the total external torque about O is zero, the total angular momentum vector about O is conserved, i.e. it is constant.

4. The **moment of inertia** of an extended body modelled as a system of n particles, about an axis, is given by

   $$I = \sum_{i=1}^{n} m_i d_i^2,$$

   where m_i is the mass of the ith particle, and d_i is the distance of the ith particle from the axis.

   If the body is rigid, and the axis is fixed relative to the body, then I is a constant. Its value depends on the size and shape of the body.

5. Suppose that a rigid body is rotating about a fixed axis with angular velocity $\boldsymbol{\omega}$. Let $\mathbf{L}$ be the **angular momentum** about a point O on the axis, and let L_{axis} be the component of $\mathbf{L}$ in the direction of the axis. Then

   $$L_{\text{axis}} = I\dot{\theta},$$

   where I is the moment of inertia of the body about the axis of rotation, and θ is the angular displacement (from some fixed line normal to the axis of rotation).

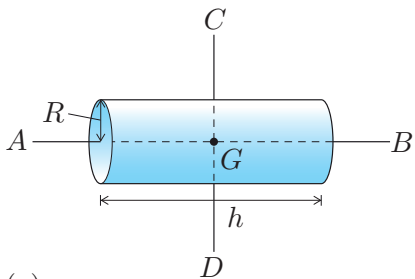
6. Suppose that a rigid body is rotating about a fixed axis, and that its angular displacement (from some fixed line normal to the axis) is θ. Then the **equation of rotational motion** is

   $$I\ddot{\theta} = \Gamma_{\text{axis}},$$

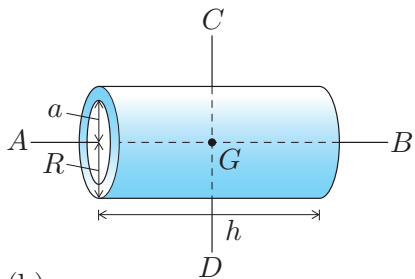
   where Γ_{axis} is the component of the total external torque on the body in the direction of the axis of rotation.

7. Moments of inertia of rigid bodies with a number of common geometric shapes are given in the table below. Each rigid body is assumed to have mass M and uniform density. In each case the axis passes through the centre of mass of the body.

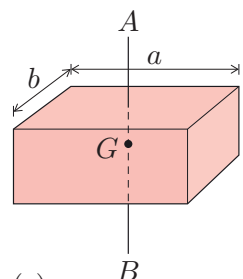
Object	Axis	Dimensions	Moment of inertia	Figure
Solid cylinder	Axis of cylinder	Radius R	$\frac{1}{2}MR^2$	(a) below, axis AB
Solid cylinder	Normal to axis of cylinder	Radius R, length h	$\frac{1}{4}MR^2 + \frac{1}{12}Mh^2$	(a) below, axis CD
Hollow cylinder	Axis of cylinder	Inner radius a, outer radius R	$\frac{1}{2}M(R^2 + a^2)$	(b) below, axis AB
Hollow cylinder	Normal to axis of cylinder	Inner radius a, outer radius R, length h	$\frac{1}{4}M(R^2 + a^2) + \frac{1}{12}Mh^2$	(b) below, axis CD
Solid rectangular cuboid	Normal to one pair of faces	Faces normal to axis have sides of lengths a, b	$\frac{1}{12}M(a^2 + b^2)$	(c) below, axis AB
Thin straight rod	Normal to rod	Length h	$\frac{1}{12}Mh^2$	
Solid sphere	Through centre	Radius R	$\frac{2}{5}MR^2$	
Hollow sphere	Through centre	Inner radius a, outer radius R	$\frac{2}{5}M\frac{R^5 - a^5}{R^3 - a^3}$	
Thin spherical shell	Through centre	Radius R	$\frac{2}{3}MR^2$	



(a)



(b)



(c)

8. Suppose that I_{AB} is the moment of inertia of a rigid body of mass M about a line AB. Let EF be a line through the centre of mass of the body that is parallel to AB, let the distance between the lines AB and EF be d, and let I_{EF} be the moment of inertia of the body about EF. Then the **parallel axes theorem** gives

$$I_{AB} = I_{EF} + Md^2.$$

9. Suppose that a rigid body is rotating with angular speed ω about a fixed axis. Let I be the moment of inertia of the body about the axis of rotation. Then the **kinetic energy** of the body is

$$T = \tfrac{1}{2} I\omega^2.$$

10. If $\mathbf{r}_i^{\text{rel}}$ denotes the position of the ith particle relative to the centre of mass, and $\mathbf{F}_i$ denotes total external force on the ith particle, then the **total external torque** on the system relative to the centre of mass is defined by

$$\boldsymbol{\Gamma}^{\text{rel}} = \sum_{i=1}^{n} \mathbf{r}_i^{\text{rel}} \times \mathbf{F}_i,$$

and the **total angular momentum** relative to the centre of mass is defined as

$$\mathbf{L}^{\text{rel}} = \sum_{i=1}^{n} \mathbf{r}_i^{\text{rel}} \times m_i \dot{\mathbf{r}}_i^{\text{rel}}.$$

11. Let $\mathbf{R}$ be the position vector of the centre of mass of a body relative to some fixed point O, let M be the total mass of the body, and let $\mathbf{F}$ be the total external force on the body.

The **angular momentum decomposition theorem** states that if the total angular momentum of the body relative to the centre of mass is $\mathbf{L}^{\text{rel}}$, then the total angular momentum of the body relative to O is given by

$$\mathbf{L} = \mathbf{R} \times M\dot{\mathbf{R}} + \mathbf{L}^{\text{rel}}.$$

The **torque decomposition theorem** states that if $\boldsymbol{\Gamma}^{\text{rel}}$ is the total external torque on the body relative to the centre of mass, then the total external torque on the body about O is given by

$$\boldsymbol{\Gamma} = \mathbf{R} \times \mathbf{F} + \boldsymbol{\Gamma}^{\text{rel}}.$$

12. The **total external torque** on an extended body relative to its centre of mass is equal to the rate of change of the total angular momentum relative to the centre of mass. So if $\boldsymbol{\Gamma}^{\text{rel}}$ is the total external torque on the body relative to the centre of mass, and $\mathbf{L}^{\text{rel}}$ is the total angular momentum of the body relative to the centre of mass then

$$\boldsymbol{\Gamma}^{\text{rel}} = \dot{\mathbf{L}}^{\text{rel}}.$$

13. Suppose that each particle in a system of n particles is subject to an external force of the form $cm_i\mathbf{k}$, where c is a constant, m_i is the mass of the ith particle, and $\mathbf{k}$ is a fixed vector, and that there are no other external forces on the system. Then a special case of the **law of conservation of angular momentum** is that $\boldsymbol{\Gamma}^{\text{rel}} = \mathbf{0}$, so $\mathbf{L}^{\text{rel}}$ is constant.

14. The **kinetic energy decomposition theorem** states that the kinetic energy of an extended body is equal to the kinetic energy of an equivalent particle that has the velocity of the body's centre of mass, plus the sum of the kinetic energies due to the motion, relative to the centre of mass, of all the body's constituent particles.

15. Consider a rigid body of mass M rotating about an axis of fixed orientation through its centre of mass, with angular velocity $\boldsymbol{\omega} = \dot{\theta}\mathbf{k}$, where $\mathbf{k}$ is a fixed unit vector and θ is the anticlockwise angular displacement relative to the axis of rotation. Let I be the moment of inertia of the body about the axis of rotation.

    The $\mathbf{k}$-component $L^{\text{rel}}_{\text{axis}}$ $(= \mathbf{L}^{\text{rel}} \cdot \mathbf{k})$ of the angular momentum of the body relative to the centre of mass is given by

    $$L^{\text{rel}}_{\text{axis}} = I\dot{\theta}.$$

    The **equation of relative rotational motion** of the body is

    $$\Gamma^{\text{rel}}_{\text{axis}} = I\ddot{\theta},$$

    where $\Gamma^{\text{rel}}_{\text{axis}}$ $(= \boldsymbol{\Gamma}^{\text{rel}} \cdot \mathbf{k})$ is the $\mathbf{k}$-component of the total external torque relative to the centre of mass.

    The kinetic energy T of the body is the sum of the kinetic energy of an equivalent particle at the centre of mass and the rotational kinetic energy relative to the centre of mass, i.e.

    $$T = \tfrac{1}{2}M\left|\dot{\mathbf{R}}\right|^2 + \tfrac{1}{2}I\omega^2,$$

    where $\mathbf{R}$ is the position vector of the centre of mass, and ω $(= |\dot{\theta}|)$ is the angular speed.

16. When a body with a circular cross-section (a disc, cylinder or sphere) rolls without slipping on a surface, the curves traced out by the instantaneous point of contact on the body and on the surface must have the same lengths (this is known as the **rolling condition**). In other words, the point on the cylinder that is in contact with the slope at any instant does not move relative to the slope.

    Consider, for example, a cylinder of radius R rolling a distance x on a plane while it turns through an angle θ about its axis in such a way that x increases as θ increases. Then, if no slipping occurs,

    $$R\theta = x,$$

    which is the rolling condition in this case.

17. In order for a body to roll without slipping, a frictional force $\mathbf{F}$ must act at the point of contact, just as in the case of a static object on an inclined plane. The condition that the body rolls rather than slips is $|\mathbf{F}| \leq \mu|\mathbf{N}|$, where μ is the coefficient of static friction, and $\mathbf{N}$ is the normal reaction at the point of contact.

Index