

Course Companion

A-Level Mathematics

Every topic handout in one volume

exercises.lol

Contents

1. Pure Mathematics 1	2
2. Pure Mathematics 2	20
3. Pure Mathematics 3	29
4. Mechanics	38
5. Probability Statistics 1	46
6. Probability Statistics 2	54

Pure Mathematics 1

A-Level Mathematics

This handout covers Topic 1: **Pure Mathematics** and calculus. 1. It is the algebra core of the course. Each ## section is one syllabus subtopic.

Quadratics



A suspension bridge: the main cable hangs in a parabola.

A **quadratic** is an expression of the form $ax^2 + bx + c$, where $a \neq 0$. The letters a , b , c are the **coefficients** (the fixed numbers). Much of this section is about solving the **equation** $ax^2 + bx + c = 0$.

Completing the square

To **complete the square** means to write the quadratic in the form

$$a(x + p)^2 + q.$$

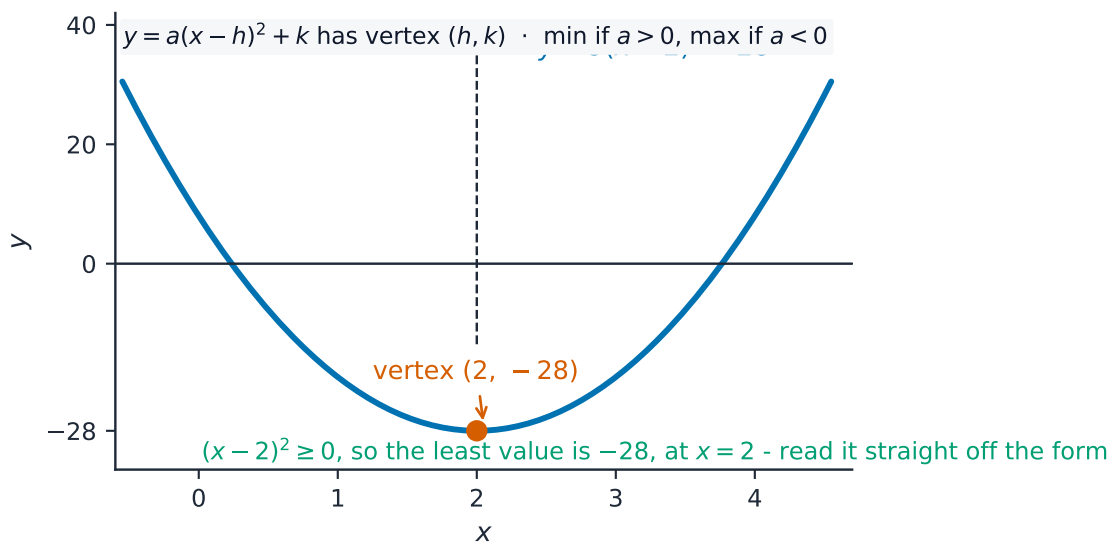
This form is useful: it shows the **vertex** (turning point) of the curve at $(-p, q)$, and it gives a quick way to solve the equation.

Worked example. Write $9x^2 - 36x + 8$ in the form $p(x + q)^2 + r$.

Take the factor 9 out of the first two terms, then complete the square inside:

$$\begin{aligned} 9x^2 - 36x + 8 &= 9(x^2 - 4x) + 8 \\ &= 9((x - 2)^2 - 4) + 8 \\ &= 9(x - 2)^2 - 36 + 8 = 9(x - 2)^2 - 28. \end{aligned}$$

So $p = 9$, $q = -2$, $r = -28$.



The completed square hands you the vertex: least value -28 at x equals 2

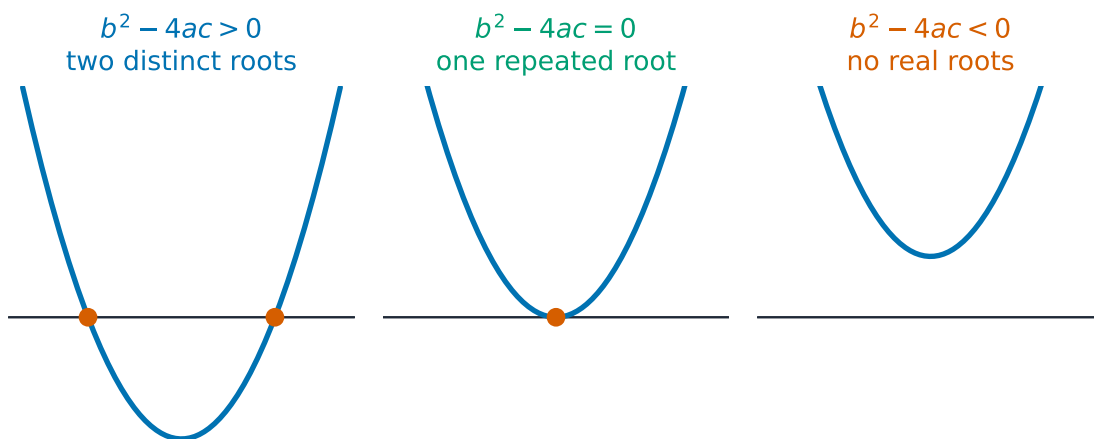
The discriminant

The **discriminant** of $ax^2 + bx + c$ is

$$\Delta = b^2 - 4ac.$$

It tells you how many **real roots** (real solutions) the equation $ax^2 + bx + c = 0$ has:

Discriminant	Roots
$b^2 - 4ac > 0$	two distinct real roots
$b^2 - 4ac = 0$	one repeated real root
$b^2 - 4ac < 0$	no real roots



$$ax^2 + bx + c = 0 \Rightarrow x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (\text{real roots need } b^2 - 4ac \geq 0)$$

The sign of $b^2 - 4ac$ decides how many times the parabola meets the x -axis.

Worked example. Find the values of the constant k for which $3kx^2 + (k+8)x + 3 = 0$ has two distinct real roots.

Here $a = 3k$, $b = k + 8$, $c = 3$. For two distinct real roots you need $b^2 - 4ac > 0$:

$$(k+8)^2 - 4(3k)(3) > 0 \Rightarrow k^2 + 16k + 64 - 36k > 0 \Rightarrow k^2 - 20k + 64 > 0.$$

Factorise: $(k-4)(k-16) > 0$, so $k < 4$ or $k > 16$. You also need $a \neq 0$, so $k \neq 0$.

Quadratic equations and inequalities

To solve a **quadratic equation** , factorise, complete the square, or use the formula

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

To solve a **quadratic inequality** such as $(k - 4)(k - 16) > 0$, find the two roots, then decide which side of each root makes the statement true. A sketch of the parabola helps: the curve is above the x -axis (positive) outside the roots and below it (negative) between them.

Simultaneous equations

To solve a pair of **simultaneous equations** where one is linear and one is quadratic, use **substitution** : rearrange the linear equation for one letter, then put that into the quadratic. This gives a single quadratic to solve.

Equations that are quadratic in disguise

Some equations are quadratic in **some function of x** . For example $x^4 - 5x^2 + 4 = 0$ is quadratic in x^2 : let $u = x^2$, solve $u^2 - 5u + 4 = 0$, then go back to x . You will use this idea again in trigonometry.

Functions

A **function** is a rule that sends each input to exactly one output. Write it as $f(x)$. The **composition of functions** $fg(x)$ means apply g first, then apply f to the result.

- The **domain** is the set of allowed inputs x .
- The **range** is the set of outputs the function actually produces.

A function is **one-one** if different inputs always give different outputs. (No output is repeated.) Only a one-one function has an **inverse function** f^{-1} , which reverses the rule.

The **composition** of two functions means doing one after the other. $fg(x)$ means "do g first, then f ": $fg(x) = f(g(x))$. The composite function fg exists only when the range of g lies inside the domain of f .

Finding an inverse

To find f^{-1} : write $y = f(x)$, make x the subject, then swap letters.

Worked example. The function $f(x) = (x + 3)^2 - 12$ is defined for $x \geq 0$. Find $f^{-1}(x)$.

Write $y = (x + 3)^2 - 12$ and solve for x :

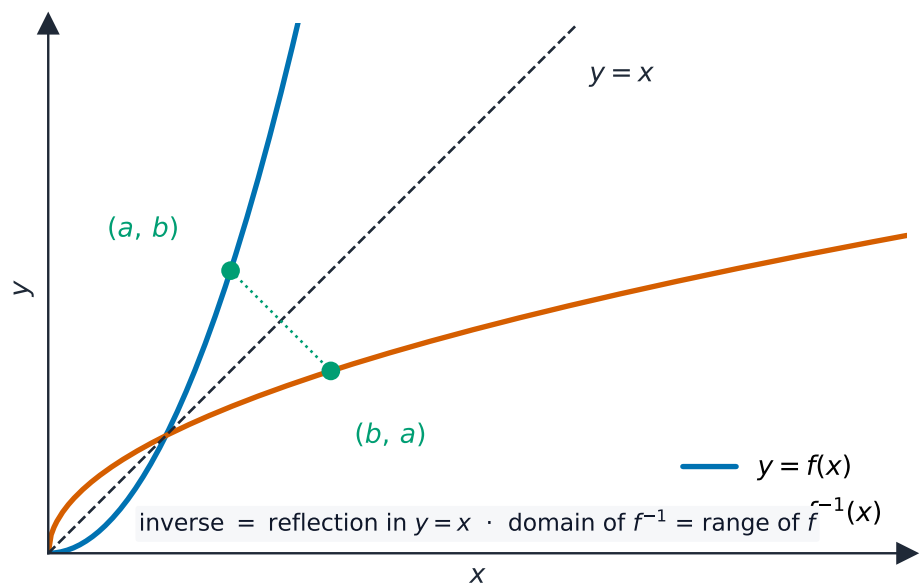
$$(x + 3)^2 = y + 12 \Rightarrow x + 3 = \sqrt{y + 12} \Rightarrow x = \sqrt{y + 12} - 3.$$

You take the **positive** square root because $x \geq 0$ means $x + 3 \geq 3 > 0$. So

$$f^{-1}(x) = \sqrt{x + 12} - 3.$$

Graphs of inverses and transformations

The graph of $y = f^{-1}(x)$ is the **reflection** of $y = f(x)$ in the line $y = x$.

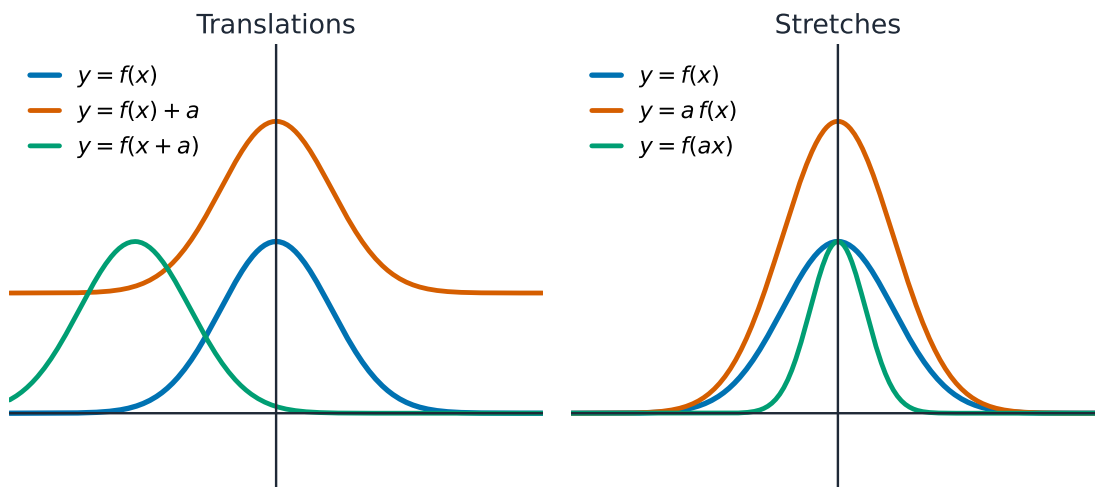


Reflecting $y = f(x)$ in the line $y = x$ gives its inverse; a point (a, b) becomes (b, a) .

You should know these **transformations** of $y = f(x)$:

New equation	Effect on the graph
$y = f(x) + a$	translation up by a
$y = f(x + a)$	translation left by a
$y = a f(x)$	stretch in the y -direction, scale factor a
$y = f(ax)$	stretch in the x -direction, scale factor $\frac{1}{a}$

When two transformations are combined, the order can matter. State each one fully (type, direction, and amount).



$f(x) + a$ moves up · $f(x + a)$ moves left · $af(x)$ stretch y · $f(ax)$ compress x if $a > 1$

Adding to the output or input slides the curve; a multiplier stretches it.

Coordinate geometry

Coordinate geometry

studies lines and circles using their equations.

Straight lines

The **gradient** (steepness) of the line joining (x_1, y_1) and (x_2, y_2) is

$$m = \frac{y_2 - y_1}{x_2 - x_1}.$$

You can write the **equation of a straight line**

in any of these forms:

$$y = mx + c, \quad y - y_1 = m(x - x_1), \quad ax + by + c = 0.$$

Two lines are **parallel** when their gradients are equal, and **perpendicular** (at right angles) when the product of their gradients is -1 .

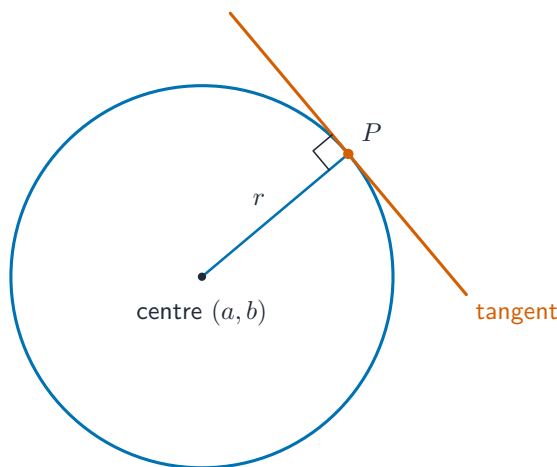
Circles

The **circle** with **centre** (a, b) and **radius** r has equation

$$(x - a)^2 + (y - b)^2 = r^2.$$

An expanded form like $x^2 + y^2 - 6x + 10y - 27 = 0$ is the same circle: complete the square in x and in y to find the centre and radius.

A **tangent** to a circle touches it at one point and is perpendicular to the radius at that point. This right-angle fact solves most circle problems.



A tangent touches the circle once and meets the radius at a right angle.

Worked example. The points $P(1, 1)$ and $Q(7, 11)$ are the ends of a **diameter** of a circle. Find the equation of the circle.

The centre is the midpoint of PQ :

$$\left(\frac{1+7}{2}, \frac{1+11}{2} \right) = (4, 6).$$

The radius is half the length of PQ :

$$r = \frac{1}{2} \sqrt{(7-1)^2 + (11-1)^2} = \frac{1}{2} \sqrt{36 + 100} = \frac{1}{2} \sqrt{136} = \sqrt{34}.$$

So the circle is $(x - 4)^2 + (y - 6)^2 = 34$.

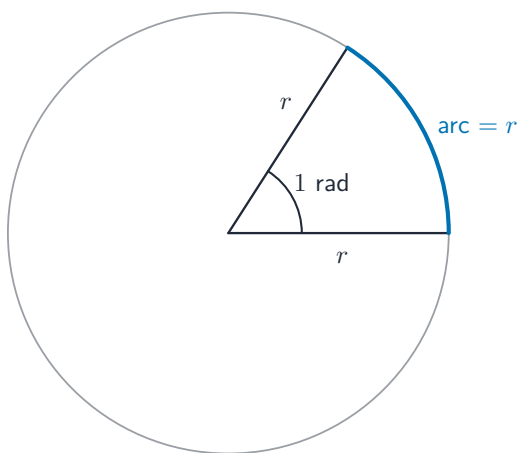
Circular measure

Radians

A **radian** is another way to measure angles. One radian is the angle at the centre of a circle that cuts off an **arc** equal in length to the radius. The link between radians and **degrees** is

$$\pi \text{ radians} = 180^\circ.$$

So to change degrees to radians, multiply by $\frac{\pi}{180}$; to change radians to degrees, multiply by $\frac{180}{\pi}$.



one **radian**: the angle whose arc is exactly one radius long

$$\pi \text{ rad} = 180^\circ, \text{ so } 1 \text{ rad} \approx 57.3^\circ$$

$$\text{degrees} \rightarrow \text{radians: } \times \frac{\pi}{180} \quad \text{radians} \rightarrow \text{degrees: } \times \frac{180}{\pi}$$

One radian: the angle whose arc equals the radius

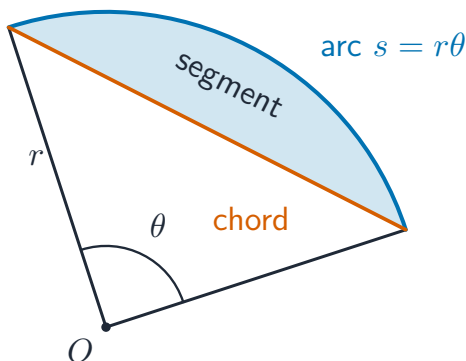
Arc length and sector area

For a **sector** with radius r and angle θ **in radians**:

$$\text{arc length} = s = r\theta, \quad \text{sector area} = A = \frac{1}{2}r^2\theta.$$

A **chord** cuts the sector into a triangle and a **segment**. The segment area is the sector minus the triangle:

$$\text{segment} = \frac{1}{2}r^2\theta - \frac{1}{2}r^2 \sin \theta = \frac{1}{2}r^2(\theta - \sin \theta).$$



The shaded segment is the part of the sector between the chord and the arc.

Worked example. A sector has centre O and the angle at O is $\frac{2}{3}\pi$ radians. Show that the segment cut off by the chord has area about $0.614r^2$.

$$\text{segment} = \frac{1}{2}r^2 \left(\frac{2}{3}\pi - \sin \frac{2}{3}\pi \right) = \frac{1}{2}r^2(2.0944 - 0.8660) = \frac{1}{2}r^2(1.2284) \approx 0.614r^2.$$

Trigonometry

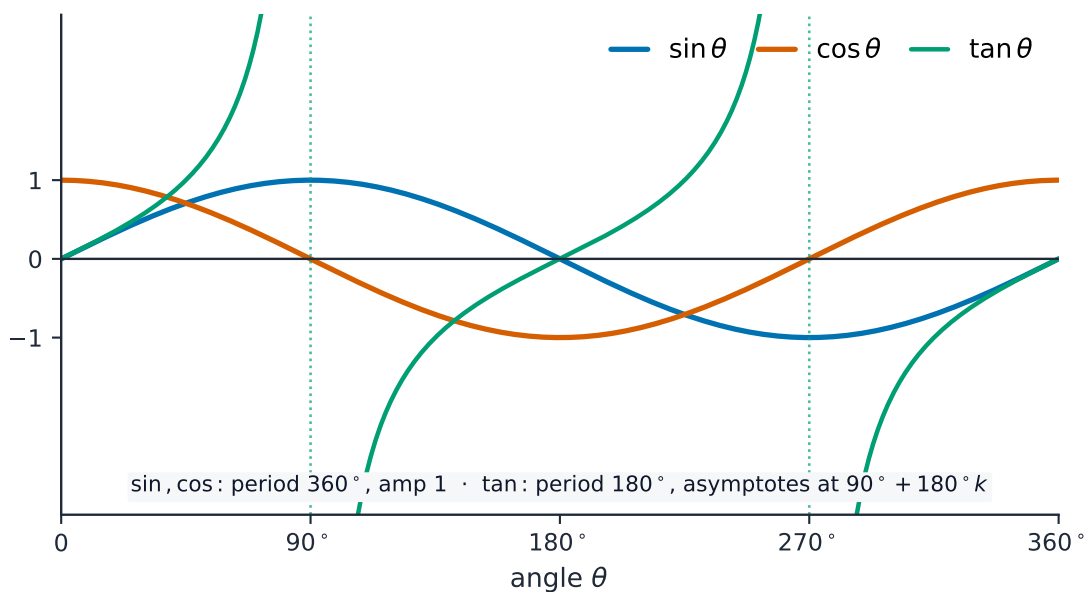


A Ferris wheel: a point on the rim rises and falls like a sine curve.

Graphs and exact values

You must know the shape of the graphs of the **sine**, **cosine** and **tangent function** (written sin, cos, tan). The sine and cosine graphs wave between -1 and 1 and repeat every 360° (2π). Learn these exact values:

θ	30°	45°	60°
$\sin \theta$	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$
$\cos \theta$	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$
$\tan \theta$	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$



Over one turn sin and cos stay between -1 and 1 ; tan shoots off at 90° and 270° .

The notations $\sin^{-1} x$, $\cos^{-1} x$, $\tan^{-1} x$ mean the **inverse** angle (the **principal value**).

Identities

An **identity** is true for every value of the angle. The two you must know are

$$\tan \theta \equiv \frac{\sin \theta}{\cos \theta}, \quad \sin^2 \theta + \cos^2 \theta \equiv 1.$$

Use them to rewrite an equation so it contains only one trig function.

Solving trigonometric equations

To solve a **trigonometric equation**, first reduce it to one function, then find every solution in the given interval.

Worked example. Solve $6 \sin \theta = 1 + \frac{2}{\sin \theta}$ for $-180^\circ < \theta < 180^\circ$.

Multiply through by $\sin \theta$ to clear the fraction. This makes a **quadratic in $\sin \theta$** :

$$6 \sin^2 \theta - \sin \theta - 2 = 0 \Rightarrow (3 \sin \theta - 2)(2 \sin \theta + 1) = 0.$$

So $\sin \theta = \frac{2}{3}$ or $\sin \theta = -\frac{1}{2}$.

- $\sin \theta = \frac{2}{3}$: $\theta = 41.8^\circ$ or $\theta = 180^\circ - 41.8^\circ = 138.2^\circ$.
- $\sin \theta = -\frac{1}{2}$: $\theta = -30^\circ$ or $\theta = -150^\circ$.

The four solutions are $\theta = -150^\circ, -30^\circ, 41.8^\circ, 138.2^\circ$.

Series

The binomial expansion

For a positive integer n , the **binomial expansion** is

$$(a + b)^n = a^n + \binom{n}{1} a^{n-1} b + \binom{n}{2} a^{n-2} b^2 + \cdots + b^n,$$

where $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ is a **binomial coefficient**.

Worked example. Find the first three terms, in ascending powers of x , of $(2 - px)^5$.

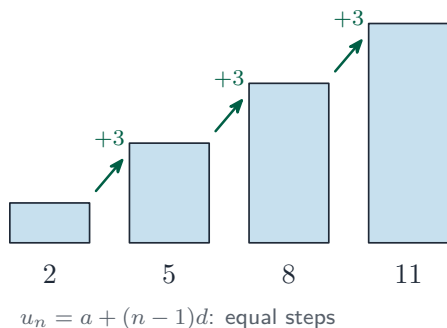
$$(2 - px)^5 = 2^5 + \binom{5}{1} 2^4 (-px) + \binom{5}{2} 2^3 (-px)^2 + \cdots = 32 - 80px + 80p^2x^2 + \cdots$$

Arithmetic and geometric progressions

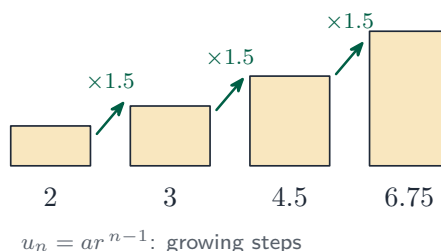
A **progression** (sequence) is a list of terms following a rule.

- An **arithmetic progression** (AP) adds a fixed **common difference** d each step. The n th **term** is $u_n = a + (n - 1)d$, and the sum of the first n terms is $S_n = \frac{n}{2}(2a + (n - 1)d)$.
- A **geometric progression** (GP) multiplies by a fixed **common ratio** r each step. The n th term is $u_n = ar^{n-1}$, and $S_n = \frac{a(1 - r^n)}{1 - r}$.

arithmetic: add d each step



geometric: multiply by r each step



An AP climbs in equal steps; a GP's steps grow by the same ratio

A GP is **convergent** — it **converges** (settles to a limit) — when $|r| < 1$. Then it has a **sum to infinity**

$$S_{\infty} = \frac{a}{1 - r}.$$

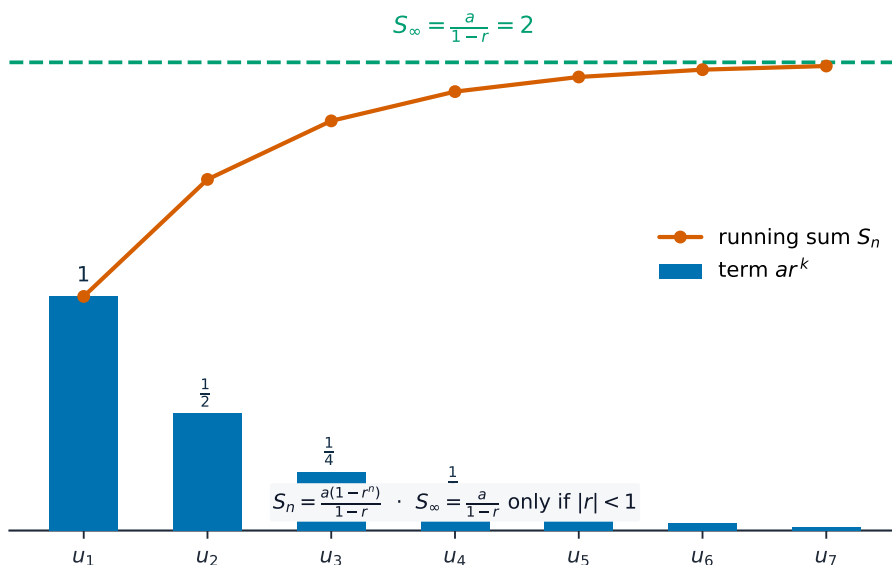
Worked example. The third term of a GP is 18 and the sum of the first three terms is 26. The common ratio is negative. Find the sum to infinity.

From $ar^2 = 18$ you get $a = \frac{18}{r^2}$. Put this into $a(1 + r + r^2) = 26$:

$$18(1 + r + r^2) = 26r^2 \Rightarrow 8r^2 - 18r - 18 = 0 \Rightarrow (4r + 3)(r - 3) = 0.$$

The ratio is negative, so $r = -\frac{3}{4}$ and $a = \frac{18}{(3/4)^2} = 32$. Then

$$S_{\infty} = \frac{32}{1 - (-\frac{3}{4})} = \frac{32}{\frac{7}{4}} = \frac{128}{7}.$$



*Convergent GP: bars shrink toward zero while the running sum approaches S infinity
 $= a/(1-r)$*

Differentiation

Differentiation finds the **gradient of a curve** at each point. The gradient is the **limit** of the gradients of shorter and shorter **chords**, called the **derivative**.

The rules

Write the derivative as $f'(x)$ or $\frac{dy}{dx}$. The basic rule is

$$\frac{d}{dx}(x^n) = nx^{n-1} \quad \text{for any rational } n.$$

Differentiate sums term by term, and use the **chain rule** for a function inside a function:

$$\frac{d}{dx} f(g(x)) = f'(g(x)) \cdot g'(x).$$

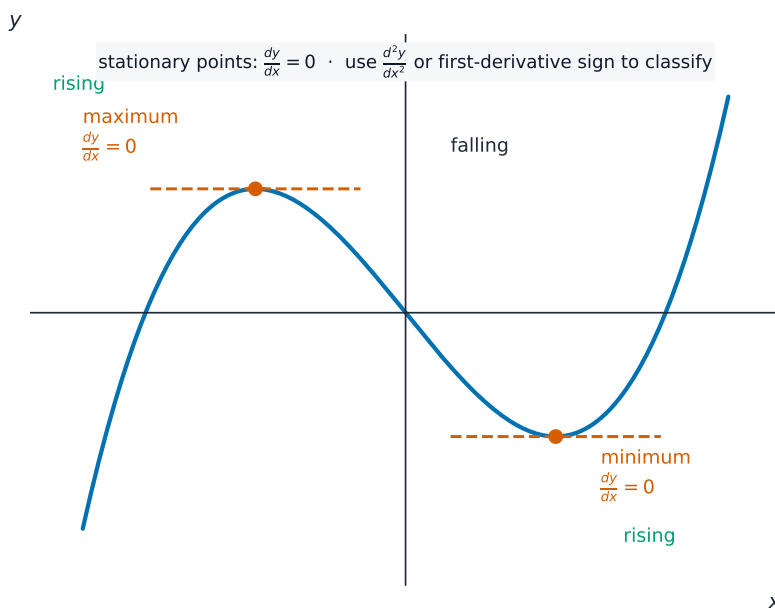
Differentiating again gives the **second derivative** $f''(x)$ or $\frac{d^2y}{dx^2}$.

Using the derivative

- **Tangent and normal.** The gradient of the curve at a point is the gradient of the **tangent** there. The **normal** is perpendicular to the tangent, so its gradient is $-\frac{1}{(\text{tangent gradient})}$.
- **Increasing or decreasing.** The function is an **increasing function** where $\frac{dy}{dx} > 0$, and a **decreasing function** where $\frac{dy}{dx} < 0$.
- **Rate of change.** A derivative is a **rate of change**. Linked rates use the chain rule, e.g. $\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt}$.

Stationary points

A **stationary point** is where $\frac{dy}{dx} = 0$. Test its nature with the second derivative: $f''(x) > 0$ gives a **minimum point**, and $f''(x) < 0$ gives a **maximum point**. When $f''(x) = 0$ the test is inconclusive: the point may be a **point of inflexion**, where the curve changes concavity (the way it bends) – check the sign of $\frac{dy}{dx}$ just before and after to decide.



At a maximum or a minimum the tangent is flat, so $\frac{dy}{dx} = 0$.

Worked example. The curve $y = 4x^{1/2} - x$ has a maximum point at $x = a$. Find a .

$$\frac{dy}{dx} = 2x^{-1/2} - 1 = 0 \Rightarrow \frac{2}{\sqrt{x}} = 1 \Rightarrow \sqrt{x} = 2 \Rightarrow x = 4.$$

So $a = 4$.

Integration

Integration is the reverse of differentiation. Reversing the power rule gives

$$\int (ax + b)^n dx = \frac{(ax + b)^{n+1}}{a(n+1)} + C \quad (n \neq -1).$$

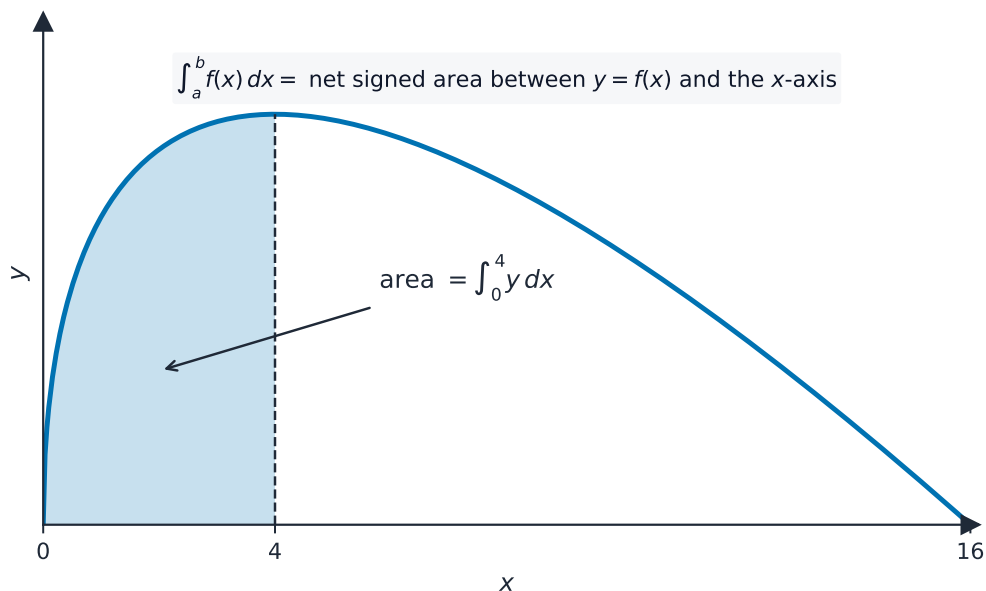
The $+C$ is the **constant of integration**. If you know one point on the curve, substitute it to find C .

Definite integrals and area

A **definite integral** has limits and gives a number:

$$\int_p^q f(x) dx = [F(x)]_p^q = F(q) - F(p).$$

The area of the **region** between a curve and the x -axis, from $x = p$ to $x = q$, is $\int_p^q y dx$. For the area between two curves, integrate (top curve $-$ bottom curve).



The definite integral $\int_0^4 y dx$ is the shaded area under the curve.

Worked example. The curve $y = 4x^{1/2} - x$ meets the x -axis again at $x = 16$. Find

the area between the curve and the x -axis from $x = 0$ to $x = 4$.

$$\int_0^4 (4x^{1/2} - x) dx = \left[\frac{8}{3}x^{3/2} - \frac{x^2}{2} \right]_0^4 = \frac{8}{3}(8) - \frac{16}{2} = \frac{64}{3} - 8 = \frac{40}{3}.$$

An **improper integral** has an infinite limit or an endpoint where the integrand is undefined; evaluate it as a **limit**. For example,

$$\int_1^\infty x^{-2} dx = \lim_{b \rightarrow \infty} \left[-\frac{1}{x} \right]_1^b = \lim_{b \rightarrow \infty} \left(1 - \frac{1}{b} \right) = 1,$$

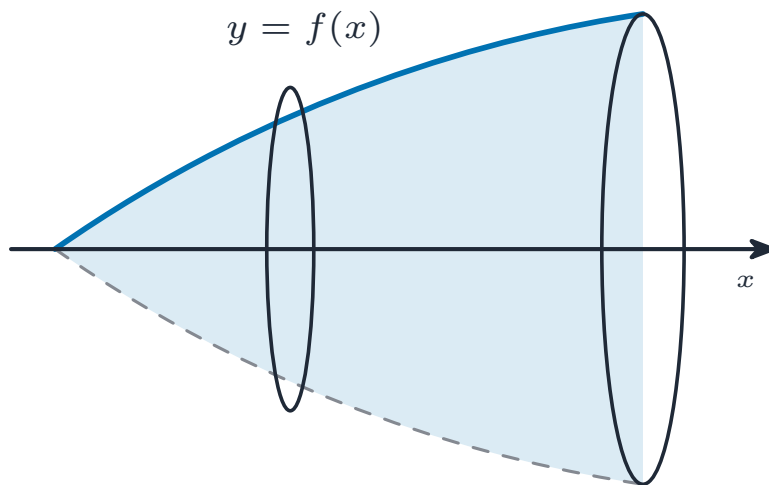
and $\int_0^1 x^{-1/2} dx = [2x^{1/2}]_0^1 = 2$, which is finite even though the integrand blows up at $x = 0$.

Volume of revolution

When a region is turned all the way around an axis it sweeps out a solid. The **volume of revolution** about the x -axis is

$$V = \pi \int_p^q y^2 dx,$$

and about the y -axis it is $V = \pi \int x^2 dy$. For example, the region under $y = \sqrt{x}$ from $x = 0$ to $x = 4$, turned about the x -axis, has volume $\pi \int_0^4 x dx = \pi \left[\frac{x^2}{2} \right]_0^4 = 8\pi$.



rotate the region around the x -axis to sweep a solid

Rotating the region under a curve around the x -axis sweeps out a solid; each thin slice is a disc of area πy^2 , so $V = \pi \int y^2 dx$

Exam tips

- Show **every line of algebra** — method marks are lost by jumping steps; use the **discriminant** $b^2 - 4ac$ to decide the number of real roots.
- Work in **radians** for circular measure (arc = $r\theta$, sector area = $\frac{1}{2}r^2\theta$) and for calculus of trig functions.
- For **differentiation**, set $\frac{dy}{dx} = 0$ for stationary points and use the second derivative to classify them.
- For **integration**, add $+c$ to an indefinite integral and use limits for area; an area below the axis gives a negative integral.

Pure Mathematics 2

A-Level Mathematics

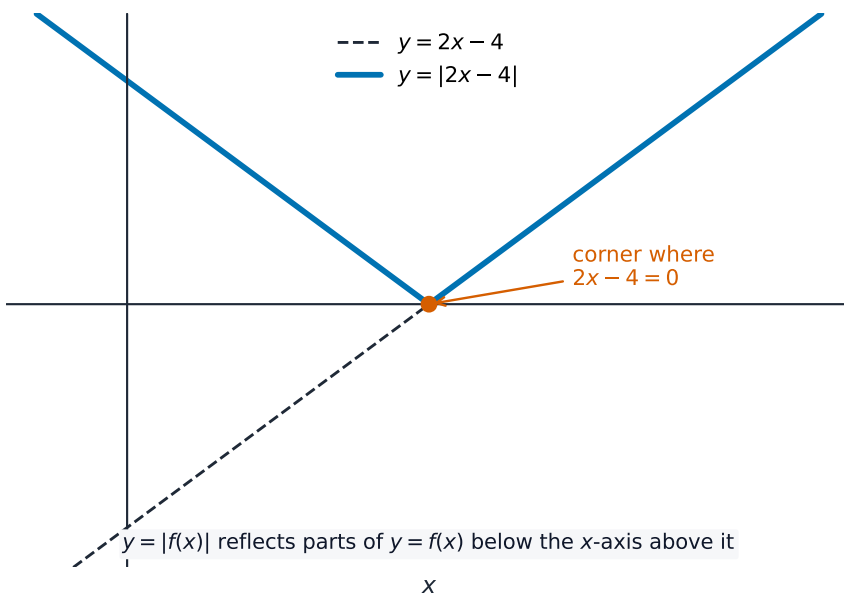
This handout covers Topic 2: **Pure Mathematics** 2. It adds the modulus and polynomial algebra, **logarithms** and the **exponential function**, more trigonometry, and new ways to differentiate and integrate.

Algebra

The modulus

The **modulus** $|x|$ is the size of a number with its sign removed, so $|x| \geq 0$ always. The graph of $y = |ax + b|$ is a "V" shape that bounces off the x -axis. Two useful rules for solving equations and inequalities are

$$|a| = |b| \Leftrightarrow a^2 = b^2, \quad |x - a| < b \Leftrightarrow a - b < x < a + b.$$



Taking the modulus folds the part of the line below the axis upward into a V.

Worked example. Solve $|3x + 8| < 9$.

Using the second rule with the inequality written as $-9 < 3x + 8 < 9$:

$$-9 < 3x + 8 < 9 \Rightarrow -17 < 3x < 1 \Rightarrow -\frac{17}{3} < x < \frac{1}{3}.$$

Polynomial division and the factor and remainder theorems

A **polynomial** is a sum of powers of x , such as $2x^4 + 3x^2 - 5$. Its **degree** is the highest power. When you divide one polynomial by another you get a **quotient** and a **remainder**.

- **Remainder theorem** : the remainder when $p(x)$ is divided by $(x - a)$ is $p(a)$.
- **Factor theorem** : $(x - a)$ is a factor of $p(x)$ exactly when $p(a) = 0$.

Worked example. The polynomial $p(x) = 2x^4 + kx^3 + kx^2 + 17x + 18$ has factor $(x + 2)$. Find k .

By the factor theorem $p(-2) = 0$:

$$2(16) + k(-8) + k(4) + 17(-2) + 18 = 0 \Rightarrow 16 - 4k = 0 \Rightarrow k = 4.$$

Logarithmic and exponential functions



Earthquake strength is measured on the logarithmic Richter scale.

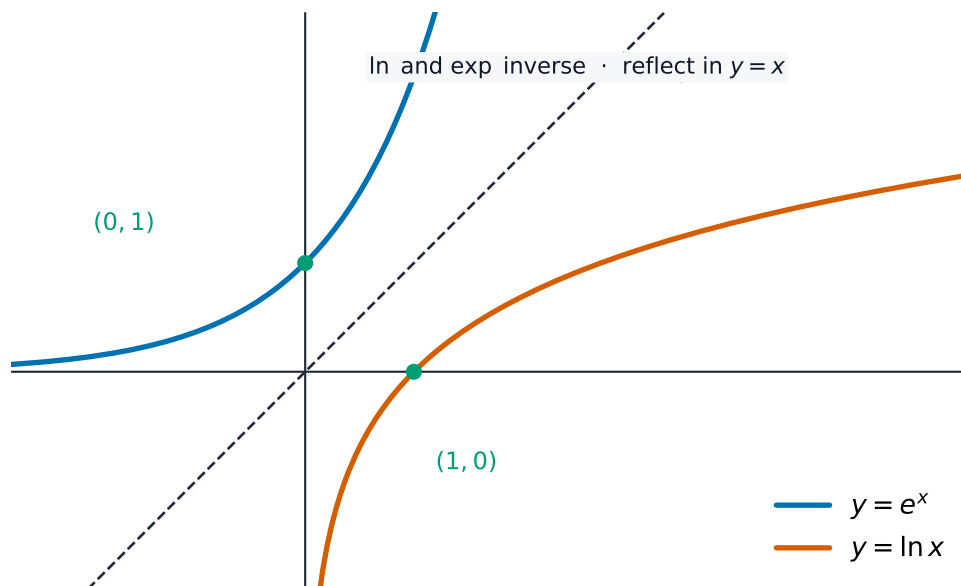


Populations can grow exponentially when resources are plentiful.

A **logarithm** answers the question "what power?". If $a^x = y$ then $x = \log_a y$. Logarithms and **indices** (powers) are reverse ideas. The **laws of logarithms** are

$$\log(mn) = \log m + \log n, \quad \log \frac{m}{n} = \log m - \log n, \quad \log(m^k) = k \log m.$$

The exponential function e^x and the **natural logarithm** $\ln x$ are inverse functions, so $\ln(e^x) = x$ and $e^{\ln x} = x$. When the unknown is in the power, take logs of both sides.



e^x and $\ln x$ undo each other, so each is the other reflected in $y = x$.

Worked example. Solve $4^x < 0.05$.

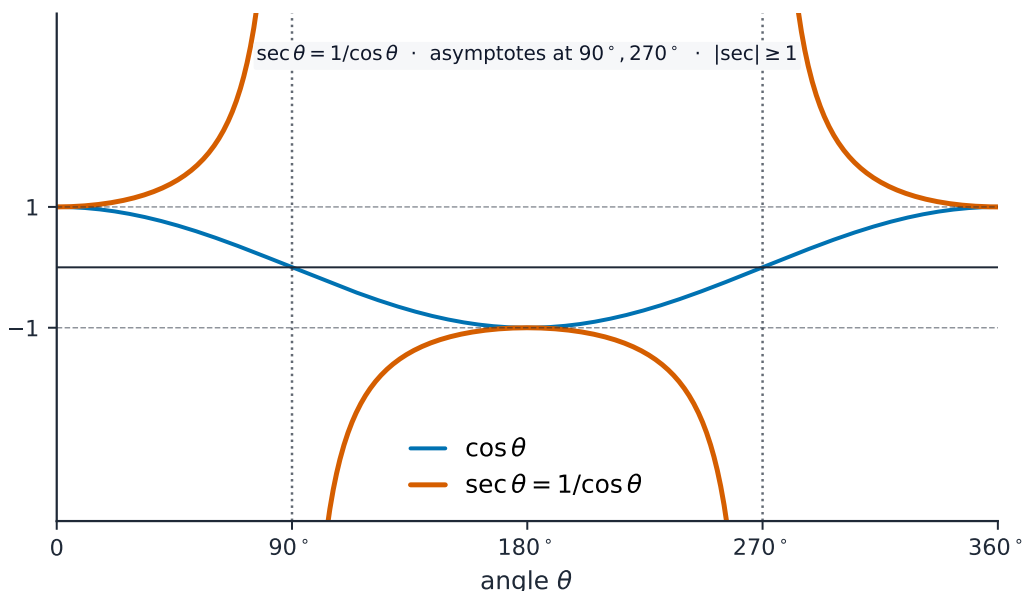
$$4^x < 0.05 \Rightarrow x \ln 4 < \ln 0.05 \Rightarrow x < \frac{\ln 0.05}{\ln 4} = -2.16 \text{ (3 s.f.)}.$$

Linear form : a relationship like $y = Ax^n$ becomes a straight line if you take logs: $\ln y = \ln A + n \ln x$. Plotting $\ln y$ against $\ln x$ gives a line with gradient n and intercept $\ln A$, so you can find the unknown constants.

Trigonometry

There are three more functions, each the reciprocal of one you know: the **secant**

$$\sec \theta = \frac{1}{\cos \theta}, \text{ the } \mathbf{cosecant} \quad \csc \theta = \frac{1}{\sin \theta}, \text{ and the } \mathbf{cotangent} \quad \cot \theta = \frac{1}{\tan \theta}.$$



$\sec \theta = 1/\cos \theta$ rises to an asymptote wherever $\cos \theta = 0$.

You must know these **trigonometric identities** and choose the right one for each problem:

$$\sec^2 \theta \equiv 1 + \tan^2 \theta, \quad \csc^2 \theta \equiv 1 + \cot^2 \theta.$$

You also use the **compound angle** formulae for $\sin(A \pm B)$, $\cos(A \pm B)$, $\tan(A \pm B)$, the **double angle** formulae

$$\sin 2A = 2 \sin A \cos A, \quad \cos 2A = 2 \cos^2 A - 1, \quad \tan 2A = \frac{2 \tan A}{1 - \tan^2 A},$$

and the **R-formula**

$$a \sin \theta + b \cos \theta = R \sin(\theta + \alpha), \text{ where } R = \sqrt{a^2 + b^2} \text{ and } \tan \alpha = \frac{b}{a}.$$

Worked example. Solve $2 \tan^2 \theta + 3 \sec \theta = 18$ for $-180^\circ < \theta < 180^\circ$.

Replace $\tan^2 \theta$ with $\sec^2 \theta - 1$ to get one function:

$$2(\sec^2 \theta - 1) + 3 \sec \theta = 18 \Rightarrow 2 \sec^2 \theta + 3 \sec \theta - 20 = 0 \Rightarrow (2 \sec \theta - 5)(\sec \theta + 4) = 0.$$

So $\sec \theta = \frac{5}{2}$ or $\sec \theta = -4$, giving $\cos \theta = \frac{2}{5}$ or $\cos \theta = -\frac{1}{4}$. The solutions are $\theta = \pm 66.4^\circ$ and $\theta = \pm 104.5^\circ$.

Differentiation

Learn these standard derivatives:

$$\frac{d}{dx}e^x = e^x, \quad \frac{d}{dx}\ln x = \frac{1}{x}, \quad \frac{d}{dx}\sin x = \cos x, \quad \frac{d}{dx}\cos x = -\sin x, \quad \frac{d}{dx}\tan x = \sec^2 x.$$

For a product or a quotient of two functions, use:

- the **product rule** : $(uv)' = u'v + uv'$;
- the **quotient rule** : $\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$.

When a curve is given by **parametric equations** $x = x(t)$, $y = y(t)$, the gradient is $\frac{dy}{dx} = \frac{dy/dt}{dx/dt}$. When y is defined **implicitly** (not made the subject), differentiate every term with respect to x , using the chain rule on the y terms, then solve for $\frac{dy}{dx}$.

Worked example. Given $y = 6x \cos(x^2 + 1)$, find $\frac{dy}{dx}$.

Use the product rule with $u = 6x$ and $v = \cos(x^2 + 1)$ (and the chain rule for v):

$$\frac{dy}{dx} = 6 \cos(x^2 + 1) + 6x \cdot (-2x \sin(x^2 + 1)) = 6 \cos(x^2 + 1) - 12x^2 \sin(x^2 + 1).$$

Integration

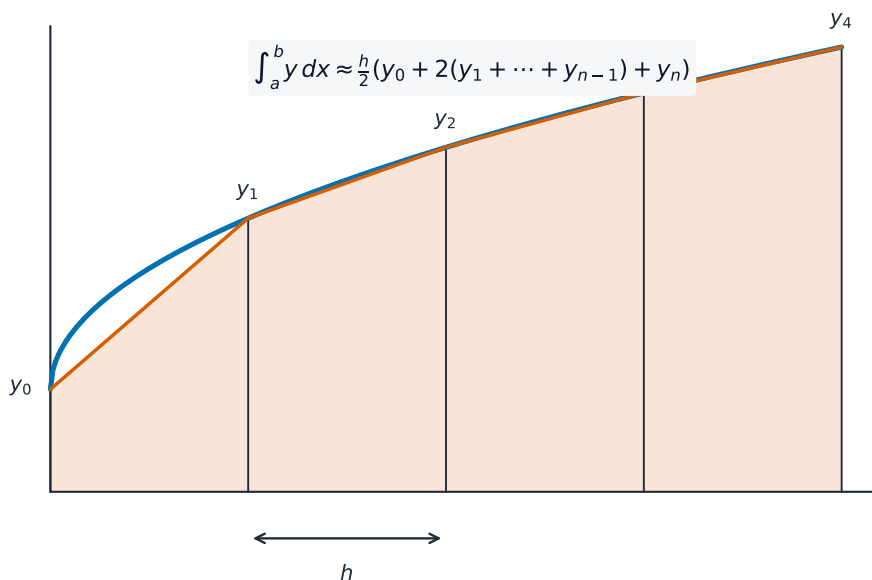
Integration is **reverse differentiation** — reverse each new derivative; harder integrals may need **integration by substitution** (developed in Pure 3). For a linear inside function $(ax + b)$:

$$\int e^{ax+b} dx = \frac{1}{a}e^{ax+b} + C, \quad \int \frac{1}{ax+b} dx = \frac{1}{a} \ln |ax+b| + C,$$

$$\int \sin(ax+b) dx = -\frac{1}{a} \cos(ax+b) + C, \quad \int \cos(ax+b) dx = \frac{1}{a} \sin(ax+b) + C, \quad \int \sec^2(ax+b) dx = \frac{1}{a} \tan(ax+b) + C.$$

To integrate a power of sin or cos, first use an identity to remove the power. When you cannot integrate exactly, the **trapezium rule** estimates a definite integral:

$$\int_a^b y dx \approx \frac{h}{2} [y_0 + y_n + 2(y_1 + y_2 + \cdots + y_{n-1})].$$



Each strip of width h is a trapezium; their areas add up to estimate the integral.

Worked example. Find $\int 6 \sin^2 x \, dx$.

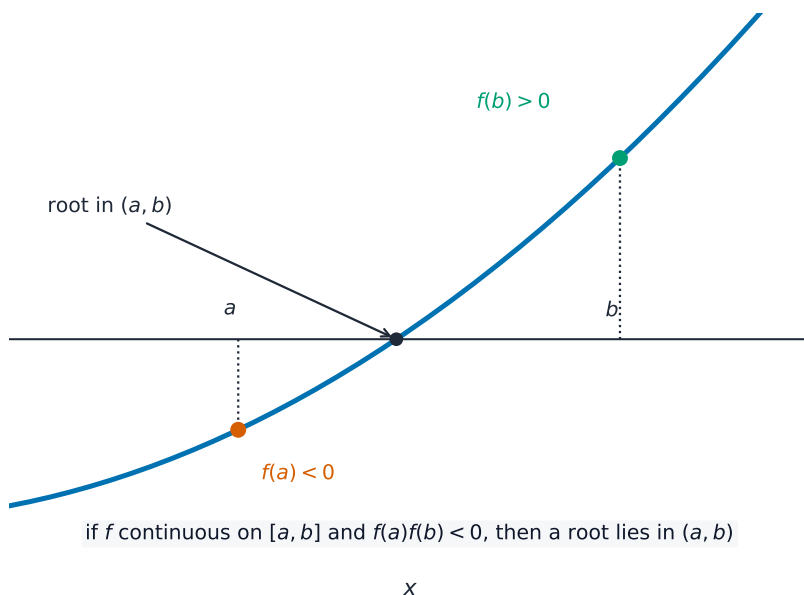
Use $\sin^2 x = \frac{1}{2}(1 - \cos 2x)$:

$$\int 6 \sin^2 x \, dx = \int (3 - 3 \cos 2x) \, dx = 3x - \frac{3}{2} \sin 2x + C.$$

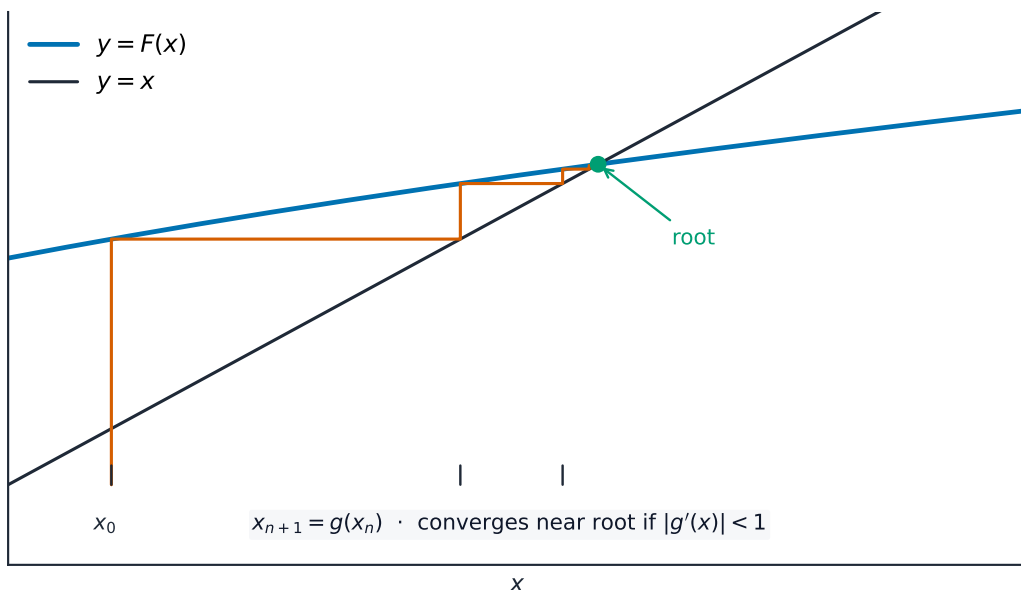
Numerical solution of equations

Many equations cannot be solved exactly. Two ideas help you find a **root** (a solution).

- **Sign change** : if $f(a)$ and $f(b)$ have opposite signs (and the graph has no break between them), a root lies between a and b .
- **Iteration** : rearrange the equation into the form $x = F(x)$, then use the **iterative formula** $x_{n+1} = F(x_n)$. Start from a first guess x_0 and repeat. If the values are **convergent** they settle down and **converge** to a root. Keep going until the answer is steady to the accuracy asked for.



$f(a)$ and $f(b)$ have opposite signs, so a root is trapped between a and b .



Each step goes up to $y = F(x)$ then across to $y = x$; the staircase closes in on the root.

Worked example. A root β of an equation satisfies $x = \sqrt[3]{-2x - 4.5}$, and $-1.4 < \beta < -1.0$. Use the iteration $x_{n+1} = \sqrt[3]{-2x_n - 4.5}$ with $x_0 = -1.2$.

$$x_1 = \sqrt[3]{-2(-1.2) - 4.5} = \sqrt[3]{-2.1} = -1.281, \quad x_2 = \sqrt[3]{-2(-1.281) - 4.5} = -1.247, \quad \dots$$

The values settle near -1.26 , so $\beta = -1.26$ (3 s.f.).

Exam tips

- Use the **laws of logarithms** to solve equations; remember $\ln$ and e^x are inverses.
- For **numerical methods**, show a **sign change** to locate a root, set out the iteration clearly, and give the answer to the stated accuracy.
- Learn the **chain, product and quotient rules** and identify which the function needs.
- Solve modulus equations $|f(x)| = g(x)$ by considering both the positive and negative cases, and sketch to check.

Pure Mathematics 3

A-Level Mathematics

This handout covers Topic 3: **Pure Mathematics** 3. Subtopics 3.1–3.6 build on the algebra, logarithms, trigonometry, differentiation, integration and numerical methods of Pure Mathematics 2, so this handout explains what is **new** in Pure 3 and then covers the three big new areas: vectors, differential equations and complex numbers.

Algebra

Two new tools join the algebra from Pure 2.

Partial fractions

A single fraction with a factorised bottom can be split into a sum of simpler fractions. This is called writing it in **partial fractions**, and it makes a **rational function** (a fraction of polynomials) easy to integrate or expand. Match the form to the bottom:

First check the top is **lower degree** than the bottom. If it is not (“top-heavy”), **divide** the polynomial first: dividing gives a **quotient** plus a **remainder** over the original bottom, and you split only that remaining proper fraction.

$$\frac{x+4}{(x+1)(x-2)} \xrightarrow{\text{split}} \frac{2}{x-2} - \frac{1}{x+1}$$

One fraction with a factorised bottom splits into a sum of simpler fractions.

$$\frac{1}{(ax+b)(cx+d)} = \frac{A}{ax+b} + \frac{B}{cx+d}, \quad \frac{1}{(ax+b)(cx+d)^2} = \frac{A}{ax+b} + \frac{B}{cx+d} + \frac{C}{(cx+d)^2}.$$

Worked example. Express $\frac{x+4}{(x+1)(x-2)}$ in partial fractions.

Write $\frac{x+4}{(x+1)(x-2)} = \frac{A}{x+1} + \frac{B}{x-2}$, so $x+4 = A(x-2) + B(x+1)$. Put $x=2$: $6=3B$, so $B=2$. Put $x=-1$: $3=-3A$, so $A=-1$. Hence

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

The binomial expansion for a rational power

The **binomial expansion**
negative, as long as $|x| < 1$:

also works when the power is a fraction or is

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \dots$$

For example $(1+x)^{1/2} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \dots$ for $|x| < 1$.

Logarithms, trigonometry and numerical methods (from Pure 2)

Subtopics 3.2, 3.3 and 3.6 are the same skills you met in Pure 2: the laws of logarithms with e^x and $\ln x$; the identities $\sec^2 \theta \equiv 1 + \tan^2 \theta$ and $\csc^2 \theta \equiv 1 + \cot^2 \theta$, the compound- and double-angle formulae, and the R -form of $a \sin \theta + b \cos \theta$; and solving an equation numerically by a **sign change** and an **iterative formula** $x_{n+1} = F(x_n)$. Use them exactly as before.

The three new trig functions are the **reciprocals** of the familiar ones: the **secant** $\sec \theta = \frac{1}{\cos \theta}$, the **cosecant** $\csc \theta = \frac{1}{\sin \theta}$, and the **cotangent** $\cot \theta = \frac{1}{\tan \theta} = \frac{\cos \theta}{\sin \theta}$. (Memory aid: match the *third* letter — sec goes with **cos**ine.) The two Pythagorean identities above come straight from dividing $\sin^2 \theta + \cos^2 \theta \equiv 1$ by $\cos^2 \theta$ or $\sin^2 \theta$. For a numerical method, an iteration $x_{n+1} = F(x_n)$ **converges** to the root when successive values get closer together.

Differentiation

The methods are those of Pure 2 (the product, quotient and chain rules, with parametric and implicit curves). One derivative is added — the **inverse tangent**

:

$$\frac{d}{dx} \tan^{-1} x = \frac{1}{1+x^2}.$$

Worked example. Differentiate $y = \tan^{-1}(3x)$.

Use the chain rule with the result above: $\frac{dy}{dx} = \frac{1}{1+(3x)^2} \times 3 = \frac{3}{1+9x^2}$.

Integration

Pure 3 adds several powerful methods.

- A new standard integral: $\int \frac{1}{x^2 + a^2} dx = \frac{1}{a} \tan^{-1} \frac{x}{a} + C$.
- **Partial fractions:** split a rational function first, then integrate each piece as a logarithm.
- The pattern $\frac{k f'(x)}{f(x)}$: this integrates to $k \ln |f(x)| + C$. For example $\int \frac{2x}{x^2 + 1} dx = \ln(x^2 + 1) + C$.
- **Integration by parts** , used for a product: $\int u \frac{dv}{dx} dx = uv - \int v \frac{du}{dx} dx$.
- **Integration by substitution** : a given change of variable turns a hard integral into an easy one.

Worked example. Find $\int x \cos x dx$.

Use integration by parts with $u = x$ and $\frac{dv}{dx} = \cos x$, so $\frac{du}{dx} = 1$ and $v = \sin x$:

$$\int x \cos x dx = x \sin x - \int \sin x dx = x \sin x + \cos x + C.$$

Vectors



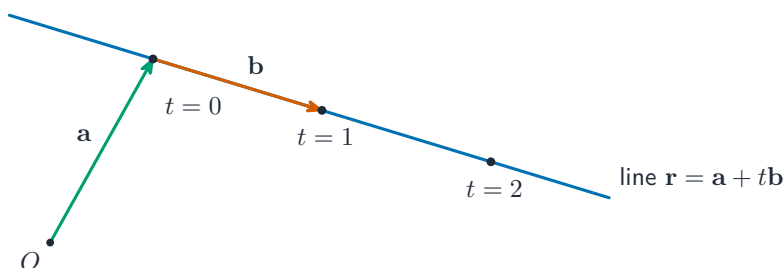
Forces like wind and water are vectors — they have both size and direction.

A **vector** has both size and direction. Write it as a column, or as $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, or as $\overrightarrow{AB}$.

- The **magnitude** (length) of $\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ is $|\mathbf{v}| = \sqrt{x^2 + y^2 + z^2}$.
- A **unit vector** has magnitude 1; divide a vector by its magnitude to make one.
- A **position vector** gives a point's place from the origin; a **displacement vector** $AB = \mathbf{b} - \mathbf{a}$ goes from one point to another. Multiplying by a **scalar** (a plain number) stretches a vector.

Lines and the scalar product

A straight line through point $\mathbf{a}$ in direction $\mathbf{b}$ has vector equation $\mathbf{r} = \mathbf{a} + t\mathbf{b}$. Two lines may be **parallel**, may **intersect** at a point, or may be **skew lines** (not parallel and never meeting).

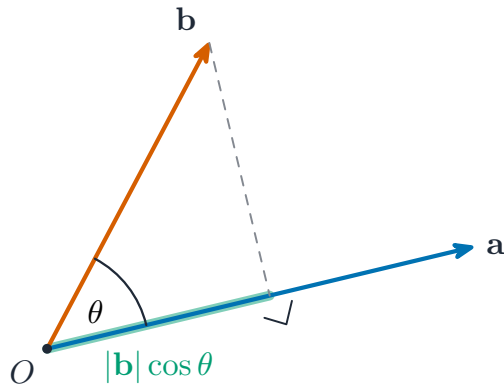


Start at the point $\mathbf{a}$, then add t copies of the direction $\mathbf{b}$ to reach any point on the line.

The **scalar product** (dot product) of $\mathbf{a}$ and $\mathbf{b}$ is

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3 = |\mathbf{a}| |\mathbf{b}| \cos \theta,$$

where θ is the angle between them. So $\mathbf{a} \cdot \mathbf{b} = 0$ means the vectors are perpendicular.



The scalar product picks out $|\mathbf{b}| \cos \theta$, how far $\mathbf{b}$ reaches along $\mathbf{a}$.

Worked example. Find the angle between $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$ and $\mathbf{b} = 2\mathbf{i} + 2\mathbf{j} + \mathbf{k}$.

$$\mathbf{a} \cdot \mathbf{b} = (1)(2) + (2)(2) + (2)(1) = 8, \quad |\mathbf{a}| = |\mathbf{b}| = 3.$$

So $\cos \theta = \frac{8}{3 \times 3} = \frac{8}{9}$, giving $\theta = 27.3^\circ$.

Differential equations

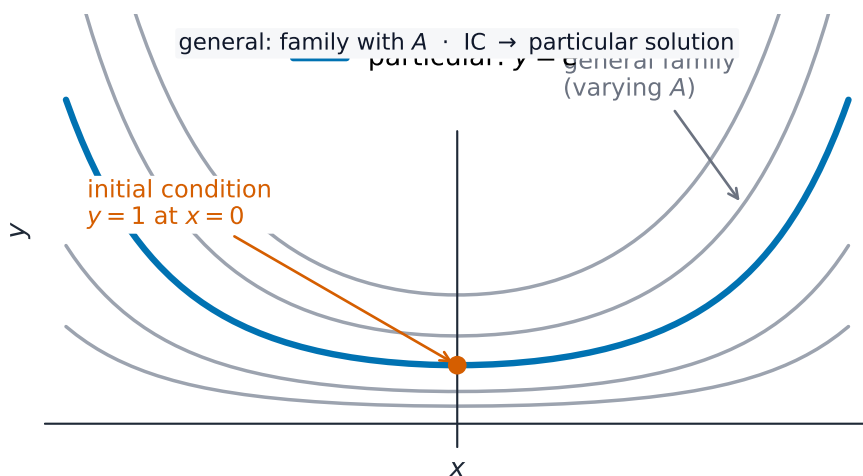
A **differential equation** links a quantity to its **rate of change**. To solve a first-order equation whose variables are **separable**, put all the y terms on one side and all the x terms on the other, then integrate both sides. This gives the **general solution**, which contains a constant. An **initial condition** (a known value) fixes the constant and gives the **particular solution**.

Worked example. Solve $\frac{dy}{dx} = xy$, given that $y = 1$ when $x = 0$.

Separate the variables and integrate:

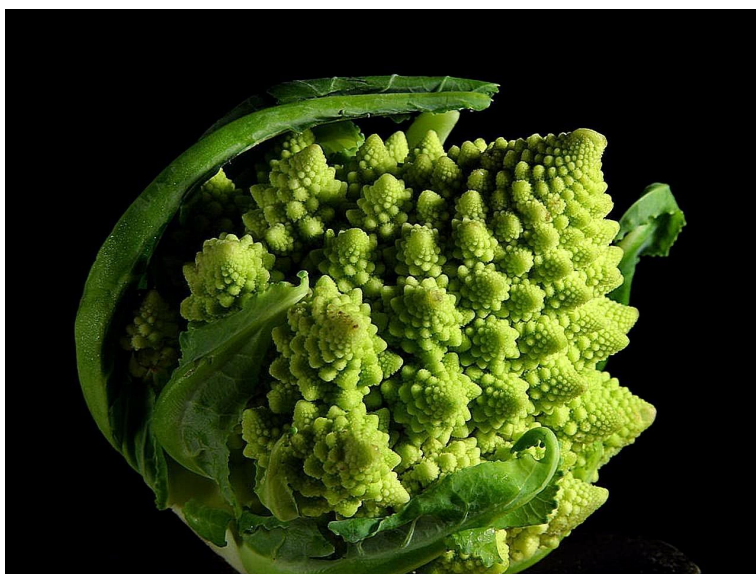
$$\int \frac{1}{y} dy = \int x dx \Rightarrow \ln y = \frac{1}{2}x^2 + c \Rightarrow y = Ae^{x^2/2}.$$

Using $y = 1$ at $x = 0$ gives $A = 1$, so $y = e^{x^2/2}$.



The constant A gives a whole family of curves; the condition $y = 1$ at $x = 0$ selects $y = e^{x^2/2}$.

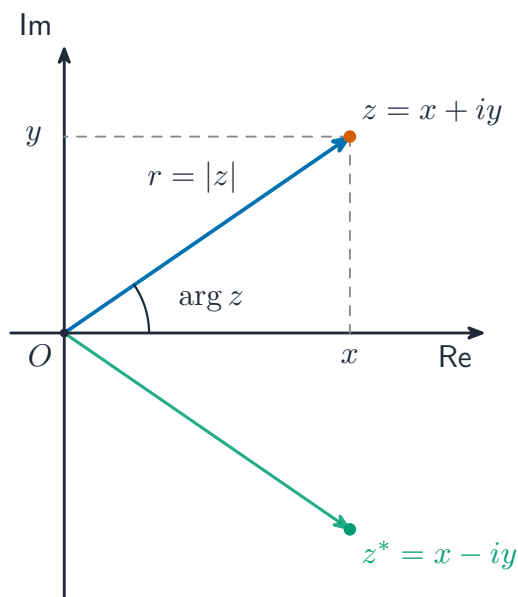
Complex numbers



Self-similar patterns like Romanesco arise from iterating functions in the complex plane.

A **complex number** has the form $z = x + iy$, where $i^2 = -1$. Here x is the **real part** and y is the **imaginary part**. This $x + iy$ is the **Cartesian form**. Two complex numbers are equal only when their real parts match and their imaginary parts match.

- The **conjugate** of $z = x + iy$ is $z^* = x - iy$. For a polynomial with real coefficients, any non-real roots come in conjugate pairs.
- The **modulus** is $|z| = \sqrt{x^2 + y^2}$ (its distance from the origin) and the **argument** is the angle the point makes, measured from the positive real axis.
- You can plot z as a point on an **Argand diagram** (the complex plane).
- The **polar form** is $z = r(\cos \theta + i \sin \theta) = re^{i\theta}$, where $r = |z|$ and θ is the argument. Multiplying multiplies the moduli and adds the arguments.
- The **loci** of points satisfying a condition on z are drawn on the Argand diagram — e.g. $|z - a| = r$ is a circle, $|z - a| = |z - b|$ a perpendicular bisector, and $\arg(z - a) = \theta$ a half-line. The **square roots** of a complex number come from solving $w^2 = z$.



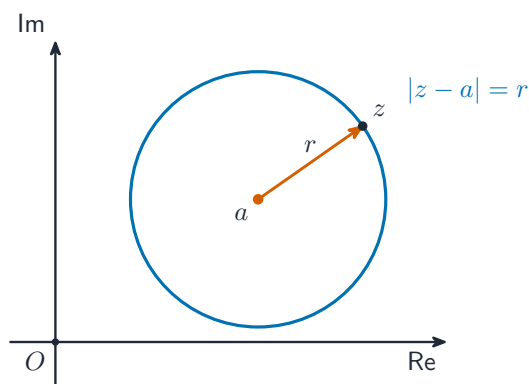
On the Argand diagram $|z|$ is the distance from O , $\arg z$ the angle, and z^* the reflection in the real axis.

To divide, multiply top and bottom by the conjugate of the bottom.

Worked example. Write $\frac{3+i}{1-i}$ in the form $x + iy$.

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

An equation or inequality in z describes a **locus** (a path or region) on the Argand diagram. For example $|z - a| = r$ is a circle of radius r centred at a .



$|z - a| = r$ is the set of points a fixed distance r from a — a circle.

Exam tips

- Split a rational function into **partial fractions** before integrating or expanding.
- Choose the right **integration technique** (substitution, by parts, or partial fractions) from the form of the integrand.
- Use the scalar (dot) product for the **angle between vectors** and to test for perpendicularity; write a line as $\mathbf{r} = \mathbf{a} + t\mathbf{b}$.
- Give **complex numbers** in the form asked for (Cartesian or modulus-argument) and show them on an Argand diagram.

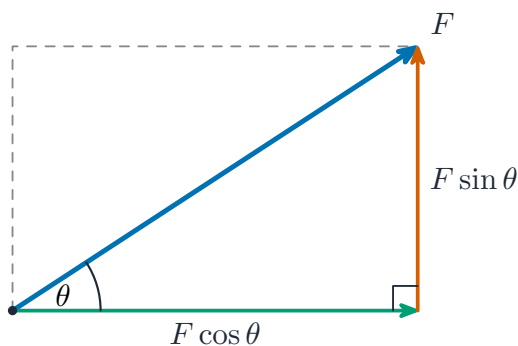
Mechanics

A-Level Mathematics

This handout covers Topic 4: **Mechanics** . It studies how forces make objects move. Throughout this topic, take the acceleration of free fall as $g = 10 \text{ m s}^{-2}$.

Forces and equilibrium

A **force** is a push or a pull. It is a vector, so it has size and direction. Because it is a vector, you can split a force into **components** (usually horizontal and vertical), and you can add several forces into one **resultant** .



A force at angle θ has a horizontal part $F \cos \theta$ and a vertical part $F \sin \theta$.

A particle is in **equilibrium** when the forces are balanced: the vector sum of the forces is zero. In practice this means the components in any direction add to zero.



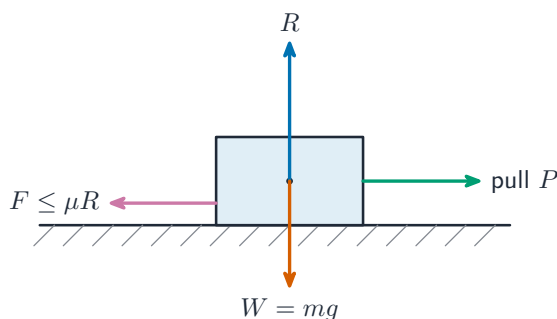
Every mechanics problem is a picture like this. Three forces act on the climber — weight straight down, tension along the rope, and a push from the rock — and because they balance, the climber hangs still in equilibrium. Resolving each into horizontal and vertical parts turns the picture into equations

Friction

When two surfaces touch, the **contact force** between them has two parts: the **normal reaction** R , at right angles to the surface, and the **friction** F , along the surface, which opposes sliding. A "smooth" surface is a model with no friction.

Friction can only grow up to a maximum. At that maximum the body is in **limiting equilibrium**, about to slip, and the friction is **limiting friction**. The maximum is set by the **coefficient of friction** μ :

$$F \leq \mu R, \quad \text{with } F = \mu R \text{ at the point of slipping.}$$



On a rough surface the contact force splits into the normal reaction R and friction F (at most μR).

By **Newton's third law**, the two surfaces push on each other with equal and opposite forces.

Worked example. A block of weight 20 N rests on a rough horizontal table with coefficient of friction $\mu = 0.4$. Find the largest horizontal force that can be applied before the block slides.

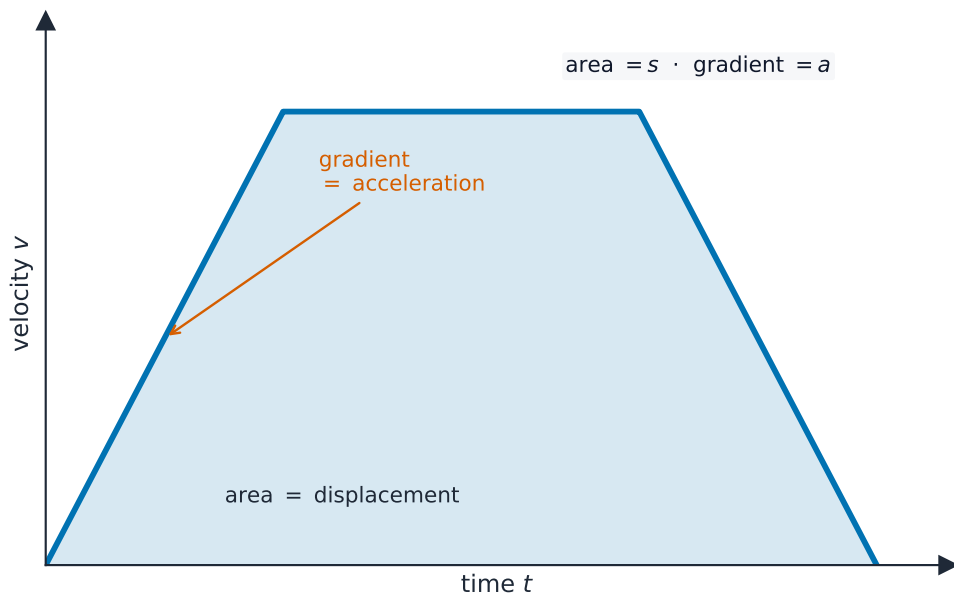
The table's normal reaction balances the weight, so $R = 20$ N. The block is on the point of slipping when friction reaches its maximum $F = \mu R = 0.4 \times 20 = 8$ N. In equilibrium the applied force equals the friction, so the largest force it can resist is 8 N.

Kinematics of motion in a straight line

Distance and **speed** are scalars (size only). **Displacement** and **acceleration** are vectors (size and direction).

On a **velocity-time graph**, the area under the graph is the displacement and the gradient is the acceleration (a negative acceleration is a **deceleration**).

). On a displacement-time graph, the gradient is the velocity. More generally, differentiate with respect to time to go from displacement to velocity to acceleration, and integrate to go back.



The shaded area gives the distance travelled; the slope of the line gives the acceleration.

For motion with **constant acceleration**, use these formulae (the “suvat” equations):

$$v = u + at, \quad s = ut + \frac{1}{2}at^2, \quad v^2 = u^2 + 2as, \quad s = \frac{1}{2}(u + v)t.$$

Worked example. A car starts from rest and accelerates at 2.5 m s^{-2} for 4 s. Find its speed and the distance travelled.

$$v = 0 + 2.5 \times 4 = 10 \text{ m s}^{-1}, \quad s = 0 + \frac{1}{2}(2.5)(4^2) = 20 \text{ m}.$$

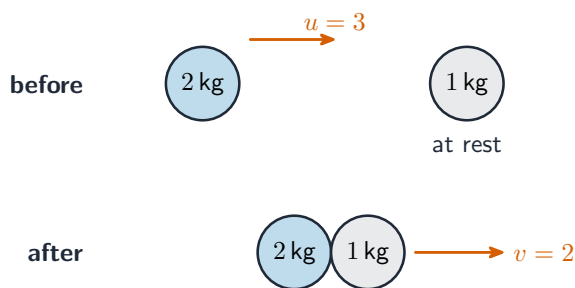
Momentum



A Newton's cradle demonstrates conservation of momentum in collisions.

The **linear momentum** of a body is mass $\times$ velocity. It is a vector. In a direct collision of two bodies, the total momentum is unchanged. This is the **conservation of linear momentum** :

$$m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2.$$



$$2(3) + 1(0) = (2 + 1)v$$

The total momentum is the same before and after the collision.

Worked example. A body of mass 2 kg moving at 3 m s^{-1} hits a stationary body of mass 1 kg, and they stick together. Find their common speed afterwards.

$$2(3) + 1(0) = (2 + 1)v \Rightarrow v = \frac{6}{3} = 2 \text{ m s}^{-1}.$$

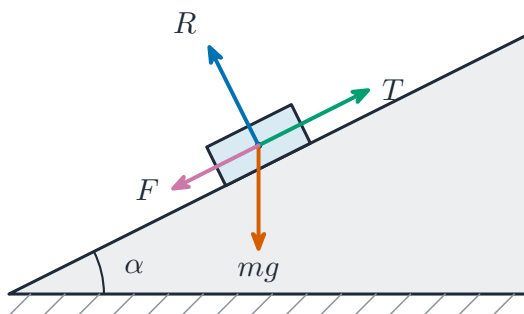
Newton's laws of motion

Newton's laws of motion connect force and acceleration. The key one is: resultant force = **mass** $\times$ acceleration,

$$F = ma.$$

The **weight** of a body is the force of gravity on it: $W = mg$. Forces in a problem may include weight, friction, **tension** in a string, the **thrust** (push) in a rod, and **air resistance**.

For motion on an **inclined plane**, split each force into a part along the slope and a part at right angles to it, then apply $F = ma$ along the slope. For **connected particles** (joined by a string), apply $F = ma$ to each body, or to the whole system.



On a slope, resolve the forces along the slope and at right angles to it.

Worked example. A block of mass 12 kg is pulled up a rough plane by a rope parallel to the slope. The plane is at 20° to the horizontal, the coefficient of friction is 0.4, and the acceleration is 2 m s^{-2} . Find the tension in the rope.

The normal reaction is $R = mg \cos 20^\circ = 120 \cos 20^\circ = 112.8 \text{ N}$, so the friction is $F = \mu R = 0.4 \times 112.8 = 45.1 \text{ N}$. Along the slope, $T - mg \sin 20^\circ - F = ma$:

$$T = ma + mg \sin 20^\circ + F = 12(2) + 120 \sin 20^\circ + 45.1 = 24 + 41.0 + 45.1 = 110 \text{ N (3 s.f.)}.$$

Energy, work and power



A roller coaster trades potential energy for kinetic energy as it rises and falls.

The **work done** by a constant force is the **scalar product** of force and displacement — the force times the distance moved in the direction of the force: $W = Fd \cos \theta$, where θ is the angle between the force and the motion. Work is measured in joules (J).

Energy comes in forms you can calculate:

- **Kinetic energy** (energy of movement): $\text{KE} = \frac{1}{2}mv^2$.
- **Gravitational potential energy** (energy of height): $\text{PE} = mgh$.

The work done by the outside forces equals the change in the total energy. When no friction acts, the total energy stays the same — the **conservation of energy**.

Power is the rate of doing work. For a force pulling in the direction of motion, $P = Fv$ (power = force $\times$ velocity). Power is measured in watts (W).

Worked example. A car engine works at 12 kW while the car moves at 20 m s^{-1} on a level road. Find the driving force.

$$P = Fv \Rightarrow F = \frac{P}{v} = \frac{12000}{20} = 600 \text{ N.}$$

Exam tips

- Draw a clear **force diagram** and resolve into perpendicular components; for equilibrium, each direction sums to zero.
- Use **SUVAT** only for constant acceleration and keep a consistent positive direction.
- Apply $F = ma$ along the direction of motion, including friction ($F = \mu R$) on a rough surface.
- State your **assumptions** (light inextensible string, smooth pulley, particle) — they are often worth a mark.

Probability & Statistics 1

A-Level Mathematics

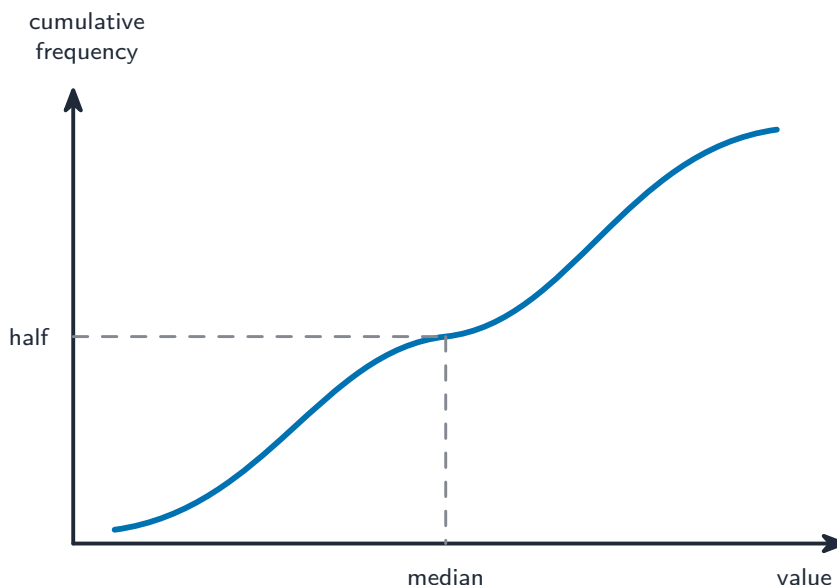
This handout covers Topic 5: **Probability & Statistics**
describing data, counting choices, and working out the chance of events.

1. It is about

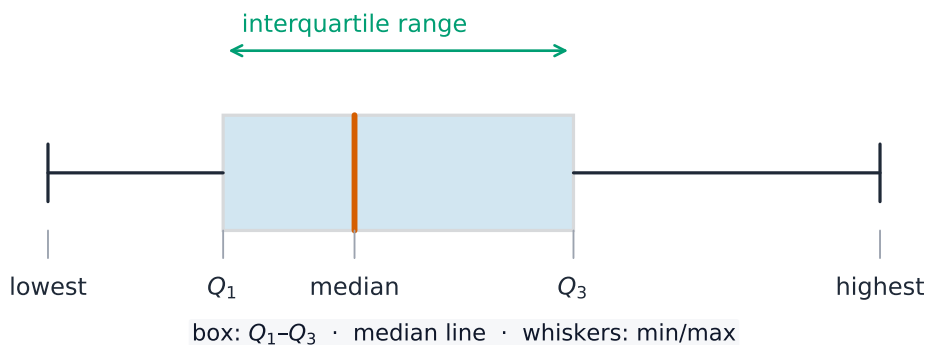
Representation of data

Choose a diagram that suits the data. You should be able to draw and read:

- a **stem-and-leaf diagram** (keeps the original values and shows the shape);
two data sets are compared with a **back-to-back** version – a shared central stem, one set's leaves increasing to the left and the other's to the right, so you can compare their medians and spreads at a glance;
- a **box-and-whisker plot** (shows the lowest value, the three quartiles, and the highest value);
- a **histogram** (for grouped data, where the **area** of each bar shows the frequency);
- a **cumulative frequency** graph (running totals, used to estimate the median and quartiles).



A cumulative frequency curve is S-shaped; read the median across from half the total frequency, and the quartiles from one and three quarters



The box spans the quartiles Q_1 to Q_3 ; the whiskers reach the lowest and highest values.

Averages and spread

A **measure of central tendency** is a single “middle” value:

- the **mean** $\bar{x} = \frac{\sum x}{n}$ (the average);
- the **median** (the middle value when the data is in order);
- the **mode** (the most common value).

A **measure of variation** shows how spread out the data is:

- the **range (of data)** (highest – lowest);
- the **interquartile range** (upper quartile – lower quartile);
- the **standard deviation** $\sigma = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$.

You often work from the totals $\sum x$ and $\sum x^2$. The square of the standard deviation is the **variance**.

Worked example. For 10 values, $\sum x = 50$ and $\sum x^2 = 300$. Find the mean and standard deviation.

$$\bar{x} = \frac{50}{10} = 5, \quad \sigma = \sqrt{\frac{300}{10} - 5^2} = \sqrt{30 - 25} = \sqrt{5} = 2.24.$$

Coding makes big numbers easier. Replace each value by $t = x - a$ for a convenient **assumed mean** a . Then $\bar{x} = a + \bar{t}$, while the standard deviation is **unchanged** (shifting every value does not spread the data). So from the

coded totals $\sum(x - a)$ and $\sum(x - a)^2$ you get $\bar{x}$ and σ directly, and two data sets can be compared through their coded totals.

Permutations and combinations

A **permutation** is an arrangement where order matters; a **combination** is a selection where order does not matter. The numbers are

$${}^nP_r = \frac{n!}{(n-r)!}, \quad {}^nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

To arrange objects in a line when some are repeated, divide by the factorial of each repeat count.

permutation (order matters)

combination (order does not)

AB

BA

{A, B}

Order matters for a permutation, but not for a combination

Worked example. How many different arrangements are there of the letters of the word NEEDLESS?

There are 8 letters, with E repeated 3 times and S repeated 2 times:

$$\frac{8!}{3!2!} = \frac{40320}{6 \times 2} = 3360.$$

Arranging 3 of 4 objects (4P_3)

$$\begin{array}{ccccc} \boxed{4} & \times & \boxed{3} & \times & \boxed{2} & = & 24 \\ \text{1st pick} & & \text{2nd pick} & & \text{3rd pick} & & \end{array}$$

$$\text{Choosing 3 of 4 (order ignored): } {}^4C_3 = \frac{{}^4P_3}{3!} = 4$$

$${}^nP_r = n!/(n-r)! \quad \cdot \quad {}^nC_r = {}^nP_r/r!$$

$${}^nP_r = n!/(n-r)!; \quad {}^nC_r = {}^nP_r/r!$$

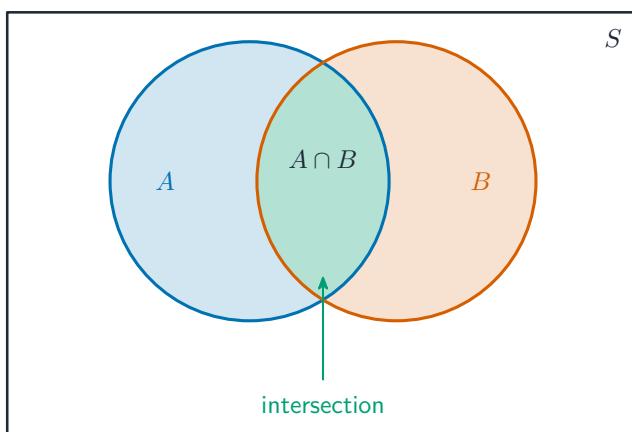
Probability



Dice: a familiar starting point for probability.

Find a probability by counting equally likely outcomes, or by using permutations and combinations. Combine probabilities with these rules:

- **addition** for "or": $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;
- **multiplication** for "and" when events are independent: $P(A \cap B) = P(A)P(B)$.



The overlap of the two circles is $A \cap B$; the addition rule subtracts it once so it is not counted twice.

Two events are **mutually exclusive events** if they cannot both happen,
and **independent events** if one happening does not change the chance

of the other. To test independence, check whether $P(A \cap B) = P(A) \times P(B)$. A **conditional probability** is the chance of A given that B has happened:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}.$$

Worked example. Events have $P(A) = 0.5$, $P(B) = 0.4$ and $P(A \cap B) = 0.2$. Are A and B independent?

Test: $P(A) \times P(B) = 0.5 \times 0.4 = 0.2 = P(A \cap B)$. The values are equal, so A and B are independent.

Discrete random variables

A **discrete random variable** X takes separate values, each with a probability. List them in a **probability distribution table**; the probabilities must add to 1. Then the **expectation** (mean) and **variance** are

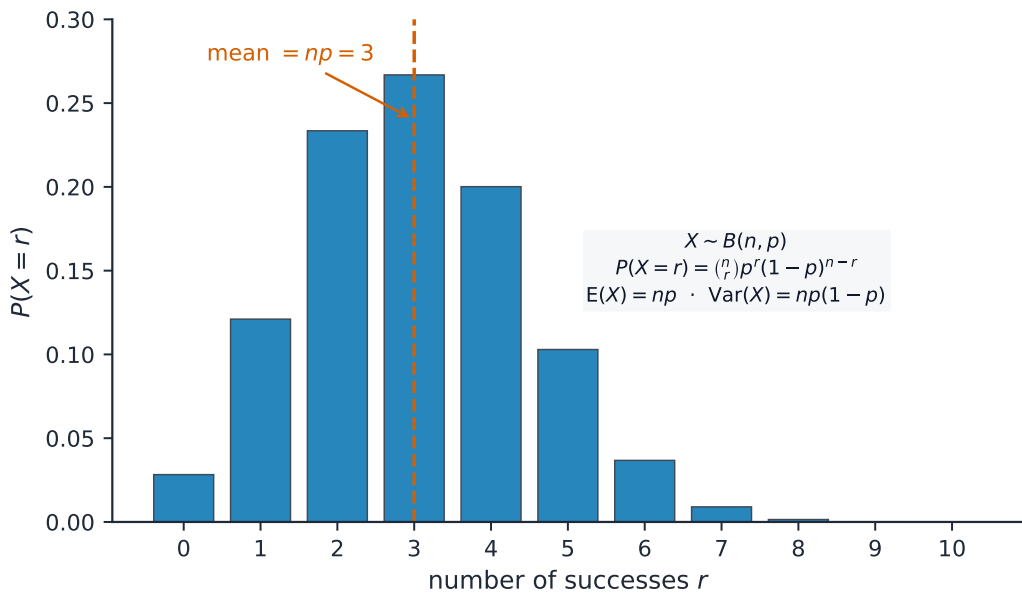
$$E(X) = \sum x P(X = x), \quad \text{Var}(X) = \sum x^2 P(X = x) - (E(X))^2.$$

Two special models:

- the **binomial distribution** $X \sim B(n, p)$, for the number of successes in n independent trials: $P(X = r) = \binom{n}{r} p^r (1 - p)^{n-r}$, with $E(X) = np$ and $\text{Var}(X) = np(1 - p)$;
- the **geometric distribution**, for the trial on which the first success happens: $P(X = r) = (1 - p)^{r-1} p$, with $E(X) = \frac{1}{p}$.

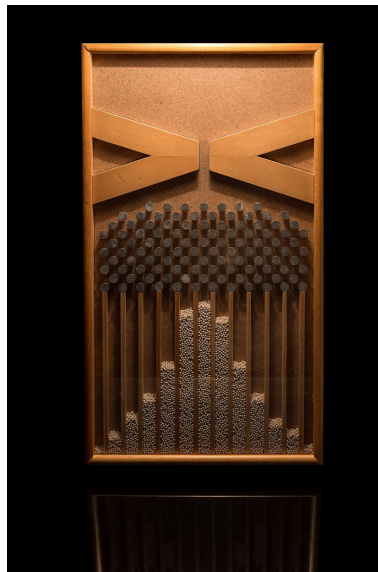
Worked example. $X \sim B(10, 0.3)$. Find $P(X = 2)$ and $E(X)$.

$$P(X = 2) = \binom{10}{2} (0.3)^2 (0.7)^8 = 45 \times 0.09 \times 0.05765 = 0.233, \quad E(X) = 10 \times 0.3 = 3.$$



The distribution $B(10, 0.3)$: each bar is $P(X = r)$, clustered around the mean $np = 3$.

The normal distribution



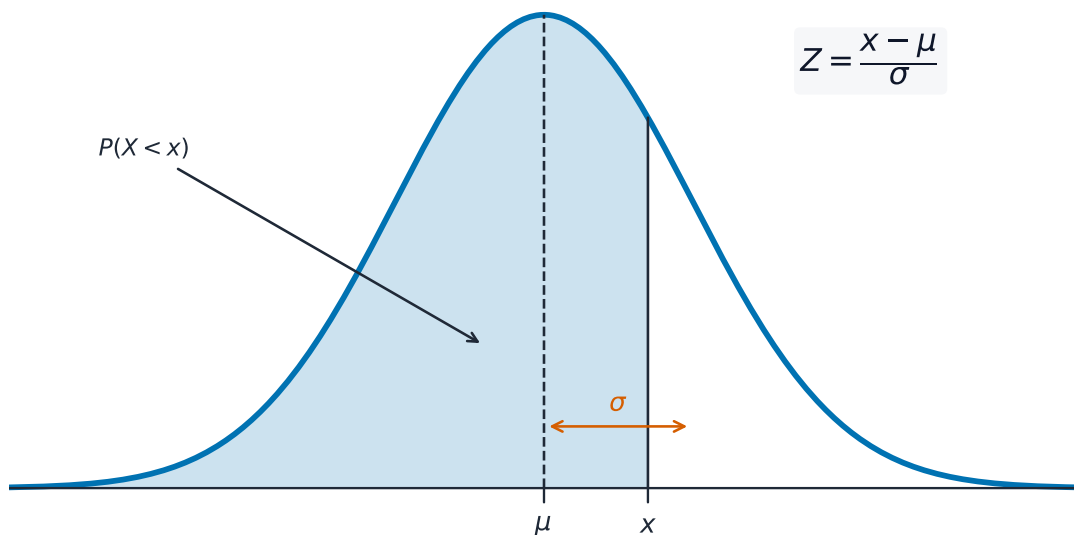
A Galton board: balls falling through pins pile up into the bell-shaped normal distribution.

The **normal distribution** models a **continuous random variable** with a symmetric bell shape. Write $X \sim N(\mu, \sigma^2)$, where μ is the mean and σ is the standard deviation. To use the tables, by **standardisation**

convert to the variable $Z \sim N(0, 1)$:

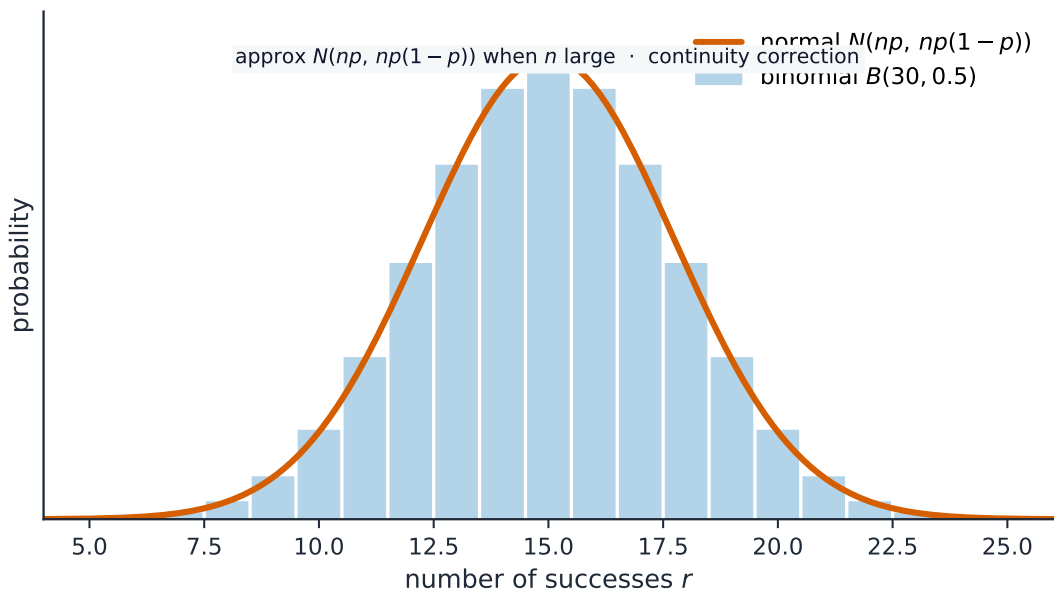
$$Z = \frac{X - \mu}{\sigma}.$$

Then $P(X < x) = P\left(Z < \frac{x - \mu}{\sigma}\right)$, which you read from the normal table Φ .



A normal probability is an area under the bell curve; standardizing rescales it to $Z \sim N(0, 1)$.

The normal distribution is also a good **approximation** to the binomial when n is large. Because you replace a discrete variable by a continuous one, apply a **continuity correction** (adjust by 0.5).



When n is large the binomial bars follow a normal curve of the same mean and variance.

Worked example. Bags of rice have mass $X \sim N(\mu, 0.14^2)$. Given that $P(X < 1.48) = 0.22$, find μ .

From the table, $P(Z < z) = 0.22$ gives $z = -0.772$. So

$$\frac{1.48 - \mu}{0.14} = -0.772 \Rightarrow \mu = 1.48 + 0.772 \times 0.14 = 1.59 \text{ kg (3 s.f.)}.$$

Exam tips

- Decide whether order matters: **permutations** (nP_r) when it does, **combinations** (nC_r) when it does not.
- For a **discrete random variable**, check the probabilities sum to 1 and use $E(X) = \sum x P(X = x)$.
- For the **normal distribution**, standardise with $z = (x - \mu)/\sigma$, **sketch and shade**, then read the table.
- Apply a **continuity correction** when approximating a discrete variable by the normal.

Probability & Statistics 2

A-Level Mathematics

This handout covers Topic 6: **Probability & Statistics**

2. It adds the Poisson model, combining random variables, continuous distributions, and the ideas of estimation and testing.

The Poisson distribution



People arriving at random in a queue follow a Poisson distribution.

The **Poisson distribution**

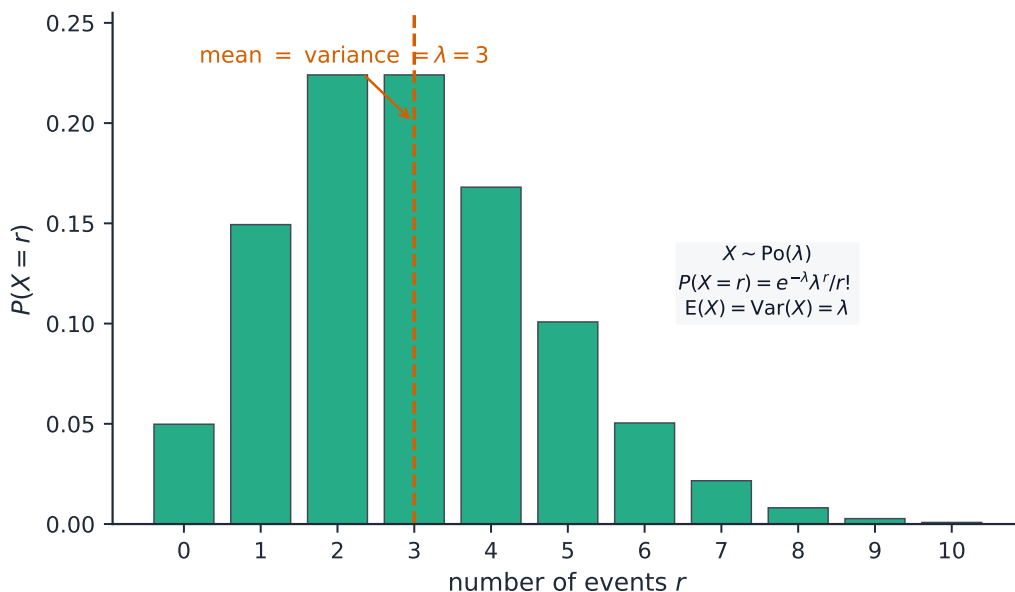
$X \sim \text{Po}(\lambda)$ models the number of random events in a fixed interval, when events happen at a steady average rate λ :

$$P(X = r) = e^{-\lambda} \frac{\lambda^r}{r!}.$$

For a Poisson variable the mean and the variance are both equal to λ . The Poisson distribution is a good **approximation** to the **binomial distribution** when n is large and p is small. The normal distribution (with continuity correction) approximates the Poisson when λ is large.

Worked example. $X \sim \text{Po}(3)$. Find $P(X = 2)$.

$$P(X = 2) = e^{-3} \frac{3^2}{2!} = e^{-3} \times 4.5 = 0.224.$$



The Poisson distribution $\text{Po}(3)$: for a Poisson variable the mean and variance both equal λ .

Linear combinations of random variables

When you change a variable by a linear rule, the **expectation** (mean) and **variance** follow these rules:

$$E(aX + b) = aE(X) + b, \quad \text{Var}(aX + b) = a^2 \text{Var}(X).$$

For two **independent** variables X and Y :

$$E(aX + bY) = aE(X) + bE(Y), \quad \text{Var}(aX + bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y).$$

Two useful facts: if X has a **normal distribution** then so does $aX + b$; and the sum of independent Poisson variables is again Poisson.

Worked example. X has mean 5 and variance 4. Find $E(3X - 1)$ and $\text{Var}(3X - 1)$.

$$E(3X - 1) = 3(5) - 1 = 14, \quad \text{Var}(3X - 1) = 3^2 \times 4 = 36.$$

Continuous random variables

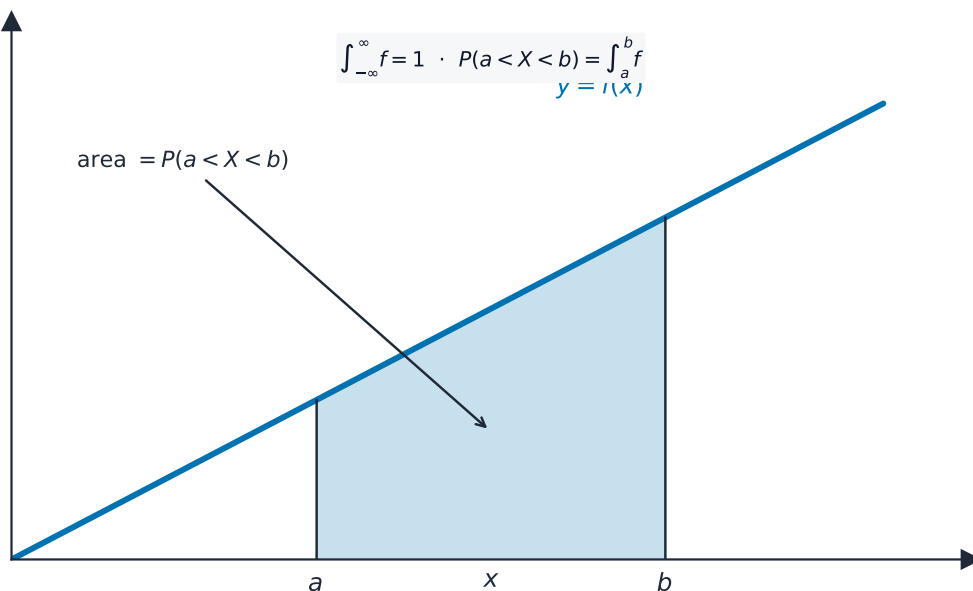
A **continuous random variable** can take any value in a range. Its probabilities come from a **probability density function** $f(x)$, with two key properties:

$$f(x) \geq 0, \quad \int_{-\infty}^{\infty} f(x) dx = 1.$$

A probability is the area under f , and the mean is found by integration:

$$P(a < X < b) = \int_a^b f(x) dx, \quad E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

The **cumulative distribution function** is $F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$; the **median** solves $F(m) = 0.5$, and other **percentiles** solve $F(x) = p$.



For a continuous variable, the probability $P(a < X < b)$ is the area under $f(x)$ between a and b .

Worked example. A continuous variable has $f(x) = \frac{1}{2}x$ for $0 \leq x \leq 2$ (and 0 elsewhere). Find $E(X)$.

$$E(X) = \int_0^2 x \cdot \frac{1}{2}x dx = \int_0^2 \frac{1}{2}x^2 dx = \left[\frac{x^3}{6} \right]_0^2 = \frac{8}{6} = \frac{4}{3}.$$

The **variance** uses the same idea, $\text{Var}(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - (E(X))^2$. For the same $f(x) = \frac{1}{2}x$ on $[0, 2]$: $\int_0^2 x^2 \cdot \frac{1}{2}x dx = \left[\frac{x^4}{8}\right]_0^2 = 2$, so $\text{Var}(X) = 2 - \left(\frac{4}{3}\right)^2 = 2 - \frac{16}{9} = \frac{2}{9}$.

Sampling and estimation



Statistics studies a sample to learn about a whole population.

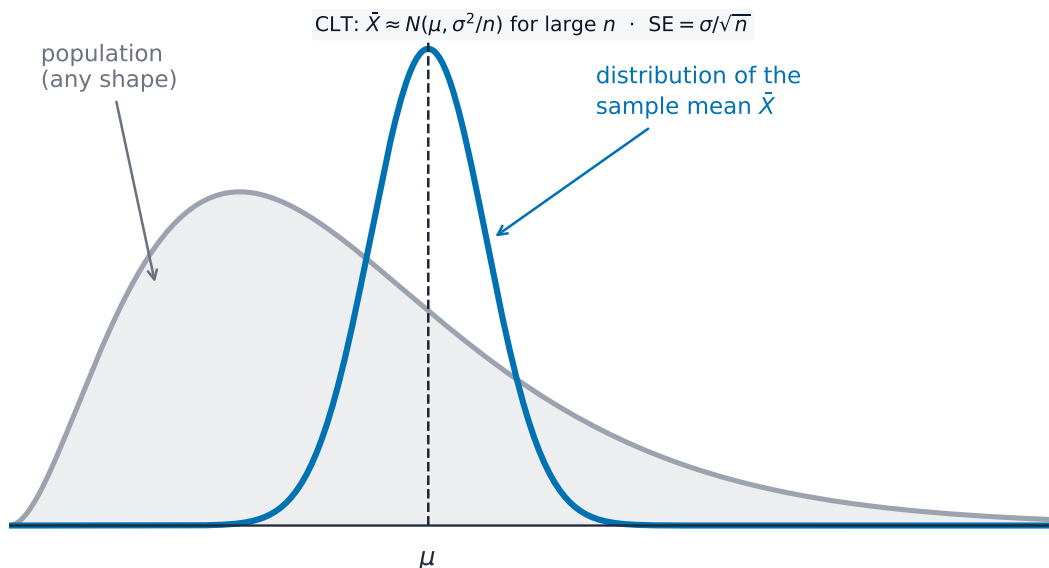
A **sample** is a small group chosen from the whole **population**. A **random sample** needs **randomness**, so that every member has a fair chance of being chosen.

Some methods are **unsatisfactory** because they are **biased**: sampling only volunteers, or the first 20 people to arrive, over-represents certain kinds of people. A genuinely random sample uses **random numbers** – number every member of the population, then draw numbers (from a table or a generator) to decide who is in the sample.

The sample mean $\bar{X}$ is itself a random variable, with

$$E(\bar{X}) = \mu, \quad \text{Var}(\bar{X}) = \frac{\sigma^2}{n}.$$

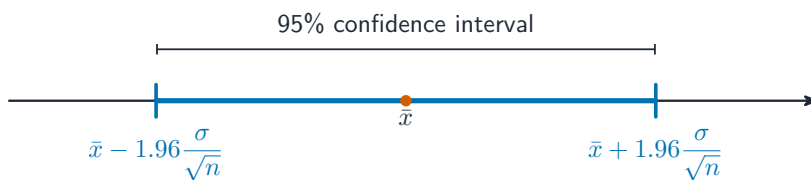
By the **Central Limit Theorem**, for a large sample $\bar{X}$ is approximately normal, whatever the shape of the population.



Whatever the population's shape, the sample mean $\bar{X}$ has a narrow, near-normal distribution centred on μ .

From a sample you can find **unbiased estimates** of the population mean and variance. A **confidence interval** gives a range that probably contains the true mean. When the population is normal with known σ (or the sample is large), a 95% interval is

$$\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}.$$



A 95% confidence interval stretches 1.96 standard errors each side of the sample mean.

You can also find a confidence interval for a **population proportion** from a large sample.

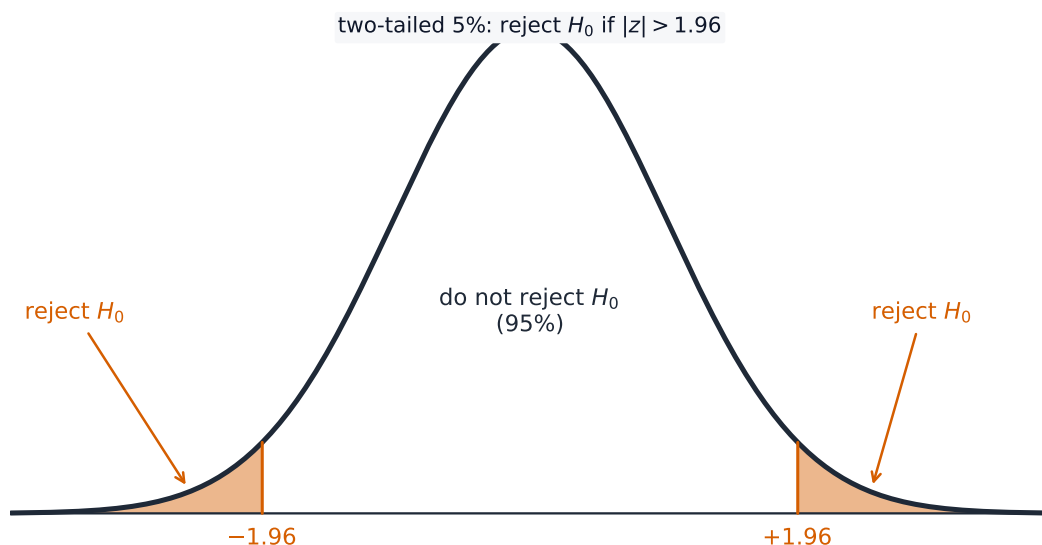
Worked example. A sample of $n = 64$ has mean $\bar{x} = 50$, from a population with $\sigma = 8$. Find a 95% confidence interval for the population mean.

$$50 \pm 1.96 \times \frac{8}{\sqrt{64}} = 50 \pm 1.96 \Rightarrow (48.0, 52.0).$$

Hypothesis tests

A **hypothesis test** uses sample data to judge a claim. You set up two statements: the **null hypothesis** H_0 (the claim being tested, usually "no change") and the **alternative hypothesis** H_1 (what you suspect instead). The test is **one-tailed** if H_1 points one way (e.g. $\mu > 50$) and **two-tailed** if it allows both ways ($\mu \neq 50$).

You fix a **significance level** (often 5%), work out a **test statistic** from the data, and see whether it lands in the **rejection region** (also called the **critical region**); if it does you reject H_0 , otherwise the statistic is in the **acceptance region**.



A two-tailed test at 5% rejects H_0 only if the test statistic falls in a shaded tail beyond ± 1.96 .

Two mistakes are possible: a **Type I error** is rejecting H_0 when it is actually true; a **Type II error** is accepting H_0 when it is actually false. You can find their probabilities from the rejection region: $P(\text{Type I}) = P(\text{statistic in the rejection region} \mid H_0)$ – this equals the significance level – and $P(\text{Type II}) = P(\text{statistic in the acceptance region} \mid H_1 \text{ true for a stated value})$, computed from the binomial, Poisson, or normal distribution.

Worked example. A population is claimed to have mean 50, with $\sigma = 8$. A sample of $n = 64$ gives $\bar{x} = 52$. Test at the 5% level whether the mean has changed.

$H_0: \mu = 50$ and $H_1: \mu \neq 50$ (two-tailed). The test statistic is

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{52 - 50}{8/8} = 2.$$

The critical value at 5% (two-tailed) is 1.96. Since $2 > 1.96$, you reject H_0 : there is evidence the mean has changed.

Worked example (a binomial test). A coin is claimed fair but suspected of landing heads too rarely: $H_0: p = 0.5$, $H_1: p < 0.5$. In $n = 30$ tosses you see $X = 9$ heads. Under H_0 , $X \sim B(30, 0.5)$, so the one-tailed tail probability is

$$P(X \leq 9) = \sum_{k=0}^9 \binom{30}{k} (0.5)^{30} \approx 0.021.$$

Since $0.021 < 0.05$, reject H_0 : the coin does seem biased against heads. (For large n the binomial is approximated by a normal; a **Poisson** test works the same way for rare events. And here, if the rule is "reject when $X \leq 9$ ", then $P(\text{Type I}) = P(X \leq 9 \mid p = 0.5) \approx 0.021$.)

Exam tips

- Use the **Poisson** distribution for rare, random, independent events; its mean equals its variance ($= \lambda$).
- When combining independent random variables, **variances add** (they never subtract).
- For a **hypothesis test**, state H_0 and H_1 , the significance level, the test statistic, and a conclusion **in context**.
- For a **confidence interval**, use the correct z (or t) value and interpret it in words.