BetterMath

IB · AA · HL

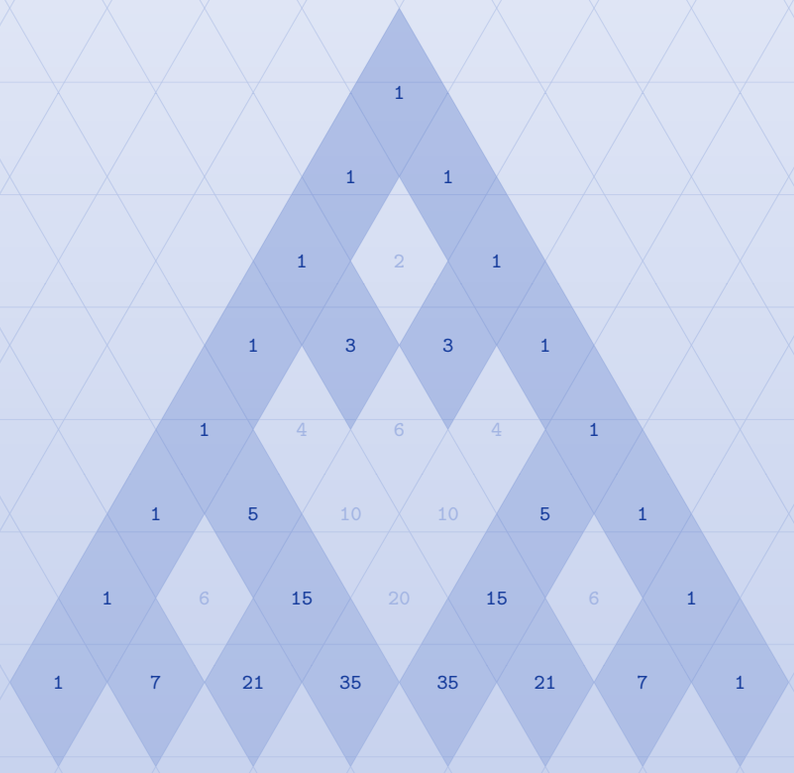
IB Diploma Programme

Mathematics

Analysis & Approaches

Higher Level

Complete course notes



$\binom{n}{k} \bmod 2$

today > yesterday

Third edition · September 2026

This page is intentionally blank

BetterMath

IB Diploma Programme

Mathematics

Analysis & Approaches

Higher Level

Complete course notes

Hasan Khan

M.Sc. Particle Physics, University of Jyväskylä · PGD Education

Third edition, September 2026

BetterMath

Complete course notes for IB Mathematics:
Analysis and Approaches, Higher Level

Third edition, September 2026

© 2026 Hasan Khan. All rights reserved.

No part of these notes may be reproduced, distributed, or used in teaching beyond personal study without the author's written permission. Requests are welcome at ibxhasan@gmail.com, and are usually granted.

These notes are designed for DP students aiming for a 7, built on a foundation of clarity and intuitive understanding. My sincere thanks go to the students and colleagues whose questions, feedback, and corrections have shaped every chapter.

Hasan Khan

Teacher of IB Mathematics and Physics · IB Examiner
M.Sc. Particle Physics, University of Jyväskylä
PGD Education · National curriculum contributor
Training with CERN, ESA and Wolfram Research

> I have taught for about ten years, in Denmark, Tanzania and now Japan, and this book grew out of those classrooms and my students' questions. Graphic design is a hobby I take far too seriously, which is why the book looks the way it does. Corrections and suggestions are always welcome.
>
> Away from teaching, I am a *pathfinder*, more interested in the route than the arrival.

---

ibxhasan.github.io/bettermath · corrections to ibxhasan@gmail.com

Set in TeX Gyre Pagella, TeX Gyre Heros and Latin Modern, using XeLaTeX.

These notes are not endorsed by, affiliated with, or published by the International Baccalaureate Organization. Syllabus references are given for the reader's convenience and follow the current IB Mathematics: Analysis and Approaches guide.

How to use these notes

These notes are written for Higher Level students who want a deep understanding of mathematics, so some pages may feel demanding. That is deliberate: the notes are written without compromise, to build strong intuition, especially if you are aiming for a 7. The badges below help you find your way through them. Each one tells you at a glance what an item is and how to use it. Each section works through its main method in the text; the gold examples give further practice and can be skipped on a first reading.

DEF	A definition to understand and use precisely.
KEY	A key result or formula worth retaining.
KEY · IB DEF · IB	Also printed in the IB formula booklet.
CAUTION	A mistake made here often enough to be worth naming.
TIP	Exam craft: what earns the mark, or what loses it.
GDC	A calculator is required, or genuinely useful.
EXACT	Leave the answer exact: a surd, a fraction, or $k\pi$.
TOK	A Theory of Knowledge question about mathematics.
IM	Mathematics in its historical and cultural setting.
RW	A real-world application or modelling context.
EXT	Beyond the syllabus: enrichment, not examinable.
HL	Additional Higher Level content: examined at HL only.
HL ONLY	A chapter whose whole content is Higher Level.

REFLECTIONS

Every chapter closes with a page of Reflections: Theory of Knowledge questions, historical and international context, real-world uses, and enrichment. None of it is examined. It is there to widen how you think about the mathematics, and to make the subject your own.

PROBLEM COLOURS

3. Your Turn. Learn each skill by practising it straight after its section.

17. End of chapter. Exam-style questions on the whole chapter.

38. Challenge. Investigations. You may leave them on a first pass, but come back to them: they prepare you for Paper 3 and can suggest an idea for your IA.

BetterMath · For IB Mathematics AA HL

Contents

NUMBER & ALGEBRA

Ch 1: Exponents & Logarithms 9

1.1 What an Exponent Counts 10
1.2 Standard Form 18
1.3 The Logarithm: an Exponent Reversed 20
1.4 The Laws of Logarithms 23
1.5 Solving Exponential Equations 28
Reflections 34
End-of-Chapter Problems 35

Ch 2: Sequences & Series 41

2.1 Arithmetic Sequences 42
2.2 Arithmetic Series 44
2.3 Geometric Sequences 48
2.4 Geometric Series 53
2.5 Sigma Notation 55
2.6 Infinite Geometric Series 58
2.7 Financial Applications 61
2.8 Counting Principles 65
2.9 The Binomial Theorem 70
2.10 The Binomial Series 74
Reflections 78
End-of-Chapter Problems 79

Ch 3: Proofs 87

3.1 Deductive Proof 88
3.2 Disproof by Counterexample 94
3.3 Proof by Contradiction 97
3.4 Proof by Induction 100
Reflections 107
End-of-Chapter Problems 109

FUNCTIONS

Ch 4: Lines & Quadratics 115

4.1 Functions 116
4.2 Linear Functions 124
4.3 Systems of Linear Equations 128
4.4 Quadratic Functions 133
4.5 Quadratic Equations 136
4.6 The Discriminant 139
4.7 Quadratic Inequalities 144
Reflections 147
End-of-Chapter Problems 148

Ch 5: Functions & Graphs 155

5.1 Composite Functions 156
5.2 Transformations of Graphs 161
5.3 Symmetry, Inverses and Related Graphs 168
5.4 Reciprocal and Rational Functions 181
5.5 Exponential and Logarithmic Functions 193
Reflections 201
End-of-Chapter Problems 202

Ch 6: Polynomials 211

6.1 Polynomial Functions and their Graphs 212
6.2 Polynomial Division 218
6.3 The Remainder and Factor Theorems 222
6.4 Solving Polynomial Equations 226
6.5 Sum and Product of Roots 228
6.6 Partial Fractions 233
Reflections 236
End-of-Chapter Problems 237

GEOMETRY & TRIGONOMETRY

Ch 7: Shapes & Measures 243

7.1 Three-Dimensional Geometry 244
7.2 Volumes and Surface Areas of Solids 248
7.3 The Sine and Cosine Rules 252
7.4 Elevation, Depression and Bearings 258
7.5 Radians, Arc Length, and Sector Area 263
Reflections 268
End-of-Chapter Problems 269

Ch 8: Trigonometry 277

8.1 The Unit Circle 278
8.2 Trigonometric Identities 284
8.3 Graphs of Trigonometric Functions 290
8.4 Reciprocal Trigonometric Functions 294
8.5 Inverse Trigonometric Functions 298
8.6 Solving Trigonometric Equations 302
Reflections 309
End-of-Chapter Problems 310

Ch 9: Vectors 317

9.1 Vectors: Magnitude and Direction 318
9.2 Position Vectors and Geometry 322
9.3 The Scalar Product 325
9.4 The Vector Equation of a Line 329
9.5 The Vector Product 333
9.6 Planes 338
9.7 Distances 345
Reflections 351
End-of-Chapter Problems 352

STATISTICS AND PROBABILITY

Ch 10: Statistics 359

10.1 Sampling and Data 360
10.2 Measures of Central Tendency 363
10.3 Measures of Dispersion 369
10.4 Visualising Distributions 374
10.5 Correlation 381
10.6 Linear Regression 386
Reflections 391
End-of-Chapter Problems 393

Ch 11: Probability 403

11.1 The Language of Probability 404
11.2 Counting the Outcomes 408
11.3 Combining Events 412
11.4 Conditional Probability and Independence 417
11.5 Tree Diagrams 422
11.6 Bayes' Theorem 425
Reflections 431
End-of-Chapter Problems 432

Ch 12: Probability Distributions 441

12.1 Discrete Random Variables 442
12.2 Expected Value and Variance 444
12.3 The Binomial Distribution 448
12.4 Continuous Random Variables 453
12.5 The Normal Distribution 459
12.6 Standardisation and the Inverse Normal 463
Reflections 467
End-of-Chapter Problems 468

CALCULUS

Ch 13: Limits & Derivatives 475

13.1 Limits 476
13.2 The Derivative 481
13.3 Differentiating Powers 487
13.4 Tangents and Normals 489
13.5 The Standard Derivatives 492
13.6 The Rules of Differentiation 496
13.7 Higher Derivatives 501
13.8 L'Hôpital's Rule 503
Reflections 506
End-of-Chapter Problems 507

Ch 14: Derivative Applications 513

14.1 Stationary Points and Their Nature 514
14.2 Points of Inflexion 520
14.3 Optimisation 525
14.4 Kinematics 529
14.5 Implicit Differentiation 533
14.6 Related Rates 536
Reflections 540
End-of-Chapter Problems 541

Ch 15: Integral Calculus 549

15.1 Antiderivatives 550
15.2 Definite Integrals and Area 552
15.3 Standard Integrals 558
15.4 Rewriting Before Integrating 562
15.5 Integration by Substitution 565
15.6 Integration by Parts 568
15.7 Areas Between Curves 573
15.8 Volumes of Revolution 576
15.9 Kinematics by Integration 579
Reflections 583
End-of-Chapter Problems 586

Ch 16: Differential Equations 595

16.1 Separable Variables 596
16.2 Homogeneous Differential Equations 600
16.3 The Integrating Factor 602
16.4 Maclaurin Series 607
16.5 Euler's Method 614
Reflections 618
End-of-Chapter Problems 619

NUMBER & ALGEBRA

Ch 17: Complex Numbers 627

17.1 Cartesian Form 628
17.2 Modulus and Argument 632
17.3 Products, Quotients, De Moivre 634
17.4 Roots of Complex Numbers 639
17.5 Trigonometric Identities from De Moivre 645
Reflections 650
End-of-Chapter Problems 651

Symbols and notation

Anything defined only inside one chapter is explained where it appears.

NUMBERS AND SETS

$\mathbb{N}$	natural numbers
$\mathbb{Z}$, $\mathbb{Z}^+$	integers, positive integers
$\mathbb{Q}$	rational numbers
$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers
$\in$	is an element of
$\cup$, $\cap$	union, intersection
$\emptyset$	the empty set
$\lvert x \rvert$	modulus of x
$0.\overline{27}$	recurring decimal, $0.2727\ldots$
$\approx$	approximately equal to
$\neq$, $\leq$, $\geq$	not equal to, at most, at least
$\pm$	plus or minus
$\propto$	is proportional to

LOGIC AND PROOF

$\Longrightarrow$	implies
$\Longleftrightarrow$	if and only if
$\equiv$	is identically equal to
■	end of proof

ALGEBRA AND COUNTING

$n!$	n factorial
${}^nC_r$	combinations: n choose r
${}^nP_r$	permutations
$\sum$	sum
u_n	the nth term
S_n, S_∞	sum of n terms, sum to infinity
d, r	common difference, common ratio
Δ	the discriminant, $b^2 - 4ac$

FUNCTIONS

$f(x)$	the value of f at x
f^{-1}	the inverse function
$f \circ g$	composite: f after g
$\lfloor x \rfloor$	the greatest integer $\leq x$
$\rightarrow$	tends to
∞	infinity

CALCULUS

$\lim_{x \to a}$	limit as x tends to a
dy/dx	derivative of y with respect to x
f', f''	first and second derivatives
d^ny/dx^n	the nth derivative
$\int$	integral
$\int_a^b$	definite integral from a to b
$+C$	constant of integration
$\dot{x}$, $\ddot{x}$	derivatives with respect to time

GEOMETRY AND ANGLES

$[AB]$	the segment with endpoints A and B
AB	the length of $[AB]$
(AB)	the line through A and B
$A\hat{B}C$	the angle at vertex B
$\triangle ABC$	the triangle ABC
$\perp$, $\parallel$	perpendicular to, parallel to
°, rad	degrees, radians

VECTORS

$\mathbf{a}$	a vector
$\lvert \mathbf{a} \rvert$	the magnitude of $\mathbf{a}$
$\overrightarrow{AB}$	the vector from A to B
$\binom{h}{k}$	a column vector; a translation
$\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$	unit vectors along the axes
$\mathbf{a} \cdot \mathbf{b}$	scalar (dot) product
$\mathbf{a} \times \mathbf{b}$	vector (cross) product

COMPLEX NUMBERS

i	the imaginary unit, $i^2 = -1$
z^*	the conjugate of z
$\lvert z \rvert$, $\arg z$	modulus and argument of z
Re, Im	real part, imaginary part
$\mathrm{cis}\,\theta$	$\cos\theta + i\sin\theta$

STATISTICS AND PROBABILITY

$\bar{x}$	the sample mean
μ, σ, σ^2	population mean, standard deviation, variance
Q_1, Q_3	lower and upper quartiles
IQR	interquartile range
r	Pearson's correlation coefficient
$P(A)$	the probability of A
A'	the complement of A
$P(A \mid B)$	the probability of A given B
$E(X)$	the expected value of X
$\mathrm{Var}(X)$	the variance of X
$\sim$	is distributed as
$B(n, p)$	binomial distribution
$N(\mu, \sigma^2)$	normal distribution

ABBREVIATIONS

AA HL	Analysis and Approaches, Higher Level
AHL	additional higher level content
GDC	graphical display calculator
s.f., d.p.	significant figures, decimal places
IA	internal assessment
TOK	Theory of Knowledge

NUMBER & ALGEBRA

Chapter 1

Exponents & Logarithms

1

SYLLABUS 1.1 1.5 1.7 2.9

Japan lies in one of the world's most active seismic zones. Suppose a seismologist asks you to plot two recordings on one axis. The first has an amplitude (the height of the wave) of 10^4. The second is 10^9, a hundred thousand times larger.

An ordinary scale, with evenly spaced marks, cannot show both. If the marks are close enough together to show the first, the second lies far beyond the page. If the scale is long enough for the second, the first sits so close to zero that you cannot see it.

So let equal spaces along the axis mean multiplication by 10 rather than addition of 1. The scale then reads $10^0, 10^1, 10^2$, and the two recordings, with exponents 4 and 9, sit five equal steps apart. That exponent is the **logarithm** of the amplitude, and it is the idea this chapter is built on.

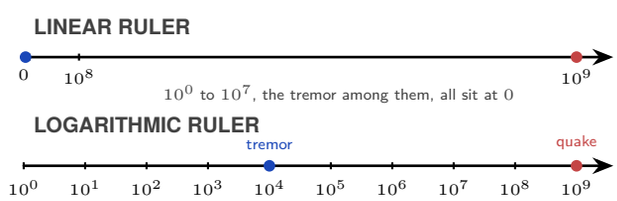


Figure 1.1 A logarithmic scale separates values that collapse near zero on a linear scale.

Earthquake magnitude works this way, and so do pH, decibels, and the brightness of stars. Writing a quantity as a power of ten turns multiplication into addition, so a range too wide to plot fits on an axis you can read. By the end of this chapter you will be able to use the laws of exponents and standard form, move between exponential and logarithmic form, apply the laws of logarithms and change the base of a logarithm, and solve equations with the unknown in the exponent.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

1.1 What an Exponent Counts

An exponent, also called a *power* or an *index*, is shorthand for repeated multiplication. In a^n, a is the **base** and n is the **exponent**. When n is a positive integer,

$$a^n = \underbrace{a \cdot a \cdots a}_{n \text{ factors}}.$$

For example, $a^3 = a \cdot a \cdot a$.

When you multiply powers with the same base, it is tempting to multiply the exponents. Writing out the factors shows what actually happens:

$$a^3 \cdot a^2 = (a \cdot a \cdot a)(a \cdot a) = a^5.$$

There are five factors of a, not six. Since the exponent counts how many times the base appears as a factor, the exponents are **added**:

$$a^m \cdot a^n = a^{m+n}.$$

Division cancels common factors, so the exponents are **subtracted**:

$$\frac{a^5}{a^2} = \frac{a \cdot a \cdot a \cdot a \cdot a}{a \cdot a} = a^3 = a^{5-2}.$$

Raising a power to another power repeats the whole power, so the exponents are **multiplied**:

$$(a^2)^3 = a^2 \cdot a^2 \cdot a^2 = a^6 = a^{2\times 3}.$$

A power outside a bracket applies to every factor inside:

$$(ab)^n = a^n b^n, \qquad \left(\frac{a}{b}\right)^n = \frac{a^n}{b^n}.$$

The same laws extend naturally to zero and negative exponents. For $a \neq 0$, the subtraction rule gives $a^3/a^3 = a^{3-3} = a^0$; since any nonzero number divided by itself is 1, this forces $a^0 = 1$. Continuing the same pattern gives $a^{-n} = 1/a^n$. These values are not imposed on us; they are the only ones that keep the laws consistent. The main laws are collected below.

DEF For $a, b \neq 0$ and $m, n \in \mathbb{Z}$:

$a^m \cdot a^n = a^{m+n}$	product of powers: add the exponents
$\dfrac{a^m}{a^n} = a^{m-n}$	quotient of powers: subtract the exponents
$(a^m)^n = a^{mn}$	power of a power: multiply the exponents
$(ab)^n = a^n b^n$	power of a product: raise each factor
$\left(\dfrac{a}{b}\right)^n = \dfrac{a^n}{b^n}$	power of a quotient: raise top and bottom
$a^0 = 1$	zero exponent: gives 1
$a^{-n} = \dfrac{1}{a^n}$	negative exponent: take the reciprocal

CAUTION Do not confuse $(a^m)^n$ with a^{m^n}. In $(a^m)^n$, the power of a power, you multiply the exponents. In a^{m^n} you start at the top: first work out m^n, then raise a to that result. The two are usually not equal: $(2^3)^2 = 2^6 = 64$, but $2^{3^2} = 2^9 = 512$.

CAUTION A negative exponent means *reciprocal*, not *negative value*. $2^{-3} = \frac{1}{8}$, not -8.

1.1.1 Worked examples

How do these rules combine when a problem has several operations at once?

Worked through. To simplify $\frac{10c^7}{5c^3}$, split the coefficient and the variable and handle them separately:

$$\frac{10c^7}{5c^3} = \frac{10}{5} \cdot c^{7-3} = 2c^4.$$

Apply one rule at a time.

EXAMPLE 1.1

Simplify $(2x^2y^{-4})^5$, giving your answer with positive exponents. It contains several operations: a bracket raised to a power and a negative exponent. Raise each part in turn: the 2 to the 5th, x^2 to the 5th, and y^{-4} to the 5th.

$$(2x^2y^{-4})^5 = 32x^{10}y^{-20} = \frac{32x^{10}}{y^{20}}.$$

CAUTION The bracket in $(2x^2)^3$ applies to everything inside it, including the 2. $(2x^2)^3 = 8x^6$, not $2x^6$.

1.1.2 Beyond counting factors

The idea of a^n as repeated multiplication only makes sense when n is a positive integer. The laws above already fixed $a^0 = 1$ and $a^{-n} = 1/a^n$. A fractional exponent such as $3^{1/2}$ needs the same treatment: we choose its value so that the laws keep working, and the next section shows that this gives $3^{1/2} = \sqrt{3}$.

1.1.3 Surds and fractional exponents

What is $9^{1/2}$? Since $\frac{1}{2}$ is not an integer, repeated multiplication cannot answer this directly. But the power-of-a-power law gives

$$\left(9^{1/2}\right)^2 = 9^{(1/2)\times 2} = 9,$$

so $9^{1/2}$ must be the non-negative number whose square is 9:

$$9^{1/2} = \sqrt{9} = 3.$$

Notice the value is 3, not ± 3: the symbol $\sqrt{9}$ means the **principal** (non-negative) square root.

DEF For $a > 0$, $n \in \mathbb{Z}^+$ and $m \in \mathbb{Z}$:

$a^{1/n} = \sqrt[n]{a}$ — unit fraction: the denominator is the root

$a^{m/n} = \left(\sqrt[n]{a}\right)^m = \sqrt[n]{a^m}$ — in general: root then power, in either order

When n is even the symbol $\sqrt[n]{a}$ means the **positive** root, exactly as $\sqrt{9}$ means 3 and not ± 3.

A useful way to remember this:

denominator n $\longrightarrow$ root,	numerator m $\longrightarrow$ power

These are not new laws; they are the meanings chosen so that the familiar exponent laws keep holding.

Using fractional exponents

For $a^{m/n}$, it is usually easier to take the root first, then apply the power:

$$64^{2/3} = \left(\sqrt[3]{64}\right)^2 = 4^2 = 16.$$

Either order gives the same result, $64^{2/3} = \sqrt[3]{64^2} = 16$, but taking the root first keeps the numbers small.

The same laws run backwards. If the unknown carries a fractional exponent, raise both sides to the reciprocal of that exponent: the two multiply to 1, and the power-of-a-power law leaves the unknown on its own. The definition above covers positive bases only, so this move is safe when the question restricts the unknown to positive values. If a question allows negative values and reads $x^{2/5}$ as $(\sqrt[5]{x})^2$, the even power discards the sign: $x^{2/5} = 4$ gives $\sqrt[5]{x} = \pm 2$, so $x = \pm 32$, while raising both sides to $\frac{5}{2}$ finds only 32. Check the restriction on x before counting solutions.

EXAMPLE 1.2

Solve $x^{3/4} = 125$ for $x > 0$.

The exponent $\frac{3}{4}$ is undone by $\frac{4}{3}$, since $\frac{3}{4} \times \frac{4}{3} = 1$. Raise both sides to that power:

$$\left(x^{3/4}\right)^{4/3} = 125^{4/3} \implies x^1 = \left(\sqrt[3]{125}\right)^4 = 5^4 = 625.$$

Check by substituting: $625^{3/4} = \left(\sqrt[4]{625}\right)^3 = 5^3 = 125$. The condition $x > 0$ is what allows the fractional exponent in the first place, since $x^{3/4}$ means $(\sqrt[4]{x})^3$ and the fourth root needs a non-negative input.

Working with surds

A **surd** is an irrational root left in exact form, such as $\sqrt{2}$ or $\sqrt[3]{5}$, rather than rounded to a decimal. On a non-calculator paper, surds should be kept exact and simplified using the root laws

$$\sqrt{ab} = \sqrt{a}\,\sqrt{b}, \qquad \sqrt{\tfrac{a}{b}} = \frac{\sqrt{a}}{\sqrt{b}},$$

valid for $a, b \geq 0$ (with $b > 0$ in the quotient). Two techniques cover most questions:

1. **Simplify** a surd by extracting its largest perfect-square factor.
2. **Combine** only like surds, the terms that have the same value under the root.

Two surds that look unrelated often become like surds once each has been simplified, so always simplify before trying to combine.

CAUTION The root laws hold for products and quotients only. $\sqrt{a} + \sqrt{b}$ is not $\sqrt{a+b}$, and

$\sqrt{a}-\sqrt{b}$ is not $\sqrt{a-b}$, so surds can be combined only when they are like surds.

EXAMPLE 1.3

Simplify $\sqrt{147}-\sqrt{48}$.

Take the largest perfect-square factor out of each root:

$$\sqrt{147}=\sqrt{49\cdot 3}=7\sqrt{3}, \qquad \sqrt{48}=\sqrt{16\cdot 3}=4\sqrt{3}.$$

Both terms now carry $\sqrt{3}$, so they are like surds and only their coefficients subtract:

$$7\sqrt{3}-4\sqrt{3}=3\sqrt{3}.$$

The root laws cover products and quotients but never sums or differences, so $\sqrt{147}-\sqrt{48}$ is not $\sqrt{99}$: one is $3\sqrt{3}\approx 5.20$, the other about 9.95.

1.1.4 Rationalising the denominator

An exact answer is usually written without a surd in the denominator; removing it is called **rationalising the denominator**.

- For a single surd, multiply top and bottom by that surd.
- For a sum or difference, multiply by its **conjugate**, formed by changing the sign between the terms.

The conjugate works because $(a+\sqrt{b})(a-\sqrt{b})=a^2-b$ is a difference of two squares, so the surd drops out of the denominator.

EXAMPLE 1.4

Rationalise $\dfrac{1}{\sqrt{3}}$ and $\dfrac{6}{2+\sqrt{3}}$.

For the first, multiply by $\dfrac{\sqrt{3}}{\sqrt{3}}$:

$$\frac{1}{\sqrt{3}}=\frac{\sqrt{3}}{3}.$$

For the second, multiply by the conjugate $2-\sqrt{3}$:

$$\frac{6}{2+\sqrt{3}}=\frac{6(2-\sqrt{3})}{(2+\sqrt{3})(2-\sqrt{3})}=\frac{6(2-\sqrt{3})}{4-3}=12-6\sqrt{3}.$$

The conjugate works because the product is a difference of two squares, $(2+\sqrt{3})(2-\sqrt{3})=2^2-(\sqrt{3})^2=1$, so the surd leaves the denominator.

TIP In the real numbers, $(-4)^{1/2}$ is undefined: no real number squares to -4. In general, the radicand, the number inside the root, must be non-negative for an even root to have a real value.

Square roots of negative numbers are complex numbers, which are introduced in Chapter 17.

1.1.5 Exponents as continuous growth

There is another way to read a non-integer exponent. Suppose a quantity doubles once in every unit of time, so that 2^t is its growth factor after time t. Then 2^3 is three complete doublings, while

$$2^{1/2} = \sqrt{2}$$

is the growth factor halfway through one doubling period. Because the growth is exponential, "halfway" refers to the time, not to half the final increase. This reading extends to every real value of t, whether negative, fractional, or irrational:

$$2^{-3}, \qquad 2^{0}, \qquad 2^{1/2}, \qquad 2^{\sqrt{2}}.$$

Read as growth over time t, each has a plain meaning. $2^0 = 1$ is no growth at all: at time zero the population is unchanged, one hundred per cent of where it started. $2^{-3} = \frac{1}{8}$ looks backwards in time, to three periods before now (three seconds ago, if one doubling takes a second), when the population was an eighth of its present size. $2^{1/2} = \sqrt{2}$ is the factor after half a period, and $2^{\sqrt{2}}$ the factor after an irrational stretch of time. For any positive base a, the expression a^x has a meaning for every real exponent x, which is where the exponential function of Chapter 5 begins.

1.1.6 The number *e*

Suppose an account pays 100% interest in a year. Paid once, at the end, an amount of 1 becomes 2. Now suppose the same 100% arrives in two instalments of 50%. After six months the balance is $1 \times 1.5 = 1.5$, and the second 50% is taken on 1.5 rather than on 1, so the year ends at $1.5 \times 1.5 = 2.25$. Splitting the same total rate into more instalments gives more, because an amount added early is itself subject to all the growth that follows.

That leads to a question with a surprising answer. If a finer split always gives more, does the year-end amount grow without limit? In n equal steps the starting amount is multiplied by $\left(1 + \frac{1}{n}\right)^n$, so the question is what that expression does as n increases.

instalments in the year, n	amount after one year
1 (once, at the end)	2.00000
2 (every six months)	2.25000
12 (monthly)	2.61304
365 (daily)	2.71457
8760 (hourly)	2.71813
$n \to \infty$ (continuous)	$2.71828\ldots = e$

It does not grow without limit. Each extra split adds less than the split before it, because the money it adds arrives later in the year and so has less of the year left in which to earn anything itself. Going from yearly to six-monthly gains 0.25; going from daily to hourly gains about 0.0036. The totals close in on one value, $2.71828\ldots$, and that number is called e. Like π, it is irrational: its decimal expansion neither ends nor repeats.

So e answers one specific question: what does 100% growth come to when it acts continuously instead of in steps? Other rates are handled in the exponent. A continuous rate k acting for time t multiplies the starting amount by e^{kt}. This model describes radioactive decay, and it is a common first model for population growth and for the gap between a cooling object's temperature and its surroundings. Chapter 13 gives a deeper reason for e: the function e^x grows at a rate equal to its own value.

EXT Formally, $e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$. Each n is a number of compounding steps per year, so letting n grow without bound is what "continuously" means. Limits are defined properly in Ch. 13.

YOUR TURN 1.1 What an Exponent Counts

Laws of exponents

1. Simplify each of the following, giving your answers with positive exponents.

(a) $\dfrac{6c^9}{2c^4}$

(b) $(3x^2y^{-3})^4$

(c) $\dfrac{12a^5b^3}{4a^2b^{-1}}$

(d) $\dfrac{(2x^3)^2 \cdot x^{-4}}{4x^2}$

2. Simplify each of the following.

(a) $\dfrac{4x^3 - 8x^2}{2x}$

(b) $\dfrac{16a^5b^2 - 12ab^4}{4ab^2}$

3. Show that $\dfrac{(a^2b^{-1})^3}{a^{-3}b^2} = \dfrac{a^9}{b^5}$.

4. Evaluate each pair of numbers without a calculator, and state whether the two numbers are equal.

(a) $(2^2)^3$ and 2^{2^3}

(b) $(3^2)^2$ and 3^{2^2}

Fractional exponents

5. Evaluate each of the following.

(a) $27^{2/3}$

(b) $81^{-3/4}$

(c) $\left(\dfrac{8}{27}\right)^{-2/3}$

(d) $32^{-2/5}$

6. Write each of the following in the form x^n.

(a) $\sqrt[4]{x^3}$

(b) $\dfrac{1}{\sqrt[3]{x^2}}$

(c) $\dfrac{\sqrt{x^5}}{x^2}$

(d) $\dfrac{x^2 \cdot \sqrt{x}}{x^{3/2}}$

Equations with a fractional exponent

7. Solve each equation for $x > 0$.

(a) $x^{3/2} = 64$

(b) $x^{2/3} = 9$

(c) $x^{-1/2} = \frac{1}{4}$

(d) $x^{5/2} = 32$

8. In this question x may be any real number, and $x^{m/n}$ means $\left(\sqrt[n]{x}\right)^m$ with n odd. Solve each equation, giving every real solution.

(a) $x^{2/5} = 4$

(b) $x^{4/3} = 81$

Surds

9. Simplify each of the following.

(a) $\sqrt{50} - \sqrt{8}$ (b) $\sqrt{12} + \sqrt{27}$

(c) $\sqrt{18} + \sqrt{32} - \sqrt{2}$ (d) $\sqrt{75} - 2\sqrt{12}$

Rationalising the denominator

10. Rationalise the denominator of each expression.

(a) $\dfrac{4}{3-\sqrt{5}}$, in the form $a + b\sqrt{5}$ (b) $\dfrac{6}{\sqrt{3}}$

(c) $\dfrac{1}{\sqrt{7}+\sqrt{5}}$ (d) $\dfrac{2+\sqrt{3}}{2-\sqrt{3}}$

11. Rationalise $\dfrac{4+\sqrt{15}}{4-\sqrt{15}}$, giving your answer in the form $a + b\sqrt{15}$.

Concept check

12. A student writes $\sqrt{9+16} = \sqrt{9} + \sqrt{16} = 7$.

(a) Identify the error, and find the correct value.

(b) State the law for square roots that does hold, and use it to simplify $\sqrt{9 \times 16}$.

1.2 Standard Form

The mass of a proton, written out in full, is

$$0.00000000000000000000000000167 \text{ kg},$$

that is, 26 zeros after the decimal point before the first significant figure. The mass of the Sun runs to the opposite extreme,

$$1{,}990{,}000{,}000{,}000{,}000{,}000{,}000{,}000{,}000 \text{ kg},$$

a 1 followed by 30 digits. Both numbers are quite difficult to write, let alone compare or calculate with.

There is a shorter way to write both. Instead of all the digits, record *the power of 10* that places the decimal point where it belongs. Mathematicians call the result **standard form** (also known as scientific notation).

DEF A number is in **standard form** when written as $a \times 10^k$, where $1 \leq a < 10$ and $k \in \mathbb{Z}$.

The proton becomes 1.67×10^{-27} kg. The Sun becomes 1.99×10^{30} kg. The exponent k tells you immediately whether you are dealing with something microscopic or cosmic.

1.2.1 Calculating with standard form

Multiplication and division follow straight from the exponent laws. For a product, multiply the coefficients and add the powers of 10; for a quotient, divide the coefficients and subtract the powers:

$$(a \times 10^m)(b \times 10^n) = (a \times b) \times 10^{m+n}.$$

One check matters afterwards. If the new coefficient falls outside the range $1 \leq a < 10$, shift the decimal point and adjust the power. For example, $(5 \times 10^3)(4 \times 10^2) = 20 \times 10^5 = 2 \times 10^6$: the coefficient 20 becomes 2, and the exponent rises by one to account for the extra factor of 10.

Addition and subtraction need one more move first. Just as you cannot add metres and kilometres without converting, you cannot add two numbers in standard form until they share the same power of 10. Rewrite the one with the smaller exponent so the powers match, then add the coefficients.

EXAMPLE 1.5

Calculate $(3.5 \times 10^4) \times (2 \times 10^{-7})$.

Group and multiply:

$$(3.5 \times 2) \times 10^{4+(-7)} = 7 \times 10^{-3}.$$

EXAMPLE 1.6

Calculate $(4.8 \times 10^5) + (3.2 \times 10^4)$.

Rewrite 3.2×10^4 as 0.32×10^5 so both terms share the power 10^5:

$$4.8 \times 10^5 + 0.32 \times 10^5 = 5.12 \times 10^5.$$

CAUTION In addition and subtraction, do **not** add the exponents. Only multiplication and division combine powers of 10.

CAUTION Calculator display notation is not acceptable in written answers. If your graphic display calculator (GDC) shows 5.2E30, you must write 5.2×10^{30}.

YOUR TURN 1.2 Standard Form

Converting between forms

13. Write 0.000047 in standard form.

14. Write each number in ordinary form.

(a) 3.2×10^5 (b) 8.06×10^{-4}

(c) 1.5×10^7 (d) 9×10^{-1}

Calculating in standard form

15. Calculate $(2.4 \times 10^3) \times (3 \times 10^{-5})$, giving your answer in standard form.

16. Calculate $\dfrac{6.6 \times 10^8}{2.2 \times 10^3}$, giving your answer in standard form.

17. Calculate each of the following, giving your answer in standard form.

(a) $(5.1 \times 10^4) + (3.9 \times 10^3)$ (b) $(7.2 \times 10^{-3}) - (5 \times 10^{-4})$

(c) $(2.5 \times 10^6) + (7.5 \times 10^6)$ (d) $(9.4 \times 10^{-2}) + (6 \times 10^{-2})$

Using a calculator

18. A GDC shows each result below in its display notation. Write each one in standard form, as it must appear in a written answer.

(a) 6.02E23 (b) 4.5E-8

19. GDC The mass of an electron is 9.11×10^{-31} kg. The mass of a proton is 1.67×10^{-27} kg. Find how many times heavier a proton is than an electron. Give your answer in standard form to 3 significant figures.

Concept check

20. A student writes $(3 \times 10^4) + (2 \times 10^3) = 5 \times 10^7$. Identify the error, and give the correct answer in standard form.

1.3 The Logarithm: an Exponent Reversed

Many familiar operations come in inverse pairs: addition with subtraction, multiplication with division. So what is the inverse of exponentiation? If $a^x = b$, and you know a and b, how do you find x? That is what a logarithm does. It is not a new idea: it is the exponential equation $a^x = b$ rearranged to make the exponent x the subject.

DEF · IB For $a > 0$, $a \neq 1$, and $b > 0$:

$$a^x = b \iff x = \log_a b$$

The **common logarithm** $(\log x)$ uses base 10. The **natural logarithm** $(\ln x)$ uses base $e \approx 2.718$, the continuous-growth constant met earlier in this chapter.

The condition $b > 0$ in that definition is not a technicality. A positive base raised to any real power gives a positive result: $10^2 = 100$, $10^0 = 1$, $10^{-3} = 0.001$. No exponent turns 10 into -5, and no exponent turns 10 into 0. So $\log(-5)$ and $\log 0$ do not exist. **Whatever appears inside a logarithm must be positive.** The same holds for every base: $\log_a x$ **is defined only when** $x > 0$**.**

The base carries conditions of its own. It must be positive: a negative base has no consistent meaning for fractional exponents, and base 0 gives only 0 or undefined powers; and it cannot be 1, because every power of 1 is 1 and no exponent would give any other value.

The word *logarithm* comes from the Greek *logos* (ratio) and *arithmos* (number).

The graph of $y = \log_a x$ is the mirror image of $y = a^x$ across the line $y = x$, with a vertical asymptote at $x = 0$. These graphs and their transformations are developed in Chapter 5 (Functions).

Because the exponential and logarithm are inverses, each undoes what the other does:

KEY For base 10 and base e:

$\log(10^x) = x$	base 10: true for every real x
$10^{\log x} = x$	the same pair reversed, but only for $x > 0$, since $\log x$ is undefined otherwise
$\ln(e^x) = x$	base e: true for every real x
$e^{\ln x} = x$	reversed, and again only for $x > 0$

The formula booklet prints the general statement $\log_a a^x = x = a^{\log_a x}$; these four lines are its base-10 and base-e cases.

Two special values follow at once from the definition, and both are used often: since $a^1 = a$ and $a^0 = 1$,

KEY For any base $a > 0$, $a \neq 1$:

$\log_a a = 1$ since $a^1 = a$

$\log_a 1 = 0$ since $a^0 = 1$

1.3.1 Working with logarithms

To evaluate a logarithm by hand, reverse the exponent question: $\log_a b$ is defined as the exponent that turns a into b. Read $\log_a b$ as **"a to what power gives b?"** and in many cases the answer is immediate.

Worked through. To find $\log_3 81$, ask the inverse question: 3 to what power gives 81? You know $3^4 = 81$, so $\log_3 81 = 4$. The logarithm is just the exponent.

EXAMPLE 1.7

Solve $\log_{10}(y+1) = 3$.
The equation says: the exponent you need to get $(y+1)$ from base 10 is 3. So $y + 1 = 10^3 = 1000$, giving $y = 999$.

EXAMPLE 1.8

Without a calculator, find $e^{3\ln 2}$.
Think of $e^{\ln 2}$ first: the exponential and natural log cancel, giving 2. So $e^{3\ln 2} = (e^{\ln 2})^3 = 2^3 = 8$.

CAUTION Solving an equation with a logarithm in it can produce a value that makes an argument negative or zero. Such a value is not a solution, however correct the algebra was. Substitute every answer back into the original equation and keep only those that leave each argument positive.

YOUR TURN 1.3 The Logarithm: an Exponent Reversed

Switching between forms

21. Rewrite each statement in the other form.

(a) Write $3^4 = 81$ in logarithmic form.

(b) Write $\log_5 x = 3$ in exponential form, and hence find x.

Evaluating logarithms

22. Evaluate each of the following without a calculator.

(a) $\log_2 32$ (b) $\log_3 \frac{1}{9}$

(c) $e^{2\ln 3}$ (d) $\ln e^5$

23. Evaluate each of the following without a calculator.

(a) $\log 10\,000$ (b) $\log 0.001$

(c) $\ln \frac{1}{\sqrt{e}}$ (d) $10^{\log 5+1}$

24. State whether each expression is defined. If it is, write down its value.

(a) $\log_5 125$ (b) $\log_5 1$

(c) $\log_5 0$ (d) $\log_5(-25)$

(e) $\log_1 7$ (f) $\log_{1/2} 8$

Logarithm and exponential as inverses

25. Each statement is either true for every value allowed by the definition, or not. State which, and correct those that are not.

(a) $\log_a(a^x) = x$ for every real x (b) $a^{\log_a x} = x$ for every real x

(c) $\log_a 1 = 0$ for every valid base a (d) $\log_a 0 = 0$ for every valid base a

Solving logarithmic equations

26. Solve each equation. Reject any value the definition does not allow.

(a) $\log_4(x+1) = 2$ (b) $\log_3(2x-1) = 4$

(c) $\log_2(x^2 - 3x) = 2$

Concept check

27. A student says: "$\log_{10} x = -2$ *has no solution, because a logarithm cannot be negative.*" Explain the error, and solve the equation.

1.4 The Laws of Logarithms

A single logarithm is just a number. What makes logarithms worth having is that they trade hard arithmetic for easy arithmetic: multiplication becomes addition, division becomes subtraction, and a power becomes a coefficient. The last of those is the one this chapter is heading towards, because it is what brings an unknown down out of an exponent.

The laws work because a logarithm is an exponent. Multiplying two powers with the same

base adds the exponents, $a^p \cdot a^q = a^{p+q}$, and reading that sentence in terms of logarithms is already the first law: multiplication becomes addition.

The first three laws below are the three exponent laws restated for logarithms. A fourth, change of base, is a different kind of rule: it rewrites any logarithm as a ratio of two logarithms in a base of your choice.

KEY · IB For $a, b, x, y > 0$ with $a, b \neq 1$:

$\log_a(xy) = \log_a x + \log_a y$	product: a product becomes a sum
$\log_a\left(\frac{x}{y}\right) = \log_a x - \log_a y$	quotient: a quotient becomes a difference
$\log_a(x^m) = m \log_a x$	power: an exponent becomes a coefficient
$\log_a x = \frac{\log_b x}{\log_b a}$	change of base: rewrite in any base you can compute

All three proofs share one setup. Write $x = a^p$ and $y = a^q$, so that $p = \log_a x$ and $q = \log_a y$ by definition. Each law is then a one-line consequence of the matching exponent law:

- **Product.** $xy = a^p \cdot a^q = a^{p+q}$, so $\log_a(xy) = p + q = \log_a x + \log_a y$.
- **Quotient.** $\frac{x}{y} = a^{p-q}$, so $\log_a\left(\frac{x}{y}\right) = p - q = \log_a x - \log_a y$.
- **Power.** $x^m = (a^p)^m = a^{mp}$, so $\log_a(x^m) = mp = m \log_a x$.

The third law is the one used to solve exponential equations. When the unknown is in an exponent, as in b^x, the power law gives $\log_a(b^x) = x \log_a b$, which moves the unknown into a coefficient. This often converts an exponential equation into a simpler algebraic equation.

1.4.1 Change of base

What do you do when your GDC only has $\log$ (base 10) and $\ln$ (base e), but you need a logarithm in a different base, or you want to rewrite a base-10 logarithm in terms of base e? The change-of-base law above rewrites any logarithm as a ratio of two logarithms you can compute. In practice you take base e, which turns the law into $\log_a x = \frac{\ln x}{\ln a}$.

To prove it, let $\log_a x = y$, so $x = a^y$. Take $\ln$ of both sides: $\ln x = y \ln a$. Solve for y.

EXAMPLE 1.9

Evaluate $\log_3 5$ on a GDC.

Convert to natural logarithms:

$$\log_3 5 = \frac{\ln 5}{\ln 3} \approx \frac{1.609}{1.099} \approx 1.46.$$

One case of the law is worth carrying separately. Take the argument to be b and compute in base b; since $\log_b b = 1$, the numerator disappears:

$$\log_a b = \frac{\log_b b}{\log_b a} = \frac{1}{\log_b a}.$$

KEY For $a, b > 0$ with $a, b \neq 1$:

$$\log_a b = \frac{1}{\log_b a}.$$

The formula booklet does not print this, so quote it with the change-of-base step above, or work from change of base directly.

Swapping a base and an argument is the move to reach for when the unknown sits in the base, because it carries that unknown into an argument, where the ordinary laws apply. Often no swap is needed at all: a logarithm with an unknown base is still an exponent question read backwards.

EXAMPLE 1.10

Solve $\log_x 81 = 4$.

The statement says that x raised to the power 4 gives 81, so

$$x^4 = 81 \implies x = 3 \text{ or } x = -3.$$

A base must satisfy $x > 0$ and $x \neq 1$. That rules out $x = -3$, leaving $x = 3$.

An unknown base is where those two conditions matter: the algebra alone offers two roots, and only the definition of a logarithm removes one.

1.4.2 When one base is a power of another

Questions often mix bases that are powers of one another: $\log_2 x$ alongside $\log_4 x$, or $\log_2 x$ alongside $\log_8 x$. Change of base rewrites them all in one base, and for this particular case the result is short enough to know on sight.

Change $\log_{a^2} x$ to base a:

$$\log_{a^2} x = \frac{\log_a x}{\log_a a^2} = \frac{\log_a x}{2} = \tfrac{1}{2}\log_a x,$$

since $\log_a a^2 = 2\log_a a = 2$ by the power law. Squaring the base halves the logarithm. The same steps with a^n in place of a^2 give $\log_a a^n = n$, so for $a > 0$, $a \neq 1$, $x > 0$ and any $n \neq 0$,

$$\log_{a^n} x = \frac{1}{n} \log_a x.$$

This result is not in the formula booklet. Show the change-of-base step that gives it.

EXAMPLE 1.11

Solve $\log_2 x + \log_4 x = 6$.

The bases are 2 and $4 = 2^2$, so rewrite the second one in base 2: $\log_4 x = \frac{1}{2}\log_2 x$. Writing $t = \log_2 x$, the equation becomes

$$t + \tfrac{1}{2}t = 6 \implies \tfrac{3}{2}t = 6 \implies t = 4,$$

so $x = 2^4 = 16$. Check: $\log_2 16 = 4$ and $\log_4 16 = 2$, and these sum to 6.

1.4.3 Applying the laws

These laws can be used in either direction: they split one logarithm into several, or gather several back into one. Both directions are needed, in simplifying expressions and in solving equations.

Worked through. To express $\log_5\left(\frac{a^3}{b}\right)$ in terms of $p = \log_5 a$ and $q = \log_5 b$, apply the quotient law, then the power law:

$$\log_5\left(\frac{a^3}{b}\right) = \log_5 a^3 - \log_5 b = 3\log_5 a - \log_5 b = 3p - q.$$

EXAMPLE 1.12

Write $2\ln a + 6\ln b$ as a single logarithm.

Use the power law in reverse first, then the product law:

$$2\ln a + 6\ln b = \ln a^2 + \ln b^6 = \ln(a^2 b^6).$$

EXAMPLE 1.13

Solve $\log_2 x + \log_2(x-2) = 3$.

The two logs on the left can combine into one product:

$$\log_2[x(x-2)] = 3 \implies x(x-2) = 2^3 = 8.$$

$$x^2 - 2x - 8 = 0 \implies (x-4)(x+2) = 0.$$

So $x = 4$ or $x = -2$. But $\log_2(-2)$ does not exist, so reject $x = -2$. The answer is $x = 4$. Always check domain when logarithms are involved. It is easy to generate solutions that look valid algebraically but are meaningless.

CAUTION $\log(x + y) \neq \log x + \log y$. The product law is about $\log(xy)$, not a sum inside the log.

YOUR TURN 1.4 The Laws of Logarithms

Combining and expanding logarithms

28. Write each expression as a single logarithm.

(a) $\log_3 5 + \log_3 7$

(b) $\log_2 24 - \log_2 3$, and evaluate it

(c) $3\log_a x + \log_a y$

(d) $2\ln a - 3\ln b + \ln c$

29. Express each logarithm in terms of the quantities given.

(a) $\log_5\left(\dfrac{a^2}{b}\right)$, in terms of $\log_5 a$ and $\log_5 b$

(b) $\log_3 18$, given $p = \log_3 2$

(c) $\log_{10} 72$, given $\log_{10} 2 = p$ and $\log_{10} 3 = q$

(d) $\log_5 75$, given $\log_5 3 = 0.682$

30. Show that $\log_a(x^2 - 1) - \log_a(x + 1) = \log_a(x - 1)$, for $x > 1$.

Change of base

31. Use the change of base formula in each of the following.

(a) Write $\log_4 9$ in terms of natural logarithms.

(b) Express $\log_8 x$ in terms of $\log_2 x$.

(c) Find the exact value of $\log_4 8$.

(d) Show that $\log_{1/2} x = -\log_2 x$.

(e) Show that $\log_a b \times \log_b(a^2) = 2$, for $a, b > 0$ and $a, b \neq 1$.

32. GDC Evaluate $\log_3 50$ to 3 significant figures.

Solving equations with the laws

33. Solve each equation, using a law of logarithms first. Reject any value the logarithms do not allow.

(a) $\log_3 x + \log_3(x + 6) = 3$

(b) $\log_2(x + 3) - \log_2(x - 1) = 2$

(c) $2\log_5 x = \log_5 9 + \log_5 4$

Bases that are powers of one another

34. Solve each equation by writing every logarithm in the same base.

(a) $\log_3 x - \log_9 x = 1$

(b) $\log_2 x + \log_8 x = 8$

(c) $\log_4 x = \log_2(x-2)$

An unknown base

35. Solve each equation for the base x. Reject any value a base does not allow.

(a) $\log_x 49 = 2$

(b) $\log_x 8 + \log_x 2 = 4$

(c) $\log_x(x+6) = 2$

Concept check

36. A student writes $\log_3(x+9) = \log_3 x + \log_3 9 = \log_3 x + 2$.

(a) Identify the law of logarithms the student has misremembered, and state it correctly.

(b) By taking $x = 18$, show that the two sides are not equal.

37. A student writes $\dfrac{\log_2 32}{\log_2 4} = \log_2 8 = 3$. Explain the error, and find the correct value.

1.5 Solving Exponential Equations

In an equation such as $5^x = 30$, ordinary algebra cannot isolate the unknown. Adding, multiplying or factorising does not isolate the x, because x is not being combined with anything: it is counting how many times the base is used. The logarithm is what brings it down, and once it is down the equation is one you already know how to solve. This section starts with equations that need no logarithm, then uses the logarithm on harder cases.

1.5.1 Same base first

Is there ever a way to solve an exponential equation without using a logarithm at all? Before anything else, look for a common base. If both sides can be written as powers of the same number, the exponents must be equal, provided the base is positive and not 1: $a^p = a^q$ forces $p = q$ for $a > 0,\ a \neq 1$. When it is available, this is usually quicker than a logarithm, and it suits non-calculator papers.

EXAMPLE 1.14

Solve $\left(\frac{1}{3}\right)^x = 9^{x+1}$.

Both sides are powers of 3: $\frac{1}{3} = 3^{-1}$ and $9 = 3^2$. Rewrite and equate exponents:

$$3^{-x} = 3^{2(x+1)} \implies -x = 2x + 2 \implies x = -\tfrac{2}{3}.$$

No logarithm was needed. Check for a shared base before anything else; look for common bases such as $2, 4, 8$ and $3, 9, 27$.

1.5.2 The general strategy

When no common base exists, take a logarithm of both sides, use $\log(x^m) = m \log x$ to bring the exponent down, then solve what is now a linear or quadratic equation. The choice of which logarithm to use ($\log$ or $\ln$) does not matter: pick whichever gives cleaner arithmetic.

Worked through. To solve $5^x = 30$, take $\ln$ of both sides:

$$x \ln 5 = \ln 30 \implies x = \frac{\ln 30}{\ln 5} \approx 2.11.$$

Or equivalently, $x = \log_5 30$.

EXAMPLE 1.15

Solve $7^{x-2} = 5^{x+3}$ exactly.

With different bases on each side, take $\ln$ of both sides and use the power law:

$$(x-2)\ln 7 = (x+3)\ln 5.$$

Collect x on one side:

$$x(\ln 7 - \ln 5) = 3\ln 5 + 2\ln 7 \implies x = \frac{\ln 5^3 + \ln 7^2}{\ln(7/5)} = \frac{\ln 6125}{\ln(7/5)}.$$

TIP Give an exact answer, like $x = \dfrac{\ln 6125}{\ln(7/5)}$, whenever a question says *exact* or on a non-calculator paper (Paper 1). When a decimal is asked for, or on a calculator paper (Paper 2), round to three significant figures unless told otherwise. Three significant figures is the IB default, and rounding too early or giving too few figures costs accuracy marks.

1.5.3 When the unknown appears twice

Taking a logarithm gets nowhere on $2^{2x} - 10 \cdot 2^x + 16 = 0$, because x sits in two exponents and a sum of powers has no log law to break it up. This section covers two common patterns. When one power is the square of another, a **substitution** such as $u = a^x$ turns it into a quadratic. Solve the quadratic for u, then convert each root back. Reject any root with $u \leq 0$, since $a^x > 0$ for every real x (with $a > 0$) and no exponent can produce it.

When the exponents differ by a constant, a common factor collapses the terms into one.

EXAMPLE 1.16

Solve $2^{2x} - 10 \cdot 2^x + 16 = 0$.

The first term is the square of the second, since $2^{2x} = (2^x)^2$. Put $u = 2^x$:

$$u^2 - 10u + 16 = 0 \implies (u-2)(u-8) = 0,$$

so $u = 2$ or $u = 8$. Now convert each value back:

$$2^x = 2 \implies x = 1, \qquad 2^x = 8 \implies x = 3.$$

Both roots are positive, so both survive. A negative or zero value of u would have to be rejected instead of converted, since $2^x > 0$ for every real x and no exponent produces it.

EXAMPLE 1.17

Solve $3^{x+2} - 3^x = 72$.

Split the first exponent so that both terms carry 3^x: by the product law, $3^{x+2} = 3^x \cdot 3^2 = 9 \cdot 3^x$. The equation becomes

$$9 \cdot 3^x - 3^x = 72 \implies 8 \cdot 3^x = 72 \implies 3^x = 9,$$

so $x = 2$. Check: $3^4 - 3^2 = 81 - 9 = 72$. A constant added in the exponent is a constant factor outside it, which is what lets the two terms be gathered into one.

1.5.4 From words to an equation

Many growth and decay problems use one model,

$$A(t) = A_0 \times r^{\,t},$$

where A_0 is the starting amount and r is the multiplier for each period. The wording gives you r. A rise of 6% per year means $r = 1.06$. A fall of 6% means $r = 0.94$. A quantity that doubles each period has $r = 2$.

When the change is continuous rather than in steps, the model uses base e:

$$A(t) = A_0\, e^{kt},$$

where k is the **continuous growth rate** met in §1.1: positive when the quantity grows, negative when it decays.

These are not two models but one, written twice. Since $r = e^{\ln r}$,

$$A_0\, r^{\,t} = A_0\, \left(e^{\ln r}\right)^{t} = A_0\, e^{(\ln r)\, t},$$

so the two forms agree exactly when

$$k = \ln r.$$

A quantity that doubles each period has $k = \ln 2 \approx 0.693$; one falling 6% a year has $k = \ln 0.94 \approx -0.0619$, the negative sign marking decay. Use whichever form the question hands you: a percentage per period gives r, a rate described as *continuous* gives k.

Once the model is written, the logarithm provides the final step.

Decay is often described not by a rate but by a **half-life**: the time in which half of the quantity disappears. One half-life multiplies the amount by $\frac{1}{2}$, and a time t contains t/T half-lives, where T is the half-life itself. So

$$A(t) = A_0\, \left(\tfrac{1}{2}\right)^{t/T}.$$

This is the same $A_0\, r^{\,t}$ model with $r = \frac{1}{2}$, and with time counted in half-lives instead of single periods. A **doubling time** works identically, with 2 in place of $\frac{1}{2}$.

EXAMPLE 1.18

Algae in a Finnish lake grow continuously at 12% per year, starting from 200 grams. When does the mass reach 500 grams?

Continuous growth uses base e, so $A(t) = 200\, e^{0.12t}$. Set it equal to the target:

$$200\, e^{0.12t} = 500 \implies e^{0.12t} = 2.5 \implies t = \frac{\ln 2.5}{0.12} \approx 7.64 \text{ years.}$$

The 12% here is a continuous rate, so it sits in the exponent as 0.12 directly, with no conversion to a yearly factor.

EXAMPLE 1.19

A medical isotope has a half-life of 6 hours: every 6 hours, half of it decays. An 80 mg sample is prepared. How long until only 12 mg remains?

Halving every 6 hours means $A(t) = 80\, \left(\frac{1}{2}\right)^{t/6}$, with t in hours. Set this equal to the target and take a logarithm:

$$80\, \left(\tfrac{1}{2}\right)^{t/6} = 12 \implies \left(\tfrac{1}{2}\right)^{t/6} = 0.15 \implies t = 6 \cdot \frac{\ln 0.15}{\ln 0.5} \approx 16.4 \text{ hours.}$$

The answer need not be a whole number of half-lives: 16.4 hours is a little over 2.7 of them, and the model runs continuously between them.

YOUR TURN 1.5 Solving Exponential Equations

Same base

38. Solve each equation, giving your answer as an exact value.

(a) $5^{2x} = 125$
(b) $4^x = 8^{x-1}$
(c) $9^{x+1} = 27^x$
(d) $2^x \cdot 4^{x+1} = 8$

Taking logarithms

39. GDC Solve each equation, giving your answer to 3 significant figures.

(a) $2^x = 10$
(b) $3^{x+1} = 20$
(c) $5^{2x} = 40$
(d) $7 \cdot 2^x = 200$

40. Solve each equation exactly. Give your answers to (a) and (b) in the form $\dfrac{\ln p}{\ln q}$.

(a) $3^{x+1} = 5^x$
(b) $2^{2x-1} = 7^{x+2}$
(c) $3 \cdot 2^x = 2 \cdot 3^x$

The unknown in two exponents

41. Solve each equation. Reject any value that no power can take. *(Hint: in (a), put $u = e^x$)*

(a) $e^{2x} - 5e^x + 6 = 0$
(b) $2^{2x} - 2 \cdot 2^x - 8 = 0$
(c) $3^{2x+1} - 8 \cdot 3^x - 3 = 0$

42. Solve each equation by taking out a common factor.

(a) $2^{x+3} - 2^x = 56$
(b) $5^{x+1} + 5^{x-1} = 130$

Growth and decay models

43. GDC The population of a town is $18\,000$ and rises by 4% each year.

(a) Write down a model for the population P after t years.

(b) Find the number of complete years after which the population first exceeds $30\,000$.

44. GDC A radioactive isotope has a half-life of 5.3 years. A sample has a mass of 200 g.

(a) Find the mass that remains after 12 years.

(b) Find how long it takes until only 15 g remains.

45. GDC A bacterial culture grows continuously. After t hours the number of bacteria is $N = 500\,e^{0.35t}$.

(a) Find the time taken for the number to reach 5000.

(b) Write the model in the form $N = 500\,r^{\,t}$, and hence find the percentage by which the culture grows each hour.

46. GDC Sea ice in a Finnish bay covers 1200 km^2 and shrinks by 6% each year. Find the number of years after which it has shrunk to 800 km^2.

Concept check

47. A student solves $3 \cdot 2^x = 48$ by writing $6^x = 48$, so $x = \dfrac{\ln 48}{\ln 6} \approx 2.16$. Identify the error, and solve the equation correctly.

48. A student solves $e^{2x} + e^x = 12$ by taking $\ln$ of each term, to get $2x + x = \ln 12$. Explain why this step is not valid, and solve the equation correctly.

Chapter 1 · Exponents & Logarithms

Reflections

TOK John Napier spent twenty years computing logarithms by hand and published his tables in 1614, chiefly to shorten the long multiplications of astronomy and navigation. He coined the word "logarithm", but his logarithms were not today's: they had no base in the modern sense and decreased as numbers increased. Henry Briggs, after visiting Napier, produced the base-10 tables that became standard (1617 and 1624). Does mathematics develop because we need it, or do we find uses for what mathematicians have already built?

TOK The exponent laws force $a^0 = 1$ for every $a \neq 0$, but 0^0 is often left undefined, because approaching it as a limit along different routes gives different answers, while algebra and combinatorics (the mathematics of counting) usually set $0^0 = 1$. Does leaving 0^0 undefined reveal a limitation of mathematics or the precision of mathematics?

IM Base-10 logarithms feel natural to us because we count in base 10. But the choice is cultural, not mathematical. The Babylonians used base 60 (traces of which survive in our 60-minute hour and 360-degree circle). The notation is cultural too. The word *algebra* comes from *al-jabr*, a technique described by al-Khwarizmi in ninth-century Baghdad, whose algebra was written entirely in words; the raised exponent used in this chapter was popularised by René Descartes only in 1637. What does it mean for a mathematical tool to be "natural"?

RW The decibel (dB) scale for sound is logarithmic: the sound level is $10 \log_{10}(I/I_0)$, comparing a sound's intensity I to a fixed reference intensity I_0, so each increase of 10 dB means ten times the intensity. The stellar magnitude scale in astronomy and the octave system in music work the same way. In each case, human perception responds approximately to the logarithm of the physical quantity, not to the quantity itself.

End-of-Chapter Problems
Chapter 1: Exponents & Logarithms

1. Solve $3^{2x+1} = 27^{x-1}$. [3]

2. Show that the equation $2^{x+3} - 2^x = k$ has a real solution only when $k > 0$, and find this solution in terms of k. [5]

3.

(a) Simplify $\left(2x^{-2}y^3\right)^3 \times \left(4x^4y^{-1}\right)^{-1}$, giving your answer with positive indices. [3]
(b) Solve $8^x = 2^{x+6}$. [2]
(c) Show that $\left(a^{1/2} + a^{-1/2}\right)^2 = a + 2 + a^{-1}$, for $a > 0$. [2]

4. GDC The mass of the Earth is 5.97×10^{24} kg and the mass of the Moon is 7.35×10^{22} kg.

(a) Find how many times heavier the Earth is than the Moon, to 3 significant figures. [2]
(b) Find their combined mass, giving your answer in standard form to 3 significant figures. [2]

5. Without a calculator, show that $2 < \log_2 5 < 3$. Hence find the integer part of $\log_2 40$. [4]

6. Given that $\log_a x = 3$ and $\log_a y = 5$, find each of the following.

(a) $\log_a(x^2y)$ [1] (b) $\log_a \sqrt{xy}$ [2]
(c) $\log_a\left(\dfrac{x}{y^2}\right)$ [1]

7. Let $\log_a 2 = p$ and $\log_a 3 = q$.

(a) Write $\log_a 6$ in terms of p and q. [1]
(b) Write $\log_a \sqrt{12}$ in terms of p and q. [2]
(c) Write $\log_a \dfrac{4}{9}$ in terms of p and q. [2]

8.

(a) Show that $\log_a \sqrt[n]{x} = \dfrac{1}{n}\log_a x$. [2]
(b) Hence solve $\log_2 \sqrt[3]{x} + \log_8 x = 4$. [4]

9. Prove that $\log_a b \cdot \log_b c \cdot \log_c a = 1$ for all valid bases and arguments. Hence evaluate

$\log_2 3 \cdot \log_3 5 \cdot \log_5 8$ without a calculator. [6]

10. Given that $\log_2 3 = a$, express each of the following in terms of a:

(a) $\log_2 9$ [1]
(b) $\log_4 3$ [2]
(c) $\log_8 27$ [2]
(d) $\log_3 8$ [2]

11. Solve $\log_3(x+5) - \log_3(x-1) = 2$. [4]

12. Solve $\log_6(x+2) + \log_6(x-3) = 1$, justifying any rejection of solutions. [4]

13. Solve the simultaneous equations:

$$\log_2 x + \log_2 y = 5, \qquad \log_2(x-y) = 3.$$

[5]

14. Solve $\log_x 8 = \log_4 x$. [5]

15. GDC Solve $\log_2(x^2 - 2x) = \log_4(x+4)$. [5]

16. GDC

(a) Express $\log_3 x - 2 = \log_9(x+6)$ as a quadratic equation in x, using change of base. [3]
(b) Hence find the value of x. [2]

17. Solve $\log_5(x^2 - 2x + 1) = 2\log_5(x-1) + 1$, or show that no solution exists. [5]

18. The reciprocal of a logarithm is a logarithm with its base and argument swapped.

(a) Let $N = a^k$, where $a > 0$, $a \neq 1$, $N > 0$ and $N \neq 1$. Write $\log_a N$ and $\log_N a$ in terms of k, and hence show that $\dfrac{1}{\log_a N} = \log_N a$. Explain why $N = 1$ must be excluded. [3]
(b) Hence evaluate $\dfrac{1}{\log_2 36} + \dfrac{1}{\log_3 36}$ without a calculator. [4]

19. Solve $\log_2 x + \log_x 2 = \dfrac{5}{2}$. *(Hint: let $t = \log_2 x$ and use question 18)* [6]

20. Consider the equation $(\log_3 x)^2 = \log_3 x^2 + 3$.

(a) Explain why the right-hand side equals $2\log_3 x + 3$, and state the restriction on x this requires. [2]

(b) Hence solve the equation. [4]

(c) The expression $\log_3 x^2$ is defined for every non-zero x, yet negative values of x cannot be solutions. Explain why. [2]

21. Solve $\log_2(\log_3(\log_2 x)) = 0$.

(a) Find x. [4]

(b) Working outwards from the innermost logarithm, state the condition each layer imposes on x, and hence give the values of x for which the left-hand side is defined at all. [4]

22. Evaluate

$$\log_2 3 \cdot \log_3 4 \cdot \log_4 5 \cdot \;\cdots\; \cdot \log_{31} 32,$$

the product of all thirty factors, without a calculator. Justify your method. [5]

23. Let a, b and c be positive, with a, b, c and ab all different from 1.

(a) Show that $\log_{ab} c = \dfrac{\log_a c \cdot \log_b c}{\log_a c + \log_b c}$. [5]

(b) Verify your identity with $a = 2$, $b = 5$ and $c = 7$, giving both sides to four decimal places. [2]

24. Solve the simultaneous equations

$$\log_x y + \log_y x = \frac{10}{3}, \qquad xy = 81,$$

where x and y are positive and neither is 1. [7]

25. Solve $3^{2x-1} = 4^{x+2}$, giving your answer in exact logarithmic form. [4]

26. Find the exact solution to $e^{2x} = 3e^x + 10$. *(Hint: let $u = e^x$)* [4]

27. Solve the equation $9^x - 4 \cdot 3^{x+1} + 27 = 0$. *(Hint: let $u = 3^x$; note $9^x = u^2$)* [5]

28. GDC The number of bacteria in a culture doubles every 3 hours. Initially there are 500 bacteria.

(a) Write an expression for the number of bacteria after t hours. [2]

(b) Find the time, to the nearest minute, when there are 50 000 bacteria. [3]

29. GDC The population of a city is modelled by $P(t) = 250\,000 \times e^{0.032t}$, where t is the number of years after the start of 2020.

(a) Write down the population in 2020. [1]

(b) Find the population in 2030, to the nearest thousand. [2]

(c) Find the year in which the population first exceeds 400 000. [3]

30. GDC The value V of a car after t years is given by $V = 32\,000 \times (0.85)^t$.

(a) Write down the initial value of the car. [1]
(b) Find the value after 5 years, to the nearest dollar. [2]
(c) Find the number of years until the value falls below \$10 000. [3]

31. GDC The pH of a solution is defined by $\text{pH} = -\log_{10}[\text{H}^+]$, where $[\text{H}^+]$ is the hydrogen ion concentration in mol/L.

(a) A solution has $[\text{H}^+] = 3.5 \times 10^{-4}$ mol/L. Find its pH to 2 decimal places. [2]
(b) A solution has $\text{pH} = 8.3$. Find its hydrogen ion concentration in standard form. [2]
(c) Solution A has $\text{pH} = 3$ and solution B has $\text{pH} = 7$. Find how many times more acidic solution A is than solution B. [2]

32. GDC A radioactive substance decays so that its mass M grams after t years satisfies $M = M_0 e^{-kt}$, where M_0 and k are positive constants. After 10 years the mass is halved.

(a) Show that $k = \dfrac{\ln 2}{10}$. [2]
(b) Find the percentage of the original mass remaining after 25 years. [3]
(c) Find the time for the mass to reduce to 10% of its original value. [3]

33. GDC This question is about writing very large and very small quantities in standard form.

(a) A human cell has diameter about 2×10^{-5} m, and a water molecule about 2.8×10^{-10} m. Find how many times wider the cell is, giving your answer in standard form to three significant figures. [3]
(b) The observable universe is about 8.8×10^{26} m across, and the Planck length, the scale at which current physical theories are expected to break down, is about 1.6×10^{-35} m. Find the ratio of the two, in standard form. [3]
(c) A sheet of paper is 0.1 mm thick. It is folded in half, then in half again, and so on. Find the thickness after 20 folds, in metres and in standard form. [3]
(d) Find the least number of folds needed for the thickness to exceed the distance to the Moon, 3.84×10^8 m. [4]
(e) Your answer to part (d) may be smaller than you expected. Explain why, referring to how the thickness grows. [2]

34.

(a) Prove the change of base rule $\log_a b = \dfrac{\log_c b}{\log_c a}$, for valid bases a and c. *(Hint: let $x = \log_a b$, so $a^x = b$, then take $\log_c$ of both sides)* [3]

(b) Hence show that $\log_{27} x = \frac{1}{3}\log_3 x$. [2]
(c) Solve $\log_3 x + \log_9 x + \log_{27} x = 11$. [4]
(d) Solve $\log_x 2 + \log_{x^2} 2 = 3$, giving an exact answer. [5]

Challenge Questions

35. GDC **Decibels.** The loudness L of a sound in decibels is $L = 10\log_{10}\left(\frac{I}{I_0}\right)$, where I is the intensity in W/m^2 and $I_0 = 10^{-12}$ W/m^2 is the threshold of hearing. The scale is logarithmic because human hearing spans a range of intensities far too wide to put on a linear axis.

(a) A concert has loudness 110 dB. Find the intensity I. [2]
(b) A whisper has loudness 30 dB. Show that the concert is 10^8 times as intense as the whisper. [2]
(c) Show that doubling the intensity of *any* sound raises its loudness by $10\log_{10} 2$ dB, whatever the original level was. Give this value to two decimal places. [4]
(d) Hence find how many identical sources, playing together, are needed to raise the level by 10 dB. [3]
(e) Intensity obeys an inverse-square law: at distance r from a source, $I \propto \frac{1}{r^2}$. Show that doubling your distance reduces the level by a fixed number of decibels, independent of where you started, and find it. [4]
(f) A workplace safety rule allows 8 hours of exposure at 85 dB, and halves the permitted time for every 3 dB above that. Find the permitted time at 100 dB, and explain what part (c) has to do with the choice of 3 dB. [4]
(g) Two sources, of levels L_1 and L_2 dB, play together, and their intensities add. Show that the combined level is

$$L = 10\log_{10}\left(10^{L_1/10} + 10^{L_2/10}\right),$$

and find the combined level of a 90 dB source and an 80 dB source. [3]
(h) Let $L_1 \geq L_2$ and $D = L_1 - L_2$. Show that $L - L_1 = 10\log_{10}\left(1 + 10^{-D/10}\right)$, tabulate this rise for $D = 0, 3, 6, 10, 20$, and decide, with a reason, how much quieter a second source must be before it can be ignored. Suggest how this could be extended into an exploration. [3]

36. GDC **Counting digits, and the first digit.** This investigation uses logarithms to count digits, and then to explain a pattern that auditors use to detect faked data. Throughout, $\lfloor t \rfloor$ is the greatest integer not exceeding t, and the fractional part of t is $t - \lfloor t \rfloor$.

(a) Let N be a positive integer. Show that N has exactly $\lfloor \log_{10} N \rfloor + 1$ digits. [3]
(b) Hence find the number of digits in 2^{1000}. [2]

(c) Let d be one of the digits $1, 2, \ldots, 9$. Prove that the first digit of N is d exactly when the fractional part of $\log_{10} N$ lies in the interval $[\log_{10} d,\ \log_{10}(d+1))$. Verify this for $N = 2^7$. [4]

(d) Prove that $\log_{10} 2$ is irrational. Hence show that no two of the numbers $\log_{10}(2^n)$, for $n = 1, 2, 3, \ldots$, have the same fractional part. [4]

(e) Assume the fractional parts of $\log_{10}(2^n)$ are spread evenly over $[0, 1)$ as n increases. (Part (d) shows that they never repeat; that they spread evenly is a further theorem, which you may use.) Show that the proportion of powers of 2 whose first digit is d is $\log_{10}\left(1 + \dfrac{1}{d}\right)$. [4]

(f) Evaluate this proportion for $d = 1$ and for $d = 9$. Then count the first digits of 2^1, 2^2, …, 2^{30} on your GDC and compare. [4]

(g) The pattern in part (e) is known as Benford's law, and it is also reported for data such as the populations of towns. Comment on how well your counts in part (f) support it, and why a sample of 30 might mislead. Then name one kind of real data you would expect to follow the law and one you would not, giving a reason for each, as the starting point of an exploration. [3]

37. GDC **Trapping e.** For each integer $n \geq 1$ let

$$a_n = \left(1 + \frac{1}{n}\right)^n \qquad \text{and} \qquad b_n = \left(1 + \frac{1}{n}\right)^{n+1}.$$

You may use *Bernoulli's inequality*: for every integer $m \geq 2$ and every real $x \geq -1$ with $x \neq 0$, $(1 + x)^m > 1 + mx$. (It can be proved by induction, the method of Ch. 3.) You may also use the fact that an increasing sequence that is bounded above has a limit, and so does a decreasing sequence that is bounded below.

(a) Find a_n and b_n for $n = 1, 2, 5, 10, 100$, giving values to four decimal places where they are not exact. Make a conjecture about how each sequence behaves. [3]

(b) Show that $\dfrac{a_{n+1}}{a_n} = \left(1 + \dfrac{1}{n}\right)\left(1 - \dfrac{1}{(n+1)^2}\right)^{n+1}$, and use Bernoulli's inequality to prove that $a_{n+1} > a_n$ for all $n \geq 1$. [5]

(c) Show that, for $n \geq 2$, $\dfrac{b_{n-1}}{b_n} = \dfrac{n}{n+1}\left(1 + \dfrac{1}{n^2 - 1}\right)^n$, and prove that $b_n < b_{n-1}$. [5]

(d) Show that $b_n - a_n = \dfrac{a_n}{n}$. Explain why $\{a_n\}$ and $\{b_n\}$ both have limits, and why the two limits are equal. This common limit is e; deduce that $a_n < e < b_n$ for every $n \geq 1$. [4]

(e) Use $n = 1000$ to write down an interval that contains e, and state its width. [2]

(f) Compute $n(e - a_n)$ for $n = 10$, 100 and 1000, and conjecture its limit. Explain what your conjecture says about how many correct decimal places of e are gained each time n is multiplied by 10, and suggest a way to obtain e more quickly that could be investigated further. [3]

Chapter 2

Sequences & Series

2

SYLLABUS 1.2 1.3 1.4 1.8 1.9 1.10

A sheet of paper is about 0.1 mm thick. Fold it once and it is 0.2 mm. Each fold doubles the thickness, so 42 folds would give 0.1×2^{42} mm, which is about $440{,}000$ km. The Moon is about $384{,}000$ km away. Nobody can fold a sheet of paper 42 times, but the arithmetic is correct.

One further fold would reach the Moon and return. The reason lies in how doubling accumulates. Each fold doubles a thickness that has already been doubled, so after 42 folds the paper is not 42 times its original thickness but 2^{42} times it.

Now stack the sheets instead of folding them. Place one sheet at a time, and after 42 sheets the pile is 4.2 mm high. The paper is the same and the number of steps is the same, yet the answer is smaller by a factor of about a hundred billion. Folding multiplies at each step; stacking adds.

Those two operations produce the two kinds of list studied in this chapter. A list built by repeatedly adding a fixed amount is *arithmetic*. A list built by repeatedly multiplying by a fixed factor is *geometric*. The two arise in the same situations, so choosing the wrong one is costly: a salary that rises by a fixed amount each year is arithmetic, a salary that rises by a fixed percentage is geometric, and the two grow far apart over a career.

The same two patterns describe the value of a savings account and the growth of a population by a fixed percentage, and the counting methods at the end of the chapter give the expansion of $(a+b)^{12}$ without multiplying out twelve brackets. By the end of this chapter you will be able to find any term, and the sum of the first n terms, of an arithmetic or geometric sequence without writing out the terms before it; use sigma notation; find the sum of an infinite geometric series and explain why it can be finite; count arrangements and selections; and expand $(a+b)^n$ with the binomial theorem.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

2.1 Arithmetic Sequences

A theatre has 20 seats in the front row, 23 in the second, and 26 in the third. Each row has 3 more seats than the one before it. To find how many seats are in row 50, counting up one row at a time would be slow. There is a clear pattern: each row adds 3 seats. A regular pattern like this can be written as a formula, and finding that formula is the point of this section.

A sequence where each term is obtained by adding a fixed constant to the previous one is called **arithmetic** (also called an **arithmetic progression**, or **AP**). That constant is the **common difference** d.

DEF An **arithmetic sequence** has the form $u_1,\ u_1+d,\ u_1+2d,\ \ldots$, where u_1 is the first term and d is the common difference.
For any two consecutive terms: $d = u_{n+1} - u_n$.

2.1.1 The general term

The n-th term is the first term plus $(n-1)$ steps of size d. Starting from u_1, you take one step to reach u_2, two steps to reach u_3, and so on.

KEY · IB

$$u_n = u_1 + (n-1)d$$

Worked through. To find u_{20} when the first term is 7 and the common difference is -3, take 19 steps of size -3 from $u_1 = 7$:

$$u_{20} = 7 + (20-1)(-3) = 7 - 57 = -50.$$

A negative common difference means the sequence is decreasing. The general term can be negative even when the first term is positive.

EXAMPLE 2.1

In an arithmetic sequence, $u_5 = 11$ and $u_9 = 27$. Find u_1 and d.
From u_5 to u_9 is exactly 4 steps:

$$4d = u_9 - u_5 = 27 - 11 = 16 \implies d = 4.$$

Then $u_1 = u_5 - 4d = 11 - 16 = -5$.

EXAMPLE 2.2

Find the number of terms in the arithmetic sequence $5, 9, 13, \ldots, 101$.

Here $d = 4$. Set $u_n = 101$:

$$5 + (n-1)(4) = 101 \implies 4(n-1) = 96 \implies n = 25.$$

CAUTION The formula is $u_1 + (n-1)d$, not $u_1 + nd$. For the first term ($n = 1$), you add zero differences, not one.

YOUR TURN 2.1 Arithmetic Sequences

Finding a term

1. Find the term indicated.

(a) u_7 of $3, 7, 11, \ldots$

(b) u_{15} of the AP with $u_1 = 12$, $d = -4$

(c) u_{10} of $-5, -1, 3, \ldots$

(d) u_{25} of the AP with $u_1 = 100$, $d = -7$

Finding the first term or the common difference

2. Find the quantity indicated.

(a) d, given $u_1 = 5$ and $u_9 = 37$

(b) u_1, given $u_6 = 23$ and $d = 3$

(c) d, given $u_6 = 20$ and $u_{14} = -4$

(d) u_1, given $u_{12} = 50$ and $d = 4$

3. In an arithmetic sequence, $u_2 + u_5 = 31$ and $u_8 = 38$. Find u_1 and d.

Number of terms

4. Find the number of terms in each AP.

(a) $4, 11, 18, \ldots, 144$

(b) $2, 9, 16, \ldots, 100$

(c) $100, 96, 92, \ldots, 4$

(d) $-7, -3, 1, \ldots, 61$

Testing whether a number is a term

5. Show that the sequence $5, 11, 17, \ldots$ has a term equal to 101. State which term it is.

6. Determine whether each number is a term of the sequence $3, 10, 17, \ldots$. If it is, state which term it is.

(a) 150 (b) 200

Concept check

7. A student finds the 20th term of $5, 8, 11, \ldots$ as $u_{20} = 5 + 20 \times 3 = 65$. Identify the error, and find the correct value.

2.2 Arithmetic Series

When you need the total of the first n terms of an arithmetic sequence, adding them one at a time is always possible but rarely necessary. A pattern reduces the whole sum to one short calculation.

Consider adding all integers from 1 to 100. Write the sum forwards and backwards:

$$1 \; + \; 2 \; + \; 3 \; + \; \cdots \; + \; 99 \; + \; 100$$

$$100 \; + \; 99 \; + \; 98 \; + \; \cdots \; + \; 2 \; + \; 1$$

Every vertical pair sums to 101, and there are 100 such pairs. So twice the sum equals 100×101, giving $S_{100} = 5050$.

The same pairing works for any arithmetic series: the first and last term always sum to the same value, $u_1 + u_n$. There are n such pairs when you write the series twice, so $2S_n = n(u_1 + u_n)$, and therefore:

KEY · IB

$$S_n = \frac{n}{2}(u_1 + u_n) \qquad = \qquad \frac{n}{2}(2u_1 + (n-1)d)$$

The first form is faster when you know u_n. Substituting $u_n = u_1 + (n-1)d$ gives the second, which works when only u_1 and d are known.

Worked through. To find the sum of the first 30 terms of $2, 7, 12, \ldots$, note that $u_1 = 2$ and $d = 5$. The last term is not given, so use the second form:

$$S_{30} = \frac{30}{2}(2(2) + 29(5)) = 15(4 + 145) = 15 \times 149 = 2235.$$

When a formula for S_n is given, recover the terms from it: $u_1 = S_1$, and $u_n = S_n - S_{n-1}$ for $n \geq 2$, because the two sums differ only by their last term.

EXAMPLE 2.3

The sum of the first n terms of an arithmetic series is given by $S_n = 3n^2 + 5n$. Find u_1, d, and the general term u_n.

$u_1 = S_1 = 3 + 5 = 8$.

$u_2 = S_2 - S_1 = (12 + 10) - 8 = 14$, so $d = u_2 - u_1 = 6$.

For the general term: $u_n = S_n - S_{n-1} = 3n^2 + 5n - 3(n-1)^2 - 5(n-1)$.

Expanding: $u_n = 3n^2 + 5n - 3n^2 + 6n - 3 - 5n + 5 = 6n + 2$.

Check: $u_1 = 8$ and $u_2 = 14$. Consistent.

CAUTION S_n is the total of the first n terms; u_n is just the n-th term. Confusing the two is a common error. A question asking for a "term" wants u_n; a question asking for a "sum" wants S_n.

2.2.1 Application: simple interest

A standard financial application of an arithmetic sequence is **simple interest**: the same fixed amount of interest is added each year, so the balance increases by the same amount each year.

Write the opening balance as u_1 and the yearly interest as d. After n years, n payments of interest have been added, so the balance is

$$u_1 + nd.$$

Note the n here rather than the $n - 1$ of the general term. The opening balance is the amount before any interest has been paid, so n years later there have been n payments, not $n - 1$.

EXAMPLE 2.4

\$2000 is deposited in an account paying 3% simple interest per year. Find the balance after 15 years.

The interest is $3\% \times 2000 = \$60$ every year, so the balances form an arithmetic sequence with $d = 60$. After 15 years, 15 payments of interest have been added:

$$2000 + 15 \times 60 = \$2900.$$

2.2.2 When the data are not quite arithmetic

Real measurements rarely sit exactly on a sequence. A model is still useful if the steps are *close* to constant, and then the common difference has to be estimated rather than read off. A simple estimate is the average step: divide the total change by the number of steps.

KEY For data that are approximately arithmetic over n terms,

$$d \approx \frac{u_n - u_1}{n - 1}.$$

This is the mean of the successive differences. Those differences telescope, so only the first and last readings survive: the estimate is quick, but it rests entirely on the two endpoints. Fitting a least-squares line, as in Ch. 10, uses every reading instead.

A model fitted to data is reliable only within the range the data covered. A prediction outside that range is an estimate, and a model of real growth will fail if it is carried far enough.

EXAMPLE 2.5

A tree's height is measured at the end of each year: 1.8, 2.5, 3.1, 3.9, 4.6 metres. Show that an arithmetic model is reasonable, estimate the common difference, and predict the height after 8 years.

The successive differences are 0.7, 0.6, 0.8, 0.7. They are not equal, but they are close, so an arithmetic model is reasonable. Over four steps the height rose from 1.8 to 4.6, so

$$d \approx \frac{4.6 - 1.8}{5 - 1} = \frac{2.8}{4} = 0.7 \text{ metres per year.}$$

With $u_1 = 1.8$ and $d = 0.7$, the eighth term is

$$u_8 = 1.8 + 7(0.7) = 6.7 \text{ metres.}$$

Treat that as an estimate, not a fact. Two things limit it. The model was built on five years and year 8 lies outside them, and no tree grows at a constant rate forever, so the model must fail eventually. A prediction is only as good as the range the data covered.

YOUR TURN 2.2 Arithmetic Series

Finding a sum

8. Find the sum indicated.

(a) S_{12} for the AP $2, 5, 8, \ldots$

(b) S_{20} for the AP with $u_1 = 100$, $d = -5$

(c) S_{15} for the AP $7, 11, 15, \ldots$

(d) S_{25} for the AP with $u_1 = -12$, $d = 5$

9. An AP has $u_1 = 15$, $u_n = -9$ and $n = 13$. Find S_n.

10. Find each sum.

(a) $6 + 10 + 14 + \cdots + 90$

(b) $3 + 8 + 13 + \cdots + 98$

(c) $50 + 46 + 42 + \cdots + 2$

(d) $1 + 1.5 + 2 + \cdots + 20$

11. Find the sum of all integers from 1 to 200 that are multiples of 6.

12. Find the sum $4 + 7 + 10 + \cdots + 100$.

Working backwards from a sum

13. Find n if $S_n = 234$ for the AP $3, 6, 9, \ldots$

14. Given $S_n = 4n^2 - n$, find u_1, d and the general term u_n.

Simple interest

15. \$3500 is deposited in an account paying 4% simple interest per year. Find the balance after 8 years.

16. An account pays simple interest. After 3 years the balance is \$2360, and after 7 years it is \$2840. Find the amount first deposited and the annual rate of interest.

Modelling with data

17. A plant's height is measured at the end of weeks 1 to 5, in centimetres: 4.2, 6.1, 7.8, 9.9, 11.8.

(a) Find the four successive differences, and explain why an arithmetic model is reasonable.

(b) Estimate the common difference from the first and last readings.

(c) Predict the height at the end of week 9, and give one reason to treat the prediction with caution.

Concept check

18. An arithmetic series has $S_n = n^2 + 3n$. Asked for the 10th term, a student writes $u_{10} = S_{10} = 130$. Identify the error, and find u_{10}.

2.3 Geometric Sequences

A geometric sequence multiplies by the same number at each step. Start at 1 and double: $1, 2, 4, 8, 16, \ldots$ Many real quantities change in this way. Bacteria double in every generation. A ball bounces to 60% of the height it fell from. Each extra filter that passes half the light reaching it halves the intensity again. In every case, each new value is the previous value multiplied by a fixed number.

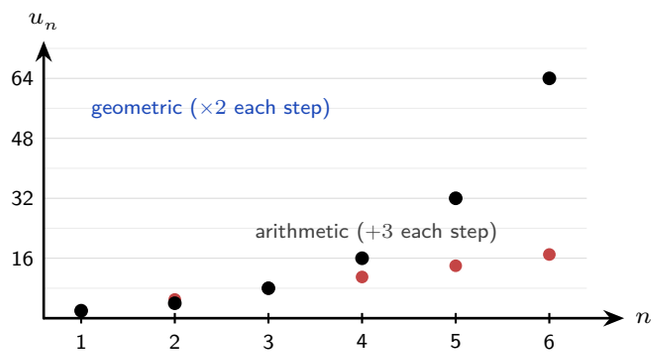


Figure 2.1 Adding a fixed amount grows in a straight line; multiplying by a fixed factor greater than 1 gives increases that themselves keep growing.

A sequence where each term is obtained by multiplying the previous term by a fixed constant is called **geometric** (also called a **geometric progression**, or **GP**). That constant is the **common ratio** r.

DEF A **geometric sequence** has the form $u_1,\ u_1 r,\ u_1 r^2,\ \ldots$, where u_1 is the first term and r is the common ratio.
For any two consecutive terms: $r = \dfrac{u_{n+1}}{u_n}$.

2.3.1 The general term

Starting from u_1, each step multiplies by r. After $(n-1)$ multiplications you reach u_n:

KEY · IB

$$u_n = u_1 \cdot r^{n-1}$$

Worked through. For the geometric sequence with $u_1 = 5$ and $r = 3$, the sixth term is

$$u_6 = 5 \cdot 3^5 = 5 \times 243 = 1215.$$

EXAMPLE 2.6

A geometric sequence has $u_3 = 28$ and $u_6 = 224$. Find u_1 and r.
From u_3 to u_6 is 3 multiplicative steps: $u_6 = u_3 \cdot r^3$, so

$$r^3 = \frac{224}{28} = 8 \implies r = 2.$$

Then $u_1 = u_3/r^2 = 28/4 = 7$.

EXAMPLE 2.7

Find the number of terms in the geometric sequence $2, 6, 18, \ldots, 4374$.
Here $r = 3$. Set $u_n = 4374$:

$$2 \cdot 3^{n-1} = 4374 \implies 3^{n-1} = 2187 = 3^7 \implies n = 8.$$

EXAMPLE 2.8

Given that k, 8 and $k + 12$ are consecutive terms of a geometric sequence, find the possible values of k.
Consecutive terms have a constant ratio, so the ratio across the first gap equals the ratio across the second:

$$\frac{8}{k} = \frac{k+12}{8} \implies 64 = k(k+12) \implies k^2 + 12k - 64 = 0.$$

Factorising, $(k + 16)(k - 4) = 0$, so $k = -16$ or $k = 4$.
Both values give a genuine sequence. With $k = 4$ the terms are $4, 8, 16$ and $r = 2$; with $k = -16$ they are $-16, 8, -4$ and $r = -\frac{1}{2}$. Do not discard the negative solution: a negative ratio makes the terms alternate in sign, which is still a geometric sequence.

2.3.2 Recursively defined sequences

Every rule so far has been **explicit**: put n in and the n-th term comes out, with no need for any term before it. A sequence can instead be given by a **recurrence**: a starting value together with a rule that builds each term from the previous one. The sequence $u_1 = 4$, $u_{n+1} = 3u_n - 4$ is defined this way, and it runs $4, 8, 20, 56, \ldots$

A recurrence produces terms one at a time, so reaching u_{50} means computing the forty-nine terms below it. A closed form removes that work. For a rule of the shape $u_{n+1} = au_n + b$, the route to one is a shift: measure the terms from the value the rule leaves unchanged, and what remains is geometric.

EXAMPLE 2.9

A sequence is defined by $u_1 = 4$ and $u_{n+1} = 3u_n - 4$. Let $v_n = u_n - 2$. Show that $\{v_n\}$ is

geometric, and find a closed form for u_n.
Subtract 2 from both sides of the recurrence:

$$u_{n+1} - 2 = 3u_n - 6 = 3(u_n - 2), \qquad \text{that is} \qquad v_{n+1} = 3v_n.$$

So $\{v_n\}$ is geometric with $r = 3$, and $v_1 = u_1 - 2 = 2$. Therefore

$$v_n = 2 \cdot 3^{n-1}, \qquad u_n = v_n + 2 = 2 \cdot 3^{n-1} + 2.$$

Check the first three terms: $2 + 2 = 4$, $6 + 2 = 8$, $18 + 2 = 20$, matching 4, $3(4) - 4$, $3(8) - 4$. The substitution is not a guess. Solving $x = 3x - 4$ gives $x = 2$, the value the rule fixes, and $v_n = u_n - 2$ measures each term from it, which is exactly what cancels the -4.

2.3.3 Where geometric growth shows up

Three contexts show the pattern, and the algebra is the same in each.

- **Population growth.** A population rising by a fixed *percentage* each year is geometric with $r = 1 + \frac{p}{100}$. A 2% annual rise gives $r = 1.02$.
- **Salary change.** An annual rise of 3% gives $r = 1.03$; a cut of 3% gives $r = 0.97$. Note that a rise then an equal cut does not return you to the start, because $1.03 \times 0.97 = 0.9991$.
- **The spread of a disease.** If each infected person passes it to r others, the number of new cases in successive rounds is geometric. That r is what public health reporting calls the reproduction number, and whether it is above or below 1 decides whether the outbreak grows or dies out.

The last one shows why the ratio matters more than the starting value. If $r > 1$ the terms grow without limit, however small u_1 is; if $0 < r < 1$ they shrink towards zero, however large u_1 is. A large outbreak with $r = 0.9$ dies out, and a single case with $r = 1.1$ does not.

Two points of reading need care. A value n periods after the start has been multiplied n times, so it is $u_1 r^n$, not $u_n = u_1 r^{n-1}$; read the question for whether it counts terms or periods. When n counts whole periods, as in the first year a value is passed, round the solution of the equation up, whatever the decimal.

EXAMPLE 2.10

A town of $40{,}000$ people grows by 2.5% each year. Find the population after 10 years, and the first year in which it exceeds $60{,}000$.
Here $u_1 = 40{,}000$ and $r = 1.025$. Ten years later is ten multiplications, so the value is $u_1 r^{10}$, not u_{10}:

$$40{,}000 \times 1.025^{10} \approx 51{,}203.$$

For the second part, solve $40{,}000 \times 1.025^n > 60{,}000$, that is $1.025^n > 1.5$. Taking logarithms

(Ch. 1),

$$n > \frac{\ln 1.5}{\ln 1.025} \approx 16.4,$$

so the population first exceeds 60,000 during year 17. Round *up* here whatever the decimal: year 16 is not yet enough.

TIP To test whether a sequence is arithmetic, check that differences are constant. To test whether it is geometric, check that ratios are constant. A sequence usually passes at most one of these tests. The exception is a constant non-zero sequence such as $5, 5, 5, \ldots$, which is arithmetic with $d = 0$ and geometric with $r = 1$.

YOUR TURN 2.3 Geometric Sequences

Finding a term

19. Find the term indicated.

(a) u_6 of the GP $2, 6, 18, \ldots$

(b) u_7 of the GP with $u_1 = 96$, $r = \frac{1}{2}$

(c) u_5 of $\frac{1}{4}, \frac{1}{2}, 1, \ldots$

(d) u_8 of the GP with $u_1 = 3$, $r = -2$

Finding the first term or the common ratio

20. Find the quantity indicated.

(a) r, given $u_1 = 5$ and $u_4 = 40$

(b) u_1, given $u_4 = 24$ and $r = 2$

(c) r, given $u_1 = 81$, $u_5 = 1$ and $r > 0$

(d) u_1, given $u_3 = 45$ and $r = 3$

21. Given $u_2 = 10$ and $u_5 = 1250$, find u_1 and r.

22. The 2nd term of a GP is 12 and the 4th term is 48. Find the two possible pairs of values of u_1 and r.

Number of terms

23. Find the number of terms in each GP.

(a) $3, 6, 12, \ldots, 768$

(b) $5, 15, 45, \ldots, 3645$

(c) $5, -10, 20, \ldots, -2560$

(d) $128, 64, 32, \ldots, \frac{1}{2}$

Three consecutive terms

24. Given that k, 10 and $k-15$ are consecutive terms of a geometric sequence, find the possible values of k and the common ratio in each case.

25. Given that x, $x+6$ and $9x$ are consecutive terms of a geometric sequence, where $x \neq 0$, find the possible values of x. For each value, write down the three terms and the common ratio.

Geometric growth in context

26. GDC Geometric growth in three contexts.

(a) A population of $25{,}000$ grows by 1.8% each year. Find the population after 12 years.

(b) A salary of \$$48{,}000$ rises by 4% each year. Find the salary in the sixth year.

(c) In an outbreak each infected person infects 1.2 others, and round 1 has 50 new cases. Find the number of new cases in round 6, and state what would happen in the long run if that figure fell to 0.8.

27. GDC A town of $32{,}000$ people grows by 2.2% each year. Find the first year in which the population exceeds $45{,}000$.

Recurrences

28. A sequence is defined by $u_1 = 5$ and $u_{n+1} = 2u_n - 3$.

(a) Write down u_2, u_3 and u_4.

(b) Let $v_n = u_n - 3$. Show that $\{v_n\}$ is geometric.

(c) Hence find a closed form for u_n.

29. A sequence is defined by $u_1 = 10$ and $u_{n+1} = \frac{1}{2}u_n + 4$.

(a) Find the value of c for which $v_n = u_n - c$ is a geometric sequence.

(b) Find a closed form for u_n, and describe what happens to u_n as n increases.

Concept check

30. A town of 2000 people grows by 5% each year. A student models the population after n years as $P = 2000 \times 0.05^{\,n}$. Identify the error, and write down the correct model.

2.4 Geometric Series

Adding the terms of a geometric sequence needs a different method. In an arithmetic series, the first and last terms have the same total as the second and second-to-last terms, so pairing works. In a geometric series these pair totals in general differ, so pairing does not help. Instead there is an algebraic method: multiply the sum by r, line the two sums up, and subtract.

Let $S = u_1 + u_1 r + u_1 r^2 + \cdots + u_1 r^{n-1}$. Multiply by r:

$$rS = u_1 r + u_1 r^2 + \cdots + u_1 r^n.$$

Every term appears in both lines except u_1 in the first line and $u_1 r^n$ in the second. Subtract rS from S:

$$S(1 - r) = u_1 - u_1 r^n = u_1(1 - r^n).$$

Divide both sides by $(1 - r)$ to get the formula.

KEY · IB

$$S_n = \frac{u_1(1 - r^n)}{1 - r} \qquad = \qquad \frac{u_1(r^n - 1)}{r - 1}, \qquad r \neq 1$$

Both forms are equivalent. When $0 < r < 1$, use the first form: both $1 - r^n$ and $1 - r$ are then positive, so no negative signs appear. When $r > 1$ the second form is tidier, for the same reason. If r is negative, either form works, but r^n changes sign with n, so take care with the arithmetic.

Worked through. To find the sum of the first 10 terms of $5 + 10 + 20 + \cdots$, take $u_1 = 5$ and $r = 2$:

$$S_{10} = \frac{5(2^{10} - 1)}{2 - 1} = 5 \times 1023 = 5115.$$

EXAMPLE 2.11

A geometric series has $u_1 = 80$ and $r = \frac{1}{2}$. Find the least value of n such that $S_n > 159$.

$$S_n = \frac{80\left(1 - \left(\frac{1}{2}\right)^n\right)}{1 - \frac{1}{2}} = 160\left(1 - \frac{1}{2^n}\right).$$

Require $160\left(1 - \frac{1}{2^n}\right) > 159$, so $\frac{1}{2^n} < \frac{1}{160}$, giving $2^n > 160$.
Since $2^7 = 128 < 160$ and $2^8 = 256 > 160$, the least value is $n = 8$.

CAUTION The formula requires $r \neq 1$. If $r = 1$, every term equals u_1 and $S_n = nu_1$. Substituting $r = 1$ into the formula gives $0/0$, which is undefined.

Solving for n in geometric series often requires taking logarithms of both sides. See Ch. 1, §1.5.

YOUR TURN 2.4 Geometric Series

Finding a sum

31. Find the sum indicated.

(a) S_8 for the GP $1, 3, 9, \ldots$
(b) S_{10} for the GP with $u_1 = 6$, $r = \frac{1}{3}$
(c) S_6 for the GP $128, 64, 32, \ldots$
(d) S_6 for the GP with $u_1 = 1$, $r = -2$

32. Find each sum.

(a) $3 + 6 + 12 + \cdots + 1536$
(b) $2 + 6 + 18 + \cdots + 486$
(c) $1 - 2 + 4 - \cdots + 256$
(d) $96 + 48 + 24 + \cdots + 3$

Working backwards from a sum

33. The sum of the first three terms of a GP is 21 and the common ratio is 2. Find the first term.

34. Find n such that $S_n = 63$ for the GP $1, 2, 4, \ldots$

35. Given $S_3 = 7$ and $S_6 = 63$, find u_1 and r.

36. A geometric sequence has $u_1 = 5$ and $S_3 = 15$.

(a) Find the two possible values of r.
(b) For each value of r, find S_{10}.

The least number of terms

37. For the geometric series $2 + 6 + 18 + \cdots$, find the least value of n such that $S_n > 10\,000$.

38. GDC A geometric series has $u_1 = 60$ and $r = 0.8$. Find the least value of n such that $S_n > 290$.

Concept check

39. A student finds S_6 for the series $2 + 6 + 18 + \cdots$ as $\dfrac{2(1-3)^6}{1-3} = \dfrac{128}{-2} = -64$. Identify the error, and find S_6.

2.5 Sigma Notation

By now you have written several sums in the form $u_1 + u_2 + \cdots + u_n$. There is a compact notation for this that appears throughout mathematics, physics, and statistics, and it should be learned properly.

The Greek letter Σ (capital sigma) is used as shorthand for "sum". The expression

$$\sum_{k=1}^{n} u_k = u_1 + u_2 + u_3 + \cdots + u_n$$

is read "the sum of u_k from $k = 1$ to n." The variable k is the **index of summation**: it takes each integer value from the lower limit to the upper limit, the corresponding term is evaluated, and all terms are added.

DEF $\displaystyle\sum_{k=m}^{n} f(k) \;=\; f(m) + f(m+1) + f(m+2) + \cdots + f(n)$, where $m \le n$.

The index k is a **dummy variable**: the result depends only on f, m, and n, not on the name chosen for the index.

2.5.1 Properties of sigma notation

These three properties are just familiar rules of arithmetic, written in sigma form. They are what you use to simplify or split a sum.

KEY For a constant c and sequences $\{a_k\}$, $\{b_k\}$:

$\displaystyle\sum_{k=1}^{n} c = nc$	a constant summed n times
$\displaystyle\sum_{k=1}^{n} c\,a_k = c\sum_{k=1}^{n} a_k$	a constant factor comes outside the sum
$\displaystyle\sum_{k=1}^{n} (a_k + b_k) = \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k$	a sum splits over addition

When the lower limit is not 1, subtract the missing terms: for $m \ge 2$,

$$\sum_{k=m}^{n} f(k) = \sum_{k=1}^{n} f(k) - \sum_{k=1}^{m-1} f(k).$$

The sum has $n - m + 1$ terms.

Worked through. To write $3 + 7 + 11 + 15 + 19$ in sigma notation, notice that the terms

form an arithmetic sequence with $u_1 = 3$, $d = 4$: the k-th term is $u_k = 4k - 1$. So

$$3 + 7 + 11 + 15 + 19 = \sum_{k=1}^{5} (4k - 1).$$

EXAMPLE 2.12

Evaluate $\sum_{k=1}^{6} 5 \cdot 2^{k-1}$.

This is a geometric series with $u_1 = 5$ and $r = 2$:

$$\sum_{k=1}^{6} 5 \cdot 2^{k-1} = S_6 = \frac{5(2^6 - 1)}{2 - 1} = 5 \times 63 = 315.$$

EXAMPLE 2.13

Evaluate $\sum_{k=3}^{10} (2k + 1)$.

The lower limit is 3, not 1. Use the fact that the full sum minus the missing part gives the partial sum:

$$\sum_{k=3}^{10} (2k + 1) = \sum_{k=1}^{10} (2k + 1) - \sum_{k=1}^{2} (2k + 1).$$

The first sum is an arithmetic series: $u_1 = 3$, $u_{10} = 21$, so $S_{10} = \frac{10}{2}(3 + 21) = 120$.

The second sum is just $3 + 5 = 8$.

Result: $120 - 8 = 112$.

CAUTION $\sum_{k=1}^{n} (a_k \cdot b_k) \neq \left(\sum a_k\right)\left(\sum b_k\right)$. You can split a sum over addition but not over multiplication. This is a common misapplication of sigma notation.

YOUR TURN 2.5 Sigma Notation

Writing a sum in sigma notation

40. Write each sum in sigma notation.

(a) $5 + 10 + 15 + 20 + 25$

(b) $4 + 8 + 16 + 32 + 64$

(c) $3 + 6 + 9 + \cdots + 30$

(d) $1 + 4 + 9 + 16 + 25$

Evaluating a sum

41. Evaluate each sum.

(a) $\displaystyle\sum_{k=1}^{6}(3k-2)$

(b) $\displaystyle\sum_{k=0}^{4}81\left(\tfrac{1}{3}\right)^{k}$

(c) $\displaystyle\sum_{k=5}^{12}(3k-4)$

42. Consider the sum $\displaystyle\sum_{k=4}^{15}(5-2k)$.

(a) State the number of terms in the sum.

(b) Evaluate the sum.

(c) Rewrite the sum using the index $j=k-3$, so that j starts at 1, and check that it gives the same value.

Using the sigma properties

43. This question uses the properties of sigma notation.

(a) By treating it as an arithmetic series, show that $\displaystyle\sum_{k=1}^{n}k=\frac{n(n+1)}{2}$.

(b) Hence find $\displaystyle\sum_{k=1}^{n}5$ and $\displaystyle\sum_{k=1}^{n}5k$.

44. Use the properties of sigma notation and the result of Question 43(a) to answer each part.

(a) Evaluate $\displaystyle\sum_{k=1}^{20}(3k+2)$.

(b) Evaluate $\displaystyle\sum_{k=11}^{30}(2k-1)$ by subtracting the terms that are missing.

(c) Evaluate $\displaystyle\sum_{k=1}^{3}k^2$ and $\displaystyle\left(\sum_{k=1}^{3}k\right)^2$. State what the two values show about sums of products.

Concept check

45. A student writes $\displaystyle\sum_{k=1}^{20}(3k+2)=3\sum_{k=1}^{20}k+2=3(210)+2=632$. Identify the error, and find the correct value.

2.6 Infinite Geometric Series

Cut a pizza in half. Eat one half and keep the other. Cut what is left in half again, eat one piece, and keep the rest. If you could continue forever, how much pizza would you eat?

The pieces you eat are $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$, ... There are infinitely many of them, and every piece is larger than zero. Adding infinitely many positive numbers sounds like it must give an infinite total. But you began with one pizza, so you can never eat more than one pizza. So the total cannot be more than 1.

That is the surprise of this section: *an addition that never ends can still have a finite answer.*

Add the pieces one at a time and record the running total after each step. Each running total is a **partial sum**:

$$S_1 = 0.5, \quad S_2 = 0.75, \quad S_3 = 0.875, \quad S_4 = 0.9375, \quad \ldots$$

Two things happen at once. Every partial sum is larger than the one before it, and every partial sum is still smaller than 1. Look at the distance from 1: it is 0.5, then 0.25, then 0.125. The distance halves at every step, so it never becomes zero, but it becomes as small as you wish. This value that the partial sums approach is called the **limit**, and here the limit is 1.

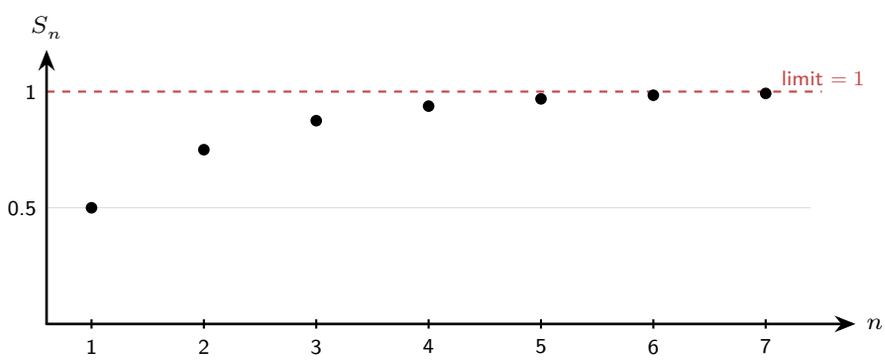


Figure 2.2 The partial sums of $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$ climb toward 1 but never pass it.

The pizza works because each piece is half the size of the piece before it. In general, $|r| < 1$ means that each term is smaller in size than the one before, by the fixed factor $|r|$, so the terms shrink toward zero.

Now change one number and the behaviour changes completely. Take $1 + 2 + 4 + 8 + \cdots$, where $r = 2$. Each term is larger than the last, the partial sums are $1, 3, 7, 15, \ldots$, and they pass every value you can name. There is no finite total.

A series whose partial sums approach a finite limit is said to **converge**. A series that does not converge is said to **diverge**: its partial sums may grow without bound, as here, or keep oscillating, as they do for $1 - 1 + 1 - 1 + \cdots$. A geometric series converges when $|r| < 1$ and diverges when $|r| \geq 1$.

To turn this into a formula, start from the finite sum and let n grow without bound. When $|r| < 1$, the power r^n shrinks toward 0:

$$S_n = \frac{u_1(1-r^n)}{1-r} \xrightarrow{n\to\infty} \frac{u_1(1-0)}{1-r} = \frac{u_1}{1-r}.$$

KEY · IB For a geometric series with $|r| < 1$:

$$S_\infty = \frac{u_1}{1-r}$$

A decimal that recurs from the decimal point is an infinite geometric series. The first term is the first copy of the repeating block, and $r = 10^{-m}$ for a block of m digits, so S_∞ gives the decimal as a fraction.

Worked through. To find the sum to infinity of $6 + 4 + \frac{8}{3} + \cdots$, first find the common ratio, $r = 4/6 = \frac{2}{3}$. Since $|r| < 1$ the series converges:

$$S_\infty = \frac{6}{1-\frac{2}{3}} = \frac{6}{\frac{1}{3}} = 18.$$

EXAMPLE 2.14

Express $0.\overline{27} = 0.272727\ldots$ as an exact fraction.

Write as an infinite geometric series with $u_1 = \frac{27}{100}$ and $r = \frac{1}{100}$:

$$0.272727\ldots = \frac{27/100}{1-1/100} = \frac{27/100}{99/100} = \frac{27}{99} = \frac{3}{11}.$$

EXAMPLE 2.15

A geometric series has first term a and common ratio r. Given $S_\infty = 8$ and $u_2 = 2$, find a and r.

From S_∞: $\dfrac{a}{1-r} = 8$, so $a = 8(1-r)$.

From $u_2 = ar = 2$: substitute a:

$$8(1-r)r = 2 \implies 8r^2 - 8r + 2 = 0 \implies 4r^2 - 4r + 1 = 0 \implies (2r-1)^2 = 0.$$

So $r = \frac{1}{2}$, and $a = 8\left(1-\frac{1}{2}\right) = 4$.

CAUTION Always verify $|r| < 1$ before applying the infinite sum formula. If $|r| \geq 1$, no finite sum to infinity exists.

YOUR TURN 2.6 Infinite Geometric Series

Finding a sum to infinity

46. Find S_∞ for each series.

(a) $12 + 4 + \frac{4}{3} + \cdots$

(b) the GP with $u_1 = 10$, $r = \frac{3}{5}$

(c) $9 + 3 + 1 + \cdots$

(d) the GP with $u_1 = 8$, $r = -\frac{1}{4}$

47. A GP has $u_1 = 16$, $u_3 = 4$ and $r > 0$. Find S_∞.

Convergence

48. State whether each series converges, giving a reason. If it converges, find its sum to infinity.

(a) $6 + 9 + 13.5 + \cdots$

(b) $20 - 15 + 11.25 - \cdots$

Recurring decimals

49. Express each recurring decimal as an exact fraction in its lowest terms.

(a) $0.\overline{153} = 0.153153\ldots$

(b) $2.1\overline{6} = 2.1666\ldots$

Finding the first term or the common ratio

50. Find the quantity indicated.

(a) u_1, given $S_\infty = 20$ and $r = \frac{1}{5}$

(b) r, given $u_1 = 6$ and $S_\infty = 9$

(c) u_1, given $S_\infty = 12$ and $r = \frac{1}{4}$

(d) r, given $u_1 = 15$ and $S_\infty = 12$

51. A geometric series has $S_\infty = 27$ and $u_2 = 6$. Find the two possible pairs of values of u_1 and r.

52. A geometric series has $S_\infty = 16$, and the sum of its first two terms is 12. Find the possible values of r and the corresponding values of u_1.

Concept check

53. A student finds the sum to infinity of $1 + \frac{3}{2} + \frac{9}{4} + \cdots$ as $\dfrac{1}{1 - \frac{3}{2}} = -2$. Explain why this answer is wrong.

2.7 Financial Applications

A bank pays you 4% per year. In the first year you earn interest on your deposit. In the second year you earn interest on the deposit *and* on the interest from the first year. The amount added is different every year, so the balances do not form an arithmetic sequence. They form a geometric one: each year the balance is multiplied by 1.04.

This is **compound interest**, and it is why money left alone for a long time grows far more than it would under simple interest. The balance after any number of years is simply a term of a geometric sequence.

Interest is often added more than once a year. If a present value PV earns interest at a nominal annual rate of $r\%$, compounded k times per year, then after n years:

KEY · IB

$$FV = PV \times \left(1 + \frac{r}{100k}\right)^{kn}$$

The quantity $\left(1 + \frac{r}{100k}\right)$ is the common ratio of the balances (here r is the interest rate, not the ratio). The total number of compounding periods is kn.

Worked through. Suppose \$5 000 is invested at 4% per year, compounded monthly. After 3 years, with $PV = 5000$, $r = 4$, $k = 12$ and $n = 3$, its value is

$$FV = 5000 \times \left(1 + \frac{4}{1200}\right)^{36} \approx \$5636.36.$$

EXAMPLE 2.16

How many complete years does it take for an investment to double at 6% per year, compounded annually?

Set $FV = 2 \cdot PV$:

$$(1.06)^n = 2 \implies n = \frac{\ln 2}{\ln 1.06} \approx 11.9.$$

After 12 complete years the investment has more than doubled.

TIP The Rule of 72: dividing 72 by the annual interest rate gives an approximate doubling time in years. At 6%, that is $72/6 = 12$ years, matching the exact calculation above.

2.7.1 Regular deposits

A single lump sum gives one term of a geometric sequence. A savings plan is different: money goes in every year, and each payment then compounds for however long it has left in

the account. The balance is a total of geometric terms, so it is a geometric *series*.

Write the payment as P and the annual rate as a decimal i. With payments made at the **end** of each year for n years, the final payment earns no interest at all, the one before it earns interest for one year, and the first earns interest for $n-1$ years. Reading the total from the last payment backwards gives a GP with first term P and common ratio $1+i$.

There is no new formula to learn: the sum is the geometric series of §2.4 with $u_1 = P$ and $r = 1+i$.

Payments made at the **start** of each year each spend one extra year in the account. Rather than remembering a correction factor, change the first term: the last payment now earns a year's interest, so $u_1 = P(1+i)$ and everything else is unchanged.

EXAMPLE 2.17

\$800 is paid into an account at the end of each year for 12 years. The account pays 5% per year, compounded annually. Find the value immediately after the twelfth payment.
The payment made at the end of year 12 has earned nothing, the one at the end of year 11 has earned interest for a single year, and the first payment has earned interest for 11 years. Totalling from the last payment backwards:

$$800 + 800(1.05) + 800(1.05)^2 + \cdots + 800(1.05)^{11}.$$

This is a geometric series with $u_1 = 800$, $r = 1.05$ and $n = 12$ terms:

$$S_{12} = \frac{800(1.05^{12} - 1)}{1.05 - 1} \approx \$12\,733.70.$$

Had the same payments been made at the *start* of each year, every one of them would earn a further year of interest, so $u_1 = 800(1.05)$ and the same sum gives

$$S_{12} = \frac{800(1.05)(1.05^{12} - 1)}{0.05} \approx \$13\,370.39.$$

Settle which end of the year the payments land on before writing down u_1; the two answers differ by a full year's interest on the whole plan.

2.7.2 Depreciation

Some things lose value instead of gaining it. A car, a laptop, or a machine may lose a fixed *percentage* of its value each year. This is called **depreciation**.

No new formula is needed. Use the same compound-interest formula with a **negative** rate r. A loss of 12% per year means $r = -12$, so the common ratio becomes

$$1 + \frac{r}{100} = 1 - \frac{12}{100} = 0.88,$$

a number less than 1. The value is still multiplied by the same number each year, so the sequence is still geometric; because that number is less than 1, the values shrink.

2.7.3 Inflation and real value

There is a second geometric process working against your savings. **Inflation** is the yearly rise in prices. Interest increases the number in your account, while inflation reduces what that number can buy. If your account grows by 2% in a year and prices rise by 3% in the same year, you hold more money but you can buy less with it.

The **real value** of an investment is its future value adjusted for the price rises over the same period. It measures the gain in purchasing power rather than the gain in the number.

Prices climb geometrically too, so the correction is a second geometric factor, this time dividing rather than multiplying.

KEY With interest producing a future value FV over n years, and prices rising at $i\%$ per year:

$$\text{real value} = \frac{FV}{\left(1+\frac{i}{100}\right)^n}.$$

Divide by the inflation factor, because the same goods cost $\left(1+\frac{i}{100}\right)^n$ times as much after n years.

EXAMPLE 2.18

\$10 000 is invested for 5 years at 4.5% per year, compounded annually. Prices rise at 2.5% per year. Find the future value and the real value of the investment.

The balance grows by a factor of 1.045 each year:

$$FV = 10\,000 \times (1.045)^5 \approx \$12\,461.82.$$

Now adjust for five years of price rises:

$$\text{real value} = \frac{12\,461.82}{(1.025)^5} \approx \$11\,014.43.$$

The account shows a gain of about \$2460, but in purchasing power the investor is only about \$1014 better off. An interest rate means little until you compare it with inflation.

TIP Your GDC's built-in Finance (TVM) solver handles compound interest, depreciation, and unknown rates or times directly: enter the known values of N, $I\%$, PV, FV, and the compounding frequency, and solve for the missing one. In examinations, state the values you entered.

Solving for n in financial problems requires exponential equations. See Ch. 1, §1.5.

YOUR TURN 2.7 Financial Applications

Compound interest

54. GDC Find the future value of each investment.

(a) \$3 000 at 4% p.a., compounded annually, for 5 years

(b) \$8 000 at 6% p.a., compounded monthly, for 3 years

(c) \$500 at 3.6% p.a., compounded monthly, for 10 years

(d) \$2 000 at 5% p.a., compounded quarterly, for 4 years

55. GDC Find the present value needed to accumulate \$15 000 in 8 years at 5% p.a., compounded quarterly.

Unknown rate or time

56. GDC Find the annual interest rate if \$5 000 grows to \$5 828.84 in 3 years, compounded annually.

57. GDC How many complete years does it take an investment to triple at 7% p.a., compounded monthly?

Depreciation

58. GDC A car valued at \$25 000 depreciates at 15% per year. Find its value after 4 years.

59. GDC A machine bought for \$80 000 is worth \$41 000 after 5 years. It loses the same percentage of its value each year. Find that percentage.

60. GDC A car bought for \$28 000 loses 12% of its value each year.

(a) Find its value after 5 years.

(b) Explain why this model never gives the car a value of zero.

Regular deposits

61. GDC \$1200 is paid into an account every year for 10 years. The account pays 4.5% p.a., compounded annually.

(a) The payments are made at the end of each year. Find the value immediately after the tenth payment.

(b) The payments are made at the start of each year instead. Find the value at the end of the tenth year.

62. GDC A saver wants to have \$50 000 after 15 years. She pays equal amounts into an account at the end of each year. The account pays 5% p.a., compounded annually. Find the size of each payment.

Inflation and real value

63. GDC $6000 is invested for 8 years at 3.8% p.a., compounded annually. Prices rise at 2.1% per year.

(a) Find the future value of the investment.

(b) Find its real value.

64. GDC Prices rise at 2.8% per year. A sum of money is kept without earning any interest. Find the number of complete years after which it has lost at least a quarter of its purchasing power.

Concept check

65. GDC A student finds the value of $5000 after 3 years at 6% p.a., compounded monthly, as 5000×1.06^{36}. Identify the error, and find the correct value.

2.8 Counting Principles HL

A four-digit PIN is built from the digits 0 to 9, and any digit may repeat. There are 10 choices for the first digit, 10 for the second, 10 for the third and 10 for the fourth, so the number of PINs is

$$10 \times 10 \times 10 \times 10 = 10\,000.$$

Writing the list out would take ten thousand lines. The total comes from the way the choices are made, and two ordinary words decide what to do with the counts.

And multiplies. When a result needs one choice *and* then another, the counts multiply. A menu offering 3 starters and 2 mains gives $3 \times 2 = 6$ meals.

A tree shows why. Each starter has the same number of branches below it, so the total is one count multiplied by the other rather than added.

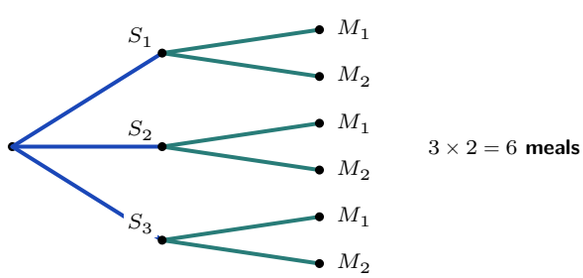


Figure 2.3 Three starters, two mains for each. Every path is a starter *and* a main, and there are $3 \times 2 = 6$ of them: that is what "and multiplies" looks like.

Add a third course and the same reasoning applies again: with 4 desserts the count becomes $3 \times 2 \times 4 = 24$. Every extra choice multiplies.

Or adds. Not every count is a chain of choices. Sometimes the answer is one thing *or* another, and the two cannot both happen: a password of length 3 *or* length 4, a team with three women *or* four. Count each possibility separately and add.

Those two words decide how the counts combine. Everything else in this section is about counting a single choice correctly once you know which word applies.

DEF HL **And multiplies** (the **multiplication principle**). If a result needs a first choice *and* a second, made in m and n ways, there are $m \times n$ results. Any number of choices chains the same way.
The condition: n must be the same number whichever way the first choice went. It may differ from m (10 people, then 9), but it may not depend on which one you picked.

Or adds (the **addition principle**). If the result is one case *or* another, add the counts.
The condition: the cases must not overlap. Nothing may be counted twice.

CAUTION Everyday English uses *or* inclusively: "students who take French or Spanish" includes those taking both. That "or" does not simply add, because the students taking both would be counted twice. In this section *or* always means the exclusive one: this case or that case, never both. The overlapping kind is handled in Ch. 11, §11.3, where $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ subtracts the double count.

Arranging n different objects in a row is one long chain of *and*s: n choices for the first position, $n-1$ for the second, and so on down to 1. The product is n **factorial**.

KEY HL

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1, \qquad 0! = 1$$

The value $0! = 1$ is set by definition, not derived. There is one way to arrange no objects at all, and fixing $0! = 1$ keeps every formula below correct at its end points.

2.8.1 Permutations and combinations

There are two standard questions about selecting r objects from n different ones. If the *order* of selection matters, the count is a **permutation**. If only the final collection matters, it is a **combination**.

For a permutation, stop the factorial early. There are n choices for the first object, $n-1$ for the second, and so on for r factors in total. Multiplying and dividing by the remaining $(n-r)$ factors turns this product into a ratio of factorials:

$$^{n}P_{r} = n(n-1)\cdots(n-r+1) = \frac{n!}{(n-r)!}.$$

The combination count follows from the permutation count. Choose 3 letters from A, B, C, D, E. The selection $\{A, B, C\}$ can be written in $3! = 6$ orders: ABC, ACB, BAC, BCA, CAB, CBA. A permutation count records six results here; a combination count records one. Every combination is therefore counted $r!$ times among the permutations, so dividing by $r!$ removes the duplicates:

$$^nC_r = \frac{^nP_r}{r!}.$$

KEY · IB HL Selecting r from n distinct objects:

$$^nP_r = \frac{n!}{(n-r)!} \quad \text{(ordered)}, \qquad ^nC_r = \frac{n!}{r!\,(n-r)!} \quad \text{(unordered)}$$

The symbol ^nC_r is read "n choose r". Other books write the same number as $\binom{n}{r}$, so both forms should be recognised; the formula booklet uses ^nC_r, and a GDC labels the two counts nPr and nCr.

The question	The count	Example
Arrange all n objects	$n!$	6 books in a row
Arrange r of the n, order matters	^nP_r	3 medals for 8 runners
Choose r of the n, order does not matter	^nC_r	a committee of 3 from 12

Three steps are enough for almost every counting question in this course.

1. Read the question for *and* or *or*. *And* multiplies, *or* adds. Check the condition each one carries.
2. Within a single choice, ask whether the order matters. If it does, use ^nP_r. If it does not, use ^nC_r.
3. If the restriction is awkward to count directly, count everything and subtract what breaks it.

Two restrictions appear often enough to name. When certain objects must stay together, treat the whole group as one object, arrange the objects, then arrange the group internally and multiply the two counts. When a question asks for *at least one* of something, step 3 is usually shorter: subtract the arrangements that have none from the total.

Worked through. When eight runners compete for gold, silver and bronze, the three medals

are distinguishable, so order matters. The number of ways to award them is

$$^8P_3 = 8 \times 7 \times 6 = 336.$$

EXAMPLE 2.19

A committee of 5 is chosen from 6 boys and 5 girls. How many committees contain exactly 2 girls?

The committee needs 2 girls *and* 3 boys, so the two counts multiply. A committee has no internal order, so each count is a combination:

$$^5C_2 \times {}^6C_3 = 10 \times 20 = 200.$$

EXAMPLE 2.20

Five people sit in a row. In how many ways can they sit if two of them refuse to sit next to each other?

Count the seatings that place the two together, then subtract. Treat the pair as one object: four objects give $4! = 24$ arrangements, and the pair can sit in $2! = 2$ orders inside their block, so

$$24 \times 2 = 48$$

seatings place them together. All five can be seated in $5! = 120$ ways, so

$$120 - 48 = 72$$

seatings keep them apart. Counting the forbidden cases and subtracting is shorter here than building the allowed ones directly.

CAUTION Ask one question before every counting problem: *does order matter?* Medals, codes and rankings: yes, so use ^nP_r. Committees, hands of cards and subsets: no, so use ^nC_r.

TIP A GDC computes both, under the labels nPr and nCr. Two kinds of counting are not required in this course: arrangements in which some of the objects are identical, and arrangements in a circle.

YOUR TURN 2.8 Counting Principles HL

Arrangements and selections

66. Find each count.

(a) arrangements of 6 different books in a row

(b) 4-digit codes using the digits 1 to 9, no digit repeated

(c) committees of 3 chosen from 12 people

(d) groups of 2 men and 2 women chosen from 7 men and 6 women

Restrictions

67. Seven different people stand in a row. Find the number of arrangements in which

(a) three particular people stand next to each other

(b) two particular people stand at the two ends of the row

68. A 4-digit code uses the digits 1 to 9 with no digit repeated. Find the number of such codes that end in an even digit.

69. A committee of 3 is chosen from 12 people, one of whom is Ana. Find the number of committees that

(a) include Ana

(b) do not include Ana

Adding separate cases

70. Each of these is an *or*: split the count into parts that cannot both happen, then add.

(a) A group of 4 is chosen from 7 men and 6 women. Find the number of groups containing at least 3 women.

(b) A password is 3 or 4 letters long, chosen from the 26 lowercase letters with no letter repeated. Find the number of possible passwords.

At least one

71. A committee of 4 is chosen from 5 men and 7 women. By subtracting the committees with no men from the total, find the number of committees with at least one man.

72. A 4-digit code uses the digits 0 to 9, and digits may be repeated. Find the number of codes that contain at least one 7.

Solving for n

73. Solve each equation for the positive integer n.

(a) ${}^{n}C_{2} = 45$

(b) ${}^{n}P_{2} = 56$

Concept check

74. A student says the number of committees of 3 that can be chosen from 10 people is ${}^{10}P_3 = 720$. Explain the error, and find the correct number.

75. A committee of 4 is chosen from 5 men and 7 women. To count committees with at least one man, a student chooses one man (5 ways) and then any 3 of the remaining 11 people, giving $5 \times {}^{11}C_3 = 825$.

(a) Explain why this overcounts.

(b) Find the correct number.

2.9 The Binomial Theorem

Expanding $(a+b)^2$ is quick. Expanding $(a+b)^3$ takes longer. Expanding $(a+b)^{12}$ by multiplying out twelve brackets is not realistic. Fortunately the answer follows a pattern, so no bracket multiplication is needed at all.

Look at the small cases:

$$(a+b)^1 = a+b, \qquad (a+b)^2 = a^2+2ab+b^2, \qquad (a+b)^3 = a^3+3a^2b+3ab^2+b^3.$$

Two features are regular. The power of a falls from n to 0 while the power of b rises from 0 to n, and in every term the two powers add to n. The coefficients form a row of **Pascal's triangle**, in which each row begins and ends with 1 and every other entry is the sum of the two entries above it.

Number the rows from zero. The top row is row 0, the next is row 1, and so on, so that **row n holds the coefficients of $(a+b)^n$**. The top row is not decoration: $(a+b)^0 = 1$, which is the single entry 1. Number the entries along a row from zero as well, so that the entry in position r of row n multiplies $a^{n-r}b^r$.

$n=0$	1	$(a+b)^0$
$n=1$	1 1	$a+b$
$n=2$	1 2 1	$a^2+2ab+b^2$
$n=3$	1 3 3 1	$a^3+3a^2b+3ab^2+b^3$
$n=4$	1 4 6 4 1	$r = 0,1,2,3,4$ along the row
$n=5$	1 5 10 10 5 1	

With both counts starting at zero, the label matches the algebra exactly: to expand $(a+b)^4$, read row 4, and its entry in position r is the coefficient of $a^{4-r}b^r$. Row n has $n+1$ entries, which is the number of terms in the expansion.

2.9.1 Why the entries are the combinations

Multiplying out $(a+b)^n$ means going through the n brackets and choosing either a or b from each one, then adding every product so formed. A product ends up as $a^{n-r}b^r$ exactly when b was chosen from r of the brackets and a from the other $n-r$. So the coefficient of $a^{n-r}b^r$ counts one thing: **how many ways there are to choose which r of the n brackets supply the b.** Choosing r objects from n, order not mattering, is what ${}^nC_r$ counts §2.8, so the coefficient *is* ${}^nC_r$.

The case $n=3$ is small enough to check by hand. Writing the three brackets as $(a+b)(a+b)(a+b)$, the term a^2b appears when exactly one bracket gives a b, and there are three ways for that to happen: the first bracket, the second, or the third. So the coefficient is 3, and ${}^3C_1 = 3$. Likewise ab^2 needs two brackets to give a b, which can be done in ${}^3C_2 = 3$ ways, while a^3 and b^3 each arise in only one way: ${}^3C_0 = {}^3C_3 = 1$.

That also explains the addition rule that builds the triangle. Pick one bracket and ask whether it supplies a b. If it does, the other $r-1$ copies of b come from the remaining $n-1$ brackets, in ${}^{n-1}C_{r-1}$ ways; if it does not, all r come from those $n-1$ brackets, in ${}^{n-1}C_r$ ways. The two cases cover every selection once, so

$$ {}^nC_r = {}^{n-1}C_{r-1} + {}^{n-1}C_r, $$

which is exactly "each entry is the sum of the two entries above it". Since both descriptions start each row with 1 and end it with 1 $({}^nC_0 = {}^nC_n = 1)$ and then grow by the same rule, the triangle and the combinations agree row for row.

KEY · IB For $n \in \mathbb{N}$:

$$ (a+b)^n = a^n + {}^nC_1\, a^{n-1}b + \cdots + {}^nC_r\, a^{n-r}b^r + \cdots + b^n, $$

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

Pascal's triangle is fastest for small n, provided you count rows from zero so that row n is the one you want. When n is large, or when a question asks for one coefficient only, use the single term ${}^nC_r\, a^{n-r}b^r$ shown in the middle of the expansion above and choose the r you need.

TIP The booklet shows that single term inside the expansion without naming it. The **general term** is

$$ T_{r+1} = {}^nC_r\, a^{n-r}b^r, $$

so $r=0$ gives the first term, $r=1$ the second, and r gives term number $r+1$. The subscript is $r+1$ because r counts how many factors of b the term contains, and the first term contains none. In an examination, write down this term, find the r that produces the power you want, then substitute.

A question asking for the **constant term**, or for the term *independent of* x, is asking for the one in which **the power of x is** 0, since $x^0 = 1$ leaves a number. Simplify the general term until x appears once, as a single power, then set that power equal to 0 and solve for r.

Worked through. To expand $(x+2)^4$, take $n = 4$ and read row 4 of the triangle, the fifth row down: $1, 4, 6, 4, 1$, with the powers of 2 increasing as the powers of x decrease:

$$(x+2)^4 = x^4 + 4x^3(2) + 6x^2(4) + 4x(8) + 16 = x^4 + 8x^3 + 24x^2 + 32x + 16.$$

EXAMPLE 2.21

Find the coefficient of x^3 in $(2x-3)^5$.
The general term is ${}^5C_r\,(2x)^{5-r}(-3)^r$. For x^3, take $5 - r = 3$, so $r = 2$:

$${}^5C_2\,(2x)^3(-3)^2 = 10 \times 8x^3 \times 9 = 720x^3.$$

The coefficient is 720. Keep the sign inside the bracket with $b = -3$: the $(-3)^r$ produces the alternating signs automatically.

EXAMPLE 2.22

Find the term independent of x in $\left(x^2 + \dfrac{2}{x}\right)^6$.
Track the power of x in the general term:

$$T_{r+1} = {}^6C_r\,(x^2)^{6-r}\left(\frac{2}{x}\right)^r = {}^6C_r\,2^r\,x^{12-3r}.$$

Independent of x means $12 - 3r = 0$, so $r = 4$:

$${}^6C_4\,2^4 = 15 \times 16 = 240.$$

TIP ${}^nC_r$ can also be read from a GDC table. To solve ${}^6C_r = 20$, generate ${}^6C_r$ for $r = 0, \ldots, 6$ and read off $r = 3$. Symmetry, ${}^nC_r = {}^nC_{n-r}$, halves the search.

YOUR TURN 2.9 The Binomial Theorem

Expanding

76. Expand or write down the terms requested.

(a) Expand $(2+y)^5$, using row 5 of Pascal's triangle.

(b) Write down the first three terms, in ascending powers of x, of $(1+2x)^7$.

(c) Expand $(2x-1)^4$.

(d) Write down the first three terms, in ascending powers of x, of $(1-3x)^6$.

A single coefficient

77. Find the coefficient requested.

(a) x^4 in $\left(3-\frac{x}{3}\right)^6$

(b) x^3 in $(2+x)^6$

(c) x^4 in $(1-2x)^7$

(d) x^5 in $(x+3)^8$

78. Find the middle term in the expansion of $(x-2y)^6$.

Term numbers and symmetry

79. Answer each part without expanding in full.

(a) Write down the 5th term of the expansion of $(1+x)^{10}$.

(b) Given that ${}^{12}C_5 = 792$, write down ${}^{12}C_7$, giving a reason.

(c) In the expansion of $(1+x)^n$, the coefficients of x^3 and x^7 are equal. Find n.

Terms with a negative power of x

80. Find the term requested.

(a) the term independent of x in $\left(x+\dfrac{1}{x^2}\right)^9$

(b) the term independent of x in $\left(2x-\dfrac{1}{x}\right)^6$

(c) the coefficient of x^3 in $\left(x-\dfrac{2}{x}\right)^7$

(d) the term independent of x in $\left(3x^2-\dfrac{1}{x}\right)^6$

Concept check

81. A student says the coefficient of x^3 in $(1-2x)^5$ is ${}^5C_3 \times 2^3 = 80$. Identify the error, and find the correct coefficient.

2.10 The Binomial Series HL

So far n has been a positive whole number, because $(a+b)^n$ meant n brackets multiplied together. But expressions such as $\sqrt{1+x} = (1+x)^{1/2}$ and $\dfrac{1}{1+x} = (1+x)^{-1}$ also have exponents. Can they be expanded too?

They can, and the same coefficient pattern still works for *any* rational n. Two things change. The expansion never ends, and it is only valid when x is small enough.

Where the extension comes from. Write the coefficient of §2.9 without factorials. In

$$ {}^nC_r = \frac{n!}{r!\,(n-r)!}, $$

the $(n-r)!$ cancels all but the first r factors of $n!$, leaving

$$ {}^nC_r = \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!}, $$

which is r factors counting down from n, divided by $r!$. Now look at what each side needs. The factorial $n!$ only makes sense for a non-negative whole number, but the product of r descending factors makes sense for *any* n: with $n = \frac{1}{2}$ and $r = 2$ it is $\frac{\left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)}{2!} = -\frac{1}{8}$, and with $n = -1$ and $r = 3$ it is $\frac{(-1)(-2)(-3)}{3!} = -1$. Keeping the coefficients in this form, and letting n be any rational number, is the whole of the extension. For a whole n nothing has changed: at $n = 3$, $r = 2$ the formula gives $\frac{3\times 2}{2!} = 3 = {}^3C_2$.

What is lost is the counting. For a whole n the coefficient counts brackets; for $n = \frac{1}{2}$ it counts nothing, and the equality with the function has to be earned another way, which is where the condition on x comes in.

KEY · IB HL For $n \in \mathbb{Q}$:

$$ (a+b)^n = a^n\left(1 + n\frac{b}{a} + \frac{n(n-1)}{2!}\left(\frac{b}{a}\right)^2 + \cdots\right) $$

The booklet omits the validity condition. When n is not a non-negative integer the series never ends, so it is valid only when $\left|\frac{b}{a}\right| < 1$: factor out the term of larger size to make this hold. When $n \in \mathbb{N}$ the series stops and needs no condition.

In practice you first factor the expression into the form $(1 + \text{something})^n$, which turns the booklet formula into the version you actually compute with:

$$ (1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \cdots, \qquad |x| < 1 \text{ when } n \notin \mathbb{N}. $$

Why does the expansion stop for whole numbers but not otherwise? Look at those descending factors $n(n-1)(n-2)\cdots$ again. If $n = 3$, the fourth factor is $n - 3 = 0$, so

every later coefficient is zero and the series ends: that is the finite theorem of §2.9. If $n = \frac{1}{2}$ or $n = -1$, no factor is ever zero, so the terms continue forever.

The condition $|x| < 1$ should look familiar. Put $n = -1$:

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \cdots$$

This is a geometric series with first term 1 and common ratio $-x$, and §2.6 says it converges to $\frac{1}{1+x}$ exactly when $|-x| < 1$. So in this case the condition is not an extra assumption: it is the convergence condition already proved, and the same bound turns out to hold for every rational n. The two infinite series of this chapter are the same idea seen from two directions.

Worked through. To expand $(1-3x)^{1/3}$ up to and including the term in x^2, apply the series with $n = \frac{1}{3}$ and x replaced by $-3x$:

$$(1-3x)^{1/3} = 1 + \tfrac{1}{3}(-3x) + \frac{\frac{1}{3}\left(-\frac{2}{3}\right)}{2}(-3x)^2 + \cdots = 1 - x - x^2 - \cdots$$

Validity requires $|-3x| < 1$, that is $|x| < \frac{1}{3}$.

When the bracket does not already start with 1, factor out the larger term first, as the booklet formula does. The expansion is then valid for $\left|\frac{b}{a}\right| < 1$.

EXAMPLE 2.23

Expand $\dfrac{1}{2+x}$ in ascending powers of x up to x^2, stating the validity.

Factor out the 2:

$$(2+x)^{-1} = \frac{1}{2}\left(1 + \frac{x}{2}\right)^{-1} = \frac{1}{2}\left(1 - \frac{x}{2} + \frac{x^2}{4} - \cdots\right) = \frac{1}{2} - \frac{x}{4} + \frac{x^2}{8} - \cdots$$

valid for $\left|\frac{x}{2}\right| < 1$, that is $|x| < 2$.

A truncated expansion gives an approximation. Choose x inside the interval of validity so that the bracket becomes the value you want, or a value it can be found from. The closer x is to 0, the faster the later terms shrink and the better the approximation.

EXAMPLE 2.24

By expanding $(1-x)^{1/2}$ up to x^2 and choosing a suitable value of x, find an approximation to $\sqrt{2}$.

The expansion is

$$(1-x)^{1/2} = 1 - \frac{x}{2} - \frac{x^2}{8} - \cdots$$

Now pick x so that $1-x$ connects to 2. Take $x = 0.02$: then $1 - x = 0.98 = \frac{49}{50}$, so

$$\sqrt{0.98} = \frac{7}{\sqrt{50}} = \frac{7\sqrt{2}}{10} \implies \sqrt{2} = \frac{10}{7}\sqrt{0.98}.$$

The series gives $\sqrt{0.98} \approx 1 - 0.01 - 0.00005 = 0.98995$, hence

$$\sqrt{2} \approx \frac{10}{7}(0.98995) \approx 1.41421,$$

agreeing with the true value to five decimal places, from just three terms. The key is small x: the closer x is to 0, the faster the series converges.

CAUTION Always state the interval of validity alongside an HL binomial expansion; the series is meaningless outside it. And the expansion needs the form $(1 + \bullet)^n$: factor first, then expand.

At HL these expansions return as Maclaurin series in Ch. 16, where the same coefficients arise from repeated differentiation.

YOUR TURN 2.10 The Binomial Series HL

Expanding $(1 + x)^n$

82. Expand each up to and including the term in x^3, stating the validity.

(a) $(1 + x)^{-1}$ (b) $(1 - x)^{-2}$

83. Expand each up to and including the term in x^2, stating the validity.

(a) $(1 + 4x)^{1/2}$ (b) $(1 + x)^{1/3}$

Factorising first

84. Expand $\dfrac{1}{3 - x}$ in ascending powers of x up to x^2, stating the validity.

85. Expand $(8 + x)^{1/3}$ in ascending powers of x up to x^2, stating the validity.

Approximations

86. Use the expansion of $(1 + x)^{1/2}$ up to the term in x^2, with $x = 0.1$, to estimate $\sqrt{1.1}$ to four decimal places.

87. This question estimates $\sqrt{3}$.

(a) Expand $(4 - x)^{1/2}$ in ascending powers of x up to x^2, stating the validity.

(b) By choosing a suitable value of x, use your expansion to estimate $\sqrt{3}$ to three decimal places.

The link with geometric series

88. This question compares the binomial series with a geometric series.

(a) Use the binomial series to expand $(1-3x)^{-1}$ up to and including the term in x^3.

(b) Show that these are the first four terms of a geometric series, and state its common ratio.

(c) Explain why the binomial expansion and the geometric series both need $|x| < \frac{1}{3}$.

Concept check

89. A student expands $(1-3x)^{-2}$ as a binomial series and states that it is valid for $|x| < 1$. Identify the error, and state the correct condition.

Chapter 2 · Sequences & Series

Reflections

TOK Around 450 BCE Zeno argued that a runner can never finish a race. First the runner must reach the halfway point, then half of what remains, and so on without end. The sum $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots=1$ seems to answer him, since infinitely many steps can have a finite total. Yet no finite number of steps ever reaches 1: we accept the answer because we define the sum to be the limit of the partial sums. Is that a discovery about infinity, or a decision we made so that the addition has an answer? And does the decision tell us anything about how a real runner moves?

IM The triangle of binomial coefficients carries a different name in almost every place it was found. Indian scholars met these numbers counting the patterns of short and long syllables in poetry and set them out as a staircase; Persian mathematicians used them to expand powers, and Iran still calls it the Khayyam triangle; China calls it Yang Hui's triangle, Italy Tartaglia's, and much of Europe names it after Pascal, whose treatise of 1654 came centuries after most of these. The same pattern was found independently again and again, and the names record who made it well known in each place, not always who reached it first. What does that tell you about who a mathematical result belongs to?

RW A patient takes 100 mg of a drug every 12 hours, and in each 12-hour interval the body removes half of the drug present. Just after each dose the amount in the body is 100, then $100+50$, then $100+50+25$ mg, and so on: the partial sums of a geometric series with $u_1=100$ and $r=\frac{1}{2}$. The amount rises towards $S_\infty=\dfrac{100}{1-\frac{1}{2}}=200$ mg and never passes it. Doctors use this steady level to choose a dose and a dosing interval, so that the drug stays above the amount that works and below the amount that harms.

EXT A geometric series with $|r|<1$ is the simplest example of a **power series**, a sum of increasing powers of x:

$$\sum_{k=0}^{\infty} x^k=\frac{1}{1-x},\qquad |x|<1.$$

This one identity produces many others. Replace x by $-x$ and integrate each term to obtain the series for $\ln(1+x)$; replace x by $-x^2$ and integrate to obtain the series for $\arctan x$. Many further functions follow in the same way, and calculus courses often introduce power series with it.

End-of-Chapter Problems
Chapter 2: Sequences & Series

1. An arithmetic sequence has first term a and common difference d. The 5th term is 17 and the sum of the first 10 terms is 155.

(a) Write down two equations in a and d. [2]
(b) Find a and d. [2]
(c) Find the first negative term of the sequence. [2]

2. An arithmetic sequence has $u_3 = 7$ and $u_9 = 25$.

(a) Find u_1 and d. [3]
(b) Find S_{20}. [2]
(c) Find the least value of n such that $S_n > 600$. [3]

3. An arithmetic series has $u_1 = 2$ and $S_{10} = 155$.

(a) Find the common difference. [2]
(b) Find the sum $\sum_{k=11}^{20} u_k$. [4]

4. An arithmetic sequence has 20 terms. The sum of the odd-positioned terms $u_1 + u_3 + u_5 + \cdots + u_{19}$ is 160, and the sum of the even-positioned terms $u_2 + u_4 + u_6 + \cdots + u_{20}$ is 140. Find the first term u_1 and the common difference d. [6]

5. The sum of the first n terms of a series is given by $S_n = n^2 + 3n$.

(a) Find u_1 and u_2. [2]
(b) Show that the series is arithmetic and state the common difference. [3]
(c) Find the value of n such that $u_n = 44$. [2]

6. An AP has $u_1 = 120$ and $d = -3$.

(a) Find the first term that is negative. [3]
(b) Find the maximum value of S_n and the value(s) of n for which it occurs. [4]

7. Consider the two arithmetic sequences $4, 7, 10, \ldots$ and $3, 7, 11, \ldots$ Both sequences contain the value 7.

(a) Find the general term of each sequence. [2]
(b) Find the next three values that appear in both sequences. [4]

(c) Show that the common terms form an arithmetic sequence and state its common difference. [2]

8. A geometric sequence has first term a and common ratio r, where $r > 0$. Given $u_3 + u_4 = 12$ and $u_5 + u_6 = 3$, find a and r. [5]

9. The first three terms of a geometric sequence are k, 6, and $k + 5$.

(a) Show that $k^2 + 5k - 36 = 0$, and hence find the possible values of k. [4]
(b) For the value of k that gives a convergent series, find S_∞. [3]

10. The first, second, and fifth terms of an arithmetic sequence with first term 3 are, in that order, the first, second, and third terms of a geometric sequence.

(a) Given the arithmetic sequence is non-constant, show that its common difference is $d = 6$ and the geometric sequence has common ratio $r = 3$. [4]
(b) Determine whether the geometric series $3 + 9 + 27 + \cdots$ has a sum to infinity, giving a reason. [2]

11. A sequence is defined by $u_1 = 3$ and $u_{n+1} = 2u_n + 1$ for $n \geq 1$.

(a) Write down u_2, u_3, and u_4. [2]
(b) Let $v_n = u_n + 1$. Show that v_n is a geometric sequence. [3]
(c) Find an explicit formula for u_n. [2]

12. An arithmetic sequence has first term a and common difference d, where $a > 0$ and $d \neq 0$. Its 1st, 4th and 13th terms form a geometric sequence.

(a) Show that $3d = 2a$. [4]
(b) Given that $a = 6$, find d and write down the first four terms of the arithmetic sequence. [3]
(c) Find the common ratio of the geometric sequence. [2]
(d) Find the sum of the first 20 terms of the arithmetic sequence. [3]
(e) Find the sum of the first 8 terms of the geometric sequence. [3]

13.

(a) By expanding $(1 - r)(1 + r + r^2 + \cdots + r^{n-1})$, show that $\displaystyle\sum_{k=1}^{n} r^{k-1} = \frac{1 - r^n}{1 - r}$ for $r \neq 1$. [3]
(b) Hence find $\displaystyle\sum_{k=1}^{10} \left(\tfrac{2}{3}\right)^{k-1}$, giving your answer as an exact fraction. [3]

14.

(a) Find $\sum_{k=1}^{n}(4k-3)$ in terms of n. [3]

(b) Find the value of n such that $\sum_{k=1}^{n}(4k-3)=780$. [3]

15. GDC A geometric series has $u_1=3$ and $S_\infty=12$.

(a) Find the common ratio. [2]

(b) Show that $S_n=12\left(1-\left(\frac{3}{4}\right)^n\right)$. [2]

(c) Find the least value of n such that $S_n>11.9$. [3]

16. Consider the geometric series $1+2x+4x^2+8x^3+\cdots$

(a) Write down the common ratio. [1]

(b) Find the values of x for which the sum to infinity exists. [2]

(c) For these values, find S_∞ as a function of x. [2]

17. A geometric sequence has all positive terms. The sum of the first two terms is 10 and the sum to infinity is 25. Find u_1 and r. [5]

18. The sum to infinity of a geometric series is $\frac{3}{2}$ times the sum of the first 4 terms. The common ratio is positive. Find the common ratio. [5]

19. The repeating decimal $0.\overline{142857}=0.142857142857\ldots$ can be written as an infinite geometric series.

(a) Write down u_1 and the common ratio r. [2]

(b) Find the exact fractional value of $0.\overline{142857}$. [3]

(c) Verify your answer by converting the fraction from part (b) back to a decimal using long division. [1]

20. A ball is dropped from a height of 10 m. After each bounce it reaches 70% of the height it fell from.

(a) Find the height reached after the 3rd bounce. [2]

(b) Find the total vertical distance travelled by the ball before it comes to rest. [4]

21. GDC A machine is bought for \$60 000 and depreciates at 18% per year.

(a) Find its value after 4 years. [2]

(b) Find the least number of complete years until the value falls below one quarter of the purchase price. [3]

22. GDC \$8000 is invested for 6 years at 5% per year, compounded quarterly. Over the

same period, inflation runs at 3% per year.

(a) Find the future value of the investment. [2]
(b) Find the real value of the investment after 6 years. [2]
(c) Comment on what your answers say about the investment's true growth. [1]

23. GDC An investment account pays 5% per year, compounded annually. A sum P is invested.

(a) Write an expression for the value of the account after n years. [1]
(b) Find the value of P if the account is worth \$20 000 after 10 years. [3]
(c) Find the number of complete years for the account to triple in value. [3]

24. GDC The value of a car decreases by the same percentage each year. A new car costs \$32 000, and after 3 years it is worth \$16 384.

(a) Show that the value falls by 20% each year. [2]
(b) Find the year in which the value first falls below \$10 000. [3]
(c) Find the total loss in value over the first 5 years. [3]

25. GDC A student can invest \$3000 in one of two accounts. Account A pays 5% per year *simple* interest; account B pays 4% per year, compounded annually.

(a) Write down an expression for the value of each account after n years. [3]
(b) Find the value of each account after 10 years. [2]
(c) Using a table or graph on your GDC, find the least whole number of years after which account B is worth more than account A. [3]

26. GDC A savings plan pays \$500 into an account at the start of each year. The account earns 4% per year, compounded annually.

(a) Show that after 3 years the total value is $500(1.04)^3 + 500(1.04)^2 + 500(1.04)$. [2]
(b) Write this as a geometric series and find the total after 10 years. [4]
(c) Find the number of complete years until the total exceeds \$8 000. [3]

27. GDC An amount P is invested. Bank A pays 4% interest compounded annually. Bank B pays 3.9% compounded monthly, so its value after t years is $P\left(1+\dfrac{0.039}{12}\right)^{12t}$.

(a) Write down an expression for the value in Bank A after t years. [2]
(b) Show that Bank B's value can be written as Pr^t, and find r to five decimal places. [3]
(c) Determine which bank gives more after 10 years, and by how much as a percentage of P. [3]
(d) Use logarithms to find how long an investment in Bank A takes to double. [4]

(e) Bank C pays 4% compounded continuously, giving $Pe^{0.04t}$. Show that Bank C always beats Bank A, and find the difference after 10 years as a percentage of P. [3]

28. A school committee of 6 is to be chosen from 9 teachers and 7 students.

(a) Find the number of committees containing exactly 4 students. [2]
(b) Find the number of committees containing at least 4 students. [3]

29.

(a) Show that ${}^nC_2 = \dfrac{n(n-1)}{2}$, and explain the connection with the sum $1 + 2 + \cdots + (n-1)$. [3]
(b) Hence find n such that ${}^nC_2 = 190$. [2]

30. In the expansion of $(1+kx)^8$, the coefficient of x^2 is 112.

(a) Find the possible values of k. [3]
(b) For $k > 0$, find the coefficient of x^3. [2]

31. Consider the expansion of $\left(x + \dfrac{2}{x^2}\right)^{12}$.

(a) Find the term independent of x. [4]
(b) Find the coefficient of x^3. [2]

32. Consider the expansion of $\left(2x - \dfrac{3}{x}\right)^{10}$.

(a) Write down an expression for the general term of the expansion. [2]
(b) Find the term that is independent of x. [4]
(c) Find the coefficient of x^4. [3]
(d) Explain why the expansion contains no term in x^3. [2]
(e) By substituting a suitable value of x, find the sum of all the coefficients. [2]
(f) State how many terms the expansion has, and which powers of x occur. [2]

33.

(a) Expand $(1+2x)^{-2}$ in ascending powers of x up to and including the term in x^3, and state the values of x for which the expansion is valid. [4]
(b) Use your expansion with $x = 0.01$ to estimate $\dfrac{1}{1.02^2}$ to 5 decimal places. [2]

34.

(a) Show that $(4+x)^{1/2} = 2\left(1 + \dfrac{x}{4}\right)^{1/2}$, and expand it in ascending powers of x up to the term in x^2, stating the validity. [4]

(b) Hence find an approximation to $\sqrt{4.2}$, giving five decimal places. [2]

Challenge Questions

35. GDC **When does a sequence settle?** A sequence is defined by a starting value u_1 and the rule $u_{n+1} = au_n + b$, where a and b are constants. This investigation finds, for every value of a, whether the terms approach a limit.

(a) Take $a = \frac{1}{2}$, $b = 3$ and $u_1 = 2$. Tabulate $u_1, u_2, \ldots, u_6$, and state the value the terms appear to approach. [2]

(b) Let $a \neq 1$. Show that if the sequence has a limit L, then $L = aL + b$, and hence $L = \dfrac{b}{1-a}$. Check this against part (a). [2]

(c) With $a \neq 1$ and $L = \dfrac{b}{1-a}$, let $v_n = u_n - L$. Show that $v_{n+1} = av_n$, and deduce that $u_n = L + (u_1 - L)\,a^{n-1}$. [4]

(d) Suppose $u_1 \neq L$. Use part (c) to describe the long-run behaviour of $\{u_n\}$ when $0 < |a| < 1$, when $a = -1$ and when $|a| > 1$, saying in each case whether the sequence converges, and whether the terms stay on one side of L or alternate about it. Describe also what happens when $a = 1$. [4]

(e) A patient takes 100 mg of a drug every 12 hours, and by the time of the next dose the body has removed 60% of the drug present. Let u_n be the amount, in mg, just after the nth dose, so $u_1 = 100$ and $u_{n+1} = 0.4u_n + 100$. Find the long-run level, and the first dose after which the amount is within 1 mg of it. Find also the largest dose that keeps the long-run level at or below 200 mg. [4]

(f) The rule $u_{n+1} = u_n^2 + c$, with $u_1 = 0$, is not of the form above. Show that a limit L can exist only if $c \leq \frac{1}{4}$. Explore the sequence for $c = -0.5,\ 0.3,\ -1$ and -2.1, describe what you see, and state a question about this rule that could be investigated further. [3]

36. GDC **Counting multiples.** Each set of multiples below is an arithmetic sequence, so its total is an arithmetic series. Throughout, $\lfloor t \rfloor$ is the greatest integer not exceeding t; in parts (a) to (d), work with the integers from 1 to 1000.

(a) Find the sum of those divisible by 3, and the sum of those divisible by 5. [3]

(b) Explain why adding your two answers counts some integers twice, and find the sum of those divisible by both. [3]

(c) Hence find the sum of those divisible by 3 or by 5. [2]

(d) Find the sum of those divisible by 3 or by 5, but not by both. [3]

(e) Let p and q be coprime integers greater than 1 (their only common positive factor is 1), and let N be a positive integer. Show that the sum of the integers from 1 to N divisible

by p or q but not both is

$$S_p + S_q - 2S_{pq}, \qquad \text{where } S_m = m\,\frac{t(t+1)}{2}, \quad t = \left\lfloor \frac{N}{m} \right\rfloor.$$

[4]

(f) Use the result of part (e) with $p = 7$, $q = 11$ and $N = 1000$. Verify your answer by summing a list on your GDC. [3]

(g) Let $T(N)$ be the sum in part (e). Show that $\dfrac{T(N)}{N^2} \to \dfrac{1}{2p} + \dfrac{1}{2q} - \dfrac{1}{pq}$ as $N \to \infty$, and evaluate this for $p = 3$, $q = 5$. Suggest how the investigation could be extended. [3]

37. A sum that is not a sum. You may have seen the claim that $1 + 2 + 3 + 4 + \cdots = -\frac{1}{12}$. Every term is positive, so the statement cannot mean what addition normally means. This question takes the popular argument apart.

Write $G = 1 - 1 + 1 - 1 + \cdots$ and $T = 1 + 2 + 3 + 4 + \cdots$

(a) Write down the first six partial sums of G. Explain why G has no sum in the usual sense. [3]

(b) Bracketing the terms as $(1 - 1) + (1 - 1) + \cdots$ gives 0; bracketing them as $1 - (1 - 1) - (1 - 1) - \cdots$ gives 1. State what this shows about applying the ordinary rules of algebra to such a series. [3]

(c) Applying $S_\infty = \dfrac{u_1}{1 - r}$ to G with $u_1 = 1$ and $r = -1$ gives $\frac{1}{2}$. State the condition under which that formula was derived, and explain why it does not hold here. [3]

(d) Write down the nth partial sum of T and show that it increases without bound. [3]

(e) The popular argument runs: assume $G = \frac{1}{2}$; deduce $1 - 2 + 3 - 4 + \cdots = \frac{1}{4}$; subtract this from T term by term to obtain $4 + 8 + 12 + \cdots = 4T$; conclude $T - \frac{1}{4} = 4T$, so $T = -\frac{1}{12}$. Identify the step that is not valid, and say why. [4]

(f) Take the partial sums of G from part (a) and average the first n of them. Compute these averages for $n = 1$ to 6, and state what value they appear to approach. [3]

(g) Part (f) shows that $\frac{1}{2}$ is not arbitrary: it is the value G receives when "sum" is defined as the limit of these averages. Show that this averaging definition gives T no value. Then explain what a statement such as $1 + 2 + 3 + \cdots = -\frac{1}{12}$ actually asserts, and why writing it with an equals sign is misleading. [3]

(h) Another definition (Abel's) replaces a series $c_0 + c_1 + c_2 + \cdots$ by $c_0 + c_1 x + c_2 x^2 + \cdots$ for $0 < x < 1$, and lets x approach 1 from below. Find the value this gives G. Given that $1 - 2x + 3x^2 - 4x^3 + \cdots = \dfrac{1}{(1 + x)^2}$ for $0 < x < 1$, find the value it gives $1 - 2 + 3 - 4 + \cdots$, and comment on the result in the light of part (e). Suggest how comparing definitions of "sum" could be taken further. [3]

38. GDC **How many terms for $\sqrt{2}$?** The binomial series for $(1 - x)^{1/2}$ converges for $|x| < 1$, but how quickly it converges depends on x. This investigation measures the error

made by stopping the series early, and uses the result to choose x well. Write

$$(1-x)^{1/2} = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \cdots .$$

(a) Show that each of the following is true:

$$\sqrt{2} = 2\left(1-\tfrac{1}{2}\right)^{1/2}, \qquad \sqrt{2} = \tfrac{10}{7}\left(1-\tfrac{1}{50}\right)^{1/2}, \qquad \sqrt{2} = \tfrac{17}{12}\left(1-\tfrac{1}{289}\right)^{1/2}.$$

[3]

(b) Find $c_0, c_1, \ldots, c_4$. Show that $\dfrac{c_{k+1}}{c_k} = \dfrac{2k-1}{2k+2}$ for $k \geq 1$, and deduce that every coefficient after c_0 is negative and that $|c_{k+1}| < |c_k|$ for $k \geq 1$. [4]

(c) Let $0 < x < 1$ and $m \geq 1$, and let E_m be the amount by which $c_0 + c_1 x + \cdots + c_m x^m$ exceeds $(1-x)^{1/2}$. Show that

$$0 < E_m < \frac{|c_{m+1}|\, x^{m+1}}{1-x}.$$

(Hint: compare the omitted terms with a geometric series) [4]

(d) For each expression in part (a), use the series up to the x^3 term to approximate $\sqrt{2}$. Give the actual error in each approximation and the bound from part (c), each to two significant figures. [4]

(e) Using $\sqrt{2} = \frac{17}{12}\left(1-\frac{1}{289}\right)^{1/2}$, find the least m for which the bound in part (c) guarantees an error in $\sqrt{2}$ below 10^{-12}. [3]

(f) The fraction $\frac{17}{12}$ satisfies $p^2 - 2q^2 = 1$. Show that whenever positive integers p and q satisfy $p^2 - 2q^2 = 1$, $\sqrt{2} = \dfrac{p}{q}\left(1 - \dfrac{1}{p^2}\right)^{1/2}$. Find the next such fraction after $\frac{17}{12}$, and estimate how many correct decimal places each extra term of the series gives for it, compared with $x = \frac{1}{2}$ and $x = \frac{1}{289}$. Suggest a direction in which this could be investigated further. [3]

NUMBER & ALGEBRA

Chapter 3

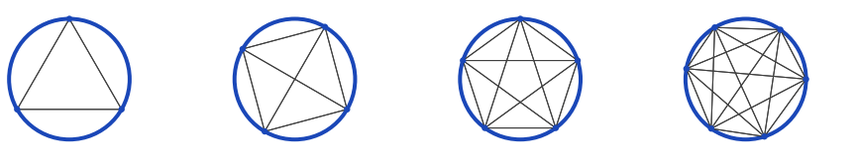
Proofs

3

SYLLABUS 1.6 1.15

Mark some points on a circle. Join every pair of points with a straight line. Then count the regions inside the circle. Place the points carefully, so that no three lines cross at the same spot. The first few cases give 1 region, then 2, then 4, then 8, then 16.

The pattern seems clear: each new point doubles the number of regions, so six points should give 32. In fact six points give **31**.



3 points, **4 regions** 4 points, **8 regions** 5 points, **16 regions** 6 points, **31 regions**

Figure 3.1 Every pair of points is joined, and no three chords meet at one point.

Five cases followed the doubling rule and the sixth did not. The first five gave no sign of it, and no number of agreeing cases could have shown the rule was true.

This is the difference between seeing a pattern and knowing that it always holds. A pattern is a guess. The word for such a guess is a **conjecture**: a statement believed true but not yet proved. The doubling rule was a conjecture, and the sixth case disproved it.

A **proof** is an argument in which each step follows from the one before, so the conclusion cannot fail, however many points you take.

By the end of this chapter you will be able to prove a statement directly from definitions, disprove one with a single counterexample, argue by contradiction, and prove a result about every positive integer by induction. Checking a rule on a few cases shows only that it has not failed yet. A proof shows that it cannot fail.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

3.1 Deductive Proof

A **proof** is a chain of logical steps. It starts from things already accepted, such as definitions, results proved earlier, and the rules of algebra, and it ends at the statement you want to establish. Each step must follow with certainty from the steps before it. If the chain is correct, the conclusion is not just likely. It cannot be false.

This is **deductive** reasoning: you apply general principles to reach a specific, guaranteed result. A familiar example is not mathematical at all:

- Every human being is mortal. (a general principle)
- Socrates is a human being. (a specific case)
- Therefore Socrates is mortal. (the conclusion)

The conclusion is not a guess, and no amount of further evidence could improve it. If the first two statements are true, the third cannot be false. Mathematics uses arguments of exactly this shape:

- Every multiple of 4 is even.
- The number 132 is a multiple of 4.
- Therefore 132 is even.

Notice that you did not divide 132 by 2 to check. The conclusion followed from the two statements above it.

The circle problem worked in the opposite direction. It used a few particular cases to guess a general rule. Reasoning in that direction, from examples to a general claim, is called induction in the everyday sense of the word, and it can lead to a false conclusion. Deduction cannot.

The two directions differ in certainty. Deduction goes from a rule about every case to one particular case. Induction in the everyday sense goes from a few cases to a rule about all of them, and only deduction is certain. Despite its name, *mathematical* induction (§3.4) is a form of deduction: it proves a rule for every positive integer from two facts that have been proved, and it is the last method this chapter builds.

3.1.1 Equality versus identity

Before you prove a statement about an equation, you must know which kind of statement it is. The equals sign is used for two very different claims.

DEF An **equation** is a statement that two expressions are equal for *some particular* values of the variable. An **identity** is a statement that they are equal for *every* value for which both sides are defined. An identity is written with the symbol $\equiv$:

$$(x+1)^2 \equiv x^2 + 2x + 1.$$

The equation $x^2 = 4$ is true only when $x = 2$ or $x = -2$. The identity $(x+1)^2 \equiv x^2 + 2x + 1$ is true for every real x. Solving an equation means finding the values that make it true. Proving an identity means showing that no value makes it false. The three-bar symbol $\equiv$ marks which of the two is meant: where it appears, the claim is "for all x", not "find x".

You already know one important identity from trigonometry:

$$\sin^2 x + \cos^2 x \equiv 1.$$

This is true for every angle x, not for some particular angles, which is why it is written with $\equiv$. It is proved and used in Ch. 8.

3.1.2 Proving an identity

The two sides of an identity have names. The expression to the left of the $\equiv$ sign is the **left-hand side**, written **LHS**. The expression to the right is the **right-hand side**, written **RHS**.

To prove an identity you take one side and transform it, step by step, into the other, using only valid algebra. Start from the more complicated side and simplify it: write down the LHS, change it one line at a time, and finish with the RHS. Write the labels LHS and RHS in your working, so the direction of the argument is clear to the marker.

CAUTION Do not start by assuming the identity is true and then manipulate both sides until you reach something obvious like $0 = 0$. That assumes what you are trying to prove. Transform *one* side into the other instead.

Worked through. To prove that $(a+b)^2 - (a-b)^2 \equiv 4ab$, the difficulty is avoiding a common error: doing the same thing to both sides. Start from the LHS instead and expand each square:

$$\text{LHS} = (a+b)^2 - (a-b)^2 = (a^2 + 2ab + b^2) - (a^2 - 2ab + b^2).$$

Now remove the brackets. Take care with the signs in the second bracket:

$$= a^2 + 2ab + b^2 - a^2 + 2ab - b^2 = 4ab = \text{RHS}. \quad \blacksquare$$

Nothing but expansion was used, and the LHS turned into the RHS. The two sides were never assumed to be equal. That they are equal is the result, not the starting point.

A quick **check** takes about five seconds, and many errors show up in it. At $a = 3$, $b = 2$ the left side is $25 - 1 = 24$ and the right side is $4(3)(2) = 24$. A check is not a proof. But if a check fails, something is certainly wrong.

TIP Check your results by substituting a number where you can. An algebra error will usually show up at one well-chosen value.

EXAMPLE 3.1

Show that $\frac{1}{4} + \frac{1}{12} = \frac{1}{3}$, and prove the algebraic generalisation

$$\frac{1}{m+1} + \frac{1}{m^2+m} \equiv \frac{1}{m}.$$

The numerical case is direct: $\frac{1}{4} + \frac{1}{12} = \frac{3}{12} + \frac{1}{12} = \frac{4}{12} = \frac{1}{3}$. Notice it is the claimed identity at $m = 3$: $m + 1 = 4$ and $m^2 + m = 12$.

For the general statement, start from the left side and factor the second denominator, $m^2 + m = m(m + 1)$:

$$\frac{1}{m+1} + \frac{1}{m(m+1)} = \frac{m}{m(m+1)} + \frac{1}{m(m+1)} = \frac{m+1}{m(m+1)} = \frac{1}{m}. \quad \blacksquare$$

One calculation confirmed a single case. The algebra covers every $m \neq 0, -1$ at once. Examination questions often use this "verify, then generalise" structure.

3.1.3 Proving a general statement

Many proofs are not about identities. They are about claims such as "the sum of two odd numbers is even". For these, the first step is to write the general object in algebra, so that one expression stands for every case at once.

The table below lists the standard forms. In every row, k and n are integers.

In words	In algebra
an even number	$2k$
an odd number	$2k+1$, or $2k-1$
two consecutive integers	n and $n+1$
three consecutive integers	$n-1$, n, $n+1$
two consecutive even numbers	$2k$ and $2k+2$
two consecutive odd numbers	$2k+1$ and $2k+3$
a multiple of 3	$3k$
a number that leaves remainder 1 when divided by 3	$3k+1$
a square number	k^2
a rational number	p/q, with p, q integers and $q \neq 0$

Choosing the right form matters. To prove a statement about *any* two odd numbers, use $2a+1$ and $2b+1$ with two different letters. Writing $2a+1$ twice would only prove the statement for two *equal* odd numbers.

When no single form works, split into cases that cover every integer, such as n even or odd, or $n = 3k$, $3k+1$, $3k+2$, and prove the claim in each case. Every integer must fall in exactly one case.

Once the general case is written in symbols, the rest is ordinary algebra.

EXAMPLE 3.2

A single example such as $3+5=8$ confirms nothing about *all* odd numbers. Prove that the sum of any two odd numbers is even.

Let the two odd numbers be $2a+1$ and $2b+1$, where a and b are integers. Their sum is

$$(2a+1)+(2b+1) = 2a+2b+2 = 2(a+b+1).$$

Since $a+b+1$ is an integer, the sum is $2 \times (\text{integer})$, which is by definition even. ■

EXAMPLE 3.3

Prove that the product of two consecutive integers is always even.

Take the consecutive integers n and $n+1$. One of any two consecutive integers is even. If n is even, the product $n(n+1)$ has an even factor. If n is odd, then $n+1$ is even, and the product has an even factor again. Either way $n(n+1)$ is a multiple of 2, so it is even. ■

To prove divisibility by a composite number, prove divisibility by each factor separately. The factors must be **coprime**, sharing no factor larger than 1: a number divisible by 2 and by 4 need not be divisible by 8, as 12 shows.

EXAMPLE 3.4

Prove that $n(n+1)(2n+1)$ is divisible by 6 for every integer n.

Since $6 = 2 \times 3$, find a factor 2 and a factor 3 separately.

A factor of 2. One of any two consecutive integers is even, so one of n and $n+1$ is even, and the product is divisible by 2.

A factor of 3. Every integer leaves remainder 0, 1 or 2 on division by 3, so n has one of the forms $3k$, $3k+1$ or $3k+2$, where k is an integer. If $n = 3k$, then the first factor is a multiple of 3. If $n = 3k+1$, then $2n+1 = 6k+3 = 3(2k+1)$. If $n = 3k+2$, then $n+1 = 3k+3 = 3(k+1)$. The three cases cover every integer, and in each one of the three factors is a multiple of 3, so the product is divisible by 3.

The product is divisible by 2 and by 3. Since 2 and 3 are coprime, these are separate factors, and the product is divisible by $2 \times 3 = 6$. ■

CAUTION One example can never prove a statement that begins "for all." It can only *disprove* one, which is the subject of the next section. The proofs above work because the algebra stands for every case at once.

YOUR TURN 3.1 Deductive Proof

Equation or identity

1. State whether each of the following is an equation or an identity.

(a) $3x = 12$

(b) $(x+2)^2 = x^2 + 4x + 4$

Proving an identity

2. Prove each identity. Start from one side and label LHS and RHS in your working.

(a) $(a+b)^2 \equiv a^2 + 2ab + b^2$

(b) $(a-b)^2 \equiv a^2 - 2ab + b^2$

(c) $\dfrac{1}{x} - \dfrac{1}{x+1} \equiv \dfrac{1}{x(x+1)}$, for $x \neq 0, -1$

(d) $(x+3)(x-3) \equiv x^2 - 9$

(e) $\dfrac{1}{x-1} - \dfrac{1}{x+1} \equiv \dfrac{2}{x^2-1}$, for $x \neq \pm 1$

3. Prove that $(x-3)^2 + 5 \equiv x^2 - 6x + 14$. Check your result at $x = 1$.

Verify, then generalise

4. This question checks one case, then proves the general result.

(a) Verify that $3^2 + 4^2 + 12^2 = 13^2$.

(b) Prove that $n^2 + (n+1)^2 + (n(n+1))^2 \equiv (n^2+n+1)^2$, and state the value of n that gives part (a).

5. Two examples of a pattern are $7 \times 9 = 8^2 - 1$ and $11 \times 13 = 12^2 - 1$.

(a) Verify both examples.

(b) Prove that the product of two consecutive odd numbers is always one less than the square of the even number between them.

Odd, even and divisibility

6. Prove each statement about odd and even integers.

(a) The sum of any two even integers is even.

(b) The sum of an even integer and an odd integer is odd.

(c) The product of any two odd integers is odd.

7. Prove each statement for every integer n.

(a) $(2n+1)^2$ is odd.

(b) $n^2 + 3n$ is even.

(c) $(2n+1)^2 - 1$ is divisible by 4.

(d) $n^3 - n$ is divisible by 6.

Proof by cases

8. Prove that $n^2 + n + 1$ is odd for every integer n, by considering the cases n even and n odd.

9. Prove that, for every integer n, the remainder when n^2 is divided by 3 is either 0 or 1. Consider the three cases $n = 3k$, $n = 3k+1$ and $n = 3k+2$, where k is an integer.

Concept check

10. A student sets out to prove that the sum of any two odd integers is even, and writes the two odd integers as $2n+1$ and $2n+3$, where n is an integer. Explain why this does not prove the claim, and give a correct proof.

3.2 Disproof by Counterexample HL

A universal statement claims something about every member of a collection. To show that such a claim is false, you do not need to discuss the whole collection. You need one member for which the claim fails.

DEF A **counterexample** is one specific case for which a general statement fails. A single counterexample is enough to prove that the statement is false.

The logic is exact. "Every n has property P" is false as soon as one value of n does not have property P. This is what happened with the circle problem. The claim "the number of regions doubles each time" was a universal statement, and $n=6$ was a counterexample to it.

Worked through. The formula n^2+n+41 is famous because it produces a prime number for $n=0,1,2,\ldots,39$, forty primes in a row. To decide whether it produces a prime for *every* non-negative integer n, test the claim at the first value the pattern does not cover. At $n=40$:

$$40^2+40+41=1600+40+41=1681=41^2.$$

This is 41×41, not prime. So $n=40$ is a counterexample, and the statement "n^2+n+41 is always prime" is false. ■

Forty cases agreed with the claim. One case was enough to make it false. Examples can make a claim look true. They can never make it certain.

EXAMPLE 3.5

Disprove the claim: "every multiple of 3 is odd."

The number 6 is a multiple of 3, and 6 is even. That single case is a counterexample, so the claim is false. ■

EXAMPLE 3.6

Show that this statement is not always true: "there are no positive integer solutions to the equation $x^2+y^2=10$."

Take $x = 1$ and $y = 3$. Both are positive integers, and

$$1^2 + 3^2 = 1 + 9 = 10.$$

So a positive integer solution does exist, and the statement is false. ■

Notice what the answer contains: the values, the substitution, and the conclusion. Giving only "$x = 1$, $y = 3$" would not earn full marks.

TIP A counterexample must be *explained*, not just named. State the case, substitute it into the claim, show the calculation, and say in words why this makes the claim false. The IB is explicit that stating the counterexample alone is not enough.

Two more claims are disproved by counterexamples. The first is "if $a^2 > b^2$ then $a > b$". Take $a = -3$ and $b = 1$: then $9 > 1$ while $-3 < 1$, because squaring removes the sign. The second is "the product of two irrational numbers is irrational". Take $\sqrt{2} \times \sqrt{2} = 2$, which is rational.

RW In 1640 Fermat studied the numbers $F_n = 2^{2^n} + 1$ and found that $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$ and $F_4 = 65\,537$ are all prime. He conjectured that every F_n is prime. Almost a century later, in 1732, Euler tested the next one:

$$F_5 = 2^{32} + 1 = 4\,294\,967\,297 = 641 \times 6\,700\,417.$$

One counterexample was enough to disprove the conjecture. No larger Fermat prime has ever been found.

CAUTION A counterexample only *disproves* a claim. Finding no counterexample after many attempts is not a proof. The circle problem at the start of this chapter shows why: five cases agreed, and the sixth did not.

YOUR TURN 3.2 Disproof by Counterexample HL

Counterexamples with integers

11. Disprove each claim by giving a counterexample. Show the substitution, and say why it makes the claim false.

(a) $n^2 + n + 1$ is prime for every positive integer n.
(b) Every prime number is odd.
(c) If n is divisible by 4 then n is divisible by 8.
(d) The sum of two prime numbers is always even.
(e) $2^n + 1$ is prime for every positive integer n.
(f) If $a > b$ then $a^2 > b^2$.

Counterexamples with algebra

12. Disprove each claim by giving a counterexample. Show the substitution, and say why it makes the claim false.

(a) $n^2 \geq 2n$ for all integers n.
(b) $x^2 > x$ for all real x.
(c) If n^2 is even then n is odd.
(d) $\sqrt{a+b} = \sqrt{a} + \sqrt{b}$ for all $a, b \geq 0$.
(e) If p is prime then $2^p - 1$ is prime.

13. Disprove each claim by giving a counterexample.

(a) If ab is rational, then a and b are both rational.
(b) $\sqrt{x^2} = x$ for every real number x.

Showing that a statement is not always true

14. Show that each statement is not always true.

(a) There is no integer n for which $n^2 + 1$ is divisible by 5.
(b) There are no positive integers x and y with $x^2 - y^2 = 15$.

Concept check

15. A student checks $n^2 - n + 41$ for $n = 1, 2, \ldots, 10$, finds a prime every time, and concludes that $n^2 - n + 41$ is prime for every positive integer n.

(a) Explain why the checks do not prove the claim.
(b) By taking $n = 41$, show that the claim is false.

3.3 Proof by Contradiction HL

Suppose a friend tells you they have written down the largest whole number that exists. You do not need to see it to know they are wrong. Whatever the number is, add 1 to it. The result is a whole number, and it is larger. The claim disproves itself.

You have just proved something without ever finding the answer. You did not search for the largest whole number. You assumed it existed and showed that the assumption leads to an impossible situation.

This is the idea of **proof by contradiction**. Assume the statement you want to prove is *false*. Reason correctly from that assumption until you reach something impossible. Correct reasoning cannot turn a true assumption into an impossible result, so the assumption must have been false. The original statement is therefore true.

DEF In **proof by contradiction**, you assume the negation of what you want to prove, then derive a logical contradiction. The contradiction forces the assumption to be false, which establishes the original claim.

The method has a Latin name, *reductio ad absurdum*, meaning "reduction to the absurd". The method is useful because the assumption itself is a starting point. A direct proof that no fraction equals $\sqrt{2}$ would mean checking infinitely many fractions. Assuming that such a fraction exists produces an equation instead, and an equation can be worked on. To prove a number irrational, assume it equals $\frac{p}{q}$, where p and q are integers with $q \neq 0$ that share no common factor. The contradiction usually comes from showing that they do share one. Two classic results, each more than two thousand years old, show the method at work.

Worked through. To prove that $\sqrt{2}$ is irrational, a classical result alluded to by Aristotle in the 4th century BCE, assume the opposite: that $\sqrt{2}$ *is* rational. Then we can write

$$\sqrt{2} = \frac{p}{q},$$

where p and q are integers **sharing no common factor**, so the fraction is fully cancelled. The contradiction at the end of the proof depends on this condition, so state it explicitly. Squaring both sides gives $2 = \dfrac{p^2}{q^2}$, hence

$$p^2 = 2q^2.$$

So p^2 is even, which forces p itself to be even (the square of an odd number is odd). Write $p = 2m$. Then $p^2 = 4m^2$, and substituting,

$$4m^2 = 2q^2 \implies q^2 = 2m^2,$$

so q^2 is even, and therefore q is even too. But now p and q are both even, sharing the factor

2, contradicting our choice of a fully cancelled fraction. The assumption that $\sqrt{2}$ is rational is impossible, so $\sqrt{2}$ is irrational. ∎

Tradition links this result to the end of the Pythagorean belief that every ratio of lengths is a ratio of whole numbers. The diagonal of a unit square has length $\sqrt{2}$, and the proof above says no fraction gives it.

Worked through. To prove that there are infinitely many prime numbers, a result from Euclid's *Elements* (c. 300 BCE), assume the opposite: that there are only finitely many primes, and list them all as $p_1, p_2, \ldots, p_n$. Build the number

$$N = p_1 \times p_2 \times \cdots \times p_n + 1,$$

the product of every prime, plus one. Dividing N by any prime on the list leaves remainder 1, so N is divisible by none of them. Yet every integer greater than 1 has at least one prime factor. That factor cannot be on our list, so our list was not complete, a contradiction. No finite list can hold all the primes, so there are infinitely many. ∎

Notice what the proof does not do: it never produces a new prime, only shows that the list was incomplete. N itself need not be prime. With $2, 3, 5, 7, 11, 13$ the number $N = 30031 = 59 \times 509$, and both factors are primes missing from the list.

The $\sqrt{2}$ proof used one fact about even numbers: if p^2 is even, then p is even. To prove that $\sqrt{5}$ is irrational in the same way, you need the matching fact about multiples of 5. A small result proved on the way to a larger one is called a **lemma**. With the matching lemma in place of the fact about even numbers, the $\sqrt{2}$ proof works for $\sqrt{5}$ step by step. The next example proves that fact directly, by checking every case; it is not itself a proof by contradiction.

EXAMPLE 3.7

Prove that if n^2 is a multiple of 5, then n is a multiple of 5.

Nothing is assumed here; every possible case is checked. Every integer lies within 2 of a multiple of 5, so n is $5k$, $5k \pm 1$ or $5k \pm 2$ for some integer k. Square each form:

$$(5k)^2 = 5(5k^2), \qquad (5k \pm 1)^2 = 5(5k^2 \pm 2k) + 1, \qquad (5k \pm 2)^2 = 5(5k^2 \pm 4k) + 4.$$

Only the first of these is a multiple of 5; the others leave remainder 1 or 4. So n^2 can be a multiple of 5 only when n has the form $5k$, which is to say only when n is a multiple of 5. ∎

With this lemma in place of the fact about even numbers, assume $\sqrt{5} = \frac{p}{q}$ in lowest terms, deduce that 5 divides p and then that 5 divides q, and reach the same contradiction.

Rational numbers have **closure**: adding, subtracting or multiplying two rational numbers, or dividing one by a non-zero rational, gives a rational number. So to show that an expression built from a known irrational is irrational, assume it is rational and use only these operations to make the known irrational the subject. The known irrational is then rational, a

contradiction.

EXAMPLE 3.8

Prove that $(1+\sqrt{2})^2$ is irrational. You may use the theorem that $\sqrt{2}$ is irrational.

Expand first, so that the known irrational stands alone in a single term:

$$(1+\sqrt{2})^2 = 1 + 2\sqrt{2} + 2 = 3 + 2\sqrt{2}.$$

Assume the opposite: that this number is rational. Then

$$3 + 2\sqrt{2} = \frac{p}{q},$$

where p and q are integers with $q \neq 0$. Isolate $\sqrt{2}$ by subtracting 3 and halving:

$$\sqrt{2} = \frac{1}{2}\left(\frac{p}{q} - 3\right) = \frac{p - 3q}{2q}.$$

Now $p - 3q$ and $2q$ are integers and $2q \neq 0$, so the right-hand side is a ratio of two integers, and $\sqrt{2}$ is rational. That contradicts the theorem that $\sqrt{2}$ is irrational, so the assumption fails and $(1+\sqrt{2})^2$ is irrational. ■

Both middle steps use closure. Subtracting 3 and dividing by 2 each move from one rational number to another, so neither could have carried a rational number to an irrational one.

The contradiction here is with a result already established, not with anything internal to $(1+\sqrt{2})^2$. Expanding made the argument possible, because it wrote the expression in terms of $\sqrt{2}$: identify the known irrational first, then rearrange to make it the subject.

CAUTION State your assumption clearly at the start ("assume …is rational") and name the contradiction when you reach it. A proof by contradiction with no clearly identified contradiction proves nothing.

YOUR TURN 3.3 Proof by Contradiction HL

No largest or smallest

16. Prove each statement by contradiction.

(a) There is no smallest positive rational number.

(b) There is no smallest positive real number.

Statements about integers

17. Let n be an integer. Prove each statement by contradiction.

(a) If n^2 is odd then n is odd.

(b) If n^2 is even then n is even.

18. Prove each statement by contradiction.

(a) There are no integers x and y with $2x + 4y = 5$.

(b) There are no integers x and y with $6x + 9y = 10$.

Irrational numbers

19. Prove each statement by contradiction.

(a) The sum of a rational number and an irrational number is irrational.

(b) $\sqrt{3}$ is irrational.

(c) If x is irrational and $x \neq 0$, then $\dfrac{1}{x}$ is irrational.

20. Prove each statement by contradiction. You may use the facts that $\sqrt{2}$ and $\sqrt{3}$ are irrational.

(a) $5 - \sqrt{2}$ is irrational.

(b) $(2 + \sqrt{3})^2$ is irrational.

Infinitely many primes

21. Suppose there were only finitely many primes, $p_1, p_2, \ldots, p_n$. Let $N = p_1 p_2 \cdots p_n - 1$. By considering the remainder when N is divided by each p_i, prove by contradiction that there are infinitely many primes.

Concept check

22. To prove "*if n^2 is even then n is even*" by contradiction, a student begins: "*Assume that n^2 is odd.*" Explain the error, and state the correct assumption.

3.4 Proof by Induction HL

Picture an endless line of dominoes. You cannot knock over infinitely many by hand. But suppose you can show two things: that the first domino falls, and that *whenever* a domino falls it topples the next. Then every domino in the line must fall, however long the line.

Mathematical induction is exactly this idea, used to prove a statement P_n for every integer n from some starting value onward. The subscript in P_n matters: it marks the n-th claim in an infinite list $P_1, P_2, P_3, \ldots$, not a function being evaluated. You verify the first

claim, then prove that each claim forces the one after it.

KEY **Principle of mathematical induction.** To prove that P_n holds for all integers $n \geq 1$:

1. Show that P_1 is true.
2. Assume that P_k is true for **some** $k \geq 1$ (the *inductive hypothesis*).
3. Show that P_k implies P_{k+1}.
4. Conclude: *since P_1 is true, and P_k true implies P_{k+1} true, by the principle of mathematical induction P_n is true for all integers $n \geq 1$.*

(If the statement starts later, at n_0, begin at P_{n_0} instead of P_1.)

An induction proof in this course follows these four steps, and the sentence in step 4 should be written out in full. It is not decoration. It is the step that turns the two facts you proved into a conclusion about every n, and markschemes usually award a mark for it.

CAUTION In step 2, assume P_k for **some** k, never for *all* k. Assuming it for all k assumes the very thing you are proving, and examiners withhold the reasoning marks for it. The hypothesis is one domino falling, not the whole line.

3.4.1 Why induction works

The domino image can be made precise. Suppose steps 1 and 3 are done, and pick any n you like, say $n = 1000$. Then P_1 is true; P_1 forces P_2; P_2 forces P_3; and after 999 applications of the inductive step, P_{1000} is true. The same finite chain reaches *any* n, so no value of n is left out, even though only two things were ever proved.

There is a second way to see it, and it connects induction to the previous section. Suppose the two steps are done, but P_n failed for some values of n. Among the failures there would be a *smallest* one, say P_m. Now $m \neq 1$, because step 1 checked P_1. So $m - 1 \geq 1$, and P_{m-1} is true, since m was the smallest failure. But step 3 says that P_{m-1} forces P_m, which contradicts the failure of P_m. So the principle of induction can itself be justified by contradiction, and the two techniques of this chapter are closely connected.

Induction is not only for sums. It also proves divisibility statements and inequalities below, De Moivre's theorem in Ch. 17, and formulas for n-th derivatives in Ch. 13. When a claim is about every integer $n \geq n_0$ and each case builds on the one before, induction is a natural method to try.

3.4.2 Induction for a sum

How to approach it. Here P_n says that a sum equals a formula. Add the next term of the sum to both sides of the assumption P_k. Then rearrange the right-hand side until it becomes the original formula with $k + 1$ written wherever n appeared. Write that target expression down before you begin, so you know what you are aiming at.

Worked through. To prove that $3 + 7 + 11 + \cdots + (4n - 1) = n(2n + 1)$ for all integers $n \geq 1$, let P_n be the statement $3 + 7 + \cdots + (4n - 1) = n(2n + 1)$.

Step 1 (P_1): the left side is 3; the right side is $1(2(1) + 1) = 3$. They agree, so P_1 is true.

Step 2: assume P_k is true for some $k \geq 1$, that is,

$$3 + 7 + \cdots + (4k - 1) = k(2k + 1).$$

Step 3: add the next term, $4(k + 1) - 1 = 4k + 3$, to both sides:

$$3 + 7 + \cdots + (4k - 1) + (4k + 3) = k(2k + 1) + 4k + 3.$$

Expand the right side and factorise it:

$$= 2k^2 + 5k + 3 = (k + 1)(2k + 3) = (k + 1)(2(k + 1) + 1).$$

This is exactly P_{k+1}, the original formula with n replaced by $k + 1$. So $P_k \Rightarrow P_{k+1}$.

Step 4: since P_1 is true, and P_k true implies P_{k+1} true, by the principle of mathematical induction P_n is true for all integers $n \geq 1$. ■

3.4.3 Induction for divisibility

The same four steps prove divisibility statements.

How to approach it. Here P_n says that an expression E_n is divisible by a fixed number d. Turn the assumption into an equation: divisible by d means $E_k = dm$ for some integer m. Then rewrite E_{k+1} so that E_k appears inside it, replace that part by dm, and take d outside as a common factor. The aim is to reach the form $d \times (\text{integer})$.

EXAMPLE 3.9

Prove that $8^n - 1$ is divisible by 7 for all integers $n \geq 1$.

Let P_n be the statement "$8^n - 1$ is divisible by 7."

Step 1 (P_1): $8^1 - 1 = 7$, which is divisible by 7. So P_1 is true.

Step 2: assume P_k is true for some $k \geq 1$, so $8^k - 1 = 7m$ for some integer m.

Step 3: consider the next case:

$$8^{k+1} - 1 = 8 \cdot 8^k - 1 = 8(8^k - 1) + 8 - 1 = 8(7m) + 7 = 7(8m + 1).$$

Since $8m + 1$ is an integer, $8^{k+1} - 1$ is divisible by 7, which is P_{k+1}. So $P_k \Rightarrow P_{k+1}$.

Step 4: since P_1 is true, and P_k true implies P_{k+1} true, by the principle of mathematical induction P_n is true for all integers $n \geq 1$. ■

Notice where the hypothesis was used: rewriting $8^k - 1$ as $7m$ is the only reason the final line factors cleanly.

EXAMPLE 3.10

In $8^n - 1$ the previous expression appeared as soon as a factor was taken out at the front. Prove that $n^5 + 4n$ is divisible by 5 for all integers $n \geq 1$, where it has to be extracted from an expansion instead.

Let P_n be the statement "$n^5 + 4n$ is divisible by 5."

Step 1 (P_1): $1^5 + 4(1) = 5$, which is divisible by 5. So P_1 is true.

Step 2: assume P_k is true for some $k \geq 1$, so $k^5 + 4k = 5m$ for some integer m.

Step 3: expand the next case and collect $k^5 + 4k$ out of it:

$$(k+1)^5+4(k+1) = k^5+5k^4+10k^3+10k^2+5k+1+4k+4 = (k^5+4k)+5k^4+10k^3+10k^2+5k+5.$$

The hypothesis replaces the bracket by $5m$, and every remaining term has a factor 5:

$$= 5m + 5(k^4 + 2k^3 + 2k^2 + k + 1) = 5(m + k^4 + 2k^3 + 2k^2 + k + 1).$$

Since $m + k^4 + 2k^3 + 2k^2 + k + 1$ is an integer, $(k+1)^5 + 4(k+1)$ is divisible by 5, which is P_{k+1}. So $P_k \Rightarrow P_{k+1}$.

Step 4: since P_1 is true, and P_k true implies P_{k+1} true, by the principle of mathematical induction P_n is true for all integers $n \geq 1$. ■

Expand fully first, then look for the previous expression inside the result. If what is left over after removing it does not carry the required factor, the grouping is wrong, not the statement.

3.4.4 Induction for an inequality

How to approach it. Here P_n says that $A_n > B_n$. Inequalities are proved by chaining: if $a > b$ and $b \geq c$, then $a > c$. Start from A_{k+1} and rewrite it so that A_k appears, then use the assumption to replace A_k by something smaller. That gives the first link. Now prove separately that this new quantity is at least B_{k+1}, which gives the second link. The two links join into a single chain from A_{k+1} down to B_{k+1}.

EXAMPLE 3.11

Prove that $2^n > n^2$ for all integers $n \geq 5$. The base case is not $n = 1$ here, and that matters: the inequality is false for $n = 2, 3, 4$.

Let P_n be the statement $2^n > n^2$.

Step 1 (P_5): $2^5 = 32$ and $5^2 = 25$, and $32 > 25$. So P_5 is true.

Step 2: assume P_k is true for some $k \geq 5$, that is, $2^k > k^2$.

Step 3: then

$$2^{k+1} = 2 \cdot 2^k > 2k^2.$$

It is now enough to show $2k^2 \geq (k+1)^2$. Expanding, this asks whether $2k^2 \geq k^2 + 2k + 1$, that is, $k^2 - 2k - 1 \geq 0$. For $k \geq 5$ this certainly holds, since

$k^2 - 2k - 1 = (k-1)^2 - 2 \geq 16 - 2 > 0$. Therefore

$$2^{k+1} > 2k^2 \geq (k+1)^2,$$

which is P_{k+1}. So $P_k \Rightarrow P_{k+1}$.

Step 4: since P_5 is true, and P_k true implies P_{k+1} true for $k \geq 5$, by the principle of mathematical induction $2^n > n^2$ for all integers $n \geq 5$. ■

3.4.5 Induction for a recurrence

How to approach it. Here the sequence is given by a rule linking each term to the one before, with no formula in n to prove. Compute the first few terms, conjecture a **closed form** for u_n, and let P_n be that closed form. In the inductive step the hypothesis is substituted into the recurrence rather than added to a sum: substitute the assumed u_k into the rule for u_{k+1}, then simplify until the closed form with $k+1$ in place of n appears.

EXAMPLE 3.12

A sequence is defined by $u_1 = 1$ and $u_{n+1} = \dfrac{u_n}{1 + 2u_n}$ for $n \geq 1$. Find u_2, u_3 and u_4, conjecture a formula for u_n, and prove it by induction.

Each term comes from the one before:

$$u_2 = \frac{1}{1+2} = \frac{1}{3}, \qquad u_3 = \frac{1/3}{1 + 2/3} = \frac{1/3}{5/3} = \frac{1}{5}, \qquad u_4 = \frac{1/5}{1 + 2/5} = \frac{1/5}{7/5} = \frac{1}{7}.$$

The terms so far are $1, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}$, which suggests the conjecture $u_n = \dfrac{1}{2n-1}$. Let P_n be that statement.

Step 1 (P_1): the definition gives $u_1 = 1$, and the formula gives $\dfrac{1}{2(1)-1} = 1$. They agree, so P_1 is true.

Step 2: assume P_k is true for some $k \geq 1$, that is, $u_k = \dfrac{1}{2k-1}$.

Step 3: substitute the hypothesis into the recurrence:

$$u_{k+1} = \frac{u_k}{1 + 2u_k} = \frac{1/(2k-1)}{1 + 2/(2k-1)}.$$

Multiply numerator and denominator by $2k-1$:

$$= \frac{1}{(2k-1) + 2} = \frac{1}{2k+1} = \frac{1}{2(k+1) - 1},$$

which is the formula with n replaced by $k+1$, so $P_k \Rightarrow P_{k+1}$.

Step 4: since P_1 is true, and P_k true implies P_{k+1} true, by the principle of mathematical induction $u_n = \dfrac{1}{2n-1}$ for all integers $n \geq 1$. ■

The four terms produced the conjecture and established nothing; the induction established it. A question of this shape marks the two jobs separately, so give the terms, state the conjecture, and then prove it.

CAUTION Two failures are common. One is forgetting the base case, so the dominoes can topple each other yet never start. The other is proving the inductive step *without using* the hypothesis. If you never used P_k, you have not built a chain, only restated the problem.

YOUR TURN 3.4 Proof by Induction HL

Sums

23. Prove each result by induction, for all $n \geq 1$.

(a) $1 + 2 + 3 + \cdots + n = \dfrac{n(n+1)}{2}$

(b) $1 + 3 + 5 + \cdots + (2n-1) = n^2$

(c) $1^2 + 2^2 + \cdots + n^2 = \dfrac{n(n+1)(2n+1)}{6}$

(d) $1^3 + 2^3 + \cdots + n^3 = \dfrac{n^2(n+1)^2}{4}$

24. Prove each result by induction, for all $n \geq 1$.

(a) $1 + 4 + 7 + \cdots + (3n-2) = \dfrac{n(3n-1)}{2}$

(b) $1 + 2 + 4 + \cdots + 2^{n-1} = 2^n - 1$

(c) $1 \cdot 2 + 2 \cdot 3 + \cdots + n(n+1) = \dfrac{n(n+1)(n+2)}{3}$

(d) $\dfrac{1}{1 \cdot 2} + \dfrac{1}{2 \cdot 3} + \cdots + \dfrac{1}{n(n+1)} = \dfrac{n}{n+1}$

Divisibility

25. Prove each divisibility result by induction, for all $n \geq 1$.

(a) $3^n - 1$ is divisible by 2.

(b) $7^n - 3^n$ is divisible by 4.

(c) $n^3 + 2n$ is divisible by 3.

(d) $2n^3 + 3n^2 + n$ is divisible by 6.

Inequalities

26.

(a) Prove by induction that $2^n \geq n + 1$ for all $n \geq 0$.

(b) Explain why the base case here is $n = 0$ rather than $n = 1$.

27. Prove each inequality by induction. Check that the base case is the first value for which the inequality holds.

(a) $n! > 2^n$ for all integers $n \geq 4$

(b) $3^n > n^3$ for all integers $n \geq 4$

Recurrences

28. A sequence is defined by $u_1 = 2$ and $u_{n+1} = 2u_n - 1$ for $n \geq 1$.

(a) Find u_2, u_3 and u_4, and conjecture a formula for u_n.

(b) Prove your conjecture by induction.

29. A sequence is defined by $u_1 = 1$ and $u_{n+1} = 2u_n + n$ for $n \geq 1$. Prove by induction that $u_n = 3 \cdot 2^{n-1} - n - 1$ for all $n \geq 1$.

Concept check

30. A student claims that $n^2 + n + 1$ is even for every integer $n \geq 1$. The student shows that, if $k^2 + k + 1$ is even, then $(k+1)^2 + (k+1) + 1$ is even.

(a) Show that the student's inductive step is correct.

(b) Explain why the claim is nonetheless false, and state which part of an induction proof the student left out.

31. In an induction proof, a student writes: "*Assume that the result is true for all positive integers* n." Explain the error, and write the correct assumption.

Chapter 3 · Proofs

Reflections

TOK A proof reduces a claim to things already accepted. But those starting points, the axioms, are assumed rather than proved. In 1931 Kurt Gödel showed that any consistent system (one that never proves a contradiction) rich enough to describe arithmetic contains true statements it cannot prove. Mathematical certainty is real, then, but not absolute. It is certainty *relative* to a chosen foundation. Does that make mathematics discovered, or invented?

IM The Pythagoreans believed that any two lengths share a common unit, so that every ratio of lengths is a ratio of whole numbers. Their own theorem contradicted that belief: the diagonal of a unit square has length $\sqrt{2}$, and no such ratio gives it. Nobody knows how the discovery was made. The classical argument, which Aristotle alludes to, is the one used in §3.3, and it is a proof by contradiction: assume $\sqrt{2} = p/q$ in lowest terms, and an impossibility follows. The discovery is credited to Hippasus in the fifth century BCE, and later writers report that his school received it badly. The result stood regardless, and Greek mathematics had to widen its idea of number.

RW Proof is not confined to pure mathematics. In software engineering, *formal verification* proves that a program meets its specification exactly. No amount of testing can give that guarantee, so the method is used where failure would be very dangerous: aircraft control, medical devices, spacecraft. Cryptography uses proved properties of prime numbers. Its security also depends on problems, such as factorising very large numbers, that are believed to be hard but have not been proved hard.

EXT The four-colour theorem says any map can be coloured with four colours so that no two neighbouring regions match. In 1976 it became one of the first major theorems whose proof needed a computer. The machine checked nearly two thousand cases, far more than a person could. That raised a new question. If a proof is too long for anyone to read, but every step has been checked by machine, is it still a proof? What is a proof *for*: certainty, or understanding?

A gallery of conjectures

Many theorems began as conjectures. The seven below are a small selection, chosen because each is about whole numbers and can be stated with nothing beyond the arithmetic of this course. Those marked open have been neither proved nor disproved; the remaining two show what settling a conjecture looks like, in opposite directions. None is examinable, and far more exist than are listed here, so you are invited to search out others of your own.

Goldbach conjecture (1742): open. Every even number greater than 2 is the sum of two primes: $28 = 5 + 23$, $100 = 3 + 97$. Checked for every even number up to 4×10^{18}, and still unproved.

Twin prime conjecture: open. There are infinitely many primes p for which $p + 2$ is also prime, such as $11, 13$ and $101, 103$. In 2013 Yitang Zhang proved there are infinitely many prime pairs at most 70 million apart. Later work brought that down to 246. Closer, but 246 is not 2.

Collatz conjecture: open. Take any positive integer. Halve it if even; triple it and add 1 if odd. Repeat. Does every starting number reach 1? Try 27 on your GDC: it takes 111 steps and climbs to 9232 on the way. Checked for a vast number of starting values, proved for none.

Odd perfect numbers: open for ∼2300 years. A number is perfect if it equals the sum of its proper divisors, such as $6 = 1 + 2 + 3$ and $28 = 1 + 2 + 4 + 7 + 14$. Every perfect number ever found is even. Whether an odd one exists is among the oldest unsolved problems in mathematics. Any example must be larger than 10^{1500}.

Legendre's conjecture: open. Between any two consecutive squares n^2 and $(n + 1)^2$ there is always at least one prime. It is true in every case ever checked, and still unproved.

Fermat's Last Theorem (1637): proved, 1995. No positive integers satisfy $x^n + y^n = z^n$ for $n \geq 3$. Fermat wrote in a book's margin that he had a proof, but that the margin was too small to hold it. The first accepted proof came 358 years later, from Andrew Wiles, and it uses mathematics far beyond anything Fermat could have known. Conjectures can be settled, and this one was.

Pólya's conjecture (1919): disproved. For every $N > 1$, at least half of the numbers up to N have an odd number of prime factors, counted with repetition. Checked for millions of cases and believed for nearly forty years, it was disproved by Haselgrove in 1958; the smallest counterexample, $N = 906{,}150{,}257$, was found by Tanaka in 1980. This is the circle problem from the start of the chapter on a larger scale: millions of cases agreed, one did not, and the claim is false.

End-of-Chapter Problems
Chapter 3: Proofs

1. Prove that the product of any three consecutive integers is divisible by 6. [5]

2. Consider the cube of a binomial.

(a) Prove the identity $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$. [3]
(b) Hence prove that $(n+1)^3 - n^3 = 3n^2 + 3n + 1$ for every integer n. [3]

3. Disprove the claim "$n^2 + 1$ is never divisible by 5" by giving two counterexamples. [3]

4. Prove by contradiction that $\sqrt{3}$ is irrational. [6]

5. Let a be a non-zero rational number and b an irrational number.

(a) Prove by contradiction that ab is irrational. [4]
(b) Hence state, with a reason, whether $\dfrac{\sqrt{2}}{5}$ is rational or irrational. [2]

6.

(a) Prove by contradiction that $x^2 - 6x + 10 = 0$ has no real solution, without using the discriminant. [3]
(b) Prove by contradiction that if n is an integer and n^2 is a multiple of 3, then n is a multiple of 3. [3]

7. Prove that $\sqrt{2} + \sqrt{3}$ is irrational. You may assume that $\sqrt{6}$ is irrational. [6]

8. Prove by contradiction that $\sqrt[3]{5}$ is irrational. [6]

9. Prove by induction that $7^n - 1$ is divisible by 6 for all $n \geq 1$. [5]

10. Prove by induction that $6^n + 4$ is divisible by 5 for all $n \geq 1$. [5]

11. Prove by induction that $n! > 2^n$ for all $n \geq 4$, stating why the base case is $n = 4$. [5]

12. Prove by induction that $5^{2n} - 1$ is divisible by 24 for all $n \geq 1$. [5]

13. Prove by induction that $2^n > n$ for all $n \geq 1$. [5]

14. Prove by induction that $9^n - 1$ is divisible by 8 for all $n \geq 1$. [5]

15. Prove by induction that $n^3 + 2n$ is divisible by 3 for all $n \geq 1$. [5]

16. Prove by induction that $\sum_{r=1}^{n} r^2 = \frac{n(n+1)(2n+1)}{6}$ for all $n \geq 1$. [6]

17. Prove by induction that $\sum_{r=1}^{n} r^3 = \frac{n^2(n+1)^2}{4}$ for all $n \geq 1$. [6]

18. Prove by induction that $\sum_{r=1}^{n} r(r+1) = \frac{n(n+1)(n+2)}{3}$ for all $n \geq 1$. [6]

19. Prove by induction that $\sum_{r=1}^{n} (3r - 2) = \frac{n(3n-1)}{2}$ for all $n \geq 1$. [6]

20. Prove by induction that $\sum_{r=1}^{n} \frac{1}{r(r+1)} = \frac{n}{n+1}$ for all $n \geq 1$. [6]

21. A student claims that $n^2 - n + 11$ is prime for every positive integer n.

(a) Verify the claim for $n = 1$ and $n = 2$. [2]

(b) Disprove the claim with a counterexample. [3]

22. Consider parity of squares.

(a) Prove that if n is odd then n^2 is odd. [2]

(b) Hence prove by contradiction that if n^2 is even then n is even. [3]

23. Prove that $1 + 3 + 5 + \cdots + (2n - 1) = n^2$ in two ways:

(a) directly, using the sum of an arithmetic sequence; [3]

(b) by induction. [3]

24. Prove that $\log_2 3$ is irrational. [6]

25. Prove by induction that $\sum_{r=1}^{n} r \cdot r! = (n+1)! - 1$ for all $n \geq 1$. [6]

26.

(a) Prove by induction that $\sum_{r=1}^{n} r = \frac{n(n+1)}{2}$ for all $n \geq 1$. [5]

(b) You may assume that $\sum_{r=1}^{n} r^2 = \frac{n(n+1)(2n+1)}{6}$. Find $\sum_{r=1}^{n} r(r-1)$, giving your answer in fully factorised form. [4]

(c) Hence evaluate $\sum_{r=1}^{20} r(r-1)$. [3]

(d) Show that $\sum_{r=1}^{n} r(r-1) = 2\,{}^{n+1}C_3$, and explain why this means the sum is always even. [3]

27.

(a) Prove that if n^2 is a multiple of 3, then n is a multiple of 3. [3]

(b) Hence prove by contradiction that $\sqrt{3}$ is irrational. [5]

(c) Explain precisely where the argument in part (b) breaks down if you attempt to use it to prove that $\sqrt{9}$ is irrational. [3]

(d) A student writes: *"$\sqrt{3}$ is irrational, so $\sqrt{3} \times \sqrt{3}$ must be irrational too."* Explain the error, and state what it shows about products of irrational numbers. [3]

28. A sequence is defined by $u_1 = 1$ and $u_{n+1} = \frac{u_n}{1+u_n}$ for $n \geq 1$.

(a) Find u_2, u_3 and u_4. [3]

(b) Write down a conjecture for u_n in terms of n. [1]

(c) Prove your conjecture by induction. [5]

(d) A student's whole proof reads: *"Assume the statement is true for $n = k$. Then it is true for $n = k+1$. Hence it is true for all n."* State what is missing, and explain why the argument establishes nothing without it. [3]

(e) Give a statement $P(n)$ for which $P(k) \Rightarrow P(k+1)$ is true for every k, but $P(n)$ is false for every n. Justify both claims. [3]

The remaining questions use the proof techniques of this chapter on material introduced later in the book. They appear in the order those topics are covered, so return to them as you go.

29. Let n and r be integers with $1 \leq r \leq n$. Starting from ${}^{n}C_r = \frac{n!}{r!\,(n-r)!}$, prove Pascal's identity

$$ {}^{n}C_r + {}^{n}C_{r-1} = {}^{n+1}C_r. $$

[5]

30. Using Pascal's identity, prove by induction that for every integer $n \geq 1$,

$$ (a+b)^n = \sum_{r=0}^{n} {}^{n}C_r a^{n-r} b^r. $$

[7]

31. You may assume the compound-angle formulae for $\cos(A+B)$ and $\sin(A+B)$. Prove by induction that

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

for every integer $n \geq 1$. [6]

32. You may assume that $\dfrac{d}{dx}(x) = 1$ and the product rule. Prove by induction that

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

for every integer $n \geq 1$. [6]

33. You may assume $\displaystyle\lim_{h\to 0}\frac{\sin h}{h} = 1$ and $\displaystyle\lim_{h\to 0}\frac{\cos h - 1}{h} = 0$. Working from the definition

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

prove that the derivative of $\sin x$ is $\cos x$. [6]

34. Let $I_n = \displaystyle\int_0^{\pi/2} \sin^n x\,dx$ for integers $n \geq 0$.

(a) Use integration by parts to prove that $I_n = \dfrac{n-1}{n} I_{n-2}$ for every integer $n \geq 2$. [5]

(b) Write down I_0 and I_1, and hence find I_4 and I_5. [4]

Challenge Questions

35. Counting the regions of a circle. This investigation completes the problem from the start of the chapter. Throughout, n points are placed on a circle so that no three chords meet at a single interior point.

(a) By drawing, verify that $n = 1, 2, 3, 4, 5$ give $1, 2, 4, 8, 16$ regions. [2]

(b) Explain why each interior intersection point corresponds to exactly one choice of 4 of the n points, and deduce that there are ${}^nC_4$ interior points. [3]

(c) The chords and arcs form a network. Show that it has $V = n + {}^nC_4$ vertices and $E = n + {}^nC_2 + 2\,{}^nC_4$ edges. *(Hint: each interior point adds one extra segment to each of the two chords through it)* [4]

(d) Euler's formula for a connected planar network states $V - E + F = 2$, where F counts all faces including the region outside the circle. Use it to prove that the number of regions inside the circle is ${}^nC_4 + {}^nC_2 + 1$. [4]

(e) Hence find the number of regions for $n = 6$ and $n = 10$. [2]

(f) If the n points are equally spaced, the count for $n = 6$ is 30 rather than 31. Explain what changes. [3]

(g) Suggest two ways in which this investigation could be extended, for example by changing where the points are placed or by moving to three dimensions. For one of them, state a conjecture and describe how you would test it. [3]

36. Fermat's conjecture. In 1640 Fermat studied the numbers $F_n = 2^{2^n} + 1$, for integers $n \geq 0$, and conjectured that all of them are prime.

(a) Show that F_0, F_1, F_2, F_3 and F_4 are prime. [3]

(b) Let m be a positive integer, and suppose m has an odd factor $d > 1$, so $m = de$. Using the factorisation

$$x^d + 1 = (x + 1)\left(x^{d-1} - x^{d-2} + \cdots - x + 1\right) \quad (d \text{ odd}),$$

with $x = 2^e$, prove that $2^m + 1$ is not prime. [5]

(c) Let m be a positive integer. Deduce that if $2^m + 1$ is prime, then m must be a power of 2. [2]

(d) Euler found in 1732 that 641 divides $F_5 = 2^{32} + 1$. This part proves it without dividing. Let $x = 5 \cdot 2^7$, and note that $641 = x + 1 = 5^4 + 2^4$. Show that 641 divides $x^4 - 1$, and that $5^4 \cdot 2^{28} + 2^{32} = 641 \cdot 2^{28}$. Deduce that 641 divides $2^{32} + 1$, and state what this shows about Fermat's conjecture. [5]

(e) Euler also proved that, for $n \geq 2$, every prime factor of F_n has the form $k \cdot 2^{n+2} + 1$ for some positive integer k. Check this for the factor 641 of F_5, and explain how the result shortens a search for a factor of F_6. Suggest a direction in which Fermat numbers could be investigated further. [3]

37. A proof that names no example. Consider the number $\sqrt{2}^{\sqrt{2}}$.

(a) Suppose $\sqrt{2}^{\sqrt{2}}$ is rational. Explain why this immediately produces an irrational number raised to an irrational power that is rational. [2]

(b) Suppose instead $\sqrt{2}^{\sqrt{2}}$ is irrational. Show that $\left(\sqrt{2}^{\sqrt{2}}\right)^{\sqrt{2}} = 2$, and explain why this also produces such a pair. [4]

(c) Conclude that there exist irrational a and b with a^b rational. [2]

(d) Given that $\log_2 3$ is irrational (Q24), show that $2\log_2 3$ is irrational and that $\sqrt{2}^{\,2\log_2 3} = 3$. Explain why this gives a second proof of the result in part (c). [3]

(e) The argument in parts (a) to (c) proves that such a pair exists without telling you which case is the true one. Compare it with the proof in part (d), and comment on what each proof gives you. [3]

(f) Investigate which positive rational numbers r can be written as a^b with a and b both irrational and $a > 0$. Justify your answer for at least one value of r that needs a different method from part (d), and suggest how the question could be taken further. [3]

38. Cassini's identity and a missing square. The Fibonacci numbers are defined by $f_1 = f_2 = 1$ and $f_{n+2} = f_{n+1} + f_n$ for $n \geq 1$.

(a) Compute $f_{n-1}f_{n+1} - f_n^2$ for $n = 2, 3, \ldots, 8$, and make a conjecture. [2]

(b) Prove by induction that $f_{n-1}f_{n+1} - f_n^2 = (-1)^n$ for all integers $n \geq 2$. [5]

(c) An 8×8 square is cut into four pieces: two right-angled triangles with perpendicular sides 3 and 8, and two right-angled trapezia with parallel sides 3 and 5 and perpendicular height 5. The pieces are rearranged to appear to fill a 5×13 rectangle. Use part (b) to explain why the two areas differ by exactly 1, and state the general version of the puzzle for an $f_n \times f_n$ square, saying when the rectangle appears larger and when smaller. [3]

(d) In the rectangle the long edges of the pieces should lie along a diagonal. Find the gradient of the hypotenuse of a triangle, the gradient of the slanted edge of a trapezium, and the gradient of the diagonal of the rectangle. Hence explain where the extra unit of area is. [4]

(e) The same puzzle made from a 21×21 square uses pieces whose slanted edges have gradients $\frac{8}{21}$ and $\frac{5}{13}$. Find the difference between these gradients, compare it with part (d), and explain why this version is harder to detect by eye. [2]

(f) Cassini's identity is the case $r = 1$ of Catalan's identity, $f_n^2 - f_{n-r}f_{n+r} = (-1)^{n-r}f_r^2$ for $1 \leq r < n$. Verify Catalan's identity for $n = 7$, $r = 3$, and suggest a direction in which this investigation could be extended. [3]

39. When does 2^n overtake n^3? Exponential growth eventually beats any power, but "eventually" can take a while. This investigation finds exactly when, and proves it.

(a) Tabulate 2^n and n^3 for $n = 1, 2, \ldots, 12$. State the values of n in this range for which $2^n > n^3$, and conjecture the least integer N such that $2^n > n^3$ for all $n \geq N$. [2]

(b) Show that $\left(1 + \dfrac{1}{k}\right)^3 < 2$ for every integer $k \geq 4$. [2]

(c) Prove by induction that $2^n > n^3$ for all integers $n \geq 10$. [5]

(d) The inductive step in part (c) is valid for every $k \geq 4$, and $2^1 > 1^3$ is true. Explain why neither fact allows the induction to start before $n = 10$. [2]

(e) Find the least integer N such that $2^n > n^4$ for all $n \geq N$, and prove your answer. [4]

(f) Let $N(k)$ be the least integer such that $2^n > n^k$ for all $n \geq N(k)$. Using your results and the notes' $N(2) = 5$, find $N(k)$ for $k = 5, 6, 7$ by calculation, and conjecture how $N(k)$ grows with k. Suggest how the conjecture could be tested or explained. [3]

Chapter 4

Lines & Quadratics

4

SYLLABUS 1.16 2.1–2.4 2.6 2.7 2.12

On 17 January 1995 an earthquake struck near Kobe, in western Japan. The Akashi Kaikyō Bridge was under construction. Its two towers already stood in the strait, but the roadway between them did not yet exist. The earthquake moved the towers apart.

The engineers did not begin the work again. They measured the new gap and changed the design to match it. The bridge opened in 1998 with a main span of 1991 metres, one metre longer than the 1990 originally planned.

The engineers could do this because the shape of the bridge follows from a few measurements through equations. The main cables of a suspension bridge carry a level roadway of even weight, and they hang in the curve called a *parabola*, the same curve as a thrown ball. If one value in the equations changes, every other measurement follows from it.

That is the value of a **function**. A function takes an input and returns exactly one output. A whole structure therefore becomes a single rule, and a rule can be graphed, studied and recalculated. This chapter builds two functions that later chapters use often: lines, which change at a steady rate, and quadratics, which give the cable its curve.

By the end of this chapter you will give the domain, range and inverse of a function, and find the equation of a line from the information you are given. You will solve systems of linear equations, write a quadratic in its factorised, vertex and expanded forms, use the discriminant to count the roots of an equation, find the sum and product of the roots without solving, and solve quadratic inequalities.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

4.1 Functions

A function is a rule you can rely on. Give it an input and it returns one definite output, and it returns the same output every time. That is what makes it possible to reason about.

DEF A **function** f assigns exactly one output $f(x)$ to each input x. The set of inputs it accepts is the **domain**; the set of outputs that actually occur is the **range**.

The phrase "exactly one" is where the restriction lies. A rule that could return two different outputs for the same input is not a function. On a graph this becomes the **vertical line test**: a curve represents a function only if no vertical line crosses it more than once.

The letter f carries no meaning of its own. A function is named for what it measures, and the letter inside the bracket for what it depends on. A height that changes with time is written $h(t)$, a velocity $v(t)$, a cost that depends on how many items you make $C(n)$. Read $C(n)$ as "the cost of making n items", not as C multiplied by n. Examination questions name their functions this way, so read the notation as a sentence rather than as algebra.

That naming points at what functions are usually for. A **mathematical model** is a function chosen to describe something real: the letters stand for measured quantities, and the rule claims that one determines the other. Modelling adds a question to every problem that pure algebra never asks, which is whether the answer means anything.

Two habits follow from that. The domain is set by the situation, not by the algebra alone: $h(t) = 20t - 5t^2$ is defined for every real t, but a stone in flight only exists from the moment it is thrown until it lands, so the model holds on $0 \le t \le 4$ and nowhere else. And a value that solves the equation can still be rejected, because a negative length or a negative time is arithmetic the situation does not allow. Every model is an approximation, so ask what it leaves out before trusting it far beyond the data it came from.

4.1.1 Reading the domain

A function need not accept every number. Three operations put limits on what you may substitute. Learn the three and you can read the domain of every function in this chapter.

Where x appears	Condition	Reason
In a denominator	$\neq 0$	you cannot divide by zero
Inside an even root	≥ 0	a negative number has no real square root
Inside a logarithm	> 0	zero and negative numbers have no logarithm

Note the difference between the last two. A square root allows zero, because $\sqrt{0} = 0$. A logarithm does not, because no power of the base gives zero. That single distinction, $\geq$ against $>$, is a common slip in this topic.

Function	Condition	Domain
$\sqrt{x+3}$	$x+3 \geq 0$	$x \geq -3$
$\sqrt{5-x}$	$5-x \geq 0$	$x \leq 5$
$\dfrac{1}{x-2}$	$x-2 \neq 0$	$x \neq 2$
$\ln(x-4)$	$x-4 > 0$	$x > 4$
$\dfrac{1}{\sqrt{x-1}}$	$x-1 > 0$, since it is both under a root and in a denominator	$x > 1$

When a function contains more than one of these, every condition must hold at once. Write them all down, then take the values that satisfy all of them.

Worked through. For $f(x) = \dfrac{\sqrt{x+4}}{x-3}$ there are two restrictions. The square root needs $x+4 \geq 0$, so $x \geq -4$. The denominator needs $x-3 \neq 0$, so $x \neq 3$. Both must hold, so the domain is

$$x \geq -4, \quad x \neq 3.$$

Notice that $x = 3$ satisfies the first condition but fails the second, so it must still be removed.

EXAMPLE 4.1

State the domain of $g(x) = \ln(9 - x^2)$.

A logarithm needs a positive argument, so $9 - x^2 > 0$, that is $x^2 < 9$. This is a quadratic inequality, solved in §4.7: the parabola $x^2 - 9$ lies below the axis between its roots. So

$$-3 < x < 3.$$

Both ends are excluded, because at $x = \pm 3$ the argument is zero and $\ln 0$ does not exist.

CAUTION Solve the condition, do not guess it. For $\sqrt{5-x}$ the domain is $x \leq 5$, not $x \geq 5$. Write the inequality down and rearrange it every time, however obvious it looks.

4.1.2 Reading the range

The range is the set of outputs that actually occur. There is no single rule, but most range arguments use one of four facts, each saying that some expression can never be negative. A square is never negative, so $x^2 \geq 0$; a modulus is never negative, so $|x| \geq 0$; the square root symbol means the non-negative root, so $\sqrt{x} \geq 0$; and a positive base raised to any power stays positive, so $a^x > 0$.

To use them, build the function up from the part you know. Ask what that part can be, then follow the operations one at a time.

Function	Reasoning	Range
$y = x^2 + 1$	$x^2 \geq 0$, then add 1	$y \geq 1$
$y = \|x\| - 3$	$\|x\| \geq 0$, then subtract 3	$y \geq -3$
$y = \sqrt{x} + 2$	$\sqrt{x} \geq 0$, then add 2	$y \geq 2$
$y = 3 - \|x + 1\|$	$\|x + 1\| \geq 0$, so subtracting it from 3 gives at most 3	$y \leq 3$
$y = -(x - 2)^2 + 5$	$(x - 2)^2 \geq 0$, and the minus sign turns the inequality round	$y \leq 5$

CAUTION A minus sign in front reverses the direction. Since $(x - 2)^2 \geq 0$, it follows that $-(x - 2)^2 \leq 0$, so $-(x - 2)^2 + 5 \leq 5$. Multiplying an inequality by a negative number always flips the sign.

For a quadratic, complete the square first. The vertex form shows the smallest or largest output immediately.

EXAMPLE 4.2

State the range of $f(x) = x^2 - 12x + 40$.

Complete the square:

$$x^2 - 12x + 40 = (x - 6)^2 - 36 + 40 = (x - 6)^2 + 4.$$

Now use the first fact. Since $(x - 6)^2 \geq 0$, adding 4 gives $f(x) \geq 4$. The range is $y \geq 4$.

The smallest value, 4, occurs at $x = 6$, which is the vertex.

EXAMPLE 4.3

State the domain and range of $g(x) = \sqrt{x-2}$.
Domain. The square root needs $x - 2 \geq 0$, so $x \geq 2$.
Range. As x runs from 2 upward, $x - 2$ runs from 0 upward, and so does its square root. So $g(x) \geq 0$.
The domain comes from the inputs the rule permits. The range comes from the outputs those inputs give. Always ask both questions.

TIP On Paper 2, check a range on your GDC. Graph the function and look at the lowest and highest points the curve reaches. The algebra is still expected in the working, but the graph catches a sign error in seconds.

4.1.3 Reading a graph

Most of what you need from a function can be read straight from its graph. Five features matter, and each has its own name.

Feature	What it is	How to find it
x-intercepts (also called zeros, or roots)	the points where the curve meets the x-axis	set $y = 0$ and solve
y-intercept	the point where the curve meets the y-axis	set $x = 0$
Turning points	the local maximum and minimum points, where the curve changes direction	GDC, or the vertex form for a quadratic
Asymptote	a line that the curve gets closer and closer to as x or y becomes very large	look for values the function cannot take
Intersection	a point that lies on both of two graphs	solve the two equations together

Your GDC finds four of these directly. The graph menu has zero, min/max and intersect tools, and evaluating the function at $x = 0$ gives the y-intercept. An asymptote comes from the equation, not from a GDC tool. On Paper 2 these are the intended method. Graph the function, read the feature, and write down what you asked the calculator to find. Many equations have no algebraic solution. For those, the GDC is the method to use, not just a faster one.

EXAMPLE 4.4

Use a GDC to find the y-intercept, the zeros and the turning points of $f(x) = x^3 - 2x^2 - 5x + 4$. Give coordinates to three significant figures where they are not exact.

The y-intercept needs no calculator. Setting $x = 0$ leaves the constant term, so the curve meets the y-axis at $(0, 4)$.

Graph the function over a window wide enough to show every feature; $-3 \le x \le 4$ works here. The curve crosses the x-axis three times, so run the zero tool three times, once on each crossing:

$$(-1.86,\ 0), \qquad (0.678,\ 0), \qquad (3.18,\ 0).$$

The maximum and minimum tools return both coordinates at once:

$$\text{local maximum } (-0.786,\ 6.21), \qquad \text{local minimum } (2.12,\ -6.06).$$

The turning points check the zeros. The local maximum lies above the axis and the local minimum below it, so the curve must cross the axis three times, and three is what the zero tool found.

TIP Know the difference between the command terms: **sketch** means show the shape with key features labelled, not to scale; **draw** means an accurate, to-scale plot. A sketch of a parabola needs at least the vertex, the intercepts, and the correct opening direction.

4.1.4 Inverse functions

An **inverse function** reverses the rule. If f sends 3 to 7, then f^{-1} sends 7 back to 3. An inverse exists only when no two inputs give the same output. If two did, the way back would have two answers, and a function may return only one.

DEF The **inverse** f^{-1} of a function f undoes it: $f^{-1}(f(x)) = x$. It exists when f is **one-to-one**, meaning no two inputs share an output (the *horizontal* line test).

The inverse swaps input and output. Two things follow. Its graph is the mirror image of f in the line $y = x$, and its domain and range are those of f exchanged. To find a formula, write $y = f(x)$, swap x and y, then solve for y.

EXAMPLE 4.5

Find the inverse of $f(x) = 2x + 3$.

Write $y = 2x + 3$, then swap x and y:

$$x = 2y + 3.$$

Solving for y,

$$y = \frac{x-3}{2}, \qquad \text{so} \qquad f^{-1}(x) = \frac{x-3}{2}.$$

Check: f doubles and adds three; f^{-1} subtracts three and halves. Each step of f is undone, in reverse order.

When x appears twice, once on top and once underneath, you cannot undo the rule step by step. Swap the letters anyway, clear the fraction, and collect the y terms.

EXAMPLE 4.6

Find f^{-1} for $f(x) = \dfrac{4x+3}{x-2}$, and state its domain.

Write $y = \dfrac{4x+3}{x-2}$ and swap x and y:

$$x = \frac{4y+3}{y-2}.$$

Multiply both sides by $y-2$ to clear the fraction, then gather every term containing y on one side:

$$x(y-2) = 4y+3 \implies xy - 2x = 4y+3 \implies xy - 4y = 2x+3.$$

Both terms on the left carry a factor of y, so factorise and divide:

$$y(x-4) = 2x+3 \implies f^{-1}(x) = \frac{2x+3}{x-4}, \qquad x \neq 4.$$

The domain of f^{-1} is the range of f, and this confirms it: $f(x) = 4$ would need $4x+3 = 4x-8$, which no x satisfies, so f never outputs 4.

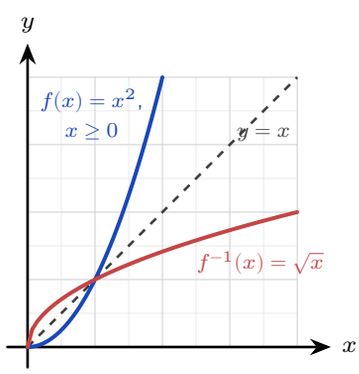


Figure 4.1 A function and its inverse are mirror images in the line $y = x$.

The inverse also gives another way to read an equation. Asking "for which x does $f(x) = 10$?" is the same as asking "what does f^{-1} send 10 to?"

KEY Solving $f(x) = 10$ is equivalent to evaluating $f^{-1}(10)$.

EXAMPLE 4.7

For $f(x) = 2x + 3$, solve $f(x) = 10$ in two ways.

Directly: $2x + 3 = 10$ gives $2x = 7$, so $x = \frac{7}{2}$.

Using the inverse found above, $f^{-1}(x) = \dfrac{x - 3}{2}$, so

$$f^{-1}(10) = \frac{10 - 3}{2} = \frac{7}{2}.$$

The two routes must agree, because they ask the same question. When a question gives you f^{-1} already, the second route is the shorter one.

CAUTION The notation f^{-1} means the inverse, *not* the reciprocal. $f^{-1}(x)$ is not $\dfrac{1}{f(x)}$.

YOUR TURN 4.1 Functions

What a function is

1. For $f(x) = x^2 - 3x$, find each value.

(a) $f(2)$ (b) $f(-1)$

(c) $f(0)$ (d) the values of x with $f(x) = 0$

2. Decide whether each rule defines y as a function of x, and justify your answer.

(a) $y^2 = x$ (b) $y = x^2$

Domain

3. State the domain of each function.

(a) $\sqrt{2x-6}$ (b) $\sqrt{7-x}$

(c) $\dfrac{1}{x+5}$ (d) $\dfrac{1}{x^2-4x}$

(e) $\ln(5-x)$ (f) $\ln(2x+1)$

(g) $\dfrac{1}{\sqrt{x+2}}$ (h) $\sqrt{x^2+1}$

4. State the domain of each function. Each has more than one restriction.

(a) $\dfrac{\sqrt{x+6}}{x-2}$ (b) $\dfrac{\sqrt{x}}{x-9}$

5. A ball is thrown upwards from ground level. Its height after t seconds is $h(t) = 15t - 5t^2$ metres, until it lands.

(a) Find the time at which the ball lands.

(b) State the domain of h in this model.

(c) The value $h(4)$ can be calculated. Explain why it has no meaning in this model.

Range

6. State the range of each function on the domain given.

(a) $x^2 - 4, \quad x \in \mathbb{R}$ (b) $2|x| + 1, \quad x \in \mathbb{R}$

(c) $5 - \sqrt{x}, \quad x \geq 0$ (d) $\dfrac{1}{x^2+1}, \quad x \in \mathbb{R}$

7. Consider $f(x) = \sqrt{2-x}$.

(a) State the domain of f.

(b) State the range of f.

(c) Compare the domain of f with the domain of $g(x) = \sqrt{x-2}$.

Reading a graph

8. GDC Consider $f(x) = x^3 - x^2 - 4x + 2$. Give coordinates to three significant figures where they are not exact.

(a) Write down the y-intercept of the graph of f.

(b) Use a GDC to find the zeros of f.

(c) Use a GDC to find the coordinates of the local maximum point and the local minimum point.

Inverse functions

9. Find the inverse of each function.

(a) $f(x) = 3x - 6$

(b) $f(x) = \dfrac{x+1}{2}$

(c) $f(x) = x^3 + 1$

(d) $f(x) = \dfrac{1}{x+2}$

(e) $f(x) = \dfrac{2x+1}{x-3}$

(f) $f(x) = \dfrac{x-4}{x+2}$

10. Consider $f(x) = \dfrac{3x+2}{x-1}$.

(a) State the domain of f.

(b) Find $f^{-1}(x)$.

(c) State the domain of f^{-1}.

(d) Verify that $f^{-1}(f(2)) = 2$.

11. Use an inverse function to solve each equation.

(a) For $f(x) = 5x - 3$, find $f^{-1}(x)$, and hence solve $f(x) = 17$ by evaluating $f^{-1}(17)$.

(b) A function g has inverse $g^{-1}(x) = \dfrac{2x+1}{x-3}$, $x \neq 3$. Solve $g(x) = 5$ without finding $g(x)$.

12. The function $f(x) = x^2$ has no inverse over all real numbers. State a restricted domain on which an inverse does exist, and give that inverse.

Concept check

13. A student writes: "$f(x) = 2x + 3$, *so* $f^{-1}(x) = \dfrac{1}{2x+3}$." Identify the error and find the correct $f^{-1}(x)$.

4.2 Linear Functions

The simplest function that actually changes is the line: its output rises at a constant rate. That rate is the **gradient**, the rise divided by the run.

KEY · IB The gradient of the line through (x_1, y_1) and (x_2, y_2) is

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

A gradient and one point on the line are enough to fix the line completely. From that fact come three ways to write its equation. All three describe the same line.

KEY · IB A straight line can be written

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

Here m is the gradient and c the y-intercept.

Use the first form when you know the gradient and the y-intercept. Use the second when you know the gradient and any point. The third is the general form, with whole-number coefficients.

Two lines are **parallel** when they rise at the same rate. They are **perpendicular** when you get one gradient from the other by turning the fraction upside down and changing the sign.

KEY For two lines with gradients m_1 and m_2:

$$m_1 = m_2 \ \text{(parallel)}, \qquad m_1 m_2 = -1 \ \text{(perpendicular)}.$$

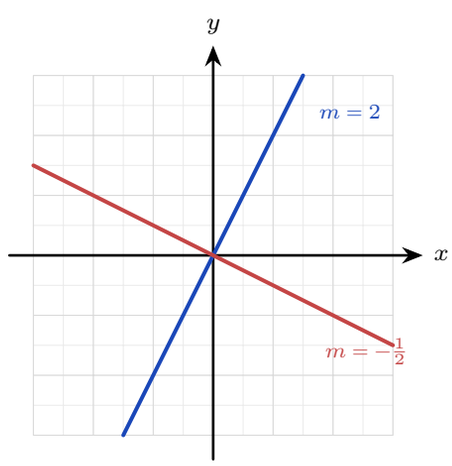


Figure 4.2 Perpendicular gradients: turn the fraction upside down and change the sign.

Worked through. To find the equation of the line through $(1, 2)$ and $(4, 8)$, first find the gradient:

$$m = \frac{8 - 2}{4 - 1} = \frac{6}{3} = 2.$$

Then use the point $(1, 2)$ in point–gradient form:

$$y - 2 = 2(x - 1) \implies y = 2x.$$

EXAMPLE 4.8

Find the equation of the line through $(0, 4)$ perpendicular to $y = 3x - 1$.

The given line has gradient 3, so the perpendicular gradient is $-\frac{1}{3}$. Through $(0,4)$,

$$y = -\tfrac{1}{3}x + 4.$$

Two points fix more than a gradient. They also have a distance between them and a point halfway along, and both come from the coordinates directly. The distance is Pythagoras applied to the horizontal and vertical gaps; the midpoint is the average of the coordinates.

KEY · IB For the points (x_1, y_1) and (x_2, y_2), the distance between them and the midpoint of the segment joining them are

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \qquad M = \left(\frac{x_1 + x_2}{2},\ \frac{y_1 + y_2}{2}\right).$$

The order of the points does not matter in either formula: the differences are squared, and the sums are symmetric. Both formulas extend to three dimensions by adding a z-term to each; see Ch. 7.

Together with a perpendicular gradient these give the **perpendicular bisector** of a segment, the line at right angles to it through its midpoint. Every point on the perpendicular bisector of AB is the same distance from A as from B, so it is also the set of all such points.

EXAMPLE 4.9

The points $A(2,-1)$ and $B(6,7)$ are given. Find the length of AB, the midpoint of AB, and the equation of the perpendicular bisector of AB.

The horizontal gap is 4 and the vertical gap is 8, so

$$AB = \sqrt{(2-6)^2 + (-1-7)^2} = \sqrt{16 + 64} = \sqrt{80} = 4\sqrt{5}.$$

The midpoint is

$$M = \left(\frac{2+6}{2},\ \frac{-1+7}{2}\right) = (4, 3).$$

The gradient of AB is $\dfrac{7-(-1)}{6-2} = 2$, so the bisector has gradient $-\frac{1}{2}$. Through $M(4,3)$,

$$y - 3 = -\tfrac{1}{2}(x - 4) \implies y = -\tfrac{1}{2}x + 5.$$

Every point of this line is the same distance from A as from B. Take $(0,5)$ on it: the distance to A is $\sqrt{4+36} = \sqrt{40}$, and to B it is $\sqrt{36+4} = \sqrt{40}$.

YOUR TURN 4.2 Linear Functions

Gradient and equation of a line

14. Find the gradient of the line through each pair of points.

(a) $(1, 3)$ and $(5, 11)$
(b) $(2, -1)$ and $(5, 8)$
(c) $(-1, 4)$ and $(3, -4)$
(d) $(0, 5)$ and $(2, 5)$

15. Find the equation of each line, giving your answer as $y = mx + c$.

(a) gradient 2, through $(0, -3)$
(b) through $(2, 1)$ and $(4, 7)$
(c) gradient $-\frac{1}{2}$, through $(4, 0)$
(d) through $(2, -1)$, perpendicular to $y = -2x + 3$

16. Write each equation as $y = mx + c$ and state the gradient.

(a) $2x + 3y = 6$
(b) $4x - 2y = 8$

17. Find the equation of each line in the form $ax + by + d = 0$, where a, b and d are integers.

(a) gradient $\frac{2}{3}$, through $\left(\frac{1}{2}, -1\right)$
(b) through $(-1, 2)$ and $(3, -4)$

Parallel and perpendicular lines

18. Answer each question about parallel and perpendicular lines.

(a) Write down the gradient of a line perpendicular to $y = 4x + 1$.
(b) Write down the gradient of a line parallel to $3x + y = 7$.
(c) Are $y = 2x + 1$ and $y = 2x - 3$ parallel? Give a reason.
(d) Are $y = 3x - 1$ and $y = -\frac{1}{3}x + 2$ perpendicular? Give a reason.

Distance, midpoint and perpendicular bisector

19. For each pair of points, find the exact distance between them and the midpoint of the segment joining them.

(a) $(1, 2)$ and $(4, 6)$
(b) $(-3, 5)$ and $(1, -1)$

20. The points $A(-1, 4)$ and $B(5, 0)$ are given.

(a) Find the midpoint M of AB.
(b) Find the equation of the perpendicular bisector of AB.
(c) The point $P(k, 8)$ lies on the perpendicular bisector. Find k, and verify that $PA = PB$.

Concept check

21. A student writes: "*The line $2x + 3y = 6$ has gradient 2, so a line perpendicular to it has gradient $-\frac{1}{2}$.*" Identify the error and find the correct perpendicular gradient.

4.3 Systems of Linear Equations HL

Two lines in a plane usually cross at one point. Finding that point means finding the values of x and y that satisfy *both* equations at once, a **system** of equations. The standard tool is **elimination**: combine the equations to cancel one variable, solve for the other, then substitute back.

Worked through. To solve $2x + y = 7$ and $x - y = 2$, add the two equations, which cancels y:

$$(2x + y) + (x - y) = 7 + 2 \implies 3x = 9 \implies x = 3.$$

Substituting into $x - y = 2$ gives $3 - y = 2$, so $y = 1$. The lines cross at $(3, 1)$.

Not every system has a single answer, and the picture shows why. Two lines can do three things. They cross once, which gives a **unique solution**. They stay parallel and never meet, so there is **no solution**. Or they are the same line drawn twice, which gives **infinitely many solutions**.

CAUTION If two equations have the same gradient but different intercepts, such as $2x + y = 3$ and $4x + 2y = 10$, the lines are parallel and the system has *no* solution. Elimination then produces a false statement such as $0 = 4$.

The same idea extends to three equations in three unknowns, where each equation is now a plane in space. You eliminate one variable at a time to reduce a 3×3 system to a 2×2 one, or read the solution directly from a GDC.

EXAMPLE 4.10

Solve the system

$$x + y + z = 6, \qquad 2x - y + z = 3, \qquad x + 2y - z = 2.$$

Eliminate z in pairs. Subtracting the first equation from the second:

$$(2x - y + z) - (x + y + z) = 3 - 6 \implies x - 2y = -3.$$

Adding the first and third equations:

$$(x + 2y - z) + (x + y + z) = 2 + 6 \implies 2x + 3y = 8.$$

Now solve this 2×2 system. From $x - 2y = -3$ we get $x = 2y - 3$; substituting,

$$2(2y - 3) + 3y = 8 \implies 7y = 14 \implies y = 2,$$

so $x = 1$. Finally the first equation gives $1 + 2 + z = 6$, hence $z = 3$. The solution is

$(x, y, z) = (1, 2, 3)$.

Eliminating one variable at a time, working down the system, is called **row reduction**. It is what your GDC's simultaneous-equation solver does inside. Use the solver to check your work. You still need the algebra in an examination, and you need it when the solver returns only an error message.

4.3.1 The augmented matrix

In the elimination above, what changed was never the letters x, y, z, only the numbers in front of them. The letters only marked the columns, so you can stop writing them. For a new system, the coefficients and constants alone form a grid called the **augmented matrix**:

$$\begin{cases} x + y + z = 6 \\ 2x - y + z = 7 \\ x + 2y - z = 3 \end{cases} \longrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 2 & -1 & 1 & 7 \\ 1 & 2 & -1 & 3 \end{array}\right)$$

The vertical bar shows where the equals signs were. Nothing new is happening: the matrix records the same elimination in a shorter form, with fewer symbols to copy incorrectly.

Elimination becomes three **row operations**. None of them changes the solution set. Each one is reversible, and each does to a whole equation only what you were always allowed to do:

KEY The three elementary row operations:

$$R_i \leftrightarrow R_j \ \text{(swap two rows)}, \qquad R_i \to kR_i \ \text{(multiply a row by } k \neq 0\text{)},$$

$$R_i \to R_i + kR_j \ \text{(add a multiple of one row to another)}.$$

Swapping reorders the equations, scaling multiplies one through by a constant, and the third adds a multiple of one equation to another: the same three moves as elimination, written compactly.

The first non-zero entry of a row is its **pivot**. The goal is a step shape, with zeros below each pivot. The last row then contains one unknown, the row above it two, and so on. This shape is called **row echelon form**. To reach it, work one column at a time, using row operations to clear every entry below the pivot; a row swap that brings a 1 into the pivot position keeps the arithmetic simple. Once you reach it, work upwards from the last row, finding one unknown at a time. That is **back-substitution**.

EXAMPLE 4.11

Solve, by row reduction,

$$x + y + z = 6, \qquad 2x - y + z = 7, \qquad x + 2y - z = 3.$$

Work on the augmented matrix, clearing the first column below the pivot. $R_2 \to R_2 - 2R_1$ and $R_3 \to R_3 - R_1$:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 2 & -1 & 1 & 7 \\ 1 & 2 & -1 & 3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & -3 & -1 & -5 \\ 0 & 1 & -2 & -3 \end{array}\right)$$

A leading entry of 1 gives simpler arithmetic than -3, so swap $R_2 \leftrightarrow R_3$. Then clear below it with $R_3 \to R_3 + 3R_2$:

$$\to \left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & -2 & -3 \\ 0 & -3 & -1 & -5 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & -7 & -14 \end{array}\right)$$

The step shape is complete. Now read it from the bottom up. The last row says $-7z = -14$, so $z = 2$. The middle row says $y - 2z = -3$, so $y = 4 - 3 = 1$. The top row gives $x = 6 - y - z = 3$. The solution is $(x, y, z) = (3, 1, 2)$.

Every step here is the elimination you already know. The matrix keeps the working short and the columns lined up, and misaligned columns are a common cause of mistakes.

Once a system is in row echelon form, its rows determine the outcome. There are only three possible outcomes, and every 3×3 system gives exactly one of them.

Rows reduce to	What it means	The three planes
last row $(0\ 0\ a \mid b)$, $a \neq 0$	one unique solution	meet at a single point
any row $(0\ 0\ 0 \mid c)$, $c \neq 0$	no solution: the row says $0 = c$, which is false	have no common point; two are parallel, or the three meet in a triangular prism
$(0\ 0\ 0 \mid 0)$, and no row of the type above	infinitely many solutions: the row says $0 = 0$, which is true but tells us nothing	share a whole line, or a whole plane if two rows vanish

First look for a row $(0\ 0\ 0 \mid c)$ with $c \neq 0$. Any such row, in any position, means there is no solution. If there is none, count the rows that become all zeros. No zero row gives one solution. One zero row gives a line of solutions, described by one free variable. Two zero rows mean all three equations describe the same plane, and the solutions form that whole plane, described by two free variables.

4.3.2 When a system has no unique solution

The table above lists the three outcomes. When one row reduces to $0 = 0$, choose an unknown in the last non-zero row as the free variable and set it equal to a parameter, such as $z = t$. Back-substitute up the echelon form to write each other unknown in terms of t. The result, (x, y, z) in terms of t with $t \in \mathbb{R}$, is the **general solution**, and it describes the line where the planes meet. In the next example a constant k in the equations decides between the outcomes.

EXAMPLE 4.12

Consider the system

$$x + y + 2z = 2, \qquad 2x - y + z = 1, \qquad x + 4y + 5z = k.$$

Find the value of k for which the system has infinitely many solutions, and give the general solution. Show that for every other value of k there is no solution.

The question asks for infinitely many solutions, so expect a row that reduces to zeros. Eliminate x using the first equation. Second minus twice the first:

$$(2x - y + z) - 2(x + y + 2z) = 1 - 4 \implies -3y - 3z = -3 \implies y + z = 1.$$

Third minus the first:

$$(x + 4y + 5z) - (x + y + 2z) = k - 2 \implies 3y + 3z = k - 2 \implies y + z = \frac{k-2}{3}.$$

The same left side, $y + z$, has appeared twice. The two remaining equations are consistent only if

$$1 = \frac{k-2}{3} \implies k = 5.$$

If $k \neq 5$, subtracting them gives $0 = \frac{k-5}{3} \neq 0$, a false statement: the system is inconsistent, and no solution exists for any such k.

If $k = 5$, the third equation gives no new information, and the system reduces to two equations. Let the free variable be $z = t$. Then $y = 1 - t$, and the first equation gives

$$x = 2 - y - 2z = 2 - (1 - t) - 2t = 1 - t.$$

The general solution is

$$(x, y, z) = (1 - t,\ 1 - t,\ t), \qquad t \in \mathbb{R},$$

a whole line of solutions: the line where the three planes meet. Check with $t = 0$: the point $(1, 1, 0)$ satisfies all three original equations. So does $(0, 0, 1)$, from $t = 1$. One free parameter gives one line, so the geometry and the algebra give the same result.

CAUTION "No unique solution" covers two different cases, so always separate them. A row that reduces to $0 = 0$ gives infinitely many solutions. A row that reduces to $0 = c$, with $c \neq 0$, gives none. Examination questions can ask for both, and for the general solution in terms of a

parameter.

YOUR TURN 4.3 Systems of Linear Equations HL

Two equations in two unknowns

22. Solve each system by elimination.

(a) $x + y = 5, \quad x - y = 1$
(b) $3x - y = 5, \quad x + y = 3$
(c) $x + y = 4, \quad 2x + y = 7$
(d) $2x + 3y = 12, \quad x - y = 1$

23. Solve each system by substitution.

(a) $y = 2x, \quad x + y = 6$
(b) $y = x - 1, \quad 2x + y = 8$

24. State how many solutions each system has, and give a reason.

(a) $2x + y = 4, \quad 4x + 2y = 9$
(b) $y = x + 1, \quad y = x + 3$
(c) $x + y = 3, \quad 2x + 2y = 6$
(d) $x + y = 3, \quad x - y = 1$

Three equations in three unknowns

25. Solve each system of three equations.

(a) $x + y + z = 6, \quad x - y + z = 2, \quad x + y - z = 0$
(b) $x + y + z = 4, \quad 2x - y + z = 8, \quad x + 2y - z = -3$
(c) $x + y + z = 2, \quad 2x - y + 3z = -1, \quad x - 2y + z = -1$
(d) $2x + y - z = -4, \quad x + 3y + 2z = 3, \quad 3x - y + z = 9$

Row reduction and the number of solutions

26. For parts (a) to (c), write the augmented matrix, reduce it to row echelon form, and state the number of solutions.

(a) $x + y + z = 6, \quad x + 2y - z = 2, \quad 2x + 3y + z = 11$
(b) $x + y + z = 3, \quad 2x + 2y + 2z = 7, \quad x - y + z = 1$
(c) $x + y + z = 6, \quad x - y + z = 2, \quad 2x + 2z = 8$
(d) For $x + y + z = 3, \quad 2x + y + 3z = 7, \quad 3x + 2y + 4z = k$, find the value of k for which the system has infinitely many solutions, and state what happens for every other value of k.

Concept check

27. A student eliminates x from $x + 2y = 4$ and $2x + 4y = 8$, obtains $0 = 0$, and writes "*no solution*".

(a) Explain why this conclusion is wrong, and state the correct number of solutions.
(b) The second equation is changed to $2x + 4y = 10$. State what elimination now gives, and what that result means.

4.4 Quadratic Functions

A **quadratic** is a function whose highest power is x^2. Its graph is the **parabola**, the curve of Galileo's cannonball: a single smooth arch, symmetric about a vertical line through its turning point.

The same quadratic can be written in three forms, and each form shows a different feature immediately.

KEY A quadratic can be written in three forms:

$$y = ax^2 + bx + c, \qquad y = a(x-p)(x-q), \qquad y = a(x-h)^2 + k.$$

Expanded form shows the y-intercept c; **factorised** form shows the roots p and q; **vertex** form shows the turning point (h, k).

The sign of a sets the direction. If $a > 0$ the parabola opens upward and the turning point is a minimum. If $a < 0$ it opens downward and the turning point is a maximum. The curve is symmetric, and its mirror line passes through the vertex.

KEY · IB The graph of $f(x) = ax^2 + bx + c$ has axis of symmetry

$$x = -\frac{b}{2a}.$$

To move from expanded form to vertex form, **complete the square**. This means rewriting $x^2 + \frac{b}{a}x$ as a perfect square, then subtracting whatever that added. It is used throughout the course.

Worked through. To write $y = x^2 - 10x + 21$ in vertex form, take the $x^2 - 10x$ part and complete the square. Half of -10 is -5, and $(-5)^2 = 25$:

$$x^2 - 10x = (x-5)^2 - 25.$$

So

$$y = (x-5)^2 - 25 + 21 = (x-5)^2 - 4.$$

The vertex is $(5, -4)$ and the axis of symmetry is $x = 5$. Because $a = 1 > 0$, this vertex is the minimum point.

EXAMPLE 4.13

Sketch $y = x^2 - 2x - 3$, showing the vertex, axis of symmetry, and all intercepts.

Factorise for the roots: $x^2 - 2x - 3 = (x-3)(x+1)$, so the x-intercepts are $x = -1$ and $x = 3$.

The y-intercept is $c = -3$. Completing the square, $y = (x-1)^2 - 4$, so the vertex is $(1, -4)$ and the axis of symmetry is $x = 1$, which is exactly halfway between the two roots.

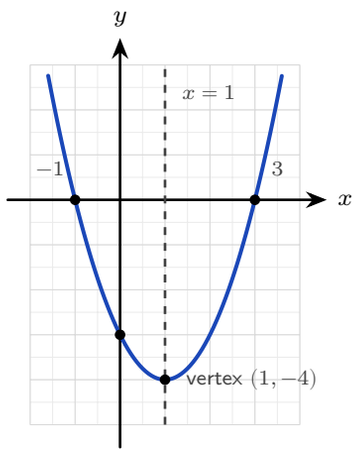


Figure 4.3 $y = x^2 - 2x - 3$, with its vertex, axis of symmetry and intercepts.

The three forms also work in reverse. Given points on a parabola, choose the form that already contains what you know, then use the remaining point to fix a. Two roots and one further point call for the factorised form.

EXAMPLE 4.14

A parabola passes through $(0, -8)$, $(-1, 0)$ and $(4, 0)$. Find its equation.
Two of the points lie on the x-axis, so they are the roots. The factorised form gives

$$y = a(x+1)(x-4),$$

with a still unknown: infinitely many parabolas share a pair of roots, and they differ only in how steeply they open. Substitute the third point, $(0, -8)$, to pin it down:

$$-8 = a(0+1)(0-4) = -4a \implies a = 2.$$

Expanding,

$$y = 2(x+1)(x-4) = 2x^2 - 6x - 8.$$

Check the three points: at $x = 0$, $y = -8$; at $x = -1$ and $x = 4$, $y = 0$. Had the curve passed through $(0, -4)$ instead, the same substitution would give $a = 1$, halving every y-coordinate while leaving the roots where they are.

YOUR TURN 4.4 Quadratic Functions

Completing the square

28. Write each quadratic in vertex form and state its vertex.

(a) $y = x^2 + 4x + 1$ (b) $y = x^2 - 8x + 10$

(c) $y = x^2 + 3x + 1$ (d) $y = x^2 + 2x + 5$

29. Write each quadratic in the form $a(x-h)^2 + k$. *(Hint: take out the factor a from the x^2 and x terms first)*

(a) $2x^2 + 12x + 5$ (b) $3x^2 - 6x + 1$

(c) $-x^2 + 6x - 4$ (d) $-2x^2 - 8x + 3$

Vertex, axis and range

30. State the vertex and the axis of symmetry of each quadratic.

(a) $y = (x-2)^2 + 5$ (b) $y = x^2 - 8x + 3$

(c) $y = -(x+1)^2 + 2$ (d) $y = 2(x-3)^2 - 1$

31. Find the range of each quadratic by completing the square. Each is defined for all real x.

(a) $x^2 - 6x + 11$ (b) $x^2 + 5x$

(c) $2x^2 - 4x + 7$ (d) $1 + 4x - 2x^2$

32. State whether each quadratic has a maximum or a minimum, give a reason, and find that maximum or minimum value.

(a) $y = -x^2 + 2x$ (b) $y = 3x^2 - x + 1$

(c) $y = x^2 - 6x + 5$ (d) $y = 5 + 4x - x^2$

Roots, intercepts and sketches

33. Consider $y = (x-1)(x+5)$.

(a) Write down its roots. (b) Find its y-intercept.

(c) Find its axis of symmetry. (d) Find its vertex.

34. Sketch each parabola, showing the vertex, the axis of symmetry and all intercepts.

(a) $y = -x^2 + 2x + 8$ (b) $y = 2x^2 - 4x - 6$

Finding the equation of a parabola

35. Find the equation of each parabola in the form $y = ax^2 + bx + c$.

(a) It has roots $x = -2$ and $x = 3$, and passes through $(0, 12)$.

(b) It passes through $(1, 0)$, $(5, 0)$ and $(3, -8)$.

(c) Its vertex is $(2, -3)$, and it passes through $(0, 5)$.

Concept check

36. A student states that the vertex of $y = (x+3)^2 - 5$ is $(3, -5)$. Identify the error and state the correct vertex.

37. A student completes the square as follows.

$$3x^2 - 12x + 5 = 3(x-2)^2 - 4 + 5 = 3(x-2)^2 + 1$$

Identify the error, correct the working, and state the vertex.

4.5 Quadratic Equations

Setting a quadratic equal to zero is the same as asking where its parabola crosses the x-axis. There are three ways to find those crossings, and all three give the same answers. Choosing the right method saves time in an examination.

Method	Use it when	Be careful with
Factorisation	the quadratic factorises with whole numbers	many quadratics do not factorise, and searching for factors takes time
Completing the square	you also want the vertex, or an exact surd answer	the arithmetic is harder when $a \neq 1$
The formula	always. It works for every quadratic	slower than factorising when factors are obvious

4.5.1 Solving by factorisation

If the quadratic factorises, you can read the roots immediately. The reason is the zero product rule: a product is zero only when one of its factors is zero.

Worked through. To solve $x^2 - 8x + 15 = 0$, find two numbers that multiply to 15 and add to -8. They are -3 and -5:

$$(x-3)(x-5) = 0.$$

A product is zero only when a factor is zero, so $x - 3 = 0$ or $x - 5 = 0$, giving $x = 3$ or $x = 5$.

CAUTION The zero product rule needs a zero on the right. From $(x-2)(x-3) = 2$ you may *not*

conclude $x - 2 = 2$. Expand, collect everything on one side, and factorise again.

4.5.2 Solving by completing the square

Completing the square works whether or not the quadratic factorises, and it gives exact answers in surd form. Rewrite the quadratic as a perfect square plus a constant, then take square roots of both sides.

EXAMPLE 4.15

Solve $x^2 + 4x - 12 = 0$ by completing the square.
Half of 4 is 2, and $2^2 = 4$, so $x^2 + 4x = (x + 2)^2 - 4$. The equation becomes

$$(x + 2)^2 - 4 - 12 = 0 \implies (x + 2)^2 = 16.$$

Take the square root of both sides, remembering both signs:

$$x + 2 = \pm 4 \implies x = 2 \text{ or } x = -6.$$

EXAMPLE 4.16

Solve $x^2 - 4x - 1 = 0$, giving exact answers.
Half of -4 is -2, and $(-2)^2 = 4$:

$$(x - 2)^2 - 4 - 1 = 0 \implies (x - 2)^2 = 5 \implies x - 2 = \pm\sqrt{5}.$$

So $x = 2 \pm \sqrt{5}$. This quadratic does not factorise with whole numbers, which is why the answer contains a surd.

TIP When $a \neq 1$, divide the whole equation by a first. For $2x^2 - 8x + 5 = 0$, divide by 2 to get $x^2 - 4x + 2.5 = 0$, then complete the square as usual.

4.5.3 Solving by formula

The formula works for every quadratic, and it is in the formula booklet.

KEY · IB The solutions of $ax^2 + bx + c = 0$, with $a \neq 0$, are

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

Substitute the three coefficients and the formula returns both roots at once.

The formula is not a rule to memorise without understanding. It is what **completing the square** gives when the coefficients are letters instead of numbers, so you can derive it

yourself from $ax^2 + bx + c = 0$ at any time.

Divide through by a, then complete the square on $x^2 + \frac{b}{a}x$:

$$x^2 + \frac{b}{a}x + \frac{c}{a} = 0 \implies \left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + \frac{c}{a} = 0.$$

Move the two constants to the right and write them over the common denominator $4a^2$:

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}.$$

Take the square root of both sides, which is where the $\pm$ enters, and then subtract $\frac{b}{2a}$:

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

The $\pm$ is the algebra recording what the picture already shows: the parabola is symmetric, with one root on each side of its axis. The expression under the root, $b^2 - 4ac$, decides how many roots there are, and that is the subject of the next section.

EXAMPLE 4.17

Solve $2x^2 - 5x - 12 = 0$.

Here $a = 2$, $b = -5$, $c = -12$. The formula gives

$$x = \frac{5 \pm \sqrt{25 + 96}}{4} = \frac{5 \pm 11}{4},$$

so $x = 4$ or $x = -\frac{3}{2}$.

YOUR TURN 4.5 Quadratic Equations

Factorisation

38. Solve each equation by factorisation.

(a) $x^2 - 5x + 6 = 0$

(b) $x^2 - x - 12 = 0$

(c) $x^2 + x - 20 = 0$

(d) $2x^2 - 5x - 3 = 0$

Rearranging to zero first

39. Rearrange each equation so that one side is 0, then solve it by factorisation.

(a) $x^2 = 7x$

(b) $x(x + 4) = 21$

(c) $(x - 2)(x - 3) = 2$

(d) $3x^2 = 2(x + 4)$

Completing the square

40. Solve each equation by completing the square, giving exact answers.

(a) $x^2 + 8x + 7 = 0$ (b) $x^2 - 2x - 5 = 0$

(c) $x^2 + 6x + 2 = 0$ (d) $2x^2 + 12x + 7 = 0$

The quadratic formula

41. Solve each equation with the quadratic formula, giving exact answers.

(a) $x^2 + 2x - 1 = 0$ (b) $x^2 + 6x + 4 = 0$

(c) $2x^2 - 3x - 2 = 0$ (d) $3x^2 + 2x - 1 = 0$

Choosing a method

42. For each equation, state which of the three methods you would choose, and give a reason. You do not need to solve the equations.

(a) $x^2 - 7x + 12 = 0$ (b) $x^2 + 4x - 2 = 0$

(c) $5x^2 - 3x - 1 = 0$ (d) $x^2 - 9 = 0$

Concept check

43. A student solves $x^2 = 5x$ by dividing both sides by x, and writes "$x = 5$". Identify the error and give the complete solution.

4.6 The Discriminant

The quadratic formula gives more information than the roots. Before you calculate a single root, the quantity under the square root sign already tells you how many real roots there are. That quantity has its own name.

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

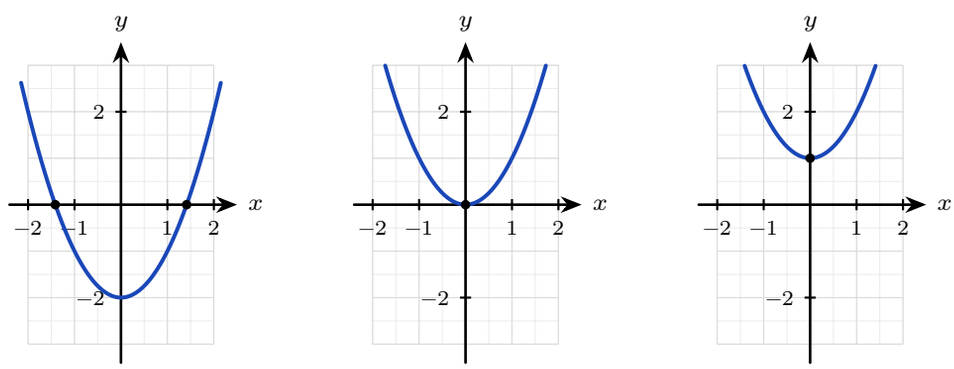
$$\Delta = b^2 - 4ac.$$

Its sign tells you how many real roots the equation has, without solving it.

The sign of Δ determines how many times the parabola meets the x-axis. The reason is that a negative number has no real square root, so $\pm\sqrt{\Delta}$ has no real value.

Discriminant	Real roots	Parabola and x-axis
$\Delta > 0$	two distinct	crosses twice
$\Delta = 0$	one repeated	touches once
$\Delta < 0$	none	does not meet it

The three cases are three positions of the same parabola relative to the x-axis.



(a) $\Delta > 0$: crosses twice (b) $\Delta = 0$: touches once (c) $\Delta < 0$: never meets

Figure 4.4 The same upward parabola in three positions: $y = x^2 - 2$, $y = x^2$ and $y = x^2 + 1$. Lowering it past the axis gives two roots, resting it on the axis gives one, and lifting it clear gives none. The discriminant measures which of the three has happened.

A downward parabola behaves the same way, with the picture turned over: the sign of a decides which way it opens, and Δ still counts the crossings.

Worked through. For $x^2 + kx + 9 = 0$ to have exactly one (repeated) root, the discriminant must be zero, $\Delta = 0$:

$$k^2 - 4(1)(9) = 0 \implies k^2 = 36 \implies k = \pm 6.$$

At these two values the parabola touches the x-axis at one point. Notice that the discriminant answers a question about the roots without finding them.

EXAMPLE 4.18

For the equation $3kx^2 + 2x + k = 0$, find the values of k that give two distinct real roots, two equal real roots, and no real roots.

Here the unknown constant appears in two places: $a = 3k$ and $c = k$, with $b = 2$. So

$$\Delta = b^2 - 4ac = 4 - 4(3k)(k) = 4 - 12k^2.$$

Now consider the three cases separately.

$$\Delta > 0: \ 4 - 12k^2 > 0 \implies k^2 < \tfrac{1}{3} \implies -\tfrac{1}{\sqrt{3}} < k < \tfrac{1}{\sqrt{3}}.$$

$$\Delta = 0: \ k^2 = \tfrac{1}{3} \implies k = \pm\tfrac{1}{\sqrt{3}}.$$

$$\Delta < 0: \ k^2 > \tfrac{1}{3} \implies k < -\tfrac{1}{\sqrt{3}} \text{ or } k > \tfrac{1}{\sqrt{3}}.$$

One value of k must be checked separately. If $k = 0$ the equation becomes $2x = 0$. That is linear, not quadratic, and it has the single root $x = 0$. So the answer for two distinct real roots is $-\frac{1}{\sqrt{3}} < k < \frac{1}{\sqrt{3}}$, *with* $k \neq 0$.

CAUTION Before using the discriminant, check that the equation really is quadratic. If the coefficient of x^2 contains the unknown constant, the value that makes it zero must be excluded and treated separately.

TIP When the equation is a quadratic, a question that says "two distinct roots", "equal roots", "a repeated root", "touches the x-axis", or "no real solutions" usually points to the discriminant. Recognising the wording tells you which method to use.

4.6.1 A line meeting a parabola

The discriminant does more than count where a parabola meets the x-axis. A point common to two graphs satisfies both equations, so setting the two expressions for y equal gives a quadratic whose real roots are the x-coordinates of the intersections. The discriminant counts them. When there is exactly one, the line touches the curve without crossing it, and is a **tangent**.

EXAMPLE 4.19

Substitute $y = mx + c$ into $y = x^2 - 2x + 5$ and read the three cases geometrically. Hence find the values of m for which $y = mx + 4$ is a tangent to that parabola.

At any intersection both equations hold, so

$$x^2 - 2x + 5 = mx + c \implies x^2 - (m+2)x + (5 - c) = 0.$$

Its coefficients are 1, $-(m+2)$ and $5 - c$, so

$$\Delta = (m+2)^2 - 4(1)(5-c) = (m+2)^2 + 4c - 20.$$

The constant term is $5 - c$, not c; the intercept c changes sign when it moves to the left-hand side. Each real root is one intersection point, so $\Delta > 0$ gives two intersections, $\Delta = 0$ gives one and the line is a tangent, and $\Delta < 0$ means the line misses the parabola entirely.

For $y = mx + 4$ take $c = 4$ and impose tangency:

$$\Delta = (m+2)^2 - 4 = 0 \implies m + 2 = \pm 2 \implies m = 0 \text{ or } m = -4.$$

Check $m = 0$: the quadratic is $x^2 - 2x + 1 = (x-1)^2$, a repeated root at $x = 1$, so the horizontal line $y = 4$ touches at the vertex $(1, 4)$. For $m = -4$ the quadratic is $x^2 + 2x + 1 = (x+1)^2$, touching at $(-1, 8)$.
A tangent shows up as a repeated root, so factorising the quadratic at each value of m both confirms the answer and gives the point of contact.

4.6.2 The sum and product of the roots HL

The roots of a quadratic are linked to its coefficients before you solve anything. If $ax^2 + bx + c$ has roots α and β, with $a \neq 0$ so that the expression really is quadratic, then it factorises as

$$a(x - \alpha)(x - \beta) = ax^2 - a(\alpha + \beta)x + a\,\alpha\beta.$$

Comparing the coefficient of x gives $b = -a(\alpha + \beta)$, and comparing the constant term gives $c = a\,\alpha\beta$. Dividing each by a, which the condition $a \neq 0$ permits, leaves the sum of the roots as $-\dfrac{b}{a}$ and their product as $\dfrac{c}{a}$. Nothing in that expansion depended on there being two brackets, so the same comparison settles a polynomial of any degree, and the result is stated properly in Ch. 6.

KEY HL If $ax^2 + bx + c = 0$ has roots α and β, then

$$\alpha + \beta = -\frac{b}{a}, \qquad \alpha\beta = \frac{c}{a}.$$

The booklet does not print this form. It prints the rule for any degree (booklet 2.12, Ch. 6), and this is its case $n = 2$.

These let you find expressions in the roots without solving the equation. Any expression that stays the same when α and β are swapped can be rewritten in terms of $\alpha + \beta$ and $\alpha\beta$; for example, $\alpha^2 + \beta^2 = (\alpha + \beta)^2 - 2\alpha\beta$.

EXAMPLE 4.20

The roots of $2x^2 - 6x + 3 = 0$ are α and β. Without solving the equation, find $\alpha^2 + \beta^2$.
Here $\alpha + \beta = -\dfrac{-6}{2} = 3$ and $\alpha\beta = \dfrac{3}{2}$. Since $(\alpha + \beta)^2 = \alpha^2 + 2\alpha\beta + \beta^2$,

$$\alpha^2 + \beta^2 = (\alpha + \beta)^2 - 2\alpha\beta = 9 - 3 = 6.$$

The identity for $\alpha^2 + \beta^2$ did the work, and the equation was never solved.

YOUR TURN 4.6 The Discriminant

The nature of the roots

44. Evaluate the discriminant of each equation and state the nature of its roots.

(a) $x^2 - 4x + 4 = 0$
(b) $2x^2 + 3x + 5 = 0$
(c) $3x^2 - 2x - 1 = 0$
(d) $x^2 + 5x + 7 = 0$

Conditions on a coefficient

45. Find the value or values of k in each case.

(a) $x^2 + 6x + k = 0$ has equal roots
(b) $x^2 + (k-2)x + 4 = 0$ has equal roots
(c) $x^2 - 3x + k = 0$ has no real roots
(d) $(k-1)x^2 + 4x + 2 = 0$ has two distinct real roots. State any value of k that must be excluded.

46. The equation $x^2 + (k+1)x + 9 = 0$ has two distinct real roots. Find the set of values of k.

A line meeting a parabola

47. Each part concerns a line and a parabola.

(a) Find the number of points at which the line $y = x + 3$ meets the parabola $y = x^2 - 3x + 5$.
(b) Show that the line $y = 2x - 1$ is a tangent to $y = x^2 - 4x + 8$, and find the point of contact.
(c) Find the value of c for which the line $y = 3x + c$ is a tangent to $y = x^2 + x + 2$.
(d) Find the set of values of m for which the line $y = mx - 1$ does not meet $y = x^2 + 3$.

The sum and product of the roots

48. Answer each part without solving the equation.

(a) The roots of $3x^2 + 6x - 2 = 0$ are α and β. Write down $\alpha + \beta$ and $\alpha\beta$.
(b) The roots of $x^2 - 4x + 1 = 0$ are α and β. Find $\alpha^2 + \beta^2$.
(c) For the same equation as in part (b), find $\dfrac{1}{\alpha} + \dfrac{1}{\beta}$.
(d) The roots of $2x^2 - 3x - 1 = 0$ are α and β. Find $(\alpha - \beta)^2$. *(Hint: expand $(\alpha - \beta)^2$ and compare it with $(\alpha + \beta)^2$)*

Concept check

49. A student says: "*the discriminant of* $2x^2 - 3x + 5$ *is negative, so the graph of* $y = 2x^2 - 3x + 5$ *lies below the* x*-axis.*" Explain why the conclusion is false.

50. Explain the difference between "*the equation has real roots*" and "*the equation has two distinct real roots*", and state the condition on the discriminant for each.

4.7 Quadratic Inequalities

A quadratic **inequality** is a different question. It does not ask where the parabola meets the axis. It asks where the parabola lies above the axis, or below it. The method has three steps.

KEY To solve a quadratic inequality:

1. Collect every term on one side, so the other side is 0.
2. Find the roots of the matching equation. These are the only places where the sign can change.
3. Sketch the parabola and read the intervals where it lies on the side you want.

With $<$ or $>$ the roots themselves are excluded; with $\leq$ or $\geq$ they are included, because the quadratic is zero there.

The sketch is the method, so here it is in full. The roots cut the axis into three stretches, and the curve keeps one sign right across each of them.

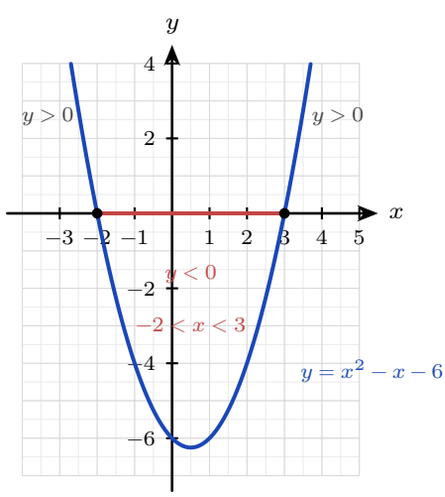


Figure 4.5 $y = x^2 - x - 6$, with roots at $x = -2$ and $x = 3$. The curve lies below the axis on the scarlet stretch between the roots and above it on both sides. So $x^2 - x - 6 < 0$ gives $-2 < x < 3$, and $x^2 - x - 6 > 0$ gives $x < -2$ or $x > 3$.

Read the picture, not a rule. An upward parabola is negative between its roots and positive

outside them; a downward one is the other way round. A quick sketch helps prevent a sign error.

If the discriminant is negative there are no roots, so the parabola lies entirely on one side of the axis. The inequality then holds for every real x or for none, and the sign of a decides which.

EXAMPLE 4.21

Solve $x^2 - x - 6 > 0$.

Factorise: $x^2 - x - 6 = (x-3)(x+2)$, so the roots are $x = -2$ and $x = 3$. The parabola opens upward. It is therefore *above* the axis outside the roots, and below the axis between them. So

$$x^2 - x - 6 > 0 \quad \text{when} \quad x < -2 \ \text{ or } \ x > 3.$$

A rough sketch of the parabola is a reliable check on the signs.

EXAMPLE 4.22

Solve $3x^2 \leq 5x + 2$.

First collect everything on one side:

$$3x^2 - 5x - 2 \leq 0.$$

Factorise: $(3x+1)(x-2) = 0$ gives roots $x = -\frac{1}{3}$ and $x = 2$. The parabola opens upward, so it lies *below* the axis between the roots. The sign $\leq$ includes the roots themselves:

$$-\tfrac{1}{3} \leq x \leq 2.$$

CAUTION Do not divide an inequality by a term containing x unless its sign is known. Dividing by a negative quantity reverses the inequality. Collect on one side and factorise instead.

EXAMPLE 4.23

Solve $x^2 + 3x + 4 > 0$.

The discriminant is $9 - 16 = -7 < 0$, so the parabola never meets the x-axis. Since $a = 1 > 0$ it opens upward, so it lies entirely above the axis. The inequality holds for *every* real x.

This is the negative-discriminant case described above, with $a > 0$.

YOUR TURN 4.7 Quadratic Inequalities

51. Solve each inequality.

(a) $x^2 - 9 < 0$ (b) $x^2 + 2x - 15 < 0$

(c) $x^2 - 4x + 3 \geq 0$ (d) $-x^2 + 4x > 0$

52. Solve each inequality. Two of them are special cases, so justify your answers.

(a) $2x^2 \leq x + 6$ (b) $x^2 \geq 16$

(c) $x^2 + x + 1 > 0$ (d) $x^2 + 4 < 0$

53. A rectangle has width x and length $x + 2$, both in centimetres.

(a) Show that its area is $x^2 + 2x$ square centimetres.

(b) Find the values of x for which the area is less than 15 square centimetres.

Concept check

54. A student solves $x^2 > 9$ by writing "$x > \pm 3$, *so* $x > 3$ *or* $x > -3$." Identify the error and give the correct solution.

Chapter 4 · Lines & Quadratics

Reflections

TOK Descartes showed that a geometric problem can be solved with algebra, and an algebraic one with a picture. A line is an equation, and a linear equation is a line. Neither form is more real than the other. Does changing the representation give you new knowledge, or only a new view of knowledge you already had?

IM Around 1800 BCE, scribes in Babylon solved problems such as "the area of a square added to its side is $\frac{3}{4}$; find the side" by a fixed sequence of steps that is completing the square, written on clay tablets with no symbols at all. In 628 CE the Indian mathematician Brahmagupta stated a general rule for such equations in words, and gave rules for calculating with negative numbers, which he called debts, and with zero. The method of §4.4 is nearly four thousand years old; only the notation is new. Does a method belong to whoever first used it, or to whoever first wrote it down in general?

RW A gradient is a rate, so any line drawn between two quantities has one: a road sign showing 12%, a wheelchair ramp of about 1 in 12 in many building codes, speed on a distance-time graph, and marginal cost (the extra cost of making one more item) on a total-cost graph. The parabola has uses of its own. It reflects every incoming ray parallel to its axis of symmetry to a single point, the focus, so a satellite dish collects weak signals there, and a headlight with its bulb at the focus sends out a parallel beam. Quadratics also model optimisation problems, such as the greatest area for a fixed length of fence or the greatest height of a thrown ball.

EXT The discriminant divides quadratics into three cases. When $\Delta < 0$ the parabola does not meet the x-axis and there are no *real* roots, but the quadratic formula still returns answers, now containing $\sqrt{-1}$. Keeping these "impossible" numbers instead of rejecting them leads to the **complex numbers** of Chapter 17, where every quadratic has exactly two roots and the discriminant only tells you whether those roots are real.

End-of-Chapter Problems
Chapter 4: Lines & Quadratics

1. Consider $f(x) = 2x - 1$.

(a) Find $f^{-1}(x)$. [2]
(b) Verify that $f(f^{-1}(x)) = x$. [2]

2. Consider $f(x) = 3 - 2x$.

(a) State the x- and y-intercepts. [2]
(b) Find $f^{-1}(x)$. [2]

3. Consider $f(x) = \dfrac{2x+1}{x-3}$.

(a) State the domain. [1]
(b) Find $f^{-1}(x)$. [4]

4. Find the equation of the line through $(3, -1)$ parallel to $2x - y = 4$. [3]

5. Consider the points $A(-1, 2)$ and $B(3, 10)$.

(a) Find the midpoint of AB. [1]
(b) Find the gradient of AB. [1]
(c) Find the length of AB. [2]

6. The line $2x + 3y = 12$ is given.

(a) Find its x- and y-intercepts. [2]
(b) State its gradient. [2]

7. A line passes through $A(1, 2)$ and $B(5, 4)$.

(a) Find the gradient of AB. [1]
(b) Find the equation of AB. [2]
(c) Find the equation of the perpendicular bisector of AB. [4]

8. Solve the system $3x + 2y = 7,\ \ x - 2y = -3$. [3]

9. Solve the system

$$x + 2y + z = 4, \qquad 2x + y - z = 2, \qquad x - y + 2z = 2.$$

[6]

10. Consider the system

$$x + y + z = 3, \qquad x + 2y + 3z = 4, \qquad 2x + 3y + 4z = k.$$

(a) Find the value of k for which the system has solutions. [3]
(b) For this value of k, find the general solution. [3]
(c) Describe the configuration of the three planes when $k = 10$. [1]

11. A quadratic has roots 2 and -5.

(a) Write it in factorised form (with leading coefficient 1). [1]
(b) Write it in expanded form. [2]

12. A parabola passes through $(0, 3)$, $(1, 0)$ and $(3, 0)$. Find its equation. [5]

13. Consider $f(x) = x^2 + 2x - 3$.

(a) Factorise $f(x)$ and state its roots. [2]
(b) Find the vertex. [2]
(c) Sketch the graph, showing all intercepts. [2]

14. Consider $f(x) = x^2 - 6x + 5$.

(a) Write $f(x)$ in vertex form. [2]
(b) State the vertex and the axis of symmetry. [2]
(c) Find the x- and y-intercepts. [3]

15. Solve the inequality $x^2 - 2x - 8 \leq 0$. [3]

16. The equation $x^2 - 4x + c = 0$ has no real roots. Find the set of values of c. [3]

17. Solve the inequality $x^2 + x - 6 > 0$. [3]

18. The equation $x^2 + kx + 4 = 0$ has two distinct real roots. Find the set of values of k. [4]

19. Find the values of k for which $x^2 + kx + (k + 3) = 0$ has equal roots. [4]

20. Two numbers have a sum of 10 and a product of 21. Find the numbers. [4]

21. The roots of $x^2 + bx + c = 0$ sum to 6 and differ by 4. Find b and c. [5]

22. Show that $x^2 + (k+1)x + k = 0$ has real roots for every real k, and find the value of k for which the roots are equal. [5]

23. The roots of $x^2 - px + q = 0$ are α and 2α. Show that $2p^2 = 9q$. [5]

24. Find the points of intersection of $y = x^2$ and $y = 2x + 3$. [4]

25. Consider the line $y = 2x - 5$ and the parabola $y = x^2 - 2x + 2$.

(a) Show that they do not intersect. [3]
(b) Explain what this means geometrically. [1]

26. A ball is thrown so that its height in metres after t seconds is $h(t) = 20t - 5t^2$.

(a) Find the time at which the ball lands. [2]
(b) Find the maximum height reached. [3]

27. Find the value of k for which the line $y = x + k$ is a tangent to the parabola $y = x^2$. [5]

28. A rectangle has length 3 more than its width w.

(a) Write its area A as a function of w. [2]
(b) Given the area is 40, find w. [3]

29. A rectangle has perimeter 20. Show that its area is $A = x(10 - x)$, where x is one side, and find the maximum possible area. [5]

30. Prove that the line $y = mx + c$ is a tangent to the parabola $y = x^2$ if and only if $c = -\dfrac{m^2}{4}$. [6]

31. GDC Consider $f(x) = x^3 - 4x + 1$.

(a) Using your GDC, sketch the graph of f for $-3 \le x \le 3$, labelling all key features. [2]
(b) Find the zeros of f, giving your answers to 3 significant figures. [2]
(c) Find the coordinates of the local maximum and local minimum points. [2]
(d) Solve $f(x) = 2x - 1$, giving your answers to 3 significant figures. [2]

32. GDC A bridge has a parabolic arch. The arch is 40 m wide at road level and 10 m high at its centre. Take the road as the x-axis, with the centre of the arch at $x = 0$.

(a) Show that the arch is modelled by $y = 10 - \dfrac{x^2}{40}$. [4]
(b) Find the height of the arch 12 m from the centre. [2]
(c) A lorry 8 m wide and 9.5 m high drives along the centre of the road. Determine whether it passes under the arch. [4]
(d) Find the greatest width of a lorry 9 m high that can pass under the arch, giving your answer to three significant figures. [4]
(e) The model gives a negative value of y when $|x| > 20$. Explain what this means for the model. [2]

33. Consider the line $y = mx + 3$ and the parabola $y = x^2 + 5x + 7$.

(a) Show that at any point of intersection, $x^2 + (5 - m)x + 4 = 0$. [3]
(b) Find the values of m for which the line is a tangent to the parabola. [4]
(c) Find the values of m for which the line does not meet the parabola. [3]
(d) For the smaller value of m found in part (b), find the coordinates of the point of tangency. [3]
(e) State the values of m for which the line meets the parabola twice. [2]

34. HL Consider the system of equations

$$x + y + z = 2, \qquad x + 2y + 3z = 5, \qquad 2x + 3y + az = b,$$

where a and b are real constants.

(a) Write the system as an augmented matrix and row reduce it. [4]
(b) Find the value of a for which the system does not have a unique solution. [3]
(c) For that value of a, find the value of b for which the system is consistent. [2]
(d) For those values of a and b, find the general solution of the system. [4]
(e) Describe how the three planes are arranged in each of the three cases: a not equal to the value in part (b); a equal to it with b consistent; and a equal to it with b not consistent. [3]

Challenge Questions

35. Tangents from a point. A non-vertical line is a tangent to the parabola $y = x^2$ when it meets the parabola exactly once. End-of-Chapter Question 30 proves that $y = mx + c$ is a tangent exactly when $c = -\dfrac{m^2}{4}$; take that result as given here. This investigation asks how many tangents pass through a given point (h, k), and where the points with a perpendicular pair of tangents lie.

(a) Show that the line through (h, k) with gradient m is a tangent to $y = x^2$ if and only if

$$m^2 - 4hm + 4k = 0.$$

[3]

(b) Find the two tangents to $y = x^2$ that pass through the point $(3, 5)$, and state the point where each one touches the parabola. [4]

(c) Hence show that the number of tangents to $y = x^2$ through (h, k) is 2 when $k < h^2$, 1 when $k = h^2$ and 0 when $k > h^2$. Describe where the point lies relative to the parabola in each case, and explain each count from the picture. [5]

(d) Suppose that $k < h^2$, and let m_1 and m_2 be the gradients of the two tangents through (h, k). Write down $m_1 + m_2$ and $m_1 m_2$ in terms of h and k. Show that the two tangents are perpendicular if and only if $k = -\frac{1}{4}$. (The line $y = -\frac{1}{4}$ is called the *directrix* of the parabola.) [4]

(e) The acute angle θ between two lines with gradients m_1 and m_2, where $m_1 m_2 \neq -1$, satisfies $\tan\theta = \left|\dfrac{m_1 - m_2}{1 + m_1 m_2}\right|$. Show that the points from which the two tangents meet at 45° satisfy $16h^2 = (1 + 4k)^2 + 16k$, and describe the curve these points form. Suggest how the investigation could continue, for example for other angles or for other parabolas. [4]

36. Three points fix a parabola. Two points fix a straight line. This investigation asks how many points are needed to fix a parabola $y = A + Bx + Cx^2$, and when a set of points fails to fix one. The unknowns are the coefficients A, B and C, and each point on the curve gives one linear equation in them.

(a) Find the parabola $y = A + Bx + Cx^2$ that passes through $(1, 1)$, $(2, 3)$ and $(3, 7)$. Write the three equations as an augmented matrix and row reduce it. [4]

(b) The third point is now replaced by $(k, 7)$, where k is a real constant. Show that row reduction leads to the last row

$$(k - 1)(k - 2)\,C = 8 - 2k,$$

and hence find C in terms of k whenever the system has a unique solution. [5]

(c) State the values of k for which the system does not have a unique solution. For each one, decide whether the system has no solution or infinitely many, and explain the result geometrically. [3]

(d) Find the value of k for which $C = 0$. Explain what the three points then have in common, and state the curve that passes through all three. [3]

(e) Conjecture how many points, with distinct x-coordinates, are needed to fix a polynomial of degree at most n. Explain your count by comparing the number of equations with the number of unknowns, and prove that two polynomials of degree at most n that agree at $n + 1$ distinct values of x are the same polynomial. *(Hint: a nonzero polynomial of degree at most n has at most n roots (Ch. 6))* [4]

(f) Without solving any equations, show that

$$p(x) = 1 \cdot \frac{(x-2)(x-3)}{(1-2)(1-3)} + 3 \cdot \frac{(x-1)(x-3)}{(2-1)(2-3)} + 7 \cdot \frac{(x-1)(x-2)}{(3-1)(3-2)}$$

passes through the three points of part (a), and that it simplifies to your answer to part (a). Explain how this construction generalises to $n+1$ points, and suggest one question about fitting a polynomial through many points that could be investigated further. [4]

Chapter 5

Functions & Graphs

5

SYLLABUS 2.5 2.8–2.11 2.13–2.16

Hold a card in front of a lamp and look at its shadow on the wall. Slide the card and the shadow slides. Move it closer to the lamp and the shadow grows. Turn it over and the shadow flips. The card itself never changes. Only what you do to it changes.

Graphs work the same way, and that makes this part of the course more manageable. There are only a few shapes to learn by heart, a line, a parabola, a square root, a reciprocal, together with a short list of moves that slide them, stretch them and reflect them. Many other curves you will meet are one of those few shapes with some of those moves applied to it.

So this chapter has two halves. The first is **composition** and the list of moves: how the output of one function becomes the input of the next, and what each move does to an equation. The second is the small set of shapes those moves are applied to, and in particular the ones whose graphs approach a line, called an asymptote. Reading that behaviour from a formula is the second skill this chapter builds.

By the end of this chapter you will be able to compose two functions and find the domain of the result, apply translations, stretches and reflections, and test a function for symmetry. You will sketch the related graphs $|f(x)|$, $f(|x|)$ and $\frac{1}{f(x)}$. You will find the asymptotes of a rational function, solve a rational inequality with a sign table, solve equations and inequalities involving $|f(x)|$, and treat the exponential and the logarithm as inverses of each other.

What ties these together is that each constant in these formulas corresponds to something you can see. Change the number and the graph moves in a way you can predict. A formula and a picture are two views of one function, and this chapter is about reading each from the other.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

5.1 Composite Functions

To work out $\sqrt{3x+1}$ you do three things, in order. Multiply by 3. Add 1. Take the square root. Many formulas in this course are built like that, from a few simple steps joined end to end. Reading a formula as a chain of steps is what lets you undo it, and later what lets you differentiate it. **Composition** is the notation for that chain.

Two functions can be joined end to end. Send a number into the first function. Take the number that comes out. Send that number into the second function. The single rule that does both steps at once is called a **composite** function.

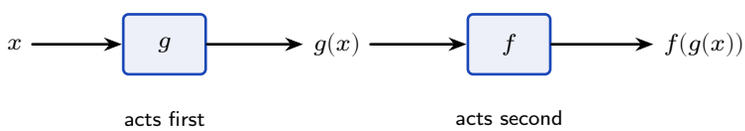


Figure 5.1 Composition sends x through g first, then sends the result into f. The inner function acts first, which is why $f \circ g$ is read from the inside out.

DEF The **composite** $f \circ g$ is the function $(f \circ g)(x) = f(g(x))$: apply g first, then f. The **identity function** $I(x) = x$ returns every input unchanged.

The symbol $\circ$ is read "of", so $f \circ g$ is "f of g". It is also read "f after g", which gives the order in which the two functions act. It is not multiplication. Read the notation from right to left: the function written closest to x acts on x first.

5.1.1 Substituting a Whole Function

Composition looks unfamiliar because a function is being written where a number normally goes. The rule for f does not change at all. Only what you write in the slot changes.

The x in $f(x)$ is a **placeholder** (a marker showing where the input goes). It is not a fixed value, and it does not have to be a single letter.

Read f as a rule with an empty slot. If $f(x) = x^2 + 1$, the rule says: square the input, then add one. Whatever you write in the slot is squared, and then one is added.

$$f(\square) = \square^2 + 1 \qquad \Longrightarrow \qquad f(3) = 3^2 + 1, \quad f(2x) = (2x)^2 + 1, \quad f(g(x)) = (g(x))^2 + 1.$$

The last one is composition. You are substituting a whole function exactly as you would substitute a number.

To build $(f \circ g)(x)$:

1. Write the formula for $f(x)$.
2. Substitute the expression $g(x)$ for **every** x in that formula.
3. Simplify.

5.1.2 Order of Composition

Putting on socks and then shoes is not the same as putting on shoes and then socks. Composition behaves in the same way, so $f \circ g$ and $g \circ f$ are usually two different functions. Deciding which function acts first is therefore a good first step in a composition question.

Worked through. Take $f(x) = x^2 + 1$ and $g(x) = 2x$. For $f \circ g$, apply g first. Its output is $2x$, so substitute $2x$ for every x in $f(x) = x^2 + 1$, then simplify:

$$(f \circ g)(x) = f(2x) = (2x)^2 + 1 = 4x^2 + 1.$$

For $g \circ f$, apply f first. Its output is $x^2 + 1$, so substitute $x^2 + 1$ for every x in $g(x) = 2x$, then simplify:

$$(g \circ f)(x) = g(x^2 + 1) = 2(x^2 + 1) = 2x^2 + 2.$$

The two answers are different. In general $f \circ g \neq g \circ f$, so always check which function acts first.

One input makes the difference concrete. Send $x = 3$ through both composites and follow it step by step.

Composite	Input	After the first function	After the second function
$f \circ g$	3	$g(3) = 6$	$f(6) = 37$
$g \circ f$	3	$f(3) = 10$	$g(10) = 20$

5.1.3 The Domain of a Composite

Each part of a composite can be defined at a number while the composite is not. The reason is that f is never applied to x. It is applied to $g(x)$, and $g(x)$ may be a value that f does not take. So two conditions must hold, and the second one is easy to miss.

KEY The **domain** of $f \circ g$ is the set of values x for which

1. x is in the domain of g, **and**
2. $g(x)$ is in the domain of f.

Find the domain from the two original functions, before simplifying. Simplification is an algebraic step, and a simplified formula may be defined at values where the composite is not.

EXAMPLE 5.1

For $f(x) = x^2$ and $g(x) = \sqrt{x - 1}$, find $(f \circ g)(x)$ and state its domain.

Apply g first, then square its output:

$$(f \circ g)(x) = f(\sqrt{x-1}) = \left(\sqrt{x-1}\right)^2 = x - 1.$$

Now take the two conditions in order.
Condition 1: the domain of g requires $x - 1 \geq 0$, so $x \geq 1$.
Condition 2: the domain of f is all real numbers, so this adds nothing here.
The domain of $f \circ g$ is therefore $x \geq 1$. The composite is the line $y = x - 1$ with everything to the left of $x = 1$ removed.

CAUTION A simplified formula is not the composite. $\left(\sqrt{x-1}\right)^2$ simplifies to $x - 1$, which is defined for every real x, but the composite exists only for $x \geq 1$. Always return to the original functions to find the domain.

EXAMPLE 5.2
For $f(x) = \dfrac{1}{x}$ and $g(x) = x - 3$, find $(f \circ g)(x)$ and state its domain.

$$(f \circ g)(x) = f(x - 3) = \frac{1}{x - 3}.$$

Condition 1: the domain of g is all real numbers, so it removes nothing.
Condition 2: 0 is not in the domain of f, so the output $g(x) = x - 3$ must not equal 0. This removes $x = 3$.
The domain of $f \circ g$ is every real number except $x = 3$.

The two examples restrict the domain for different reasons, which is why both conditions have to be checked every time.

The functions	1. x in the domain of g	2. $g(x)$ in the domain of f	Domain of $f \circ g$
$f(x) = x^2$ $g(x) = \sqrt{x-1}$	$x \geq 1$	no extra condition	$x \geq 1$
$f(x) = \dfrac{1}{x}$ $g(x) = x - 3$	every real x	$x - 3 \neq 0$, so $x \neq 3$	$x \in \mathbb{R}$, $x \neq 3$

5.1.4 Composition and Inverses

Some pairs of functions undo each other. Composition is how that relationship is written down, and composing the two is how you check that an inverse is correct.

The **inverse** of f, written f^{-1}, is the function that reverses f. If f sends 3 to 10, then f^{-1} sends 10 back to 3. Apply f and then apply f^{-1}, and you arrive back at the number you

started with. The composite is therefore the identity function.

KEY For a function f with inverse f^{-1},

$$(f^{-1} \circ f)(x) = x \qquad \text{and} \qquad (f \circ f^{-1})(x) = x.$$

An inverse exists exactly when f is **one-to-one**: no two inputs share an output.

To find an inverse, use the method of Ch. 4, where inverses were first met:

1. Write $y = f(x)$.
2. Exchange x and y.
3. Solve the new equation for y. The result is $f^{-1}(x)$.
4. State the domain of f^{-1}, which is the range of f.

Composition is also how you check the answer. Take $f(x) = 2x + 1$, whose inverse is $f^{-1}(x) = \dfrac{x-1}{2}$. Substituting one into the other gives

$$f^{-1}(f(x)) = \frac{(2x+1)-1}{2} = \frac{2x}{2} = x,$$

which confirms the pair. An important pair of inverse functions in this chapter is the exponential and the logarithm; see §5.5.

A quotient of two linear expressions needs one extra move at step 3. The variable appears twice, once above the line and once below it, so it has to be collected into a single term before it can be isolated. The range of such a function, which is the domain of its inverse, is worked out in §5.4.1.

EXAMPLE 5.3

Find the inverse of $f(x) = \dfrac{3x-1}{x-1}$, and state its domain.

Write $y = f(x)$ and exchange the letters:

$$x = \frac{3y-1}{y-1}.$$

Clear the fraction by multiplying both sides by $y - 1$:

$$x(y-1) = 3y - 1 \implies xy - x = 3y - 1.$$

Now y is in two terms. Gather those on one side and everything else on the other, then factorise and divide:

$$xy - 3y = x - 1 \implies y(x-3) = x - 1 \implies f^{-1}(x) = \frac{x-1}{x-3}, \qquad x \neq 3.$$

The excluded value is not guesswork. The domain of f^{-1} is the range of f, and f never produces 3: written as $f(x) = 3 + \dfrac{2}{x-1}$, its second term is never zero, so the output is never

exactly 3.
Check one point: $f(0) = \dfrac{-1}{-1} = 1$, and $f^{-1}(1) = \dfrac{0}{-2} = 0$, which returns the input.

YOUR TURN 5.1 Composite Functions

Forming a composite

1. For $f(x) = x + 3$ and $g(x) = 2x$, find each composite.

(a) $(f \circ g)(x)$
(b) $(g \circ f)(x)$
(c) $(f \circ f)(x)$
(d) $(g \circ g)(x)$

2. For $f(x) = x^2$ and $g(x) = x - 1$, find each of the following.

(a) $(f \circ g)(x)$
(b) $(g \circ f)(x)$
(c) $(f \circ g)(3)$
(d) $(g \circ f)(3)$

3. Let $f(x) = x^2$ and $g(x) = x + 2$.

(a) Find $(f \circ g)(x)$ and $(g \circ f)(x)$, and show that they are not the same.
(b) Find the value of x for which the two composites are equal.

The domain of a composite

4. For $f(x) = \dfrac{1}{x}$ and $g(x) = x + 1$, find each composite and state its domain.

(a) $(f \circ g)(x)$
(b) the domain of $f \circ g$
(c) $(g \circ f)(x)$
(d) the domain of $g \circ f$

5. For each pair of functions, find $(f \circ g)(x)$ and state the domain of $f \circ g$. Find the domain from f and g before you simplify.

(a) $f(x) = \sqrt{x}, \quad g(x) = 3 - x$
(b) $f(x) = x^2, \quad g(x) = \sqrt{x + 2}$
(c) $f(x) = \dfrac{1}{x - 2}, \quad g(x) = x^2 - 2$
(d) $f(x) = \sqrt{x - 1}, \quad g(x) = x^2$

Composites and inverses

6. Find the inverse of each function, and state the domain of the inverse.

(a) $f(x) = \dfrac{x+3}{x-2}$ (b) $f(x) = \dfrac{3x}{2x-1}$

7. Let $f(x) = 2x + 1$.

(a) Find $f^{-1}(x)$.
(b) Show that $(f \circ f^{-1})(x) = x$.
(c) Show that $(f^{-1} \circ f)(x) = x$.
(d) State what these two results tell you about f^{-1}.

Concept check

8. Let $f(x) = x + 3$ and $g(x) = x^2$. A student writes: "$(f \circ g)(x) = (x+3)^2$." Identify the error and find $(f \circ g)(x)$.

9. Let $f(x) = x^2$ and $g(x) = \sqrt{x+2}$. A student writes: "$(f \circ g)(x) = \left(\sqrt{x+2}\right)^2 = x + 2$, *so the domain of* $f \circ g$ *is* $x \in \mathbb{R}$." Identify the error and state the correct domain.

5.2 Transformations of Graphs

Plotting points is slow, and for most curves it is unnecessary. A small change to an equation moves the graph in one predictable way. Once you know which change produces which move, you can sketch a new curve from a familiar one in a few seconds. You can also work the other way, and write down the equation of a curve you are shown.

There are three moves: a **translation** (a slide), a **reflection** (a flip) and a **stretch** (a scaling). Each one matches one change in the equation, and each one sends the point (p, q) to a new point you can write down.

Move	Equation	Acts on	(p, q) becomes	Effect
Translation	$y = f(x) + k$	output	$(p,\ q + k)$	k up
	$y = f(x - h)$	input	$(p + h,\ q)$	h right
Reflection	$y = -f(x)$	output	$(p,\ -q)$	in the x-axis
	$y = f(-x)$	input	$(-p,\ q)$	in the y-axis
Stretch	$y = a\,f(x)$	output	$(p,\ aq)$	vertical, factor a
	$y = f(bx)$	input	$(\frac{p}{b},\ q)$	horizontal, factor $\frac{1}{b}$

The table is three pairs, one pair for each move. Read the **Acts on** column and a single pattern accounts for all six rows. Everything depends on whether the change sits outside the bracket or inside it.

> KEY **Outside** the bracket, the change acts on the output: it is **vertical**, and it does what it says. **Inside** the bracket, the change acts on the input: it is **horizontal**, and it does the **opposite**.

The first line behaves as you would expect. Adding k raises every output by k, so the graph moves up by k. Multiplying by a multiplies every output by a, so the graph stretches vertically by a.

The second line needs an explanation, because the reversal is not obvious. Take $y = f(x - h)$ and evaluate it at $x = p + h$:

$$f((p + h) - h) = f(p),$$

which is the value the original graph had at $x = p$. So the point that was at $x = p$ now appears at $x = p + h$. Every point moves h units to the *right*, not left. The same calculation works for $y = f(bx)$: at $x = \frac{p}{b}$ it returns $f(b \cdot \frac{p}{b}) = f(p)$, so the point at p moves to $\frac{p}{b}$. That is a horizontal stretch of factor $\frac{1}{b}$, which brings every point closer to the y-axis whenever $b > 1$.

5.2.1 Translations

A translation slides the whole graph. Its shape and its size do not change, only its position.

Vertically, $y = f(x) + k$ adds k to every output, so every point moves up by k. A negative k moves the graph down.

Horizontally, $y = f(x - h)$ moves the graph right by h. This sign is easy to reverse. Use the

reason above rather than memorising the result. The change is inside the bracket, so it acts on the input, and its effect is the opposite of its sign.

In examination questions a translation is usually written as a vector. The graph of $y = f(x - h) + k$ is the graph of $y = f(x)$ translated by $\binom{h}{k}$.

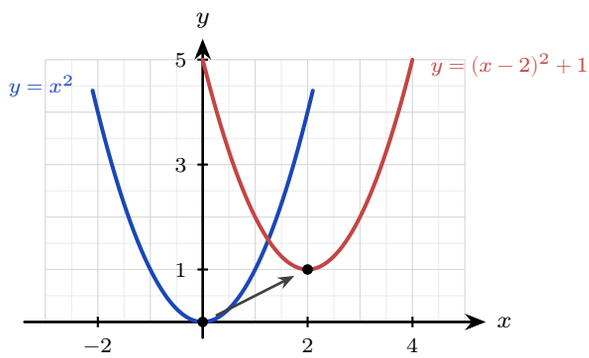


Figure 5.2 The graph of $y = (x - 2)^2 + 1$ is the graph of $y = x^2$ translated 2 units right and 1 unit up. The vertex moves from $(0, 0)$ to $(2, 1)$, and the shape is unchanged.

Worked through. To obtain $y = (x - 2)^2 + 1$ from $y = x^2$, start with the bracket. The -2 sits inside it, so it acts on the input and moves the parabola 2 units *right*. Now take the rest. The $+1$ sits outside the bracket, so it acts on the output and moves the parabola 1 unit *up*. Together these send the vertex from $(0, 0)$ to $(2, 1)$, which is exactly what the completed-square form predicts. Transformations and the completed-square form are two descriptions of the same thing.

5.2.2 Reflections

A reflection flips the graph across an axis. Every point keeps its distance from that axis and moves to the other side of it.

For $y = -f(x)$ the sign of every output changes, so the point (p, q) moves to $(p, -q)$. That is a reflection in the x-axis. A maximum becomes a minimum, and any point already on the x-axis stays where it is.

For $y = f(-x)$ the sign of every input changes, so the point (p, q) moves to $(-p, q)$. That is a reflection in the y-axis. The y-intercept does not move, because changing the sign of 0 leaves it as 0.

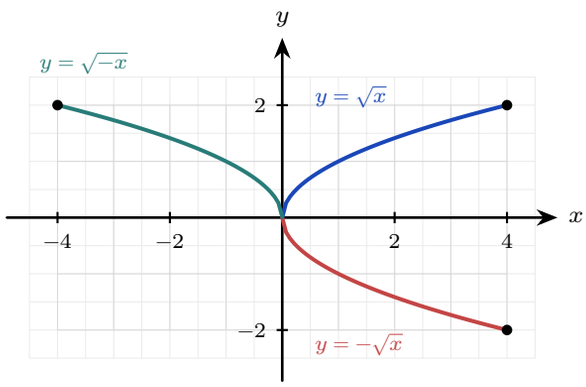


Figure 5.3 Reflections of $y = \sqrt{x}$. Changing the sign of the output gives $y = -\sqrt{x}$, and the point $(4, 2)$ moves to $(4, -2)$: a reflection in the x-axis. Changing the sign of the input gives $y = \sqrt{-x}$, and $(4, 2)$ moves to $(-4, 2)$: a reflection in the y-axis.

EXAMPLE 5.4

Describe how the graph of $y = -(x-1)^2$ is obtained from $y = x^2$.

There are two changes. The -1 is inside the bracket, so it is horizontal: translate 1 unit right.

The minus sign is outside the bracket, so it is vertical: reflect in the x-axis.

One change is horizontal and the other is vertical, so they do not affect each other and may be done in either order.

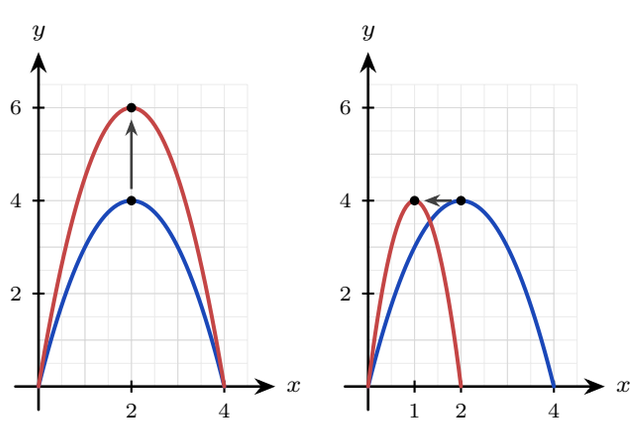
The vertex moves from $(0, 0)$ to $(1, 0)$, and the parabola now opens downward.

5.2.3 Stretches

A stretch changes the size of the graph in one direction only. One line stays fixed while every other point moves away from it or towards it.

For $y = a\,f(x)$ every output is multiplied by a, so (p, q) moves to (p, aq). This is a vertical stretch of factor a. Points on the x-axis do not move, because $a \times 0 = 0$.

For $y = f(bx)$ every input is divided by b, as shown above, so (p, q) moves to $(\frac{p}{b}, q)$. This is a horizontal stretch of factor $\frac{1}{b}$. Points on the y-axis do not move.



(a) $y = 1.5\,f(x)$ (b) $y = f(2x)$

Figure 5.4 The original curve $y = f(x)$ is cobalt and the stretched curve is scarlet. In (a) every y-coordinate is multiplied by 1.5, so the maximum moves from $(2, 4)$ to $(2, 6)$ while both roots stay on the x-axis. In (b) every x-coordinate is halved, so the maximum moves from $(2, 4)$ to $(1, 4)$ and the root at $x = 4$ moves to $x = 2$.

CAUTION A factor smaller than 1 is still called a stretch. The graph of $y = \frac{1}{2}f(x)$ is a vertical stretch of factor $\frac{1}{2}$, even though it makes the graph shorter. Write "stretch, factor $\frac{1}{2}$", not "squash".

EXAMPLE 5.5

The graph of $y = f(x)$ has a maximum point at $(6, 4)$. State the coordinates of that maximum point on the graph of (a) $y = 3f(x)$, (b) $y = f(3x)$.

(a) The 3 is outside the bracket, so it is a vertical stretch of factor 3. It multiplies the y-coordinate and leaves the x-coordinate alone: $(6, 12)$.

(b) The 3 is inside the bracket, so it is a horizontal stretch of factor $\frac{1}{3}$. It divides the x-coordinate and leaves the y-coordinate alone: $(2, 4)$.

Each stretch changes one coordinate and leaves the other unchanged. Checking which coordinate moved is a quick way to confirm you have used the right one.

5.2.4 Composite Transformations

Many graphs in this chapter are several moves away from a familiar shape, applied one after another. This is a **composite transformation**. Describing one is a matter of taking the changes in a reliable order.

Start by writing the target equation in one standard form:

$$y = a\,f\big(b(x - h)\big) + k.$$

Then a is the vertical stretch factor, b gives the horizontal stretch factor $\frac{1}{b}$, h is the horizontal translation and k is the vertical translation. Every value can be read straight from the equation.

For a quadratic this standard form is the **completed-square form**. Writing $y = a(x - h)^2 + k$ names the vertex (h, k) and the two moves that put it there, so completing the square and describing a transformation are two ways of doing the same work.

KEY To describe a composite transformation:

1. Split the changes into two groups: those **inside** the bracket, which are horizontal, and those **outside** it, which are vertical.
2. Do the two groups in either order. A horizontal move and a vertical move do not affect each other.
3. Inside each group the order matters. Apply the **stretch before the translation**.

Step 3 has a reason, and understanding it is safer than memorising it. A stretch multiplies every value it is applied to, including a translation that has already been carried out. Do the stretch first and there is no translation for it to multiply. The second example below shows what goes wrong in the other order.

Whatever sequence you write down, check it by tracking a single point. Choose a point you know on the original graph, such as a vertex, an endpoint or an intercept. Apply your moves to its coordinates. Then check that the result satisfies the target equation.

CAUTION Factorise inside the bracket first. In $y = f(2x - 6)$ the translation is *not* 6. Written as $y = f(2(x - 3))$, it is a horizontal stretch of factor $\frac{1}{2}$ followed by a translation of 3 to the right. Forgetting to factorise is a common error here.

EXAMPLE 5.6

Describe a sequence of transformations that maps $y = \sqrt{x}$ onto $y = 3\sqrt{x - 2} + 1$.

Take the change inside the bracket first. The -2 acts on the input, so it translates the graph 2 units *right* and gives $\sqrt{x - 2}$.

Now take the vertical changes, stretch before translation. The factor 3 stretches the graph vertically by scale factor 3 and gives $3\sqrt{x - 2}$. The $+1$ then translates it 1 unit *up*.

So the sequence is: translate 2 right, stretch vertically by factor 3, translate 1 up.

Check it by tracking the endpoint. The graph of $y = \sqrt{x}$ begins at $(0, 0)$. The translation carries it to $(2, 0)$. The stretch leaves it there, because $3 \times 0 = 0$. The final translation carries it to $(2, 1)$. That is exactly where $y = 3\sqrt{x - 2} + 1$ begins.

Two moves **commute** (give the same result in either order) when swapping them leaves the final graph unchanged. A horizontal move and a vertical move always commute, because they change different coordinates. A stretch and a translation in the *same* direction do not commute, unless the stretch factor is 1 or the translation is 0, so that one of the two moves does nothing. The next example shows the two different answers they produce.

EXAMPLE 5.7

Starting from $y = x^2$, apply a vertical stretch by factor 2 and a translation 1 unit up, in both orders. Are the results the same?

Stretch first, then translate. The stretch doubles every value, giving $2x^2$; the translation then

adds 1:

$$x^2 \longrightarrow 2x^2 \longrightarrow 2x^2 + 1.$$

Translate first, then stretch. The translation adds 1, giving $x^2 + 1$; the stretch then doubles everything, *including* the 1:

$$x^2 \longrightarrow x^2 + 1 \longrightarrow 2(x^2 + 1) = 2x^2 + 2.$$

The two curves are different. A vertical stretch multiplies whatever is already there, so a translation applied first is multiplied as well. A vertical stretch and a vertical translation commute only when one of them does nothing, that is when the stretch factor is 1 or the translation is 0. So state the moves in the order that builds the target curve.

TIP In examinations, "describe the sequence of transformations" expects named moves with their values: translation by $\binom{-3}{0}$, vertical stretch scale factor 2, translation by $\binom{0}{-1}$. Vague words such as "moved" or "squashed" do not earn the marks.

YOUR TURN 5.2 Transformations of Graphs

Single transformations

10. Describe the single transformation that takes $y = x^2$ to each graph.

(a) $y = x^2 + 3$
(b) $y = (x + 4)^2$
(c) $y = -x^2$
(d) $y = (x - 1)^2$

11. Describe the single transformation that takes $y = f(x)$ to each graph.

(a) $y = f(-x)$
(b) $y = -f(x)$
(c) $y = 3f(x)$
(d) $y = f(2x)$

The image of a point

12. The point $(2, 5)$ lies on $y = f(x)$. Find the coordinates of its image on each graph.

(a) $y = f(x - 3) + 1$
(b) $y = 3f(x)$
(c) $y = f(2x)$
(d) $y = -f(x) + 2$

13. The point $(-2, 3)$ lies on $y = f(x)$. Find the coordinates of its image on each graph.

(a) $y = f(2x)$
(b) $y = f(x) - 4$
(c) $y = f(-x)$
(d) $y = \frac{1}{2}f(x)$

Combined transformations

14. Describe the sequence of transformations that maps $y = \sqrt{x}$ onto each graph.

(a) $y = \sqrt{x+3}$
(b) $y = 2\sqrt{x}$
(c) $y = 2\sqrt{x+3} - 1$
(d) $y = \sqrt{3x+6}$ *(Hint: factorise inside the root first)*

15. Start from $y = x^2$.

(a) Translate 2 units up, then reflect in the x-axis. Write the new equation.
(b) Apply the same two moves in the opposite order. Write the new equation.

16. Consider $y = f(4x - 8)$.

(a) Write it in the form $y = f(b(x - h))$.
(b) Describe the sequence of transformations from $y = f(x)$.

Concept check

17. A student describes two transformations of $y = f(x)$. Identify the error in each description and give the correct one.

(a) "$y = f(x+3)$ *is a translation of* 3 *units to the right.*"
(b) "$y = f(2x)$ *is a horizontal stretch with scale factor* 2."

5.3 Symmetry, Inverses and Related Graphs HL

Half of a sketch is half the work, and for some graphs half is all you need. A graph has a **symmetry** when some motion leaves it looking exactly as it did. Reflect it, rotate it or slide it, and the picture is unchanged. One half of the curve then fixes the other half, and the symmetry gives you a way to check a sketch against the formula.

Three motions are common enough to be named, and each one has a short algebraic test.

Type	Test	Motion that leaves the graph unchanged	Examples
Even	$f(-x) = f(x)$	reflection in the y-axis	$x^2,\quad x^4,\quad \lvert x\rvert,\quad \cos x$
Odd	$f(-x) = -f(x)$	rotation of 180° about the origin	$x,\quad x^3,\quad \dfrac{1}{x},\quad \sin x$
Periodic	$f(x+p) = f(x)$	translation of p along the x-axis	$\sin x,\quad \cos x,\quad \tan x$
Neither	no test holds	none of the three	$x^2 + x,\quad x + 1,\quad e^x$

Most functions have none of these symmetries, so "neither" is a complete and correct answer when the tests fail.

The third row is the one you meet again in trigonometry. A function is **periodic** if its graph repeats at regular intervals, and the smallest positive p with $f(x+p) = f(x)$ is called the **period**. The graphs of $\sin x$ and $\cos x$ both have period 2π.

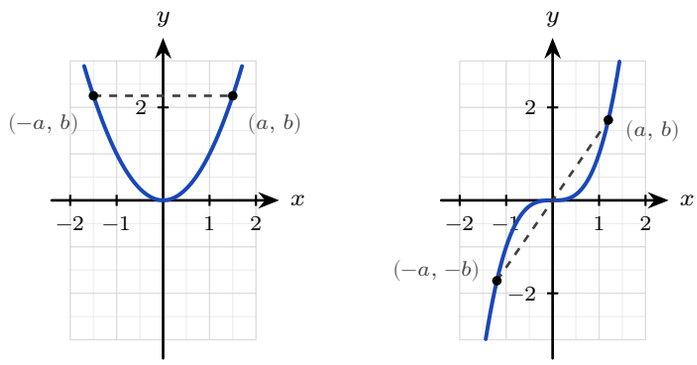
5.3.1 Testing for Even and Odd

The names come from the powers. The functions $x^2, x^4, \ldots$ have even exponents and satisfy $f(-x) = f(x)$. The functions $x, x^3, \ldots$ have odd exponents and satisfy $f(-x) = -f(x)$.

> DEF A function is **even** if $f(-x) = f(x)$ for every x in its domain, and **odd** if $f(-x) = -f(x)$ for every x in its domain. An even graph is symmetric in the y-axis; an odd graph has rotational symmetry of order two about the origin.

To test a function, substitute $-x$ for x, simplify, and compare the result with the original. There are three possible outcomes:

- the result is $f(x)$, so the function is **even**;
- the result is $-f(x)$, so the function is **odd**;
- the result is neither, so the function has **no symmetry** of this kind.



(a) $y = x^2$ is even (b) $y = x^3$ is odd

Figure 5.5 In (a) the point (a, b) has a partner at $(-a, b)$, directly across the y-axis, so reflecting in that axis leaves the curve unchanged. In (b) the partner is at $(-a, -b)$, directly opposite through the origin, so a half turn about the origin leaves the curve unchanged.

Worked through. To decide whether $f(x) = x^3 - x$ is odd, even or neither, substitute $-x$ for x, simplify, then look for a factor of -1:

$$f(-x) = (-x)^3 - (-x) = -x^3 + x = -(x^3 - x) = -f(x).$$

The result is $-f(x)$, so by the definition the function is **odd**. Its graph therefore has rotational symmetry of order two about the origin.

5.3.2 Self-Inverse Functions

A fourth symmetry belongs to functions that undo themselves, and it is usually met as an algebraic expression rather than as a graph. Recall that f^{-1} reverses f. If applying f twice already returns the input, then f reverses itself.

DEF A function is **self-inverse** if $f(f(x)) = x$ for every x in its domain. Then $f = f^{-1}$, and the graph of f is symmetric in the line $y = x$.

That symmetry has a simple reason. Reflecting any graph in the line $y = x$ produces the graph of its inverse, because the reflection swaps the two coordinates of every point. If f is its own inverse, the reflection has to return the same curve.

To test for self-inverse, substitute the whole expression for $f(x)$ in place of x, simplify, and check whether the result is x.

EXAMPLE 5.8

Show that $f(x) = \dfrac{1}{x}$ is self-inverse.

Apply the function twice. Substitute the whole expression $\frac{1}{x}$ in place of x:

$$f(f(x)) = \frac{1}{\frac{1}{x}} = x.$$

Applying f twice returns the input, so f reverses itself and $f = f^{-1}$. Its graph must therefore be symmetric in the line $y = x$, which the figure below confirms.

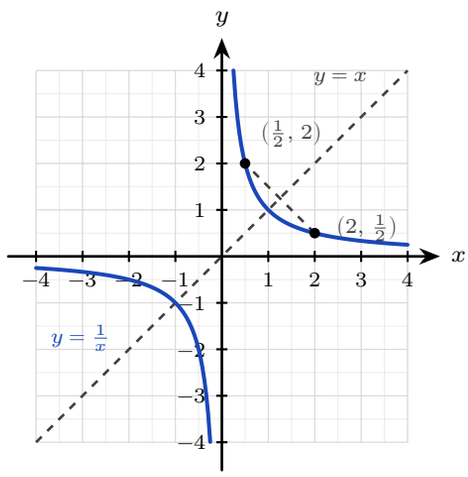


Figure 5.6 The reciprocal is symmetric in the line $y = x$. Each point on the curve has a partner with its two coordinates swapped, so reflecting in $y = x$ returns the same curve.

Self-inverse functions are not rare, and they are not all this simple. The reflection $f(x) = 6 - x$ is self-inverse, and so is $f(x) = \dfrac{2x+1}{x-2}$, which you can check by the same substitution.

5.3.3 Restricting the Domain to Get an Inverse

An inverse exists only when no two inputs share an output. Many familiar functions fail that test on their full domain, and a smaller domain makes them one-to-one.

For example, $f(x) = x^2$ is not one-to-one on all of $\mathbb{R}$, since $f(2) = f(-2)$, so it has no inverse there. Restricting the domain to $x \geq 0$ makes it one-to-one, and the inverse is $f^{-1}(x) = \sqrt{x}$.

Restricting a domain is a standard examination technique, and the restriction you choose decides which branch of the inverse you get. Two facts hold for any inverse. The domain of f^{-1} is the range of f, and the range of f^{-1} is the domain of f.

EXAMPLE 5.9

Let $f(x) = (x-3)^2$ for $x \geq 3$. Find $f^{-1}(x)$, stating its domain.

First check that an inverse exists. On $x \geq 3$ the parabola only rises, so no two inputs share an output and f is one-to-one.

Now find the inverse. Write $y = f(x)$, exchange the letters, and solve for y:

$$x = (y-3)^2 \implies y - 3 = \pm\sqrt{x} \implies y = 3 \pm \sqrt{x}.$$

Two branches appear, so the restriction must decide between them. The range of f^{-1} is the

domain of f, so $y \geq 3$. Then $y - 3$ cannot be negative, and only the $+$ branch is possible. Hence

$$f^{-1}(x) = 3 + \sqrt{x}, \qquad x \geq 0,$$

where the domain of f^{-1} is the range of f. Had the restriction been $x \leq 3$ instead, the same algebra would give $f^{-1}(x) = 3 - \sqrt{x}$. The algebra is the same until that final choice of sign, but the inverse is different.

5.3.4 The Modulus Graph $|f(x)|$

The next two graphs, $y = |f(x)|$ and $y = f(|x|)$, are also produced by reflections. None of the remaining graphs in this section needs a table of values. Each is built from a function f by one operation, and each has a single visual rule.

The modulus of a number is its size without its sign, so $|f(x)|$ is never negative. That gives the rule at once.

KEY To sketch $y = |f(x)|$: keep every part of $y = f(x)$ that lies on or above the x-axis, and reflect every part below the axis upward in the x-axis. The x-intercepts do not move, because $|0| = 0$. A root where f *crosses* the axis (where the gradient of f is not zero) becomes a corner; a root where f merely touches the axis, as in $y = x^2$, stays smooth.

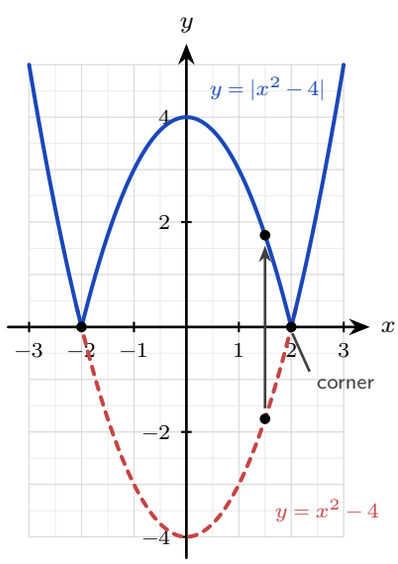


Figure 5.7 $y = |x^2 - 4|$ against $y = x^2 - 4$. The part of the parabola below the x-axis is reflected upward, shown by the arrow, which carries the point $(1.5, -1.75)$ to $(1.5, 1.75)$; everything already above the axis is untouched. The two roots stay at $x = \pm 2$ and become corners.

EXAMPLE 5.10

Sketch $y = |f(x)|$ for $f(x) = x^2 - 6x + 8$, stating the coordinates of the intercepts, the corners

and the turning point.

Find the key points of f first. Factorising gives $f(x) = (x-2)(x-4)$, so the roots are $x = 2$ and $x = 4$, and the vertex lies halfway between them at $x = 3$, where $f(3) = 9 - 18 + 8 = -1$. So f has roots $(2, 0)$ and $(4, 0)$ and a minimum at $(3, -1)$.

Only the arc between the roots lies below the x-axis, so only that arc is reflected. Everything outside $2 \leq x \leq 4$ is already above the axis and is left alone.

Now take the key points one at a time. The roots $(2, 0)$ and $(4, 0)$ do not move, because $|0| = 0$, and f *crosses* the axis at each of them, so each becomes a **corner**. The minimum $(3, -1)$ is reflected to $(3, 1)$, and it is now a local **maximum**. The y-intercept is $f(0) = 8$, which is already positive and so is left where it is, at $(0, 8)$.

So the curve falls from the upper left to the corner $(2, 0)$, rises to the peak $(3, 1)$, falls to the corner $(4, 0)$, then rises again to the upper right, and no part of it lies below the x-axis.

The peak sits at height $|-1| = 1$: reflecting in the x-axis changes the sign of the y-coordinate and leaves the x-coordinate where it was.

This shape is also what makes modulus equations solvable, and §5.3.8 uses it.

5.3.5 The Graph of $f(|x|)$

The previous rule applied the modulus to the *output*. Applying it to the *input* instead gives a different graph and a different rule, and questions on this pair often test whether you can tell them apart.

Since $|x|$ and $|-x|$ are the same number, f receives the same input at x and at $-x$, so the two outputs are equal. The graph is therefore its own mirror image in the y-axis.

KEY To sketch $y = f(|x|)$: keep the part of $y = f(x)$ with $x \geq 0$, and reflect it in the y-axis to make the part with $x < 0$. The original graph to the left of the y-axis is discarded. The result is always an even function. The two halves meet on the y-axis. Unless f has zero gradient at $x = 0$, they meet at a corner, which is a local maximum or minimum of $f(|x|)$.

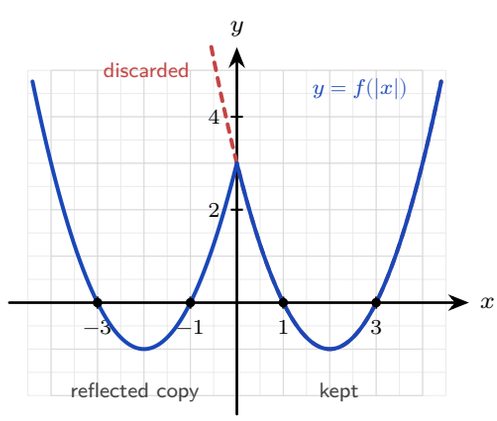


Figure 5.8 $y = f(|x|)$ for $f(x) = x^2 - 4x + 3$. The half with $x \geq 0$ is copied across the y-axis, so the roots at $x = 1$ and $x = 3$ gain partners at $x = -1$ and $x = -3$. The short dashed piece is the part of f that

is discarded.

EXAMPLE 5.11

Sketch $y = f(|x|)$ for $f(x) = x^2 - 6x + 8$, stating the coordinates of the intercepts, the corner and the turning points.

Keep the half of $y = f(x)$ with $x \geq 0$ and discard the rest. That half carries the y-intercept $(0, 8)$, the roots $(2, 0)$ and $(4, 0)$, and the minimum $(3, -1)$.

Reflecting it in the y-axis copies each of those points across. The point $(0, 8)$ is on the axis and stays where it is; $(2, 0)$ gains a partner at $(-2, 0)$; $(4, 0)$ gains a partner at $(-4, 0)$; and $(3, -1)$ gains a partner at $(-3, -1)$.

So the graph meets the x-axis at $(-4, 0)$, $(-2, 0)$, $(2, 0)$ and $(4, 0)$, and it has two minima, at $(-3, -1)$ and $(3, -1)$, both of value -1. The y-intercept is $(0, 8)$.

The y-axis is where the two halves join, and they do not join smoothly. Just to the right of it the kept half is falling, with gradient -6 at $x = 0$; just to the left the mirror image is rising, with gradient $+6$. The two pieces therefore meet at a **corner** at $(0, 8)$, which is a local maximum of value 8. The result is a W shape, symmetric in the y-axis.

The formula agrees: $f(|x|) = |x|^2 - 6|x| + 8 = x^2 - 6|x| + 8$, which is unchanged when x is replaced by $-x$, so it is even, and the discarded half of f leaves no trace in it.

5.3.6 The Graph of $[f(x)]^2$

Squaring the output also removes the sign, so $[f(x)]^2$ and $|f(x)|$ share a rule: neither is ever negative. They differ in size, because squaring changes the size and a modulus does not.

Every output is squared, so this rule has two parts.

KEY To sketch $y = [f(x)]^2$: no output is negative, so the whole curve lies on or above the x-axis. Each root of f becomes a point where the new curve **touches** the axis instead of crossing it. Where $|f(x)| > 1$ the squared value is larger, so the curve rises. Where $|f(x)| < 1$ the squared value is smaller, so the curve moves closer to the axis.

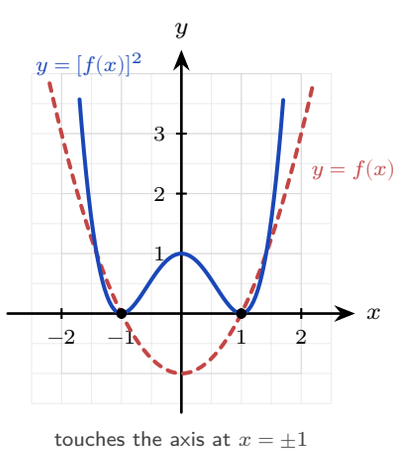


Figure 5.9 $y = [f(x)]^2$ for $f(x) = x^2 - 1$. The roots at $x = \pm 1$ become points where the squared curve

touches the x-axis. For $|x| < \sqrt{2}$, $|f(x)| \leq 1$, so the squared curve lies on or below $y = |f(x)|$; for $|x| > \sqrt{2}$, $|f(x)| > 1$, so it rises faster. The curves $y = |f(x)|$ and $y = [f(x)]^2$ meet where $|f(x)|$ is 0 or 1.

EXAMPLE 5.12

Sketch $y = [f(x)]^2$ for $f(x) = x^2 - 6x + 8$, stating where the curve meets the x-axis and the coordinates of its turning points.

The key points of f are again the roots $(2, 0)$ and $(4, 0)$ and the minimum $(3, -1)$. Squaring acts on the output only, so each of them keeps its x-coordinate.

Each root gives $0^2 = 0$, so the new curve still meets the x-axis at $(2, 0)$ and $(4, 0)$. It cannot cross there, because no output is negative, so at each root it **touches** the axis and turns back: both are minima, of value 0.

The minimum $(3, -1)$ gives $(-1)^2 = 1$, so the curve passes through $(3, 1)$. Between the two roots the squared value rises from 0 to 1 and falls back to 0, which makes $(3, 1)$ a local **maximum**.

Outside $2 \leq x \leq 4$ the size $|f(x)|$ passes 1 and keeps growing, so the squared curve climbs far faster than f does: at $x = 5$, $f(5) = 3$ while $[f(5)]^2 = 9$.

So the sketch drops steeply to touch at $(2, 0)$, arches up to $(3, 1)$, drops back to touch at $(4, 0)$, then climbs steeply away. Inside that arch $|f(x)| \leq 1$, so squaring pulls the curve closer to the axis than $|f(x)|$ lies, and on $2 \leq x \leq 4$ the two curves coincide only where $|f(x)|$ is 0 or 1, at $x = 2$, $x = 3$ and $x = 4$. Beyond the arch they meet once more on each side, wherever $|f(x)| = 1$ again: solving $x^2 - 6x + 8 = 1$ gives $x = 3 \pm \sqrt{2}$, one point below $x = 2$ and one above $x = 4$. Squaring and taking a modulus agree exactly at height 0 and height 1, and nowhere else.

5.3.7 The Graph of $f(ax + b)$

The last of these graphs uses the horizontal rules of §5.2.4. Everything in $f(ax + b)$ happens inside the bracket, so every change is horizontal. Factorise the bracket first, so that the stretch and the translation can be read off separately.

KEY $f(ax + b) = f\left(a\left(x + \frac{b}{a}\right)\right)$: a horizontal stretch of scale factor $\frac{1}{a}$, then a horizontal translation of $-\frac{b}{a}$.

Tracking a few key points, as below, confirms the order.

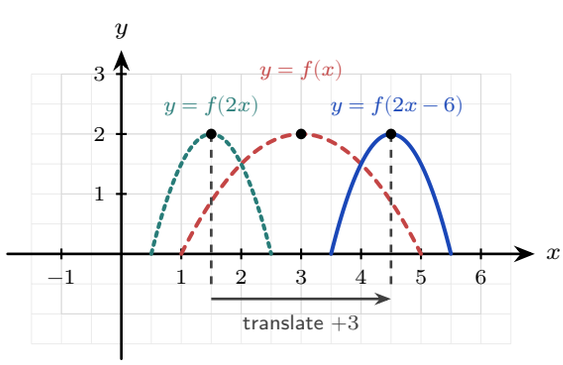


Figure 5.10 $y = f(2x-6)$, read as $f(2(x-3))$. The stretch acts first: $f(2x)$ halves the width and pulls the peak from $x=3$ to $x=1.5$. The translation of 3 to the right then carries it to $x=4.5$. Reading the -6 as a shift of 6 is a common error: the shift is $-b/a=3$, and it is applied after the stretch, not before.

EXAMPLE 5.13

Sketch $y = f(2x-6)$ for $f(x) = x^2 - 6x + 8$, stating the images of the roots and of the vertex.
Factorise inside the bracket before reading anything off:

$$f(2x-6) = f(2(x-3)).$$

Here $a=2$ and $b=-6$, so the moves are a horizontal stretch of scale factor $\frac{1}{2}$ followed by a translation of $-\frac{b}{a}=3$ to the right. Both are horizontal, so no y-coordinate changes.
Now track the key points of f: the roots $(2,0)$ and $(4,0)$ and the vertex $(3,-1)$. The stretch halves each x-coordinate and the translation then adds 3:

$$2 \to 1 \to 4, \qquad 4 \to 2 \to 5, \qquad 3 \to \tfrac{3}{2} \to \tfrac{9}{2}.$$

So the roots move to $(4,0)$ and $(5,0)$, and the vertex moves to $(\frac{9}{2},-1)$: still a minimum, still of depth -1, and still halfway between the roots. The curve is a parabola through those three points, half as wide as f, with y-intercept $f(-6) = 36+36+8 = 80$.
Expanding confirms every number:
$f(2x-6) = ((2x-6)-2)((2x-6)-4) = (2x-8)(2x-10) = 4(x-4)(x-5)$, whose roots are 4 and 5 and whose value at $x=\frac{9}{2}$ is $4(\frac{1}{2})(-\frac{1}{2}) = -1$.

5.3.8 Modulus Equations and Inequalities

The modulus removes the sign of a number. To solve a modulus equation, consider both possible signs; what sits on the other side decides how.

There are five forms, and they must be separated before any algebra starts. Two are safe, one needs a check at the end, one is read from the V-shaped graph, and the last, where both sides contain x, may have no algebraic route at all.

A modulus equal to a number. If $|A| = k$ and k is positive, then A has size k, so it is k or $-k$. Both cases must be written down.

KEY For $k>0$, $|A| = k$ means $A = k$ or $A = -k$.
For $k=0$ it means $A=0$. For $k<0$ there is no solution, because a modulus is never negative.

EXAMPLE 5.14

Solve $|2x-1| = 5$.

The right side is positive, so there are two cases:

$$2x-1=5 \implies x=3, \qquad 2x-1=-5 \implies x=-2.$$

So $x=3$ or $x=-2$.
The graph agrees. The V-shaped curve $y=|2x-1|$ meets the line $y=5$ once on each arm.

A modulus equal to a modulus. Two sizes are equal when the numbers are equal or opposite.

KEY $|A|=|B|$ means $A=B$ or $A=-B$.
Squaring both sides gives the same two answers, because $|A|^2=A^2$.

EXAMPLE 5.15

Solve $|2x-1|=|x+4|$.

$$2x-1=x+4 \implies x=5, \qquad 2x-1=-(x+4) \implies 3x=-3 \implies x=-1.$$

Both work: $|9|=|9|$ and $|-3|=|3|$. So $x=5$ or $x=-1$.

A modulus equal to an expression in x. This form looks like the first, but it needs an extra check. A modulus is never negative, so the other side cannot be negative either. Solving both cases can therefore produce a value that solves neither.

CAUTION For $|A|=B$, solve $A=B$ and $A=-B$ as usual, then **substitute each answer back** and discard any that makes B negative. These are called **extraneous solutions**. An extraneous solution can come from $A=B$ or from $A=-B$, and it is rejected because it makes B negative. For example, $|2x|=x-3$ gives $x=-3$ from $A=B$ and $x=1$ from $A=-B$; both make $x-3$ negative, so the equation has no solution.

EXAMPLE 5.16

Solve $|x-1|=2x-4$.
Two cases, as before:

$$x-1=2x-4 \implies x=3, \qquad x-1=-(2x-4) \implies 3x=5 \implies x=\tfrac{5}{3}.$$

Now test each against the right side. At $x=3$: $2(3)-4=2$, which is positive, and $|3-1|=2$. That works.
At $x=\frac{5}{3}$: $2(\frac{5}{3})-4=-\frac{2}{3}$, which is negative. A modulus cannot equal a negative number, so this value is rejected.
The only solution is $x=3$.

Inequalities. Read these from the V rather than memorising them. The graph of $y = |A|$ is a V sitting on the x-axis, and a horizontal line $y = k$ cuts it twice. Below the line is the piece between the crossings; above it are the two pieces outside.

KEY For $k > 0$:

$$|A| < k \text{ means } -k < A < k, \qquad |A| > k \text{ means } A < -k \text{ or } A > k.$$

One interval for $<$, two for $>$. Use $\leq$ and $\geq$ to include the endpoints.

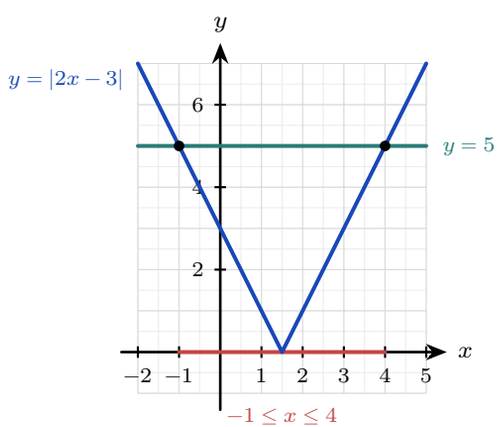


Figure 5.11 The V meets the line $y = 5$ at $x = -1$ and $x = 4$. Between those crossings the V is below the line, so $|2x - 3| \leq 5$ holds on the scarlet stretch. Outside them the V is above the line, so $|2x - 3| > 5$ holds there instead.

EXAMPLE 5.17

Solve (a) $|2x - 3| \leq 5$ and (b) $|3x + 1| > 7$.

(a) One interval, so write the double inequality and work on all three parts at once:

$$-5 \leq 2x - 3 \leq 5 \implies -2 \leq 2x \leq 8 \implies -1 \leq x \leq 4.$$

(b) Two intervals, so write two separate inequalities:

$$3x + 1 > 7 \implies x > 2, \qquad 3x + 1 < -7 \implies x < -\tfrac{8}{3}.$$

So $x < -\frac{8}{3}$ or $x > 2$.

The same two rules work when the expression inside is not linear, such as $|x^2 - 4| < 3$.

When both sides contain x. An inequality such as $|f(x)| > g(x)$ has no rule of its own, because the number of intervals depends on the two curves. The method is the one from §5.4.5: find where the two sides are equal, and those points cut the axis into intervals on which the answer cannot change.

KEY To solve $|f(x)| > g(x)$ or $|f(x)| < g(x)$:

1. Sketch $y = |f(x)|$ and $y = g(x)$ on the same axes.
2. Find every point where they meet, by solving $|f(x)| = g(x)$ or from a calculator.
3. Read off the intervals where the curve you want is the higher one.

TIP Many modulus inequalities have no algebraic solution at all. The syllabus example $|3x \arccos x| > 1$ is one. Graph both sides on your GDC, use intersect to find the crossing points, and read the intervals from the picture. State the intersection coordinates you used, to three significant figures.

The five forms are summarised below.

Form	Method	Watch for
$\|A\| = k$	$A = k$ or $A = -k$	no solution if $k < 0$
$\|A\| = \|B\|$	$A = B$ or $A = -B$	nothing; both roots are valid
$\|A\| = B$	two cases, then substitute back	reject any root making B negative
$\|A\| \lessgtr k$	$-k < A < k$, or two intervals	one interval for $<$, two for $>$
$\|f(x)\| \lessgtr g(x)$	find where the sides are equal, then test each interval	often needs a GDC

YOUR TURN 5.3 Symmetry, Inverses and Related Graphs HL

Odd and even functions

18. Determine whether each function is odd, even, or neither.

(a) $f(x) = x^4 - 3x^2$

(b) $f(x) = x^3 + x$

(c) $f(x) = x^2 + x$

(d) $f(x) = x^5$

(e) $f(x) = 5$

(f) $f(x) = \dfrac{1}{x}$

Self-inverse functions

19. A function f is self-inverse when $f(f(x)) = x$ for every x in its domain.

(a) Show that $f(x) = \dfrac{1}{x}$ is self-inverse.

(b) Show that $f(x) = 6 - x$ is self-inverse.

(c) Show that $f(x) = \dfrac{2x+1}{x-2}$ is self-inverse.

(d) Explain why $f(x) = 2x$ is not self-inverse.

(e) Explain why the graph of a self-inverse function is symmetric in the line $y = x$.

(f) Find the value of a for which $f(x) = \dfrac{4x+1}{2x+a}$ is self-inverse.

Restricting a domain

20. In each part, a domain is restricted so that an inverse exists.

(a) State a domain on which $f(x) = x^2$ has an inverse, and give that inverse.

(b) Do the same for $f(x) = (x-1)^2$.

(c) State the domain of the inverse you found in part (b).

(d) Explain why $f(x) = x^3$ needs none.

The graphs of $|f(x)|$, $f(|x|)$ and $[f(x)]^2$

21. The graph of $y = f(x)$ crosses the x-axis at $x = 2$, has no other root, and has a minimum point at $(0, -4)$.

(a) State where $y = |f(x)|$ meets the x-axis.

(b) State where $y = [f(x)]^2$ meets the x-axis.

(c) State the point that $(0, -4)$ becomes on $y = |f(x)|$.

(d) State the roots of $y = f(|x|)$ and the coordinates of its minimum point.

22. Let $f(x) = x^2 - 2x - 3$. On separate axes, sketch each graph, stating the coordinates of every intercept and turning point.

(a) $y = |f(x)|$

(b) $y = f(|x|)$

(c) $y = [f(x)]^2$

Modulus equations and inequalities

23. Solve each equation.

(a) $|x-3|=2$ (b) $|2x+1|=7$
(c) $|3-x|=1$ (d) $|x|=5$

24. Solve each equation. One part has no solution.

(a) $|x-2|=|2x+1|$ (b) $|3x|=|x+4|$
(c) $|x-1|=3x-5$ (d) $|2x+1|=x-1$

25. Solve each inequality.

(a) $|x-4|<3$ (b) $|2x+5|\geq 9$
(c) $|3-x|\leq 1$ (d) $|x^2-9|<7$

26. GDC Both sides now contain x.

(a) Sketch $y=|x-2|$ and $y=x$ on the same axes, and hence solve $|x-2|<x$.
(b) Solve $|2x|>x+3$ algebraically.
(c) Use a GDC to solve $|x^3-3x|>1$ for $0\leq x\leq 3$, giving answers to three significant figures.

27. Solve $|x^2-4|<3$.

Concept check

28. A student writes: "$f(x)=x^3+1$, *so* $f(-x)=-x^3+1\neq f(x)$. *The function is not even, so it is odd.*" Identify the error and classify f correctly.

29. A student solves $|x-3|=2x$ and gives the answer "$x=-3$ *or* $x=1$". Identify the error and give the correct solution.

5.4 Reciprocal and Rational Functions

A line, a parabola and a cubic have a value at every input, and each one rises or falls without limit. Divide by a variable and both of those stop being true. There are now inputs where the function has no value, and there are lines, called asymptotes, that the curve approaches. Finding those lines from the formula is what makes a rational curve quick to sketch. The same work also tells you where the function is positive and where it is negative.

The simplest function of this kind is the **reciprocal** $y=\dfrac{1}{x}$, and it shows all of the new behaviour at once.

It is undefined at $x=0$. As x approaches 0, the output grows without bound, towards $+\infty$

on one side and $-\infty$ on the other. For large $|x|$ the opposite happens: the output becomes very small, and the curve approaches zero. So the reciprocal has two asymptotes, the vertical line $x = 0$ and the horizontal line $y = 0$.

It is defined for every input except $x = 0$, so its **domain** is all real numbers with $x \neq 0$. It produces every output except 0, so its **range** is all real numbers with $y \neq 0$. It is also self-inverse, so its graph is symmetric in the line $y = x$. See §5.3.

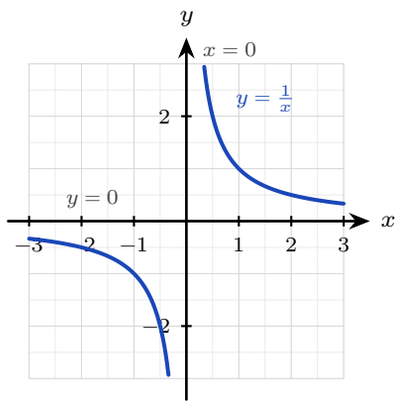


Figure 5.12 The reciprocal $y = \frac{1}{x}$. Its asymptotes are the two axes: the curve approaches $x = 0$ and $y = 0$ without meeting either. The domain is every real number except 0, and so is the range.

DEF An **asymptote** is a line that the graph of a function approaches arbitrarily closely (as close as you like). A **vertical asymptote** occurs at a value of x where the function is undefined and grows without bound. A **horizontal asymptote** is a value the function approaches as $x \to \pm\infty$.

5.4.1 Linear over Linear

Every function of the form $\dfrac{ax + b}{cx + d}$ is a reciprocal curve that has been shifted and stretched, so all of them have the same two-branch shape. Dividing the numerator by the denominator shows that directly: it rewrites the function as a constant plus a multiple of $\dfrac{1}{x - p}$, and the two asymptotes then cross at the centre of that reciprocal curve. Two conditions are needed for that to be true. If $c = 0$ the function is linear and has no asymptote at all. If $ad - bc = 0$ the numerator is a multiple of the denominator and the function is constant.

KEY For $f(x) = \dfrac{ax + b}{cx + d}$, with $c \neq 0$ and $ad - bc \neq 0$:

$$\text{vertical asymptote at } x = -\frac{d}{c}, \qquad \text{horizontal asymptote at } y = \frac{a}{c}.$$

The vertical one is where the denominator is zero. The horizontal one is the ratio of the leading coefficients.

Worked through. To sketch $f(x) = \dfrac{2x+1}{x-1}$, find the vertical asymptote first. The denominator is zero at $x = 1$, so the line $x = 1$ is the vertical asymptote. Next the horizontal asymptote. The ratio of the leading coefficients gives $y = \dfrac{2}{1} = 2$. Rewriting the function confirms the shape. Divide the numerator by the denominator:

$$\frac{2x+1}{x-1} = 2 + \frac{3}{x-1},$$

a reciprocal curve centred on the crossing of the two asymptotes, $(1, 2)$.

Now find the intercepts. Substitute $x = 0$ for the y-intercept: $f(0) = \dfrac{1}{-1} = -1$. Set the numerator to zero for the x-intercept: $2x + 1 = 0$, so $x = -\frac{1}{2}$. Both are marked on the figure below. Two asymptotes and two intercepts fix the whole sketch.

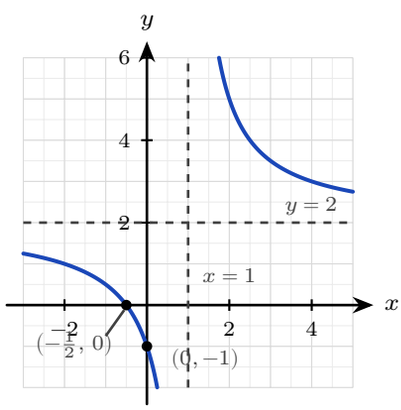


Figure 5.13 The curve $y = \frac{2x+1}{x-1}$, with its two asymptotes dashed and its two intercepts marked. These four labelled features are what a sketch must show.

TIP Marks for a sketch usually depend on labelled asymptotes and intercepts. Write the equation of each asymptote beside its dashed line, and write the coordinates beside each intercept.

The divided form answers a second question, and it answers it for any quotient of two linear expressions. A sketch shows which heights the curve reaches, but a question asking for the range wants a reason rather than a picture, and writing the function as a constant plus a fraction supplies one. Written that way, the function can never equal the constant, because the fraction is never zero. Every other value is reached, which you show by setting $f(x) = k$ and solving for x. So the range of $\dfrac{ax+b}{cx+d}$, with $c \neq 0$ and $ad \neq bc$, is every real number except the horizontal asymptote, $y = \dfrac{a}{c}$. The function below is the one whose inverse was found in §5.1.4.

EXAMPLE 5.18

State the range of $f(x) = \dfrac{3x-1}{x-1}$, $x \neq 1$, justifying your answer.

Divide the numerator by the denominator, the same first move as in the sketch above:

$$f(x) = 3 + \frac{2}{x-1}.$$

The second term is a fraction with numerator 2, so it is never 0, whatever x is. Adding something that is never 0 to 3 never gives 3, so $f(x) \neq 3$ for every x in the domain. Every other value is produced. Set $f(x) = k$ and solve back for x:

$$\frac{2}{x-1} = k - 3 \implies x - 1 = \frac{2}{k-3} \implies x = 1 + \frac{2}{k-3},$$

which is a real input, and never equal to 1, whenever $k \neq 3$. So the range is $y \in \mathbb{R}$, $y \neq 3$. The one excluded value is the horizontal asymptote. A linear-over-linear curve never takes that value, so this curve never reaches $y = 3$.

5.4.2 Reading the Asymptotes from the Degrees

The two rules above are one case of a general pattern. Which asymptote a rational function has as $x \to \pm\infty$ depends only on how the degree of the numerator compares with the degree of the denominator.

Degrees	Asymptote as $x \to \pm\infty$	Example
numerator $<$ denominator	horizontal, $y = 0$	$\frac{x+1}{x^2-4}$
numerator $=$ denominator	horizontal, at the ratio of leading coefficients	$\frac{2x+1}{x-1}$
numerator $=$ denominator $+1$	oblique, a slanted line	$\frac{x^2-2x+3}{x-1}$

In every case the vertical asymptotes sit at the zeros of the denominator, once any common factor has been cancelled.

A horizontal or oblique asymptote describes the curve only for large $|x|$. Nearer the origin the curve may cross it. A curve cannot cross a vertical asymptote, because the function is undefined there.

The syllabus names one more shape: a linear numerator over a quadratic denominator, $\frac{ax+b}{cx^2+dx+e}$. The denominator has the higher degree, so the horizontal asymptote is $y = 0$, and each real root of the denominator gives its own vertical asymptote. This form can therefore have two.

With two asymptotes and an intercept, the branches can still go several ways. The sign of f on each interval shows which way its branch goes. A rational function changes sign only where it is zero or where it is undefined, so those values cut the x-axis into intervals, and the sign is the same right across each one. Testing one point per interval is therefore

enough.

The sign beside a vertical asymptote gives the direction of that branch. Where f is positive the curve climbs to $+\infty$ as it nears the asymptote; where f is negative it falls to $-\infty$. §5.4.5 sets the same work out as a table.

EXAMPLE 5.19

Sketch the key features of $f(x) = \dfrac{x+1}{x^2-4}$.

Factorise the denominator to find the vertical asymptotes: $x^2 - 4 = (x-2)(x+2)$, so there are *two*, at $x = 2$ and $x = -2$.

The numerator has lower degree than the denominator, so the horizontal asymptote is $y = 0$.

For the intercepts, set the numerator to zero to get $x = -1$, and substitute $x = 0$ to get $f(0) = -\frac{1}{4}$.

Now the signs. The three values -2, -1 and 2 cut the axis into four intervals. Take the sign of each factor of $\dfrac{x+1}{(x-2)(x+2)}$, one interval at a time:

$$x < -2 : \ \tfrac{-}{(-)(-)} < 0, \qquad -2 < x < -1 : \ \tfrac{-}{(-)(+)} > 0,$$

$$-1 < x < 2 : \ \tfrac{+}{(-)(+)} < 0, \qquad x > 2 : \ \tfrac{+}{(+)(+)} > 0.$$

Read the branches from those signs. Where f is positive next to a vertical asymptote the curve rises to $+\infty$; where it is negative it falls to $-\infty$. The middle branch is negative to the right of $x = -1$ and positive to the left of it, and it passes through both intercepts. In particular it crosses its horizontal asymptote $y = 0$ at $x = -1$: a horizontal asymptote only describes the behaviour for large $|x|$.

A zero of the denominator does not always produce an asymptote. That is what the phrase "once any common factor has been cancelled" is guarding against. If the same factor divides the numerator too, it cancels, and all that is left at that value of x is one missing point: a **hole** (a removable discontinuity). The function still has no value there, so the domain still excludes it, but the curve on either side lines up.

EXAMPLE 5.20

Describe the behaviour of $f(x) = \dfrac{x^2-9}{x-3}$ at $x = 3$, and state its domain. Compare it with $g(x) = \dfrac{x^2-9}{x+4}$ at $x = -4$.

Factorise the numerator first, because that is what decides the answer:

$$x^2 - 9 = (x+3)(x-3).$$

For f the factor $x - 3$ appears above and below the line, so it cancels:

$$f(x) = \frac{(x+3)(x-3)}{x-3} = x + 3, \qquad x \neq 3.$$

The graph of f is therefore the straight line $y = x + 3$ with a single point removed. That point is at $x = 3$, where the line has height $3 + 3 = 6$, so the hole is at $(3, 6)$. There is no vertical asymptote at all, and the domain is $x \in \mathbb{R}$, $x \neq 3$.

For g the denominator is zero at $x = -4$. Substitute that into the numerator: $(-4 + 3)(-4 - 3) = 7$, which is not 0, so nothing cancels. As x approaches -4 the numerator approaches 7 while the denominator approaches 0, so $|g(x)|$ grows without bound and $x = -4$ is a vertical asymptote.

That substitution decides which of the two you have: put the denominator's zero into the numerator, and if the result is not 0 then no common factor exists and the asymptote stands.

5.4.3 Oblique Asymptotes HL

When the numerator has a higher degree than the denominator, there is no horizontal asymptote. The curve approaches a slanted line instead, and finding that line is a division problem.

The useful case is when the numerator's degree is exactly one more than the denominator's. **Polynomial division** then splits the function into a linear part plus a remainder, and that remainder approaches zero for large $|x|$. The linear part is the **oblique** (slant) asymptote. Polynomial division is treated in full in Ch. 6; all that is needed here is the quotient-plus-remainder form.

The simplest case is $y = x + \dfrac{1}{x}$. It is undefined at $x = 0$, so $x = 0$ is a vertical asymptote. For large $|x|$ the term $\dfrac{1}{x}$ approaches zero, so the curve approaches the oblique asymptote $y = x$.

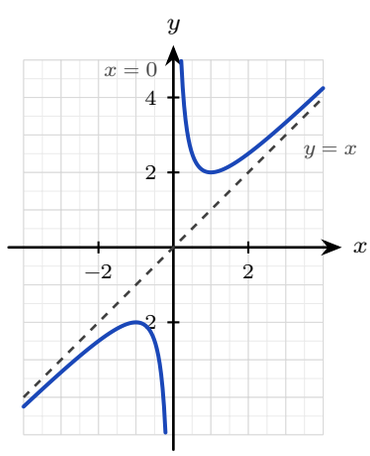


Figure 5.14 The two branches of $y = x + \frac{1}{x}$, each approaching the vertical asymptote $x = 0$ on one side and the oblique asymptote $y = x$ on the other.

There the split into a line plus a remainder was already written out. In general you produce it yourself by **polynomial division**.

EXAMPLE 5.21

Find all asymptotes of $f(x) = \dfrac{x^2 - 2x + 3}{x - 1}$.

The denominator is zero at $x = 1$, so there is a vertical asymptote there. The numerator's degree exceeds the denominator's by one, so there is an oblique asymptote, found by dividing:

$$\frac{x^2 - 2x + 3}{x - 1} = x - 1 + \frac{2}{x - 1},$$

since $(x - 1)(x - 1) + 2 = x^2 - 2x + 3$. As $|x|$ grows the remainder term $\frac{2}{x-1}$ approaches zero, so the curve approaches the line

$$y = x - 1,$$

which is the oblique asymptote.

TIP Show the division, not only the answer. The quotient *is* the asymptote, and the division itself earns method marks.

5.4.4 The Reciprocal of Any Function HL

Every reciprocal so far has been the reciprocal of something simple: of x, of a linear expression, of one polynomial divided by another. The same reading works for the reciprocal of any function at all. Given a function f, its **reciprocal** is

$$y = \frac{1}{f(x)},$$

defined at every x where f is defined and $f(x) \neq 0$.

You do not need to calculate points to sketch it. Every feature of the reciprocal can be read from the graph of f, one feature at a time.

Where f does this	The reciprocal does this	Reason
$f(x) = 0$	vertical asymptote	you cannot divide by zero
$\lvert f(x)\rvert$ **is large**	the curve approaches the x-axis	a large number has a small reciprocal
$f(x) = 1$ **or** -1	the two graphs meet	1 and -1 are their own reciprocals
$f(x) > 0$	$\frac{1}{f(x)} > 0$	the sign is unchanged
a local maximum	a local minimum	and a minimum becomes a maximum

The last row needs a second look. Where f is large the reciprocal is small, so a high point of f becomes a low point of $\frac{1}{f}$. This holds wherever f is not zero, whatever the sign of f.

CAUTION $\dfrac{1}{f(x)}$ is *not* $f^{-1}(x)$. The reciprocal divides 1 by the output; the inverse undoes the function. For $f(x) = x + 3$ the reciprocal is $\dfrac{1}{x+3}$ and the inverse is $x - 3$, which are not the same. The notation f^{-1} never means "one over f".

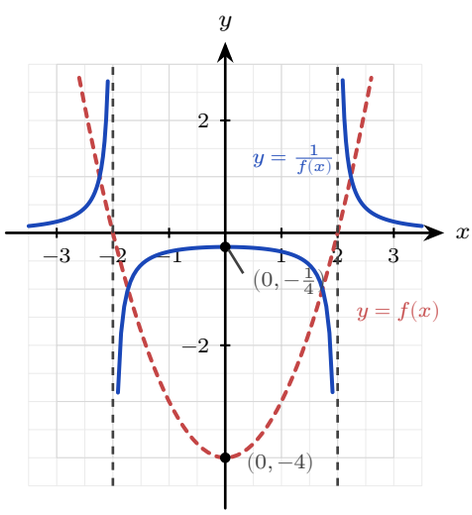


Figure 5.15 $y = \dfrac{1}{x^2 - 4}$ against $y = x^2 - 4$. Zeros of the original become vertical asymptotes at $x = \pm 2$, the minimum $(0, -4)$ becomes the maximum $(0, -\frac{1}{4})$ of the middle branch, and large values of f become values close to zero.

EXAMPLE 5.22

The graph of $y = f(x)$ touches the x-axis at $x = 5$ and has no other root, it has a local minimum at $(2, 4)$, and $f(x) \to +\infty$ as $x \to \pm\infty$. Describe the corresponding features of $y = \dfrac{1}{f(x)}$.

Take the features one at a time.

The root at $x = 5$ makes the denominator zero, so $y = \dfrac{1}{f(x)}$ has a **vertical asymptote at** $x = 5$. Since the curve only touches the axis there, f is positive on both sides of $x = 5$, so the reciprocal runs up to $+\infty$ on both sides of that asymptote.

The minimum at $(2, 4)$ has $f(2) = 4$, so the reciprocal passes through $(2, \frac{1}{4})$. A minimum of f becomes a **maximum** of the reciprocal, so this is a local maximum at $(2, \frac{1}{4})$.

Since $f(x) \to \infty$ at both ends, $\dfrac{1}{f(x)} \to 0$, so $y = 0$ is a **horizontal asymptote**.

No algebra was needed. Every answer came from one feature of f.

TIP Sketching a reciprocal in an examination: mark the roots of f as vertical asymptotes first, then the points where $f = 1$ and $f = -1$, then the turning points with their y-coordinates inverted. Draw the curve through those, one branch at a time.

5.4.5 Rational Inequalities HL

An inequality asks the same question as a sketch, in a shorter form: on which intervals is the expression positive, and on which is it negative?

One move gets it into a shape you can answer. Collect every term on one side, so that the inequality compares a single fraction with zero.

The rest is a **sign table**. A rational expression changes sign only where it is zero or where it is undefined, so those **critical values** cut the x-axis into intervals on which the sign never changes. Give the table one row for each factor and one column for each interval. Fill in the sign of every factor, then multiply down each column to get the sign of the whole expression. The last row is the answer, read off interval by interval.

KEY HL To solve a rational inequality:

1. Collect every term on one side, combine into a single fraction and factorise.
2. Mark the **critical values**, where the numerator or the denominator is zero.
3. Build a sign table on the intervals between them, and read off the intervals with the sign you need.
4. Include a zero of the numerator when the inequality is $\geq$ or $\leq$; never include a zero of the denominator.

CAUTION Do not multiply both sides by an expression whose sign you do not know. Multiplying an inequality by a negative number reverses it, so multiplying by $x-2$ would split the question into two cases, one for each sign of $x-2$. Move everything to one side instead.

The same table works when there is no fraction at all, so the quadratic inequalities of Ch. 4 can be solved this way too.

EXAMPLE 5.23

Solve $\dfrac{x+3}{x-2} \geq 2$.

Collect everything on one side and combine into a single fraction:

$$\frac{x+3}{x-2} - 2 \geq 0 \implies \frac{(x+3)-2(x-2)}{x-2} \geq 0 \implies \frac{7-x}{x-2} \geq 0.$$

Now the critical values. The numerator is zero when $7-x=0$, that is $x=7$. The denominator is zero when $x=2$. These two values cut the axis into three intervals.

Interval	$x < 2$	$2 < x < 7$	$x > 7$
$7 - x$	$+$	$+$	$-$
$x - 2$	$-$	$+$	$+$
$\frac{7-x}{x-2}$	$-$	$+$	$-$

Only the middle interval gives the sign required. Now decide the two endpoints. Exclude $x = 2$, because the fraction is undefined there. Include $x = 7$, because the inequality allows equality and the fraction is zero there. So

$$2 < x \leq 7.$$

A graph agrees: $y = \dfrac{x+3}{x-2}$ lies above the line $y = 2$ only between the asymptote and the crossing point. The boundary of the solution is a vertical asymptote at one end and a crossing at the other, and a sign table finds both.

TIP A GDC can solve an equation or inequality of this kind numerically once you move everything to one side. Rewrite $f(x) = g(x)$ as $f(x) - g(x) = 0$, graph the single function $y = f(x) - g(x)$, and read its zeros; for an inequality, read the intervals where that curve is above or below the axis. Reading zeros is another method, an alternative to locating two intersection points, and either method works when an equation has no exact solution, such as $e^x = 3 - x$.

YOUR TURN 5.4 Reciprocal and Rational Functions

Linear over linear

30. State the vertical and horizontal asymptotes of each function.

(a) $f(x) = \dfrac{1}{x-3}$

(b) $f(x) = \dfrac{3}{x+2}$

(c) $f(x) = \dfrac{2x}{x+1}$

(d) $f(x) = \dfrac{x+1}{x-2}$

(e) $f(x) = \dfrac{5}{2x-1}$

(f) $f(x) = \dfrac{4x+1}{2x-3}$

31. For each function, state the domain, the range, both asymptotes and both intercepts. Where an intercept does not exist, say why.

(a) $f(x) = \dfrac{1}{x-5}$

(b) $f(x) = \dfrac{1}{x} + 4$

(c) $f(x) = \dfrac{x-1}{x+2}$

(d) $f(x) = \dfrac{2x+3}{x-1}$

32. Consider $f(x) = \dfrac{3x-6}{x+1}$.

(a) Write $f(x)$ in the form $A + \dfrac{B}{x+1}$.

(b) Hence state the point at which the two asymptotes cross.

(c) Sketch the curve, showing both asymptotes and both intercepts.

Asymptotes and holes

33. For each function, state every vertical asymptote, and state the kind and the equation of the asymptote as $x \to \pm\infty$.

(a) $f(x) = \dfrac{1}{(x-1)(x+3)}$

(b) $f(x) = \dfrac{x}{x^2-9}$

(c) $f(x) = \dfrac{x^2+1}{x^2-4}$

(d) $f(x) = \dfrac{x^2+1}{x-2}$

(e) $f(x) = \dfrac{x^2-9}{x+2}$

34. For each function, describe the graph at every value of x where the denominator is zero, and state the domain.

(a) $f(x) = \dfrac{x^2-4}{x-2}$

(b) $f(x) = \dfrac{x^2+x-6}{x+3}$

(c) $f(x) = \dfrac{x^2-1}{x^2+x}$

Oblique asymptotes

35. Write each function as a polynomial plus a remainder, and state the oblique asymptote.

(a) $f(x) = \dfrac{x^2+3}{x}$

(b) $f(x) = \dfrac{x^2}{x-1}$

(c) $f(x) = \dfrac{2x^2-x+3}{x-1}$

(d) $f(x) = \dfrac{x^2+4x+1}{x+1}$

36. Consider $f(x) = x + \dfrac{2}{x}$.

(a) State the vertical asymptote.

(b) State the oblique asymptote.

(c) Explain why the curve never crosses the oblique asymptote.

(d) Sketch the two branches.

37. Consider $f(x) = x + \dfrac{1}{x}$.

(a) State the vertical asymptote.

(b) State the oblique asymptote.

(c) State whether each branch lies above or below the oblique asymptote, and give a reason.

The reciprocal of a function

38. For each function f, sketch $y = \dfrac{1}{f(x)}$, showing every asymptote and the coordinates of any turning point.

(a) $f(x) = x - 4$

(b) $f(x) = x^2 - 1$

(c) $f(x) = x^2 + 1$

(d) $f(x) = (x-1)(x-3)$

39. The graph of $y = f(x)$ crosses the x-axis at $x = -2$ and at $x = 2$ and has no other root. It has a local maximum at $(0, 4)$, and $f(x) \to -\infty$ as $x \to \pm\infty$.

(a) State the equations of the vertical asymptotes of $y = \dfrac{1}{f(x)}$.

(b) Find the coordinates of the turning point of $y = \dfrac{1}{f(x)}$ at $x = 0$, and state its type.

(c) Describe the behaviour of $y = \dfrac{1}{f(x)}$ as $x \to \pm\infty$.

(d) Given that $f(x) = 4 - x^2$, find the x-coordinates of the points where $y = f(x)$ and $y = \dfrac{1}{f(x)}$ meet.

40. Let $f(x) = x - 2$.

(a) On one set of axes, sketch $y = f(x)$ and $y = \dfrac{1}{f(x)}$.

(b) Find the coordinates of the two points where the graphs meet. *(Hint: they meet where $f(x) = \pm 1$)*

(c) State the interval on which $\dfrac{1}{f(x)} > f(x)$ for $x > 2$.

Rational inequalities

41. Solve each inequality. HL

(a) $x^2 \leq x + 6$

(b) $x^2 > 2x + 3$

(c) $\dfrac{x-4}{x+1} \geq 0$

(d) $\dfrac{3}{x-2} > 1$

(e) $\dfrac{x+2}{x-1} \geq 3$

(f) $\dfrac{2x-1}{x+2} \leq 1$

42. Solve $x^2 \leq 3x + 10$ using a sign table.

Concept check

43. A student solves $\dfrac{x+1}{x-3} \geq 2$ by multiplying both sides by $x - 3$: "$x + 1 \geq 2x - 6$, *so* $x \leq 7$." Identify the error and give the correct solution.

44. A student says: "*a graph never crosses its horizontal asymptote.*" Give a counter-example, and explain why the claim is false.

5.5 Exponential and Logarithmic Functions

Pour a cup of coffee at $90\,^\circ$C and set it down in a room at $20\,^\circ$C. It cools quickly at first, then more slowly, then so slowly that you can barely measure the change. Look at the numbers rather than the coffee. In each fixed interval of time the difference between the cup and the room falls by the same *fraction*, not by the same amount. Isaac Newton measured that and found the rule. No polynomial behaves this way.

An **exponential function** $y = a^x$ does, because it places the variable in the exponent rather than in the base. That one change has a large consequence: an exponential grows at a rate proportional to its own size, so a quantity twice as large grows twice as fast. The same function therefore describes cooling coffee, growing populations and decaying atoms. Ch. 1 sets up the algebra of exponents and logarithms. This section is about their graphs.

For $a > 1$ the graph rises, and rises ever more steeply: this is **growth**. For $0 < a < 1$ it falls towards zero: this is **decay**. Either way the graph passes through $(0, 1)$, stays positive everywhere, and approaches the horizontal asymptote $y = 0$ on one side. The domain is all real numbers and the range is $y > 0$.

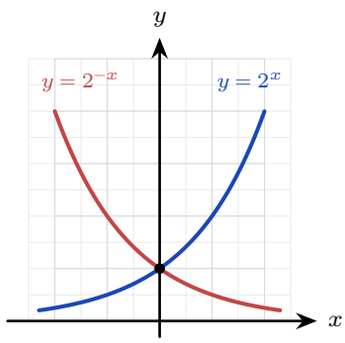


Figure 5.16 Growth and decay are reflections of each other in the y-axis. Both pass through $(0, 1)$ and approach $y = 0$, but on opposite sides.

One base is used most often in calculus: $e \approx 2.718$, the base whose curve has steepness equal to its height at every point. That property makes the calculus of e^x the simplest of any exponential. It is also why every other exponential is usually rewritten in terms of e. See Ch. 13.

KEY · IB Any exponential can be written with base e, and exponential and logarithm undo each other:

$$a^x = e^{x \ln a}, \qquad \log_a(a^x) = x, \qquad a^{\log_a x} = x \quad (a, x > 0,\ a \neq 1).$$

Worked through. For $f(x) = 5(0.4)^x$, look at the base first. Here $a = 0.4$, and $0 < 0.4 < 1$, so the function **decays**. The factor 5 stretches the graph vertically and changes nothing else, and the outputs of $(0.4)^x$ are always positive. So the domain is all real numbers, the range is $y > 0$, and the horizontal asymptote is $y = 0$.

5.5.1 Transforming an Exponential Graph

An exponential often appears in a model as

$$y = k\,a^x + c,$$

and each constant has a separate effect. The factor k stretches the graph vertically. The constant c moves it up or down.

The vertical shift is the one that changes the asymptote. The graph of a^x approaches $y = 0$, so adding c makes it approach $y = c$ instead. **The horizontal asymptote moves with the graph.**

EXAMPLE 5.24

Sketch $y = 3(2)^x - 5$, stating the asymptote and both intercepts.

Start with the asymptote. The -5 shifts the whole graph down by 5, so the horizontal

asymptote is $y = -5$, not $y = 0$.
For the y-intercept, substitute $x = 0$:

$$y = 3(2)^0 - 5 = 3 - 5 = -2.$$

For the x-intercept, set $y = 0$ and solve:

$$3(2)^x = 5 \implies 2^x = \tfrac{5}{3} \implies x = \log_2 \tfrac{5}{3} \approx 0.737.$$

The curve therefore rises from just above $y = -5$ on the left. It crosses the axes at $(0, -2)$ and at $(0.737, 0)$, then rises steeply.

CAUTION A shifted exponential does *not* have the asymptote $y = 0$. In $y = k\,a^x + c$ the asymptote is $y = c$. Reading it as $y = 0$ without checking is a common error.

5.5.2 Exponential Models

Many real quantities are modelled with base e and time as the variable:

$$A(t) = A_0\, e^{kt},$$

where A_0 is the amount at $t = 0$. The sign of k decides the behaviour, $k > 0$ for growth and $k < 0$ for decay, and its size decides the speed. Ch. 1 solves these models algebraically; what is added here is reading them off a graph.

Two data points fix the whole model. If one point is at $t = 0$, it gives A_0 directly. For the second, substitute its t and A into $A = A_0 e^{kt}$, divide by A_0 and take logarithms:

$$k = \frac{1}{t} \ln\left(\frac{A}{A_0}\right).$$

CAUTION A model fitted to two points assumes the fractional rate of change never varies. Where the real rate is itself changing, the prediction becomes less reliable, so report such a figure as what would happen if nothing changed rather than as a forecast.

EXAMPLE 5.25

Japan's population was about 128.1 million in 2010 and about 126.1 million in 2020. Model the population as $P(t) = P_0 e^{kt}$, with t in years after 2010, and use it to estimate the population in 2030.
Take $P_0 = 128.1$, the value at $t = 0$, then use the 2020 figure to find k. Substitute $t = 10$ and $P = 126.1$:

$$126.1 = 128.1\, e^{10k} \implies k = \frac{1}{10} \ln\left(\frac{126.1}{128.1}\right) \approx -0.00157.$$

The sign is negative, which confirms decay: the population falls by about 0.16% each year. For 2030, substitute $t = 20$:

$$P(20) = 128.1\, e^{-0.00157 \times 20} \approx 124.1 \text{ million.}$$

Japan's rate of decline is itself increasing, so official projections for 2030 are lower than this figure.

5.5.3 The Logarithm as an Inverse

Since $y = a^x$ is one-to-one, it has an inverse, and that inverse is the **logarithm** $y = \log_a x$. It answers the reversed question: not "what is a to the x?" but "to what power must I raise a to get x?" Everything about its graph follows from being the mirror image of the exponential in the line $y = x$.

Reflecting swaps the axes' roles. The exponential's horizontal asymptote $y = 0$ becomes the logarithm's *vertical* asymptote $x = 0$; its domain of all reals becomes the logarithm's range, and its range of positive numbers becomes the logarithm's domain. So $\log_a x$ has domain $x > 0$ and range all real numbers; for $a > 1$ it passes through $(1, 0)$ and rises without bound, ever more slowly.

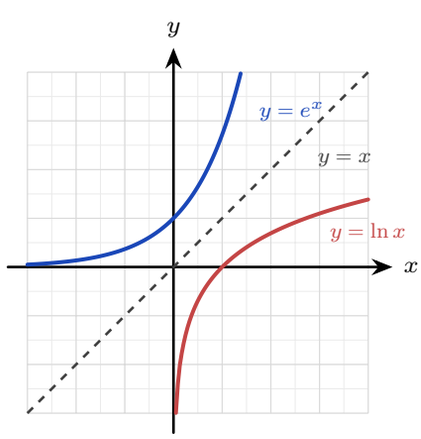


Figure 5.17 The logarithm is the exponential reflected in the line $y = x$. Reflecting exchanges the two axes, which is why one asymptote is horizontal and the other vertical.

CAUTION The logarithm is only defined for positive inputs. Writing $\ln(x - 3)$ requires $x > 3$; forgetting this domain restriction is a common error.

FINDING the domain

Because a logarithm accepts only positive inputs, every logarithmic function carries a hidden condition. To find the domain, set whatever sits inside the logarithm greater than zero, and solve.

EXAMPLE 5.26

State the domain of $f(x) = \ln(x-3)$ and of $g(x) = \log(5-2x)$.
For f, the input to the logarithm is $x-3$. Require it to be positive:

$$x - 3 > 0 \implies x > 3.$$

For g, the input is $5-2x$. Require that to be positive:

$$5 - 2x > 0 \implies 5 > 2x \implies x < \tfrac{5}{2}.$$

Note the direction. Dividing by -2 would have reversed the inequality, so it is safer to move the term across as above.

Transforming a logarithmic graph

The same shifts apply as for any function. A horizontal shift moves the *vertical* asymptote; a vertical shift leaves it where it is.

In $y = \ln(x-h)$ the change is inside the bracket, so it moves the graph h units right. The asymptote travels with it, from $x = 0$ to $x = h$.

The x-intercept moves with it too, and one fact locates it. A logarithm is zero exactly when its input is 1, so $\ln(x-h) = 0$ when $x - h = 1$, that is at $x = h + 1$.

EXAMPLE 5.27

Sketch $y = \ln(x-5)$, stating its asymptote, its domain and its x-intercept.
The input must be positive, so $x - 5 > 0$ and the domain is $x > 5$.
That same condition gives the asymptote. As x approaches 5 from above, $x - 5$ approaches 0 and $\ln(x-5)$ decreases without bound, so $x = 5$ is a vertical asymptote.
For the x-intercept, set $y = 0$. A logarithm is zero when its input is 1:

$$x - 5 = 1 \implies x = 6.$$

So the curve rises from the asymptote at $x = 5$, crosses the axis at $(6, 0)$, and continues upward ever more slowly.

Using the inverse relationship

Because the exponential and the logarithm undo each other, you can use a logarithm to solve for a variable in an exponent, and an exponential to solve for a variable inside a logarithm. That is the step that makes the method of §5.1 work here: at step 3, where you solve for y, apply a logarithm if y sits in an exponent and an exponential if y sits inside a logarithm.

EXAMPLE 5.28

Find the inverse of $f(x) = 2e^{3x} - 5$, and state its domain.
Write $y = f(x)$ and exchange the letters:

$$x = 2e^{3y} - 5 \implies \frac{x+5}{2} = e^{3y}.$$

The variable is now in an exponent, so apply the natural logarithm to both sides:

$$3y = \ln\left(\frac{x+5}{2}\right) \implies f^{-1}(x) = \tfrac{1}{3}\ln\left(\frac{x+5}{2}\right).$$

Now the domain. The domain of f^{-1} is the range of f. Since $e^{3x} > 0$ for every x, we have $2e^{3x} > 0$ and so $f(x) > -5$. The domain of f^{-1} is therefore $x > -5$, which agrees with requiring $\frac{x+5}{2} > 0$ inside the logarithm.

TIP State the domain whenever a question asks for an inverse involving a logarithm. It is a separate mark, and the quickest check is that **the domain of f^{-1} is the range of f**.

YOUR TURN 5.5 Exponential and Logarithmic Functions

Exponential and logarithmic graphs

45. Evaluate each expression.

(a) 2^{-2}
(b) e^0
(c) $\log_2 8$
(d) $\log_5 125$
(e) $\ln 1$
(f) $\ln e$

46. State whether each function grows or decays, and give a reason.

(a) $y = (0.5)^x$
(b) $y = 3^x$
(c) $y = 5e^{-0.2x}$
(d) $y = 2^{-x}$

47. For each graph, state its asymptote and the coordinates of every point where it crosses an axis.

(a) $y = 5 \cdot 2^x$
(b) $y = 2^x + 5$
(c) $y = 4(3)^x - 1$
(d) $y = \ln x$
(e) $y = \ln(x - 2)$
(f) $y = \log(x + 5)$

48. Sketch each graph, stating its asymptote and the exact coordinates of both intercepts.

(a) $y = 2(3)^x - 6$
(b) $y = 8 - 2^{x+1}$

(c) $y = 5e^{-x} - 1$

Base e and exponential equations

49. Write each in the required form.

(a) 2^x in terms of base e

(b) 5^x in terms of base e

50. Solve each equation, giving exact answers.

(a) $2^x = 8$

(b) $\log_3 x = 2$

(c) $5 \cdot 3^x = 45$

(d) $\ln x = 3$

(e) $3(2)^x = 4$

(f) $\log_2(x-1) = 4$

Domains and inverses

51. State the domain of each function.

(a) $y = \ln(3x+1)$

(b) $y = \log(10 - 2x)$

(c) $y = \ln(2x - 6)$

(d) $y = \ln(x+1) + \ln(3-x)$

(e) $y = \ln(x^2 - 4x)$

(f) $y = \ln(x^2 - 1)$

52. Find the inverse of each function, and state its domain.

(a) $f(x) = e^x + 2$

(b) $f(x) = \ln(x-1)$

(c) $f(x) = 3e^{2x} + 1$

Exponential models

53. GDC Two exponential models.

(a) A population is modelled by $P(t) = 800e^{0.03t}$, with t in years. State the initial population $P(0)$, and say whether the population grows or decays.

(b) Find the population after 10 years.

(c) Find how long it takes the population to reach 1200.

(d) A radioactive sample decays according to $A(t) = 50e^{-0.2t}$, with t in days and A in grams. Find $A(6)$, correct to three significant figures.

(e) Find the half-life of the sample.

54. GDC A colony of bacteria numbers 500 at $t = 0$ and 1800 after 4 hours. It is modelled by $N(t) = N_0 e^{kt}$, with t in hours.

(a) Find the exact value of k, and give it correct to three significant figures.

(b) Estimate the size of the colony after 6 hours.

(c) Find how long it takes the colony to reach $10\,000$.

55. GDC A machine is worth $24\,000$ dollars when new and $15\,000$ dollars three years later. Its value after t years is modelled by $V(t) = V_0 e^{kt}$.

(a) Find k, correct to three significant figures, and explain what its sign tells you.
(b) Find the value of the machine after 5 years.
(c) Find the time it takes for the value to halve.

56. GDC After a sauna is switched off, its temperature in °C after t minutes is $T(t) = 20 + 70e^{-0.12t}$.

(a) State the temperature at the moment it is switched off.
(b) Find the temperature after 5 minutes.
(c) Find how long it takes to cool to $40\,°\text{C}$.
(d) State the temperature the sauna approaches, and explain what it represents.

Equations that need a graph

57. GDC Consider the equation $e^x = 3 - x$.

(a) Explain why it has exactly one real solution.
(b) Find that solution, correct to three significant figures.
(c) State why a graph is the appropriate method for this equation.

Concept check

58. A student states that the graph of $y = 3 \times 2^x - 5$ has the asymptote $y = 0$ and the y-intercept $(0, -5)$. Identify both errors, and give the correct asymptote and y-intercept.

Chapter 5 · Functions & Graphs

Reflections

TOK Slide the graph of $y = x^2$ two units right and you get $y = (x-2)^2$. Is this the same function or a different one? It has the same shape, the same name ("a parabola"), yet a different rule and different values everywhere. When we call two things "the same" in mathematics, we are choosing what to ignore. The transformation makes that choice visible.

IM In 1872 Felix Klein, aged just twenty-three, proposed in his *Erlangen Programme* a radical reorganisation of geometry: a geometry *is* a set of transformations together with the properties they leave unchanged. Euclidean geometry studies what is unchanged by translations, rotations and reflections; other geometries allow other moves. The transformations of this chapter are the same idea in miniature, and the question "what stays the same when I move this graph?" is a central one in mathematics.

RW Exponential decay is used to measure time in physics and archaeology. Every radioactive isotope falls by half over a fixed *half-life*, from carbon-14's $5{,}730$ years, used to date ancient wood and bone, to isotopes that decay in seconds. The same exponential describes unrestricted population growth and compound interest. When growth is limited by a maximum, it becomes the logistic curve of a spreading rumour or an epidemic. Rational functions model saturation: reaction rates, drug concentrations, and average costs that flatten toward a limit.

EXT Some functions undo themselves: $f(f(x)) = x$ for every x in the domain, so $f^{-1} = f$, and the graph is its own reflection in $y = x$. A function like this is called an **involution**. Examples are $f(x) = \dfrac{1}{x}$, $f(x) = 5 - x$ and every $f(x) = \dfrac{ax + b}{cx - a}$ with $c \neq 0$ and $a^2 + bc \neq 0$. The quarter circle $f(x) = \sqrt{1 - x^2}$, $0 \leq x \leq 1$, is another. Can you explain from its graph why it must be its own inverse, and find an involution of your own that is neither a line nor a hyperbola?

End-of-Chapter Problems
Chapter 5: Functions & Graphs

1. Let $f(x) = 2x - 3$ and $g(x) = x^2$.

(a) Find $(f \circ g)(x)$. [2]
(b) Find $(g \circ f)(x)$. [2]
(c) Solve $(f \circ g)(x) = 5$. [2]

2. The graph of $y = f(x)$ has a maximum point at $(1, 4)$. State the coordinates of the maximum of:

(a) $y = f(x) + 2$; [1]
(b) $y = f(x - 3)$; [1]
(c) $y = 2f(x)$; [1]
(d) $y = -f(x)$ (state the resulting minimum). [1]

3. For each quadratic, write $f(x)$ in completed-square form and describe the transformations that map $y = x^2$ onto $y = f(x)$.

(a) $f(x) = x^2 - 4x + 5$; state also the range of f. [5]
(b) $f(x) = x^2 - 6x$; state also the vertex. [5]

4. Find the inverse of each function, and state its domain.

(a) $f(x) = \dfrac{x+1}{x-2}, \quad x \neq 2.$ [4]
(b) $f(x) = x^2 + 2, \quad x \geq 0.$ [4]
(c) $f(x) = e^x.$ [3]

5. Solve each inequality.

(a) $x^2 + 2 \leq 3x$ [4]
(b) $|2x - 1| < 5$ [3]
(c) $\dfrac{2x-1}{x+2} \leq 1$ [5]

6. Determine whether each function is odd, even, or neither.

(a) $f(x) = x^3 - 2x$ [1]
(b) $f(x) = |x|$ [1]
(c) $f(x) = x^2 + x^3$ [1]
(d) $f(x) = \dfrac{x}{x^2+1}$ [1]
(e) $f(x) = \dfrac{1}{x^2+1}$ [1]
(f) $f(x) = x + 1$ [1]

7. Let $f(x) = x^2 - 2x$.

(a) Find the roots and the vertex of f. [3]
(b) Sketch $y = |f(x)|$, showing the intercepts and the turning point. [3]
(c) Sketch $y = |x^2 - 1|$, showing the x- and y-intercepts and the shape. [4]

8. Composites of two functions.

(a) Let $f(x) = x + 2$ and $g(x) = x^2$. Solve $(f \circ g)(x) = (g \circ f)(x)$. [4]
(b) For $f(x) = \dfrac{x}{1+x}$, find $(f \circ f)(x)$ and simplify fully. [5]

9. Describing and applying transformations.

(a) The graph of $y = f(x)$ passes through $(0, 1)$, $(2, 0)$ and $(3, 4)$. Find the images of these three points on $y = f(x + 1) - 2$. [3]
(b) The graph of $y = x^2$ is transformed to $y = -(x - 1)^2 + 2$. Describe the sequence of transformations and state the coordinates of the vertex. [4]

10. Inverses of three functions.

(a) Let $f(x) = 3x + 2$. Find $f^{-1}(x)$, then solve $f(x) = f^{-1}(x)$. [5]
(b) Let $g(x) = \sqrt{x - 1}$. State the domain and range of g, then find $g^{-1}(x)$ and state its domain. [5]
(c) Let $h(x) = \dfrac{2}{x - 3}$. State both asymptotes, then find $h^{-1}(x)$. [5]

11. Solve each equation.

(a) $|x - 2| = |x + 4|$ [4]
(b) $|x| = 2 - x$, and state the coordinates of the point where $y = |x|$ meets $y = 2 - x$. [3]

12. Reasoning from the definitions of odd and even.

(a) The function f is even with $f(3) = 7$, and g is odd with $g(2) = -5$. Write down the values of $f(-3)$ and $g(-2)$. [2]
(b) Prove that if f and g are both odd functions, then $f \circ g$ is also odd. [5]

13. Finding unknown constants from given conditions.

(a) The function $f(x) = ax + b$ satisfies $f(1) = 5$ and $f(3) = 11$. Find a and b, then find $f^{-1}(x)$. [5]
(b) The quadratic $g(x) = x^2 + bx + c$ has vertex $(2, -1)$. Find b and c, then describe the transformation of $y = x^2$ that gives $y = g(x)$. [5]

14. Let $f(x) = (x + 1)^2 - 4$ for $x \geq -1$.

(a) Explain why f has an inverse on this domain. [1]
(b) Find $f^{-1}(x)$, stating its domain. [4]

15.

(a) Solve $|x-2| = |3x+1|$. [4]
(b) Hence, or otherwise, solve $|x-2| \geq |3x+1|$. [3]

16. The function $f(x) = ax + b$ satisfies $f(f(x)) = x$ for all real x. Find all possible values of a and b. [6]

17. Prove that every function $f : \mathbb{R} \to \mathbb{R}$ can be written as the sum of an even function and an odd function. [6]

18. Show that $f(x) = \dfrac{2x+3}{x-2}$ is self-inverse, that is, $(f \circ f)(x) = x$. [5]

19. Let $f(x) = x^2 - 4x + 3$.

(a) Write $f(x)$ in completed-square form, and state the coordinates of the vertex. [3]
(b) Find the roots of f and the y-intercept. [3]
(c) Sketch $y = f(x)$, showing the vertex and both intercepts. [2]
(d) On separate axes, sketch $y = |f(x)|$ and $y = f(|x|)$, showing the key points of each. [4]
(e) The graph of $y = f(x)$ is translated 2 units to the right and 3 units up to give $y = g(x)$. Find $g(x)$ in the form $ax^2 + bx + c$. [3]

20. Let $f(x) = \dfrac{3x-1}{x+2}$ for $x \neq -2$, and let $g(x) = x + 2$.

(a) State the range of f, justifying your answer. [2]
(b) Find $f^{-1}(x)$, and state its domain. [4]
(c) Find $(f \circ g)(x)$, giving your answer as a single fraction. [3]
(d) State the domain of $f \circ g$. [2]
(e) Solve $(f \circ g)(x) = 2$. [4]

21. HL Let $f(x) = x^2 - 4$.

(a) Sketch $y = f(x)$, showing the intercepts. [2]
(b) Write down the equations of the vertical asymptotes of $y = \dfrac{1}{f(x)}$, and explain how you found them. [3]
(c) Find the coordinates of the maximum point of $y = \dfrac{1}{f(x)}$ on the interval $-2 < x < 2$, and explain why it is a maximum even though the value is negative. [3]
(d) Describe the behaviour of $y = \dfrac{1}{f(x)}$ as $x \to \pm\infty$, and state the horizontal asymptote. [2]

(e) Sketch $y = \dfrac{1}{f(x)}$, showing all asymptotes and the point found in part (c). [4]

22. For each function, state both asymptotes and the coordinates of the y-intercept.

(a) $f(x) = \dfrac{1}{x+2}$ [3]

(b) $f(x) = \dfrac{2x+3}{x-1}$ [3]

(c) $f(x) = \dfrac{4x-1}{2x+1}$ [3]

(d) Sketch the graph in part (b), showing both asymptotes and both intercepts. [3]

23. For each function, state the asymptotes and find the inverse.

(a) $f(x) = \dfrac{x-2}{x+3}$; state also the domain of f. [6]

(b) $f(x) = \dfrac{3x}{x-2}$; show that $f^{-1}(x) = \dfrac{2x}{x-3}$. [6]

24. For each function, state every asymptote, and say what happens at the value of x where the denominator is zero.

(a) $f(x) = \dfrac{x^2-1}{x}$; state also the x-intercepts. [5]

(b) $f(x) = \dfrac{x^2+x-2}{x-1}$; simplify the expression first. [5]

25. A bacterial population is modelled by $P(t) = 500 \cdot 2^t$, where t is in hours.

(a) State the initial population. [1]

(b) Find the population after 3 hours. [2]

(c) Find when the population reaches 8000. [3]

26. GDC Two decay models.

(a) A radioactive sample has mass $M(t) = 80 \cdot (0.5)^{t/5}$ grams, with t in days. State the initial mass and the half-life. [3]

(b) Find the mass after 10 days. [2]

(c) The value of a car after t years is $V(t) = 20000(0.85)^t$. State the initial value, and find the value after 4 years. [3]

(d) State the long-term behaviour of $V(t)$, and name the feature of the graph it corresponds to. [2]

27. Solve each equation.

(a) $e^{2x} - 3e^x + 2 = 0$ [5]

(b) $4^x - 2^{x+1} = 8$ [5]

28. Sketch, on the same axes, $y = 2^x$ and $y = 2^{-x}$, stating the horizontal asymptote and the common point. [4]

29. Show that $f(x) = \dfrac{ax+b}{cx+d}$ is self-inverse (that is, $f(f(x)) = x$) if and only if $a + d = 0$, where $c \neq 0$ and $ad - bc \neq 0$. [6]

30. The curve $y = \dfrac{x^2 + ax + b}{x - 1}$ has oblique asymptote $y = x + 3$.

(a) Find a. [3]

(b) Using your value of a, find the value of b for which the zero of the denominator gives a hole at $x = 1$ instead of a vertical asymptote, and describe the resulting graph. [3]

31. HL Consider $f(x) = \dfrac{x^2 - 4x + 5}{x - 2}$.

(a) Show that $f(x) = (x - 2) + \dfrac{1}{x - 2}$. [3]

(b) Write down the equations of the vertical and oblique asymptotes. [3]

(c) Find the y-intercept, and show that the curve has no x-intercept. [3]

(d) Show that the curve never meets its oblique asymptote. [2]

(e) Sketch the curve, showing both asymptotes and the y-intercept. [4]

32. GDC The number of people who have heard a rumour t days after it starts is modelled by

$$N(t) = \frac{5000}{1 + 49e^{-0.8t}}.$$

(a) Find the number of people who had heard the rumour at $t = 0$. [2]

(b) Find $N(5)$, giving your answer to the nearest whole number. [2]

(c) Find the time at which 2500 people have heard the rumour, giving an exact answer and a value to three significant figures. [4]

(d) State the value N approaches as t becomes large, and interpret it in context. [3]

(e) Show that $\dfrac{5000 - N}{N} = 49e^{-0.8t}$, and hence explain why plotting $\ln\left(\dfrac{5000 - N}{N}\right)$ against t gives a straight line. State its gradient and its vertical intercept. [5]

33. Let $f(x) = \ln(2x - 4)$.

(a) State the domain of f. [2]

(b) Find the exact x-intercept of the graph of f. [3]

(c) State the equation of the vertical asymptote, and describe the behaviour of $f(x)$ near it. [3]

(d) Find $f^{-1}(x)$, and state its domain and range. [4]

(e) Describe a sequence of transformations that maps $y = \ln x$ onto $y = \ln(2x - 4)$. [3]

34. Let $f(x) = \log_3(2x - 1)$ and $g(x) = \frac{1}{2}(3^x + 1)$.

(a) State the domain and the range of f. [2]

(b) Show that g is the inverse of f, by showing that $(f \circ g)(x) = x$ for every real x. [3]

(c) State the equation of the line in which the graphs of f and g are reflections of each other, and write down the coordinates of the point where the graph of g crosses the y-axis. [2]

35. Let $f(x) = x^2 - 1$.

(a) Sketch $y = |f(x)|$ for $-3 \le x \le 3$, showing the coordinates of both corners. [3]

(b) Hence solve $|x^2 - 1| = 3$. [3]

(c) Solve $|x^2 - 1| < 3$. [3]

(d) Solve $|x - 1| = 3x + 5$, showing clearly why one candidate is rejected. [4]

Challenge Questions

36. Self-inverse rational functions. Throughout, let

$$f(x) = \frac{ax + b}{cx + d}, \qquad c \neq 0, \quad ad - bc \neq 0.$$

The function is called *self-inverse* when $(f \circ f)(x) = x$ for every x in its domain. This investigation finds exactly which of these functions are self-inverse, and then which return to x after three or four applications.

(a) Show that $f(x) = \dfrac{x + 1}{x - 1}$ is self-inverse. [2]

(b) Show that if $ad - bc = 0$ then f is constant, and explain why the condition $ad - bc \neq 0$ is therefore assumed. [3]

(c) Show that

$$(f \circ f)(x) = \frac{(a^2 + bc)\,x + b(a + d)}{c(a + d)\,x + (bc + d^2)}.$$

[4]

(d) Hence prove that f is self-inverse if and only if $a + d = 0$. *(Hint: two expressions in x are equal for all x only when their coefficients match)* [4]

(e) Let $g(x) = \dfrac{1}{1 - x}$. Show that g is not self-inverse, but that $(g \circ g \circ g)(x) = x$ for every $x \neq 0, 1$. [3]

(f) Explain why $(f \circ f \circ f)(x) = x$ for every x in the domain is equivalent to $(f \circ f)(x) = f^{-1}(x)$, where $f^{-1}(x) = \dfrac{dx - b}{-cx + a}$. By comparing this with part (c), prove that, when $a + d \neq 0$, the function f returns to x after three applications if and only if

$$(a + d)^2 = ad - bc.$$

Verify the condition for g. *(Hint: two such fractions, each with a nonzero value of "$ad - bc$", are equal for every x exactly when the four coefficients of one are a constant multiple of the four coefficients of the other)* [5]

(g) Show that $h(x) = \dfrac{x-1}{x+1}$ returns to x after four applications but not after two. Compare $(a+d)^2$ with $ad - bc$ for the self-inverse functions, for g and for h, and suggest a condition for f to return to x after n applications. [3]

37. When does the order of two transformations matter? Section 5.2 showed that a vertical stretch and a vertical translation give different curves in different orders. This investigation finds the exact condition for any two vertical moves, and a way to read that condition from the graph.

A *vertical move* with parameters $a \neq 0$ and k replaces $y = f(x)$ by $y = af(x) + k$: a vertical stretch of factor a followed by a translation of k units up. Throughout, start from $y = f(x)$.

(a) Apply a vertical stretch of factor a and then a translation of k units up, and write down the result. Do the same in the opposite order. Deduce the exact condition on a and k under which the two agree for every function f. [3]

(b) Two vertical moves have parameters (a_1, k_1) and (a_2, k_2). Show that the first followed by the second gives $y = a_1 a_2 f(x) + a_2 k_1 + k_2$, and find the result of the opposite order. Deduce that the two orders give the same curve for every f if and only if

$$k_1(a_2 - 1) = k_2(a_1 - 1).$$

[5]

(c) Check the condition in part (b) for two translations ($a_1 = a_2 = 1$), for two stretches ($k_1 = k_2 = 0$), and for a stretch with a translation, comparing the last with part (a). Then show that the moves with parameters $(2, 1)$ and $(3, 2)$ may be applied in either order, although neither is a pure stretch or a pure translation. [4]

(d) Show that when $a \neq 1$ the move with parameters (a, k) leaves every point on the line $y = \dfrac{k}{1-a}$ at the same height. Show that, when $a_1 \neq 1$ and $a_2 \neq 1$, the condition in part (b) says exactly that the two moves leave the same horizontal line fixed, and check this for the pair in part (c). [4]

(e) Explain why any vertical move and any horizontal move give the same curve in either order. *(Hint: consider which variable each move acts on)* [2]

(f) A horizontal move replaces x by $bx + c$, where $b \neq 0$. Investigate when two horizontal moves may be applied in either order, and whether a fixed-line rule like the one in part (d) still holds. Suggest one way to extend the investigation further. [3]

38. Counting the solutions of $|f(x)| = k$. Let $f(x) = x^3 - 3x$ and let $k \geq 0$ be a constant. The graph of $y = f(x)$ has a local maximum at $(-1, 2)$ and a local minimum at $(1, -2)$.

(a) Sketch $y = |f(x)|$, showing the x-intercepts in exact form and the coordinates of the

local maxima. [3]

(b) Find the number of real solutions of $|f(x)| = k$ for each of $k = 0$, $k = 1$, $k = 2$ and $k = 3$. *(Hint: for $k = 2$, each of $x^3 - 3x - 2$ and $x^3 - 3x + 2$ has a repeated linear factor)* [4]

(c) Explain, using your sketch, why the number of solutions changes at $k = 0$ and $k = 2$ and nowhere else. [3]

(d) State the number of real solutions of $|f(x)| = k$ for every $k \geq 0$, giving your answer as a list of cases. [2]

(e) Now let $g(x) = x^3 - 3x + 1$, whose graph has a local maximum at $(-1, 3)$ and a local minimum at $(1, -1)$. Give the corresponding list of cases for $|g(x)| = k$, and explain why it has more cases than the list in part (d). [4]

(f) State a rule that predicts, from the graph of $y = |g(x)|$, the values of k at which the number of solutions changes. Suggest how you would test your rule on a quartic. [2]

39. GDC **Newton's law of cooling.** A body at temperature T in a room at temperature R cools so that the difference $T - R$ decays exponentially:

$$T(t) = R + (T_0 - R)\,e^{-kt}, \qquad k > 0,$$

where T_0 is the temperature at $t = 0$ and t is measured in minutes.

(a) Show that the temperature difference $T(t) - R$ falls by the same factor over any interval of fixed length, and that this factor does not depend on T_0. [4]

(b) Deduce that the difference halves every $\dfrac{\ln 2}{k}$ minutes. [3]

(c) Three readings T_1, T_2 and T_3 are taken at times t, $t + h$ and $t + 2h$. Use part (a) to show that $(T_2 - R)^2 = (T_1 - R)(T_3 - R)$, and hence that

$$R = \frac{T_1 T_3 - T_2^2}{T_1 + T_3 - 2T_2},$$

provided the denominator is not zero. [5]

(d) A cup of coffee is left in a room whose temperature is not known. It reads $95\,^\circ$C, $65\,^\circ$C and $47\,^\circ$C at $t = 0$, 10 and 20 minutes. Find the room temperature R, and find k in exact form. [4]

(e) Find, to the nearest minute, how long the coffee takes to reach $40\,^\circ$C. [2]

(f) Explain why the model predicts that the coffee never reaches room temperature, and state which feature of the graph of T against t shows this. [2]

(g) Suppose the third reading had been recorded as $48\,^\circ$C instead of $47\,^\circ$C. Find the new estimate of R, and comment on how sensitive the method in part (c) is to a small error in one reading. Suggest how more than three readings could be used to estimate R more reliably. [3]

40. The range of a quadratic over a linear function. Section 5.4 finds the range of $\dfrac{ax + b}{cx + d}$ by writing it as a constant plus a fraction. When the numerator is a quadratic, that method no longer gives the range directly. This investigation uses the discriminant instead.

Throughout,

$$f(x) = \frac{x^2 + bx + c}{x + d}, \qquad x \neq -d,$$

where b, c and d are real constants, and $N(x) = x^2 + bx + c$ is the numerator.

(a) Let $f(x) = \dfrac{x^2 + 3}{x - 1}$. Show that $f(x) = y$ leads to $x^2 - yx + (y + 3) = 0$. By requiring this equation to have a real root, show that the range of f is $y \leq -2$ or $y \geq 6$, and check that the excluded input $x = 1$ removes no value of y. [5]

(b) For the function in part (a), find the value of x at which $f(x) = -2$ and the value at which $f(x) = 6$, and explain why these points are turning points of the graph. Write $f(x)$ as $x + 1 + \dfrac{4}{x - 1}$ and state the equations of the asymptotes. [4]

(c) Find the range of $\dfrac{x^2 - 4x + 3}{x - 2}$. [3]

(d) For $\dfrac{x^2 + x - 2}{x - 1}$, show that the discriminant method gives $(y - 3)^2 \geq 0$, which holds for every y. Explain why the range is nevertheless not every real number, and state the range. [4]

(e) Return to the general f. Show that $f(x) = y$ has a real solution when

$$\Delta(y) = y^2 + (4d - 2b)\,y + b^2 - 4c \geq 0,$$

and that $\Delta(y)$, as a quadratic in y, has discriminant $16N(-d)$. Deduce that when $N(-d) < 0$ the range is every real number, and that when $N(-d) > 0$ the range misses an interval. Interpret the condition $N(-d) < 0$ in terms of where the vertical asymptote lies relative to the zeros of N. [5]

(f) Show that when $N(-d) > 0$ the missing interval has width $4\sqrt{N(-d)}$, and check this against part (a). Suggest how the investigation could be extended, for example to $\dfrac{x^2 + bx + c}{x^2 + px + q}$. [3]

Chapter 6

Polynomials

6

SYLLABUS 1.11 2.12 HL ONLY

A car's speed is recorded once every second of a test run. Nobody wrote down a formula for it, yet a polynomial drawn through those readings predicts the speed between them.

Many quantities change smoothly rather than in sudden jumps. Temperature, water level and the position of a moving object are examples, and polynomials are among the simplest functions that change in the same way. The degree matches the situation: distance travelled at a constant speed is linear, a thrown ball follows a parabola, and the volume of a sphere grows with the cube of its radius. Over a short enough interval almost any smooth curve can be matched closely by a polynomial, so measurements are often fitted with one when no exact formula is known. See Ch. 16.

That smoothness comes from the way a polynomial is built. It uses one variable and a few constants, combined by addition and multiplication only. There is no division by the variable, no square root and no trigonometry. A polynomial is therefore defined for every real number, and its graph is a smooth, unbroken curve with no gaps, no asymptotes and no corners.

Describing a polynomial is easy. Solving one exactly can be hard. A general formula for the roots, built from the coefficients using arithmetic and roots such as $\sqrt{}$ and $\sqrt[3]{}$, exists for degree 2, degree 3 and degree 4. Niels Abel proved in 1824 that for degree 5 and above no general formula of this kind exists, although some particular equations of higher degree can still be solved exactly. By the end of this chapter you will be able to read a polynomial's graph from its factorised form, divide one polynomial by another, test whether a value is a root and factorise once you have found one, and give the sum and the product of the roots from the coefficients alone.

One idea runs through most of the chapter: a polynomial's coefficients and its roots describe the same object, and most techniques here move between the two. The chapter closes with **partial fractions**, which splits a fraction by factorising its denominator and giving each factor a numerator of its own. That is why it belongs here rather than with the rational functions of §5.4, and it makes many rational functions straightforward to integrate later.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

See Ch. 15.

6.1 Polynomial Functions and their Graphs

Much of the graph of a polynomial can be described before a single point is plotted, because the degree and the coefficients already carry that information. Reading those features from the equation is faster than building a table of values, and they are what a sketch has to show.

DEF A **polynomial function** of **degree** n has the form

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \qquad a_n \neq 0,$$

where the coefficients are constants and n is a non-negative integer.

6.1.1 Degree, Coefficients and End Behaviour

The graph features in this subsection are read from two things: the degree, and the sign of the leading coefficient.

Polynomials are named by their degree. Degree 1 is **linear**, degree 2 is **quadratic**, degree 3 is **cubic**, degree 4 is **quartic**, and degree 5 is **quintic**. The degree also limits the shape: how many times the curve can meet the x-axis, and how often it can change direction.

The constants $a_n, a_{n-1}, \ldots, a_1, a_0$ are the **coefficients** of the powers of x. The coefficient of the highest power is a_n, called the **leading coefficient**, and the term $a_n x^n$ that contains it is the **leading term**.

For a factorised polynomial, neither the degree nor the leading coefficient needs the brackets multiplied out. The degree is the sum of the degrees of the factors, which for linear factors is simply how many there are. The leading coefficient is the product of the leading coefficients of the factors.

A coefficient may take any value, with one restriction: the leading coefficient cannot be zero. If a_n were zero, the term $a_n x^n$ would be zero, and what remained would be a polynomial of degree $n-1$ instead.

The sign of the leading coefficient determines which way the ends of the curve point. It says nothing about the sign of the function elsewhere: $y = x^2 - 4$ has a positive leading coefficient, yet y is negative for $-2 < x < 2$.

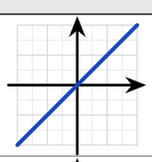
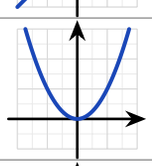
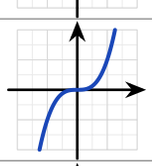
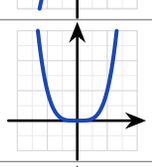
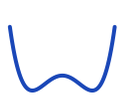
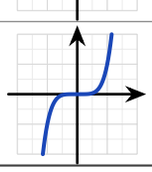
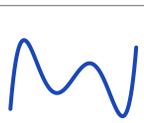
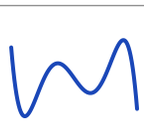
Three features can be read from the equation.

The **end behaviour**, meaning the direction of the curve as $x \to \pm\infty$, is determined entirely by the leading term. If n is odd, the two ends point in opposite directions. If n is even, they point the same way, and the sign of a_n determines whether that direction is upward or downward.

The **zeros** of f are the values of x for which $f(x) = 0$; they are also called the **roots** of the equation $f(x) = 0$, and this chapter uses both words. The real zeros are the x-intercepts of the graph, and the number of distinct real zeros can never exceed the degree. A cubic meets the axis at most three times, a quartic at most four.

A **turning point** is a place where the curve stops rising and begins to fall, or stops falling and begins to rise. A polynomial of degree n has at most $n - 1$ turning points. The possible numbers then decrease in steps of two: $n - 1$, $n - 3$, and so on. The number changes by two rather than by one because the degree fixes the direction of the two ends. Removing a maximum without also removing a minimum would reverse one end, so the number of turning points is always odd when n is even, and even when n is odd.

The table sets out all three features for the degrees you meet most often.

n	$y = x^n$	**Shape when** $a_n > 0$	**Shape when** $a_n < 0$	**Times it meets the** x**-axis**	**Turning points**
1				1	0
2				0, 1 or 2	1
3				1, 2 or 3	0 or 2
4				0, 1, 2, 3 or 4	1 or 3
5				1, 2, 3, 4 or 5	0, 2 or 4

EXAMPLE 6.1

State the degree, the leading coefficient and the two end behaviours of

(a) $p(x) = 5 - 4x + 6x^3 - x^2$ and (b) $q(x) = (1 - 2x)(x + 3)^2(x - 4)$.

(a) The terms are not written in order of power, so find the highest power rather than reading the first term. It is $6x^3$, so the degree is 3 and the leading coefficient is 6. An odd degree points the two ends in opposite directions, and a positive leading coefficient sends the right-hand end upward:

$$p(x) \to +\infty \text{ as } x \to +\infty, \qquad p(x) \to -\infty \text{ as } x \to -\infty.$$

(b) Nothing needs to be multiplied out. Count the linear factors: $(1 - 2x)$, two copies of $(x + 3)$, and $(x - 4)$, so the degree is 4. The leading coefficient is the product of the four leading coefficients,

$$(-2) \times 1 \times 1 \times 1 = -2.$$

An even degree points both ends the same way, and the negative leading coefficient sends them down: $q(x) \to -\infty$ as $x \to \pm\infty$.

The answer to (b) depends on the -2, which comes from the factor written backwards as $(1 - 2x)$. Reading the coefficient of x in that bracket as 2 and losing the minus sign gives a leading coefficient of $+2$, which reverses both ends.

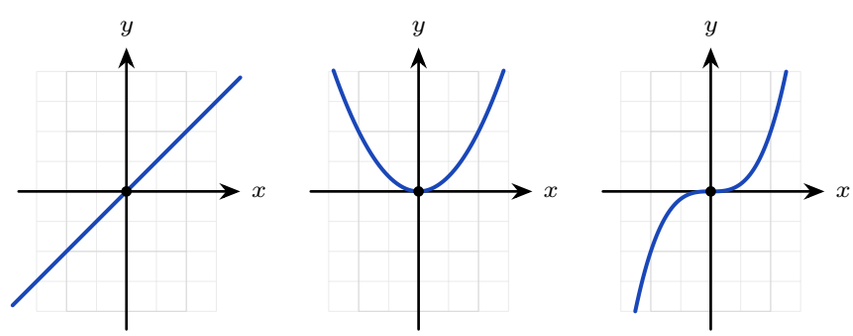
6.1.2 Roots and Multiplicity

The table gives the shape in outline. What fixes it exactly is what the curve does at each root, and that is determined by whether a factor is repeated.

The **multiplicity** of a root is the number of times its factor appears in the factorised polynomial. In $(x - 1)(x - 2)^2$ the root $x = 1$ has multiplicity one, and the root $x = 2$ has multiplicity two. A root of multiplicity one is called a **simple root**, and one of multiplicity two a **double root**.

KEY At a root of **multiplicity one** the curve crosses the axis with a non-zero gradient. At a root of **even** multiplicity it touches the axis and turns back. At a root of **odd** multiplicity greater than one it crosses the axis, and the axis is the tangent there, so the curve is flat as it crosses.

Whether the multiplicity is odd or even determines crossing or touching. Its size determines how flat the curve is where it meets the axis.



(a) $y = x$ crosses (b) $y = x^2$ touches (c) $y = x^3$ flattens

Figure 6.1 The three behaviours at a root, shown at $x = 0$. In (a) the multiplicity is one and the curve crosses with a non-zero gradient. In (b) it is two, an even number, so the curve touches and turns back. In (c) it is three, odd, so the curve crosses, with the axis as its tangent at the root.

KEY **Sketching a factorised polynomial.**

1. At each x-intercept, read the behaviour from the multiplicity of the root.
2. Find the y-intercept by substituting $x = 0$.
3. Find the two ends from the leading term.

The case below goes from the factors to the sketch; §6.1.3 works in the other direction, from a sketch back to the factors.

Worked through. To sketch $y = (x + 1)(x - 2)^2$, read the roots and their multiplicities from the factors. The factor $(x + 1)$ is single, so the graph crosses the axis at $x = -1$. The factor $(x - 2)^2$ is squared, so the graph touches the axis at $x = 2$ and turns back. Find the y-intercept. Substitute $x = 0$: $y = (1)(-2)^2 = 4$, the point $(0, 4)$. Fix the ends from the leading term. Multiplying out leads with x^3: an odd degree with a positive coefficient, so the curve runs from $-\infty$ at the left to $+\infty$ at the right. These three facts fix the shape.

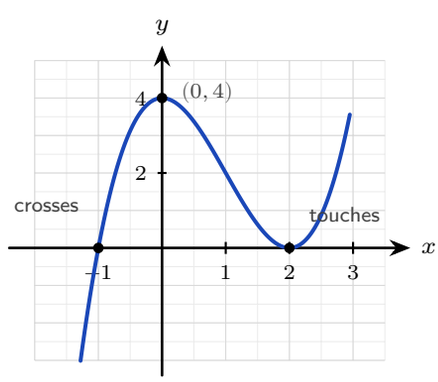


Figure 6.2 A sketch of $y = (x + 1)(x - 2)^2$. The curve crosses the axis at the simple root $x = -1$ and touches it at the double root $x = 2$. The y-intercept is $(0, 4)$, and the ends follow x^3.

When the polynomial is written out in full rather than factorised, the y-intercept is even quicker to read. Every term except the constant contains x, so substituting $x = 0$ makes all of them zero. The constant term a_0 is what remains.

KEY The graph of $y = a_n x^n + \cdots + a_1 x + a_0$ crosses the y-axis at $(0, a_0)$. *The constant term is the y-intercept.*

For $y = 2x^3 - 3x + 5$, the constant term is 5, so the graph crosses the y-axis at $(0, 5)$.

6.1.3 From the Graph Back to the Equation

The reverse problem is also examined. Given a sketch, you are asked for the polynomial that produced it. The intercepts give you almost everything you need.

Each x-intercept contributes a factor, and the behaviour there gives its multiplicity: a crossing with a non-zero gradient gives a single factor, a touch gives a squared factor, and a flat crossing gives a cubed factor. That fixes the polynomial up to one unknown number, the leading coefficient, and any known point that is not an x-intercept determines it. Check that the factors account for the whole degree of the curve; if they do not, an intercept has been missed or a multiplicity is higher than it looks.

KEY **Finding a polynomial from its graph.** Write one factor for each x-intercept, with the multiplicity the graph shows at that point. Then substitute the coordinates of any known point that is not an x-intercept to find the leading coefficient a_n.

EXAMPLE 6.2

The graph below shows a cubic. It touches the x-axis at $x = -1$, crosses it at $x = 2$, and passes through $(0, -1)$. Find its equation.

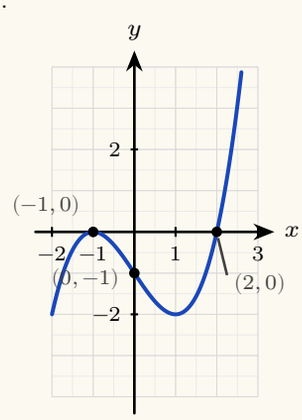


Figure 6.3 The curve touches the x-axis at $x = -1$, crosses it at $x = 2$, and passes through $(0, -1)$. These three facts determine the equation.

Write the factors from the x-intercepts. The curve touches at $x = -1$, so that root is repeated and contributes $(x + 1)^2$. It crosses at $x = 2$, so that root is single and contributes $(x - 2)$. Those factors already account for degree three, so

$$y = a(x + 1)^2(x - 2)$$

for some constant a.

Find a from the remaining point. Substitute $x = 0$ and $y = -1$:

$$-1 = a(1)^2(-2) = -2a,$$

so $a = \frac{1}{2}$ and

$$y = \tfrac{1}{2}(x + 1)^2(x - 2).$$

Check the answer against the graph. The leading term is $\frac{1}{2}x^3$, which is positive, so the curve rises to the right, and the sketch agrees.

The intercepts fix the factors; one further point fixes the leading coefficient.

YOUR TURN 6.1 Polynomial Functions and their Graphs

Degree, leading coefficient and end behaviour

1. State the degree and the leading coefficient of each polynomial.

(a) $f(x) = 3x^4 - 2x + 7$
(b) $f(x) = 5 - x^3$
(c) $f(x) = (x-1)(x+2)(x-3)$
(d) $f(x) = -2x^5 + x^2$
(e) $f(x) = (2x-1)(x+3)^2$
(f) $f(x) = (3-x)^2(2x+1)^3$

2. Describe the end behaviour of each graph as $x \to +\infty$ and as $x \to -\infty$.

(a) $y = x^3$
(b) $y = -x^4$
(c) $y = 2x^5 - x$
(d) $y = 3 - x^2$

3. Answer each question about degree and shape.

(a) State the maximum number of real roots of a degree-5 polynomial.
(b) State the maximum number of turning points of a quartic.
(c) List every possible number of turning points a cubic can have.
(d) Explain why a cubic must have at least one real root.

Roots and multiplicity

4. State the roots of each polynomial, and say at each one whether the graph crosses or touches the x-axis.

(a) $y = (x-2)(x+3)(x-5)$
(b) $y = (x-1)^2(x+4)$
(c) $y = x(x-3)^2$
(d) $y = (x+1)^2(x-2)^2$
(e) $y = (x-2)^3(x+1)$
(f) $y = x^3(x-4)^2$

5. Consider $y = (x+1)(x-3)^2$.

(a) State the roots and their multiplicities.
(b) Find the y-intercept.
(c) State the end behaviour.
(d) Sketch the curve.

From the graph to the equation

6. Find the equation of each polynomial in factorised form.

(a) A cubic crosses the x-axis at $x = -2$, $x = 1$ and $x = 3$, and passes through $(2, -8)$.
(b) A quartic touches the x-axis at $x = 1$, crosses it at $x = -1$ and $x = 3$, and passes through $(0, 6)$.

Concept check

7. A student says: "*the graph of* $y = (x+1)^3(x-4)$ *touches the* x*-axis at* $x = -1$*, because the factor* $(x+1)$ *is repeated.*" Explain why this is false.

6.2 Polynomial Division

You can already add, subtract and multiply polynomials. Addition and subtraction collect like terms, and multiplication expands every product before collecting. Division is the operation that still needs a method. Factorising depends on it: once one factor of a polynomial is known, the other is found by dividing, and §6.3 and §6.4 are built on that step.

6.2.1 The Division Algorithm

Before dividing anything it helps to know what the answer should look like, and whole-number division already shows that. When you divide 47 by 5 you write

$$47 = 5 \times 9 + 2,$$

which records a divisor, a quotient, and a remainder smaller than the divisor. Polynomial division works the same way, with "smaller" measured by degree instead of size.

KEY **The division algorithm.** For polynomials $f(x)$ and $d(x)$, with $d(x)$ not the zero polynomial, there are unique polynomials $q(x)$ and $r(x)$ such that

$$f(x) = d(x)\,q(x) + r(x),$$

where the degree of r is less than the degree of d. *Dividend equals divisor times quotient, plus remainder.*

Fix the four names now, because the methods that follow use them. The polynomial being divided, f, is the **dividend**. The polynomial you divide by, d, is the **divisor**. The result q is the **quotient**, and the part that remains, r, is the **remainder**. When $r = 0$ the division is exact, and d is a factor of f.

6.2.2 Long Division

Long division works for a divisor of any degree. It is three steps repeated until the remainder has lower degree than the divisor.

KEY **Divide, multiply, subtract.**

1. **Divide** the leading term of the dividend by the leading term of the divisor. The result is the next term of the quotient.
2. **Multiply** the whole divisor by that term.
3. **Subtract** the product from the dividend, and repeat on what is left.

Set the work out in columns, one for each power of x: the divisor to the left of the bracket, the dividend inside it, the quotient written on the top line, and the remainder on the bottom line. Write the result as $f(x) = d(x)\,q(x) + r(x)$ when you finish, because multiplying that right-hand side out checks the whole division.

Worked through. Dividing $2x^3 - 3x^2 + x - 5$ by $(x - 2)$ sets out like this:

$$
\begin{array}{rrrrrl}
 & & 2x^2 & +x & +3 & \text{quotient} \\
\cline{2-5}
x-2\,) & 2x^3 & -3x^2 & +x & -5 & \text{dividend} \\
 & 2x^3 & -4x^2 & & & \\
\cline{2-3}
 & & x^2 & +x & & \\
 & & x^2 & -2x & & \\
\cline{3-4}
 & & & 3x & -5 & \\
 & & & 3x & -6 & \\
\cline{4-5}
 & & & & 1 & \text{remainder}
\end{array}
$$

Each stage repeats the same three steps: divide, multiply, subtract.

Stage 1. Divide the leading term of the dividend by the leading term of the divisor: $2x^3 \div x = 2x^2$. Write $2x^2$ in the quotient. Multiply the divisor by it: $2x^2(x - 2) = 2x^3 - 4x^2$. Subtract that from the dividend, which leaves $x^2 + x$.

Stage 2. Repeat on $x^2 + x$. Divide: $x^2 \div x = x$. Multiply: $x(x - 2) = x^2 - 2x$. Subtract, which leaves $3x - 5$.

Stage 3. Repeat on $3x - 5$. Divide: $3x \div x = 3$. Multiply: $3(x - 2) = 3x - 6$. Subtract, which leaves 1.

The degree of 1 is lower than the degree of $x - 2$, so the division stops.

The quotient is $2x^2 + x + 3$ and the remainder is 1:

$$2x^3 - 3x^2 + x - 5 = (x - 2)(2x^2 + x + 3) + 1.$$

CAUTION Write the dividend in descending powers and insert a zero term for every missing power before dividing. Divide $x^3 - 13x + 12$ as $x^3 + 0x^2 - 13x + 12$. Without that zero term the columns shift, and the subtractions that follow go wrong.

EXAMPLE 6.3

Divide $x^3 - 13x + 12$ by $(x - 1)$.

There is no x^2 term, so write it as $x^3 + 0x^2 - 13x + 12$ and keep a column for it.

$$\begin{array}{rr|llll} & & x^2 & +x & -12 \\ \hline x-1\ ) & x^3 & +0x^2 & -13x & +12 \\ & \underline{x^3} & \underline{-x^2} & & \\ & & x^2 & -13x & \\ & & \underline{x^2} & \underline{-x} & \\ & & & -12x & +12 \\ & & & \underline{-12x} & \underline{+12} \\ & & & & 0 \end{array}$$

The remainder is 0, so the division is exact and $(x - 1)$ is a factor. The quotient $x^2 + x - 12$ factorises further:

$$x^3 - 13x + 12 = (x - 1)(x^2 + x - 12) = (x - 1)(x + 4)(x - 3).$$

A remainder of zero shows that the divisor is a factor. The next two sections are built on that fact.

A divisor of higher degree changes none of the three steps. What it changes is where the division stops: the remainder must have degree lower than the divisor, so dividing by a quadratic can leave a remainder with an x in it. Set the columns out in the same way, and keep a column for every power of x from the highest down to the constant, whether or not the dividend has a term there.

EXAMPLE 6.4

Divide $x^4 + 2x^2 - 3x + 1$ by $(x^2 + x - 1)$.

There is no x^3 term, so write the dividend as $x^4 + 0x^3 + 2x^2 - 3x + 1$ and give that power a column of its own.

$$\begin{array}{rr|llll} & & & x^2 & -x & +4 \\ \hline x^2+x-1\ ) & x^4 & +0x^3 & +2x^2 & -3x & +1 \\ & \underline{x^4} & \underline{+x^3} & \underline{-x^2} & & \\ & & -x^3 & +3x^2 & -3x & \\ & & \underline{-x^3} & \underline{-x^2} & \underline{+x} & \\ & & & 4x^2 & -4x & +1 \\ & & & \underline{4x^2} & \underline{+4x} & \underline{-4} \\ & & & & -8x & +5 \end{array}$$

Stage 1. Divide the leading terms: $x^4 \div x^2 = x^2$. Multiply the divisor by it: $x^2(x^2 + x - 1) = x^4 + x^3 - x^2$. Subtract, which leaves $-x^3 + 3x^2 - 3x$.
Stage 2. Repeat. Divide: $-x^3 \div x^2 = -x$. Multiply: $-x(x^2 + x - 1) = -x^3 - x^2 + x$. Subtract, which leaves $4x^2 - 4x + 1$.
Stage 3. Repeat. Divide: $4x^2 \div x^2 = 4$. Multiply: $4(x^2 + x - 1) = 4x^2 + 4x - 4$. Subtract, which leaves $-8x + 5$.
The remainder $-8x + 5$ has degree one, the divisor has degree two, so the division stops:

$$x^4 + 2x^2 - 3x + 1 = (x^2 + x - 1)(x^2 - x + 4) - 8x + 5.$$

Multiplying the right side out returns the dividend, which checks every stage at once. A remainder with an x in it is the normal outcome against a quadratic divisor, not a sign that a stage has gone wrong.

YOUR TURN 6.2 Polynomial Division

Dividing by a linear factor

8. Divide, and state the quotient and the remainder.

(a) $x^2 + 5x + 6$ by $(x + 2)$
(b) $x^3 - 3x^2 + 2x$ by $(x - 1)$
(c) $x^3 + 2x^2 - x + 1$ by $(x - 1)$
(d) $x^3 - 8$ by $(x - 2)$

9. Divide, and state the quotient and the remainder. Insert a zero term for each missing power first.

(a) $2x^3 - x^2 - 3$ by $(x + 1)$
(b) $x^3 + 3x^2 - 4$ by $(x + 2)$
(c) $2x^4 - 3x^3 + x - 6$ by $(x - 2)$

Dividing by a quadratic

10. Divide by the quadratic divisor, and state the quotient and the remainder.

(a) $x^4 - 1$ by $(x^2 - 1)$
(b) $x^4 - 1$ by $(x^2 + 1)$
(c) $x^3 + 2x^2 - 3x + 1$ by $(x^2 + x - 1)$
(d) $2x^4 - x^3 + 3x - 2$ by $(x^2 - 2x + 2)$

The division algorithm

11. Use $f(x) = d(x)\,q(x) + r(x)$ in each part.

(a) When $f(x)$ is divided by $x^2 - 1$, the quotient is $2x + 3$ and the remainder is $x - 4$. Find $f(x)$.
(b) Divide $x^4 + x^3 - 2x + 5$ by $x^2 + 2$, and write the result in the form $f(x) = d(x)\,q(x) + r(x)$.
(c) When $x^3 + ax^2 + bx + 3$ is divided by $x^2 + 1$, the remainder is $2x + 1$. Find a and b.

Concept check

12. A student divides $x^3 + 2x^2 + x - 1$ by $x^2 + 1$ and writes: "*quotient* x*, remainder* $2x^2 - 1$." Explain why this cannot be the final answer, and find the correct quotient and remainder.

6.3 The Remainder and Factor Theorems

6.3.1 The Remainder Theorem

Here are two operations that appear unrelated. *Evaluating* f at 2 gives the height of the graph above one point. *Dividing* f by $(x - 2)$ is the long calculation of the previous section. The remainder theorem states that the two produce the same number, and it is quickly proved.

Take the division algorithm with a linear divisor $(x - a)$:

$$f(x) = (x - a)\,q(x) + r,$$

where the remainder r must be a *constant*, because its degree is less than the degree of $x - a$, and less than one means zero.

Now substitute $x = a$ into that identity. The factor $(x - a)$ becomes zero, so the whole first term is zero, leaving $f(a) = r$. The height of the graph at a *is* the remainder, because everything else in the polynomial is divisible by $(x - a)$ and so contributes nothing at that point.

KEY **Remainder theorem.** When a polynomial $f(x)$ is divided by $(x - a)$, the remainder is $f(a)$. *One substitution replaces the division.*

The theorem holds for any linear divisor, not only for one written as $(x - a)$. Divide by $(bx - c)$ and the remainder is f evaluated at the **zero of the divisor**, meaning the value of x that makes the divisor equal to zero. For $(2x - 1)$ that value is $\frac{1}{2}$, never 1 or 2.

Read the theorem in either direction. Given the coefficients it returns the remainder; given two remainders it returns two linear equations in the unknown coefficients.

Worked through. When $f(x) = x^3 - 2x + 5$ is divided by $(x - 2)$, the remainder theorem gives the remainder as

$$f(2) = 2^3 - 2(2) + 5 = 8 - 4 + 5 = 9.$$

No division needed; the answer is a single evaluation.

EXAMPLE 6.5

Find the remainder when $f(x) = 8x^3 + 4x^2 - 6x + 5$ is divided by $(2x - 1)$.
Evaluate at the zero of the divisor. Here $2x - 1 = 0$ when $x = \frac{1}{2}$, so the remainder is

$$f\left(\tfrac{1}{2}\right) = 8 \cdot \tfrac{1}{8} + 4 \cdot \tfrac{1}{4} - 3 + 5 = 1 + 1 - 3 + 5 = 4.$$

EXAMPLE 6.6

The polynomial $f(x) = x^3 + ax + b$ leaves remainder 7 when divided by $(x - 2)$ and remainder -5 when divided by $(x + 1)$. Find a and b.
Each remainder condition is one evaluation, so two conditions give two equations:

$$f(2) = 8 + 2a + b = 7 \implies 2a + b = -1,$$

$$f(-1) = -1 - a + b = -5 \implies -a + b = -4.$$

Subtracting, $3a = 3$, so $a = 1$ and $b = -3$. The remainder theorem has turned two divisions into a pair of linear equations; questions of this kind, working from the remainders back to the coefficients, are common.

The theorem is stated for a linear divisor, but the idea behind it reaches further. Divide by a quadratic and the remainder has degree at most one, so it is $rx + s$: two unknowns instead of one. The division algorithm still records the whole calculation as an identity, and each zero of the divisor makes the divisor term vanish, leaving one equation in r and s. Two distinct zeros give two equations, which is exactly what two unknowns need. The method needs the two zeros to be distinct. Against a repeated factor such as $(x - 1)^2$ only one substitution is available, so take the second equation from comparing coefficients.

EXAMPLE 6.7

Find the remainder when $f(x) = x^3 + 2x^2 - 5x + 3$ is divided by $(x - 1)(x - 2)$.
The divisor has degree two, so the remainder has degree at most one. Write it as $rx + s$ and record the division as an identity:

$$f(x) = (x - 1)(x - 2)\,q(x) + rx + s.$$

This holds for every value of x, so choose the two values that make the product term disappear. Substituting $x = 1$,

$$f(1) = 1 + 2 - 5 + 3 = 1 \implies r + s = 1,$$

and substituting $x = 2$,

$$f(2) = 8 + 8 - 10 + 3 = 9 \implies 2r + s = 9.$$

Subtracting the first equation from the second gives $r = 8$, and then $s = -7$. The remainder is

$$8x - 7.$$

Long division confirms it: the quotient is $x + 5$, and $(x-1)(x-2)(x+5) + 8x - 7$ returns $f(x)$. Both substitutions were available because the zeros 1 and 2 are distinct.

6.3.2 The Factor Theorem

More follows when the remainder is zero. If dividing by $(x - a)$ leaves no remainder, then $(x - a)$ divides $f(x)$ exactly: it is a factor. That is what turns a test for a root into a method for factorising.

KEY **Factor theorem.** $(x - a)$ is a factor of $f(x)$ if and only if $f(a) = 0$.

To factorise an unfamiliar cubic, search for a root by testing small values. The factors of the constant term are the natural candidates, and the reason is short. If a is an integer root of $a_n x^n + \cdots + a_1 x + a_0$ with integer coefficients, then substituting and moving a_0 across gives

$$a\left(a_n a^{n-1} + a_{n-1} a^{n-2} + \cdots + a_1\right) = -a_0,$$

so a non-zero integer root a divides a_0. When the leading coefficient is not 1, the candidates are the fractions $\frac{p}{q}$ with p a factor of a_0 and q a factor of a_n. Once one root a is found, $(x - a)$ is a factor, and dividing it out leaves a polynomial one degree lower; for a cubic, that is a quadratic you can finish by hand. For a root $\frac{p}{q}$ the factor is $(qx - p)$.

That division need not be long division. To **compare coefficients**, write the quotient with unknown coefficients, one degree lower than the dividend, multiply out, and equate the coefficients of matching powers. Each unknown comes from one comparison, and any coefficient left over is a check.

EXAMPLE 6.8

Show that $(x - 1)$ is a factor of $f(x) = x^3 - 7x^2 + 14x - 8$, and factorise f completely.
Test $x = 1$:

$$f(1) = 1 - 7 + 14 - 8 = 0,$$

so by the factor theorem $(x - 1)$ is a factor.
Now divide it out. Instead of long division, *compare coefficients*. The quotient must be a quadratic, so write

$$x^3 - 7x^2 + 14x - 8 = (x - 1)(x^2 + bx + c),$$

then compare the coefficients of matching powers on the two sides. Compare the constant terms: $-c = -8$, so $c = 8$. Compare the coefficients of x^2: $b - 1 = -7$, so $b = -6$. The coefficient of x has not been used, so use it as a check: on the right it is $c - b = 8 + 6 = 14$,

which matches the left. ✓ Hence

$$f(x) = (x-1)(x^2-6x+8) = (x-1)(x-2)(x-4).$$

Had the factor not been given, the constant term -8 would give the candidates $\pm1, \pm2, \pm4, \pm8$, and $x = 1$ is the first to test.

TIP Comparing coefficients is often quicker than long division for a linear divisor. Long division is equally acceptable if you prefer it.

YOUR TURN 6.3 The Remainder and Factor Theorems

The remainder theorem

13. Find the remainder in each case.

(a) $x^3 + 2x - 1$ divided by $(x-1)$

(b) $x^3 - 4x^2 + x + 6$ divided by $(x-2)$

(c) $2x^3 - 3x + 1$ divided by $(x+1)$

(d) $x^3 + x^2 + x + 1$ divided by $(x+2)$

14. Find the remainder in each case. The divisor is not of the form $(x-a)$.

(a) $4x^3 - 4x^2 + 5x - 1$ divided by $(2x-1)$

(b) $2x^3 + x^2 - 2$ divided by $(2x+1)$

15. Use the remainder and factor theorems to answer each part.

(a) The remainder when $f(x)$ is divided by $(x-3)$ is 0. State what this tells you about f.

(b) The remainder when $f(x)$ is divided by $(x-2)$ is 7. State $f(2)$.

(c) $f(-1) = 0$. Write down a factor of $f(x)$.

(d) $x^4 - 1$ is divided by $(x-1)$. Find the remainder without dividing.

16. The polynomial $f(x) = x^3 + ax^2 + bx + 6$ has $(x+1)$ as a factor, and leaves remainder 12 when divided by $(x-2)$.

(a) Write down two equations in a and b.

(b) Solve them to find a and b.

The remainder on division by a quadratic

17. The remainder on division by a quadratic $(x-a)(x-b)$ has the form $rx + s$. Find it in each case.

(a) $x^3 - 2x^2 + 4x - 1$ divided by $(x-1)(x-3)$

(b) $x^4 + x - 3$ divided by $(x+1)(x-2)$

The factor theorem

18. Decide whether the given linear expression is a factor. Show your working.

(a) Is $(x-2)$ a factor of x^3-5x^2+8x-4?

(b) Is $(x+3)$ a factor of x^3+x^2-5x+3?

(c) Is $(x-1)$ a factor of x^3+x^2+x+1?

(d) Is $(x+2)$ a factor of x^3+8?

19. Factorise each polynomial completely.

(a) x^3-x

(b) x^3-4x

(c) $x^3-3x^2-4x+12$

(d) $2x^3+3x^2-11x-6$

20. In each part, $(x-3)$ is a factor. Find the quadratic factor by comparing coefficients, then factorise completely.

(a) x^3-2x^2-5x+6

(b) $2x^3-5x^2-4x+3$

21. Use the factor theorem in each part.

(a) One factor of x^2-5x+6 is $(x-2)$. State the other.

(b) Show that x^3-3x^2+4 has a repeated factor.

Concept check

22. Each student below has made an error in using the remainder theorem. Identify the error and find the correct remainder.

(a) To find the remainder when $f(x)=x^3-2x+5$ is divided by $(x+3)$, a student calculates $f(3)=26$.

(b) To find the remainder when $g(x)=4x^3-2x+3$ is divided by $(2x-1)$, a student calculates $g(1)=5$.

6.4 Solving Polynomial Equations

You already know how to solve linear and quadratic equations. A linear equation is rearranged directly. A quadratic is solved by factorising, by completing the square, or by the quadratic formula.

For degrees three and four, general formulas exist but are far too long to use by hand, and from degree five upward no general formula exists. The factor theorem takes their place when there is a rational root to find.

KEY To solve a polynomial equation of degree 3 or more:

1. Test the factors of the constant term, or the fractions $\frac{p}{q}$ of §6.3.2 if the leading coefficient is not 1, until one value a gives $f(a) = 0$.
2. Divide $f(x)$ by $(x - a)$, either by long division or by comparing coefficients.
3. Repeat on the quotient until a quadratic remains, then solve that in the usual way.
4. If no candidate is a root, find the roots with a GDC.

Every real root is an x-intercept of the graph, so a sketch or a GDC graph shows how many real roots to look for.

Worked through. To solve $x^3 - 3x^2 - x + 3 = 0$, test the factors of 3. At $x = 1$: $1 - 3 - 1 + 3 = 0$, so $(x - 1)$ is a factor. Dividing,

$$x^3 - 3x^2 - x + 3 = (x - 1)(x^2 - 2x - 3) = (x - 1)(x - 3)(x + 1).$$

The roots are $x = 1,\ 3,\ -1$.

A cubic with real coefficients always has at least one real root, because one end goes to $+\infty$ and the other to $-\infty$, so the curve must cross the axis. If none of the candidates is a root, the cubic has no rational roots, so every real root is irrational; the equation is then solved with a GDC.

YOUR TURN 6.4 Solving Polynomial Equations

Equations that factorise at sight

23. Solve each equation. Each is already factorised or factorises at sight.

(a) $(x - 1)(x + 2)(x - 4) = 0$

(b) $(x - 3)^2(x + 1) = 0$

(c) $x^3 = 4x$

(d) $x^3 + x^2 - 2x = 0$

Using the factor theorem

24. Solve each equation, using the factor theorem to find the first root.

(a) $x^3 + 2x^2 - 5x - 6 = 0$

(b) $x^3 - 3x^2 - 10x + 24 = 0$

(c) $2x^3 - 3x^2 - 3x + 2 = 0$

25. Solve $x^3 - 1 = 0$, giving only the real root, and explain why there is exactly one.

Rational roots

26. The leading coefficient is not 1, so test the candidates $\frac{p}{q}$. Find one root, then solve the equation.

(a) $2x^3 - x^2 + 4x - 2 = 0$
(b) $12x^3 - 16x^2 - 5x + 3 = 0$

Quartic equations

27. Solve each equation.

(a) $x^4 - 5x^2 + 4 = 0$ *(Hint: substitute $u = x^2$)*
(b) $x^4 - 2x^3 - 7x^2 + 8x + 12 = 0$ *(Hint: use the factor theorem twice)*

When no candidate is a root

28. GDC For each equation, show that none of the candidate integer roots is a root. Then use a GDC to find every real root, correct to three significant figures.

(a) $x^3 - 4x + 1 = 0$
(b) $x^3 + 2x^2 - 2 = 0$

Concept check

29. To solve $x^3 + 6x^2 + 11x + 6 = 0$, a student tests $x = 1, 2, 3$ and 6, finds that none is a root, and concludes: "*there are no integer roots, so a GDC is needed.*" Identify the error, and solve the equation.

6.5 Sum and Product of Roots

The coefficients of a polynomial determine the sum and the product of its roots. You can therefore find those two values *without solving the equation at all*. This matters whenever the roots are irrational, or when the equation is too hard to solve, because the sum and the product remain easy to obtain.

6.5.1 The Quadratic Case

Ch. 4 showed that if $ax^2 + bx + c = 0$ has roots α and β, then comparing $ax^2 + bx + c$ with $a(x - \alpha)(x - \beta)$ gives

$$\alpha + \beta = -\frac{b}{a}, \qquad \alpha\beta = \frac{c}{a}.$$

The minus sign on the sum comes from the brackets, one $-\alpha$ and one $-\beta$.

6.5.2 The General Case

The same comparison works for every degree, using the counting argument of Ch. 2. Expanding

$$a_n(x-r_1)(x-r_2)\cdots(x-r_n)$$

means choosing, from each bracket, either the x or the root, exactly as in the binomial theorem. To build the x^{n-1} term you take x from all brackets but one, so each root appears exactly once, each carrying a single minus sign: the coefficient equals $-(r_1+r_2+\cdots+r_n)$, the *sum*. To build the constant term you take the root from *every* bracket: all n roots multiplied together, carrying n minus signs, which is where the $(-1)^n$ comes from. So the leading coefficient, the next coefficient and the constant term give the sum and the product of the roots.

KEY · IB For the equation $a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0=0$,

$$\text{sum of roots} = -\frac{a_{n-1}}{a_n}, \qquad \text{product of roots} = (-1)^n\,\frac{a_0}{a_n}.$$

Divide by the leading coefficient. The sum is the next coefficient with its sign changed; the product is the constant term, with sign $(-1)^n$.

These are the first and last of a family of relations (Vieta's formulas), but the sum and product are the two the syllabus requires, used for checking answers, for building equations with required roots, and for finding an unknown coefficient.

One more relation fills the gap between them. Building the x^{n-2} term means taking the root from exactly two brackets and x from the rest, so that coefficient is the sum of the roots multiplied together in pairs. For a cubic $a_3x^3+a_2x^2+a_1x+a_0=0$ with roots α,β,γ, the three relations together are

$$\alpha+\beta+\gamma=-\frac{a_2}{a_3}, \qquad \alpha\beta+\beta\gamma+\gamma\alpha=\frac{a_1}{a_3}, \qquad \alpha\beta\gamma=-\frac{a_0}{a_3}.$$

The syllabus requires only the first and the last. The middle relation is a short extension, and it is needed for expressions such as $\alpha^2+\beta^2+\gamma^2$, which involve the roots in pairs.

EXAMPLE 6.9

The equation $2x^2-5x+3=0$ has roots α and β. Find $\alpha+\beta$ and $\alpha\beta$ without solving, then verify.

Here $a=2$, $b=-5$, $c=3$, so

$$\alpha+\beta=-\frac{-5}{2}=\frac{5}{2}, \qquad \alpha\beta=\frac{3}{2}.$$

Solving confirms it: the roots are 1 and $\frac{3}{2}$, whose sum is $\frac{5}{2}$ and product is $\frac{3}{2}$. The sum and product were available from the coefficients before the roots were found.

Worked through. The cubic $x^3 - 2x^2 - 9x + 18 = 0$ has roots α, β, γ. With $a_3 = 1$, $a_2 = -2$, $a_0 = 18$ and degree $n = 3$,

$$\alpha + \beta + \gamma = -\frac{-2}{1} = 2, \qquad \alpha\beta\gamma = (-1)^3 \frac{18}{1} = -18.$$

(The roots happen to be $-3, 2, 3$: sum 2, product -18, as promised.)

6.5.3 Symmetric Combinations of the Roots

The sum and the product are more useful than they appear, because many other expressions in the roots can be rebuilt from them. That is what makes the method valuable when the roots are irrational: the examples below answer precise questions about the roots without ever finding one.

A **symmetric combination** of the roots is an expression whose value does not change when the roots are exchanged. Both $\alpha^2 + \beta^2$ and $\frac{1}{\alpha} + \frac{1}{\beta}$ are symmetric; $\alpha - \beta$ is not. Every symmetric combination of two roots can be written in terms of $\alpha + \beta$ and $\alpha\beta$. Two rearrangements cover most cases, and a sum of reciprocals reduces by putting it over a common denominator, $\frac{1}{\alpha} + \frac{1}{\beta} = \frac{\alpha + \beta}{\alpha\beta}$.

KEY

$$\alpha^2 + \beta^2 = (\alpha + \beta)^2 - 2\alpha\beta, \qquad (\alpha - \beta)^2 = (\alpha + \beta)^2 - 4\alpha\beta.$$

Neither is a new formula. Each is the expansion of $(\alpha \pm \beta)^2$ rearranged.

EXAMPLE 6.10

The equation $x^2 - 3x + 1 = 0$ has roots α and β. Without solving, find $\alpha^2 + \beta^2$ and $\frac{1}{\alpha} + \frac{1}{\beta}$.

The roots are irrational, but their sum and product are not: $\alpha + \beta = 3$ and $\alpha\beta = 1$. For the sum of squares, use the first rearrangement above:

$$(\alpha + \beta)^2 = \alpha^2 + 2\alpha\beta + \beta^2 \implies \alpha^2 + \beta^2 = (\alpha + \beta)^2 - 2\alpha\beta = 9 - 2 = 7.$$

For the sum of reciprocals, combine over a common denominator:

$$\frac{1}{\alpha} + \frac{1}{\beta} = \frac{\alpha + \beta}{\alpha\beta} = \frac{3}{1} = 3.$$

Neither answer required a root. Both were rebuilt from the sum and the product alone.

Reading the relations backwards builds an equation from its roots. A quadratic whose roots have sum S and product P is

$$x^2 - Sx + P = 0,$$

so an equation whose roots are a transformation of α and β is found by transforming the sum and the product. The roots themselves are never needed.

EXAMPLE 6.11

With α, β as above, find a quadratic equation with integer coefficients whose roots are 2α and 2β.

A quadratic is determined by the sum and product of its roots: $x^2 - (\text{sum})x + (\text{product}) = 0$. The new sum and product follow from the old ones:

$$2\alpha + 2\beta = 2(\alpha + \beta) = 6, \qquad (2\alpha)(2\beta) = 4\alpha\beta = 4.$$

So the required equation is

$$x^2 - 6x + 4 = 0.$$

No root was ever computed: transforming an equation's roots means transforming their sum and product, nothing more.

EXAMPLE 6.12

The roots of $x^2 + kx + 12 = 0$ differ by 1. Find the possible values of k.

Let the roots be α and β with $\alpha - \beta = 1$. The coefficients give $\alpha + \beta = -k$ and $\alpha\beta = 12$. Connect the difference to these totals by squaring it:

$$(\alpha - \beta)^2 = (\alpha + \beta)^2 - 4\alpha\beta \implies 1 = k^2 - 48,$$

so $k^2 = 49$ and $k = \pm 7$. Check with $k = 7$: the equation $x^2 + 7x + 12 = 0$ has roots -3 and -4, which indeed differ by 1. A condition on the roots has become an equation in the coefficient.

YOUR TURN 6.5 Sum and Product of Roots

Sum and product from the coefficients

30. State the sum and the product of the roots of each equation.

(a) $x^2 - 7x + 10 = 0$

(b) $2x^2 + 3x - 5 = 0$

(c) $x^2 + 4x + 4 = 0$

(d) $3x^2 - 2x - 8 = 0$

31. For each cubic or quartic, state the sum and the product of the roots.

(a) $x^3 - 6x^2 + 11x - 6 = 0$

(b) $2x^3 - 4x^2 + x - 3 = 0$

(c) $x^4 - 3x^3 + 2x - 5 = 0$

Symmetric expressions in the roots

32. The roots of $2x^2 - 7x + 6 = 0$ are α and β. Without solving the equation, find each value.

(a) $\alpha + \beta$ and $\alpha\beta$

(b) $\dfrac{1}{\alpha} + \dfrac{1}{\beta}$

(c) $\alpha^2 + \beta^2$

(d) $(\alpha - \beta)^2$

33. The roots of $x^3 - 4x^2 + x + 6 = 0$ are α, β and γ.

(a) Write down $\alpha + \beta + \gamma$ and $\alpha\beta\gamma$.

(b) Write down $\alpha\beta + \beta\gamma + \gamma\alpha$.

(c) Hence find $\alpha^2 + \beta^2 + \gamma^2$.

Building an equation from its roots

34. Build the required equation.

(a) Write a quadratic with leading coefficient 1 whose roots sum to 5 and multiply to 6.

(b) Write a cubic with leading coefficient 1 whose roots are -1, 2 and 4.

35. The roots of $x^2 + 5x + 3 = 0$ are α and β.

(a) Write down $\alpha + \beta$ and $\alpha\beta$.

(b) Find the sum and product of $\alpha + 2$ and $\beta + 2$.

(c) Hence find a quadratic equation with integer coefficients whose roots are $\alpha + 2$ and $\beta + 2$.

Conditions on the roots

36. In each part, a condition on the roots fixes an unknown coefficient.

(a) The roots of $x^2 + bx + 12 = 0$ are 3 and one other. Find the other root and b.

(b) The roots of $x^2 - 7x + k = 0$ differ by 3. Find k.

(c) One root of $x^2 + kx + 18 = 0$ is twice the other. Find the possible values of k.

Concept check

37. A student writes: "*the product of the roots of* $2x^3 - 5x^2 + x + 6 = 0$ *is* $\dfrac{6}{2} = 3$, *as for a quadratic.*" Identify the error and state the correct product.

6.6 Partial Fractions

Adding two simple fractions produces one more complicated fraction. **Partial fractions** reverse that process: they split a single fraction back into a sum of simpler ones, one for each factor of the denominator. The split form is easier to use, and it makes a rational function easier to integrate. See Ch. 15.

The syllabus limits the problem to a denominator of **at most two distinct linear factors**, with the numerator of lower degree than the denominator. That is the case treated below. The larger theory is sketched at the end of the section for reference.

Each linear factor contributes one term with a constant numerator.

KEY For a proper fraction over two distinct linear factors,

$$\frac{px+q}{(x-\alpha)(x-\beta)} = \frac{A}{x-\alpha} + \frac{B}{x-\beta}.$$

The same form holds when a factor is written as $(ax+b)$, for example $\dfrac{A}{2x-1}$; only the value of x that makes that factor zero changes.

To find A and B, multiply through by the denominator, then substitute the values of x that make one factor zero. Each substitution isolates one unknown.

Worked through. To express $\dfrac{3x+5}{(x-1)(x+2)}$ in partial fractions, write

$$\frac{3x+5}{(x-1)(x+2)} = \frac{A}{x-1} + \frac{B}{x+2},$$

and multiply through by $(x-1)(x+2)$:

$$3x+5 = A(x+2) + B(x-1).$$

Setting $x=1$ isolates A: $8 = 3A$, so $A = \frac{8}{3}$. Setting $x=-2$ isolates B: $-1 = -3B$, so $B = \frac{1}{3}$. Therefore

$$\frac{3x+5}{(x-1)(x+2)} = \frac{8}{3(x-1)} + \frac{1}{3(x+2)}.$$

Choosing the two values that each make one factor zero is what makes the method quick. It turns a pair of simultaneous equations into two one-line calculations.

EXAMPLE 6.13

Express $\dfrac{x+5}{x^2-x-2}$ in partial fractions.

This denominator is not factorised, so factorise it first:

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

Now proceed as before:

$$\frac{x+5}{(x-2)(x+1)} = \frac{A}{x-2} + \frac{B}{x+1} \implies x+5 = A(x+1) + B(x-2).$$

Setting $x = 2$: $7 = 3A$, so $A = \frac{7}{3}$. Setting $x = -1$: $4 = -3B$, so $B = -\frac{4}{3}$. Hence

$$\frac{x+5}{x^2-x-2} = \frac{7}{3(x-2)} - \frac{4}{3(x+1)}.$$

TIP A question may give the denominator in expanded form. The factorising step is part of the answer, so write it out.

EXT **Beyond the syllabus.** Three things extend the method. None of them is on the AA HL syllabus, although the first is short enough to try in Your Turn Q42. A fraction whose numerator has degree equal to or higher than its denominator is **improper**, and must be divided first: $\frac{x^2}{x^2-1} = 1 + \frac{1}{x^2-1}$, and only the proper part is split. A **repeated** factor $(x-\alpha)^2$ needs one term for each power, $\frac{A}{x-\alpha} + \frac{B}{(x-\alpha)^2}$. An **irreducible quadratic**, one with negative discriminant, takes a linear numerator $\frac{Bx+C}{x^2+bx+c}$. One rule covers the last two: give each factor a numerator of degree one less than itself, and an extra term for every power of a repeated factor.

Factor in the denominator	Term(s) it contributes
linear $(x-\alpha)$	$\frac{A}{x-\alpha}$
repeated linear $(x-\alpha)^2$	$\frac{A}{x-\alpha} + \frac{B}{(x-\alpha)^2}$
irreducible quadratic (x^2+bx+c)	$\frac{Bx+C}{x^2+bx+c}$

YOUR TURN 6.6 Partial Fractions

Two linear factors

38. Express each fraction in partial fractions.

(a) $\dfrac{1}{(x-1)(x-2)}$ (b) $\dfrac{5}{(x-2)(x+3)}$

(c) $\dfrac{3}{(x+1)(x+4)}$ (d) $\dfrac{x}{(x-1)(x+1)}$

39. Express each fraction in partial fractions.

(a) $\dfrac{2x+1}{x(x+1)}$ (b) $\dfrac{x+4}{x(x-2)}$

(c) $\dfrac{2x}{(x-3)(x+1)}$ (d) $\dfrac{5x-1}{(x-1)(x+2)}$

Factorising the denominator first

40. Factorise the denominator first, then express each fraction in partial fractions.

(a) $\dfrac{4x}{x^2-4}$ (b) $\dfrac{x+7}{x^2+x-2}$

A factor of the form $ax+b$

41. Express each fraction in partial fractions. Factorise the denominator first where it is not already factorised.

(a) $\dfrac{5}{(2x+1)(x-2)}$ (b) $\dfrac{5x+1}{(3x-1)(x+1)}$

(c) $\dfrac{4x-7}{2x^2-5x+3}$

An improper fraction

42. Consider $\dfrac{x^2}{x^2-1}$.

(a) Explain why this fraction is improper.

(b) Divide to write it as $1+\dfrac{1}{x^2-1}$.

(c) Hence express the whole fraction in partial fractions.

Concept check

43. A student writes $\dfrac{2x^2+1}{(x-1)(x+2)} = \dfrac{A}{x-1}+\dfrac{B}{x+2}$, substitutes $x=1$ and $x=-2$, and obtains $A=1$ and $B=-3$. Explain why the result is wrong, and give the correct partial fractions.

Chapter 6 · Polynomials

Reflections

TOK Abel and Galois proved something surprising. The roots of a general quintic *exist*, but they cannot be written using only addition, subtraction, multiplication, division and roots. The equation $x^5 - x - 1 = 0$ has a real solution. That number is completely determined, and a computer can calculate it to a thousand decimal places. Even so, no formula of that kind can express it exactly. What does it mean, then, to "know" a number? Is such a root any less real than one you can write down?

IM The history behind this chapter is unusual. In Italy in the 1530s, mathematicians competed in public problem-solving contests, and a loser could lose his university post, so methods were kept secret. Scipione del Ferro solved one type of cubic around 1515, Niccolò Tartaglia rediscovered the method in 1535, and Gerolamo Cardano published it in 1545, after promising Tartaglia he would not, together with his student Ferrari's solution of the quartic. Almost three hundred years later Niels Abel proved that no such formula exists for the general quintic, and Évariste Galois, killed in a duel at twenty, explained exactly which equations a formula can solve. To do so he created *group theory*, the mathematical study of symmetry, which is now used in particle physics and cryptography.

RW Polynomials are widely used in computing, because they need only multiplication and addition, which processors carry out very quickly. The curved shapes in fonts and animation are drawn with *Bézier curves*, which are polynomial curves whose shape is controlled by a few chosen points. *Error-correcting codes* are built from polynomials as well. These are methods of adding extra information to a message so that it can still be read after part of it is damaged, which is how a scratched disc still plays and how a faint signal from a spacecraft is recovered. Polynomials are also used to fit smooth models to scattered measurements in science and finance.

EXT The factor theorem raises a question it cannot answer. How many roots does a polynomial really have? Over the real numbers, $x^2 + 1$ has none. The **Fundamental Theorem of Algebra** gives the answer. Over the *complex* numbers of Chapter 17, every polynomial of degree n has exactly n roots, counted with multiplicity. The sum and product formulas then hold for every polynomial, because the roots they describe always exist.

End-of-Chapter Problems

1. Consider $f(x) = x^3 - 2x^2 - 5x + 6$.

(a) Show that $x = 1$ is a root. [2]
(b) Factorise $f(x)$ completely. [3]
(c) Solve $f(x) = 0$. [1]

2. Consider $f(x) = 2x^3 + 3x^2 - 11x - 6$.

(a) Show that $(x - 2)$ is a factor. [2]
(b) Factorise $f(x)$ completely. [3]

3. Find the unknown coefficients in each polynomial.

(a) When $x^3 + ax + b$ is divided by $(x - 1)$ the remainder is 4, and $(x + 1)$ is a factor. Find a and b. [5]
(b) The polynomial $x^3 + ax^2 + bx - 2$ has $(x - 1)$ and $(x - 2)$ as factors. Find a and b. [5]
(c) When $4x^3 + cx + d$ is divided by $(2x - 1)$ the remainder is 6, and $(x + 1)$ is a factor. Find c and d. [5]

4. The polynomial $x^3 - 3x^2 + kx - 8$ has $(x - 2)$ as a factor. Find k. [3]

5. Sketch each graph, showing all intercepts and describing the behaviour at each root.

(a) $y = (x + 1)(x - 1)(x - 3)$ [4]
(b) $y = x^2(x - 3)$ [4]

6. Build the equation described in each part.

(a) A quadratic $x^2 + px + q$ has roots 2 and -5. Find p and q. [3]
(b) The cubic $x^3 + px^2 + qx + r$ has roots -1, 2 and 4. Find p, q and r. [5]
(c) The quadratic equation with roots α and β is $x^2 - 4x + 1 = 0$. Find a quadratic with leading coefficient 1 whose roots are $\alpha + 1$ and $\beta + 1$. [5]

7. Consider $f(x) = 12x^3 - 16x^2 - 7x + 6$.

(a) Any rational root of f has the form $\frac{p}{q}$ in lowest terms. State the conditions on p and q. [2]
(b) Show that $x = \frac{1}{2}$ is a root. [2]
(c) Hence factorise $f(x)$ completely. [3]

(d) Solve $f(x) = 0$. [1]

8. Solve each equation.

(a) $x^3 - 7x + 6 = 0$ [4]
(b) $2x^3 + 3x^2 - 3x - 2 = 0$ [5]

9. Each equation has roots α and β. Answer without solving the equation.

(a) $x^2 - 5x + 3 = 0$: find $\alpha^2 + \beta^2$. [4]
(b) $3x^2 - x - 2 = 0$: find $\dfrac{1}{\alpha} + \dfrac{1}{\beta}$. [3]
(c) $x^2 - 6x + 7 = 0$: find $\alpha^2 + \beta^2$ and $(\alpha - \beta)^2$. [5]

10. Consider $f(x) = x^3 - 3x + 2$.

(a) Show that $(x - 1)$ is a factor, and factorise $f(x)$ completely. [4]
(b) Sketch $y = f(x)$, showing all intercepts and the behaviour at each root. [3]
(c) Hence solve the inequality $f(x) > 0$. [2]

11. Consider $f(x) = x^4 - x^2 - 12$.

(a) Use the substitution $u = x^2$ to find the real solutions of $f(x) = 0$. [3]
(b) Explain why $f(x) = 0$ has no other real solutions. [2]
(c) Factorise $f(x)$ as far as possible over the real numbers. [2]

12. Show that $x = 1$ is a root of $x^3 + 4x^2 + x - 6 = 0$, and find the other roots. [4]

13. One root of $x^2 - 9x + k = 0$ is twice the other. Use the sum and product of the roots to find k. [4]

14. Consider $f(x) = 2x^3 - x^2 - 13x - 6$.

(a) Show that $(x + 2)$ is a factor. [2]
(b) Factorise $f(x)$ completely and solve $f(x) = 0$. [4]

15. The polynomial $x^3 - kx + 2$ has $(x - 2)$ as a factor.

(a) Find k. [2]
(b) Find the other two roots, in exact form. [4]

16. The polynomial $x^3 - 2x^2 + x - 2$ factorises as $(x - 2)(x^2 + 1)$.

(a) Verify this by expansion. [2]
(b) State the only real root, and explain why there are no others. [2]

17. When a polynomial $p(x)$ is divided by $(x-1)$ the remainder is 3, and by $(x-2)$ the remainder is 5. Find the remainder when $p(x)$ is divided by $(x-1)(x-2)$. [6]

18. The roots of $x^3 - 6x^2 + 11x - 6 = 0$ are α, β, γ. Find $\alpha^2 + \beta^2 + \gamma^2$ without solving the equation. [5]

19. Prove that if a polynomial with integer coefficients has a rational root $\frac{p}{q}$ in lowest terms, then p divides the constant term and q divides the leading coefficient. [6]

20. The cubic $x^3 + px + q = 0$ has a repeated root. Show that $4p^3 + 27q^2 = 0$. [6]

21. Find all real values of k for which $x^3 - 3x + k = 0$ has three distinct real roots. *(Hint: use Q20 to find the values of k that give a repeated root)* [6]

22. Consider $f(x) = 2x^3 - 3x^2 - 11x + 6$.

(a) Show that $(x-3)$ is a factor of $f(x)$. [2]
(b) Factorise $f(x)$ completely. [4]
(c) Solve $f(x) = 0$. [2]
(d) State the y-intercept, and describe the behaviour of $f(x)$ as $x \to \pm\infty$. [3]
(e) Sketch $y = f(x)$, showing all intercepts. [2]
(f) Hence solve the inequality $f(x) \leq 0$. [3]

23. The polynomial $p(x) = x^3 + ax^2 + bx + 6$ leaves a remainder of 12 when divided by $(x-2)$, and $(x+1)$ is a factor.

(a) Show that $4a + 2b = -2$ and $a - b = -5$. [4]
(b) Hence find a and b. [2]
(c) Factorise $p(x)$ as far as possible over the real numbers. [4]
(d) State how many real roots $p(x) = 0$ has, giving a reason. [2]
(e) Find the remainder when $p(x)$ is divided by $(x-1)(x+1)$. *(Hint: the remainder has the form $rx + s$)* [4]

24. The cubic equation $x^3 + px^2 + qx + r = 0$ has roots α, β and γ.

(a) Write down $\alpha + \beta + \gamma$, $\alpha\beta + \beta\gamma + \gamma\alpha$ and $\alpha\beta\gamma$ in terms of p, q and r. [3]
(b) The equation $x^3 - 6x^2 + 3x + 10 = 0$ has roots α, β, γ. Find $\alpha^2 + \beta^2 + \gamma^2$ without solving the equation. [4]
(c) Find $\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma}$. [3]
(d) Given that one root is -1, find the other two. [3]
(e) Write down a cubic equation with integer coefficients whose roots are 2α, 2β and 2γ. [3]

25. Express each fraction in partial fractions.

(a) $\dfrac{7x-1}{(x-1)(x+2)}$ [4]

(b) $\dfrac{1}{x^2-4}$ [4]

(c) $\dfrac{x+5}{x^2+x}$ [4]

Challenge Questions

26. Power sums of the roots. The equation $x^2 - px + q = 0$ has roots α and β. For each integer $k \geq 0$ define

$$S_k = \alpha^k + \beta^k.$$

(a) Write down the values of S_0 and S_1. [2]

(b) Show that $S_2 = p^2 - 2q$. [2]

(c) Find S_3 in terms of p and q. [3]

(d) Explain why $\alpha^k = p\,\alpha^{k-1} - q\,\alpha^{k-2}$ for every $k \geq 2$, and deduce that $S_k = p\,S_{k-1} - q\,S_{k-2}$. [4]

(e) The equation $x^2 - 3x + 1 = 0$ has roots α and β. Use the recurrence to find S_2, S_3 and S_4. [4]

(f) The roots of $x^2 - 3x + 1 = 0$ are irrational. Explain why S_k is nevertheless an integer for every k. [3]

(g) For $x^2 - x - 1 = 0$ the numbers S_k are the *Lucas numbers* $2, 1, 3, 4, 7, \ldots$. Now let the cubic $x^3 - px^2 + qx - r = 0$ have roots α, β and γ, and let $S_k = \alpha^k + \beta^k + \gamma^k$. Find a recurrence for S_k in the style of part (d), and suggest how the investigation could continue. [3]

27. Palindromic polynomials. A polynomial is *palindromic* if its list of coefficients reads the same forwards as backwards. For example $x^4 + x^3 - 4x^2 + x + 1$ has coefficients $1,\ 1,\ -4,\ 1,\ 1$.

(a) Explain why $x = 0$ is never a root of a palindromic polynomial. [2]

(b) Let $P(x)$ be palindromic of degree n. Show that $x^n P\!\left(\dfrac{1}{x}\right) = P(x)$, and deduce that if α is a root of P then so is $\dfrac{1}{\alpha}$. [4]

(c) Divide $x^4 + x^3 - 4x^2 + x + 1 = 0$ through by x^2 and substitute $u = x + \dfrac{1}{x}$. Show that the equation becomes $u^2 + u - 6 = 0$. [4]

(d) Hence find all four roots of $x^4 + x^3 - 4x^2 + x + 1 = 0$ in exact form. [5]

(e) Show that every palindromic polynomial of odd degree has $x = -1$ as a root. [3]

(f) Show that $x^3 + \dfrac{1}{x^3} = u^3 - 3u$, where $u = x + \dfrac{1}{x}$. Explain why dividing a palindromic polynomial of degree $2m$ by x^m turns the equation $P(x) = 0$ into an equation of degree

m in u. Suggest how the investigation could continue, for example with palindromic equations of degree 6 or with polynomials whose coefficients reverse with a change of sign. [3]

28. Splitting a sum with partial fractions. Partial fractions are usually a step towards integration. This investigation uses them for something else: adding infinitely many terms exactly.

(a) Show that $\dfrac{1}{n(n+1)} = \dfrac{1}{n} - \dfrac{1}{n+1}$ for every positive integer n. [2]

(b) Hence write out the sum $\displaystyle\sum_{n=1}^{N} \frac{1}{n(n+1)}$ term by term, and show that almost every term cancels. Deduce that the sum equals $\dfrac{N}{N+1}$. *(Hint: write the first three and the last two terms)* [4]

(c) State the value the sum approaches as $N \to \infty$, and explain what that value means for the graph of the partial sums. [2]

(d) Show that $\dfrac{1}{n(n+2)} = \dfrac{1}{2}\left(\dfrac{1}{n} - \dfrac{1}{n+2}\right)$, and hence find $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+2)}$. [4]

(e) Investigate $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+k)}$ for a positive integer k. Find its value for $k=3$ and $k=4$, then state a general formula and justify it. [6]

(f) Show that $\dfrac{1}{n(n+1)(n+2)} = \dfrac{1}{2}\left(\dfrac{1}{n(n+1)} - \dfrac{1}{(n+1)(n+2)}\right)$, and hence find $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+1)(n+2)}$. Then use $\dfrac{1}{n^2} < \dfrac{1}{n(n-1)}$ for $n \geq 2$ to show that $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2} < 2$, and suggest how the exact value of this sum could be investigated. [4]

29. Dividing by a product of two linear factors. End-of-Chapter Questions 17 and 23 find a remainder on division by $(x-a)(x-b)$ from two given values. This investigation finds that remainder in general, applies it to x^n, and then asks what happens when the two factors are the same. Throughout, $P(x)$ is a polynomial and a and b are real constants.

(a) Explain why the remainder when $P(x)$ is divided by $(x-a)(x-b)$ has the form $px+q$. Show that, when $a \neq b$, the remainder is

$$R(x) = P(a)\,\frac{x-b}{a-b} + P(b)\,\frac{x-a}{b-a}.$$

[4]

(b) Hence show that, for every positive integer n, the remainder when x^n is divided by x^2-3x+2 is $(2^n-1)x+2-2^n$. Find also the remainder when x^n is divided by x^2-1, treating even and odd n separately. [5]

(c) The formula in part (a) cannot be used when $a=b$. By long division or otherwise, find the remainder when x^n is divided by $(x-1)^2$ for $n=2$, 3 and 4, and conjecture the remainder for every positive integer n. [3]

(d) Prove your conjecture. *(Hint: $x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \cdots + x + 1)$)* [5]

(e) Compare the line $y = nx + 1 - n$ with the graph of $y = x^n$ near $x = 1$, and suggest what the remainder on division by $(x - a)^2$ tells you about the graph of any polynomial P. Suggest how the investigation could be extended to division by $(x - a)(x - b)(x - c)$. [3]

Chapter 7

Shapes & Measures

7

SYLLABUS 3.1 3.2 3.3 3.4

You cannot pace out the whole Earth. So how did anyone find its size? The answer came from something much smaller: an angle.

Check the fact yourself. On flat ground the horizon is level with your eye. Climb a hill and it drops slightly below your eye line, and the higher you climb, the further it drops. That drop is the **dip**, and it is small: about half a degree from a hill a few hundred metres high, the width of the Moon in the sky.

Your line of sight to the horizon is a tangent to the Earth: it touches the surface at one point, and there it meets the radius at a right angle. So the hilltop, the horizon and the centre of the Earth form one right-angled triangle, with hill height h and dip θ:

$$\cos\theta = \frac{R}{R+h}, \qquad \text{so} \qquad R = \frac{h\cos\theta}{1-\cos\theta}.$$

Measure the hill, measure the dip, and the radius comes out close to the modern 6371 km, from one hill, in one afternoon. Al-Bīrūnī did exactly this a thousand years ago, from the fort of Nandana in what is now Pakistan.

Size matters as well as shape. Double every length of an object and its surface grows fourfold while its volume grows eightfold, so bigger things have less surface for what they hold. That is why crushed ice melts faster than a block, and it is one reason insects do not grow to the size of a horse.

By the end of this chapter you will find the angle between a line and a plane in a solid, work out the volume and surface area of standard solids, solve any triangle with the sine and cosine rules, ambiguous case included, handle elevation, depression and bearings, and measure in **radians**, which turn arc length and sector area into one multiplication each.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

A calculator is assumed throughout this chapter. The EXACT badge marks the questions whose answer must be left exact.

7.1 Three-Dimensional Geometry

The measurements you need from a solid are often not the ones marked on its faces. The diagonal of a room, or the angle a roof strut makes with the floor: neither of these lies along a face. Both run through the inside of the object. Right-angled triangles drawn inside the solid, and sometimes coordinates, are how you reach them.

Before any measuring, a word on names. Examination papers use a fixed notation for points, segments, lines and angles, and they use it without explanation.

Notation	Meaning
$A(x, y)$	the point A with coordinates x and y
$[AB]$	the line segment with endpoints A and B
AB	the length of $[AB]$
(AB)	the line through A and B, continuing both ways
$A\hat{B}C$	the angle at B, between $[BA]$ and $[BC]$
$\triangle ABC$	the triangle with vertices A, B and C

The middle letter names the vertex. So $A\hat{B}C$ and $C\hat{B}A$ are the same angle, while $B\hat{A}C$ is a different one. Note too that the length of a segment carries no bars. Bars are reserved for the magnitude of a vector, which Chapter 9 introduces.

Coordinates in the plane extend to space by adding a third axis, z. The distance formula extends with them. In two dimensions it is Pythagoras' theorem. In three dimensions it is Pythagoras used *twice*, as shown below the formula.

KEY · IB The distance between (x_1, y_1, z_1) and (x_2, y_2, z_2) is

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

The midpoint is the average of the coordinates, $\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}, \frac{z_1 + z_2}{2}\right)$.

Across the base, Pythagoras gives

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Treat that as one side of a new right triangle. Its other side is the vertical rise, $(z_2 - z_1)$, and Pythagoras on this second triangle gives the formula.

The equation of a sphere comes from the same formula. A sphere is the set of points at a fixed distance r from a centre (a, b, c), so its equation is $(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2$. The sphere equation is beyond the syllabus, but it follows from the distance formula in one line.

7.1.1 The angle between a line and a plane

A line that is not parallel to a plane meets it at an angle. Many angles can be measured at that meeting point, so one of them has to be chosen, and mathematics chooses the smallest. Drop a perpendicular from each point of the line to the plane. The feet of those perpendiculars form a line in the plane, called the **projection** of the line onto the plane (its 'shadow': what a light directly overhead would cast). The angle between the line and its projection is the smallest angle available.

DEF The **angle between a line and a plane** is the angle between the line and its projection on the plane. To find it, drop a perpendicular from a point on the line down to the plane. The line, its projection and that perpendicular form a right triangle, and right-angled trigonometry solves it.

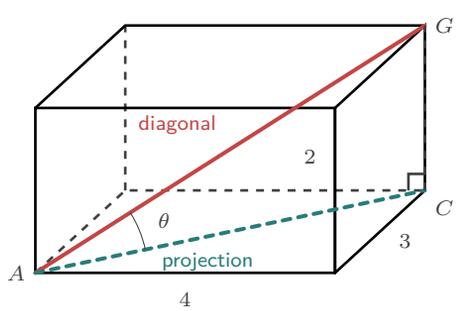


Figure 7.1 The space diagonal AG and its projection AC on the base. The vertical edge CG meets the base at a right angle, so ACG is a right triangle and θ is the angle between the diagonal and the base.

Worked through. To find the angle between the space diagonal AG and the base of the cuboid in the figure, with its 4×3 base and height 2, start with the diagonal of the base: $AC = \sqrt{4^2 + 3^2} = 5$. This is the projection of AG onto the base, and $CG = 2$ is vertical, meeting the base at a right angle. In the right triangle ACG,

$$\tan\theta = \frac{CG}{AC} = \frac{2}{5} \implies \theta = \arctan(0.4) \approx 21.8^\circ.$$

The order is the same each time. Find the projection on the plane, then the perpendicular

height, then use $\tan$.

7.1.2 The angle between two intersecting lines

Two lines that meet also make an angle, and inside a solid those two lines are often edges from the same corner. Take the two edges, then join their far ends. The three segments form a triangle, and the angle you want sits at the corner. Measure the three sides with the distance formula, then find the angle with the cosine rule of §7.3. When the triangle is isosceles, as in the example below, halving it gives a right triangle instead.

EXAMPLE 7.1

A square-based pyramid has base corners $A(0,0,0)$, $B(6,0,0)$, $C(6,6,0)$, $D(0,6,0)$ and apex $V(3,3,4)$. Find the angle between the slant edges VA and VB.

The two edges meet at V and span the triangle VAB, so the angle we want is $A\hat{V}B$. First measure the triangle. By the distance formula,

$$VA = \sqrt{(3-0)^2 + (3-0)^2 + (4-0)^2} = \sqrt{34},$$

and $VB = \sqrt{34}$ by the same calculation, while $AB = 6$ along the base.

The triangle is isosceles, so cut it in half. The midpoint of AB is $M(3,0,0)$, and VM meets AB at a right angle. In the right triangle VMA, the side $AM = 3$ is opposite the half-angle at V, and the hypotenuse is $VA = \sqrt{34}$:

$$\sin\frac{\theta}{2} = \frac{3}{\sqrt{34}} \implies \frac{\theta}{2} \approx 31.0^\circ \implies \theta \approx 61.9^\circ.$$

The solution uses the two formulas of this section: the distance formula for the three sides, the midpoint to cut the triangle in half, then right-angled trigonometry for the angle.

YOUR TURN 7.1 Three-Dimensional Geometry

Distance and midpoint

1. Find the distance between each pair of points.

(a) $(0, 0, 0)$ and $(2, 3, 6)$
(b) $(2, -1, 0)$ and $(-2, 3, 7)$
(c) $(-3, 4, 1)$ and $(3, 1, -1)$
(d) $(1.5, 0, -2)$ and $(-0.5, 4, 2)$

2. Answer each part.

(a) Find the midpoint of $(1, 2, 3)$ and $(5, 4, 1)$.
(b) Find the midpoint of $(2, -1, 4)$ and $(6, 3, 0)$.
(c) $M(2, -1, 5)$ is the midpoint of $[AB]$, where A is $(-1, 3, 4)$. Find the coordinates of B.

Angle between a line and a plane

3. A cuboid has base 3×4 and height 12.

(a) Find the length of the diagonal of the base.
(b) Find the length of the space diagonal.
(c) Find the angle between the space diagonal and the base.

4. A square-based pyramid has base side 6, and its apex is 5 vertically above the centre of the base.

(a) Find the distance from a corner of the base to the centre of the base.
(b) Find the angle between a slant edge and the base.

Angle between two lines

5. The points $O(0, 0, 0)$, $P(4, 3, 0)$ and $Q(0, 3, 2)$ are three corners of a cuboid.

(a) Find the lengths OP, OQ and PQ.
(b) Use the cosine rule to find the angle $P\hat{O}Q$ between the lines (OP) and (OQ).

6. A pyramid has square base $A(0, 0, 0)$, $B(4, 0, 0)$, $C(4, 4, 0)$, $D(0, 4, 0)$ and apex $V(2, 2, 6)$.

(a) Show that $VA = VB = 2\sqrt{11}$.
(b) Find the angle $A\hat{V}B$ between the slant edges VA and VB. *(Hint: the midpoint of $[AB]$ cuts triangle VAB into two right triangles)*

Concept check

7. A cuboid has a base 6 by 8 and height 5. To find the angle between a space diagonal and the base, a student writes $\tan\theta = \dfrac{5}{8}$, so $\theta \approx 32.0°$.

(a) Identify the error.
(b) Find the correct angle.

7.2 Volumes and Surface Areas of Solids

You already know the cuboid, the prism and the cylinder. The formula booklet adds the sphere, the cone and the pyramid. Knowing where each formula comes from makes it harder to misapply.

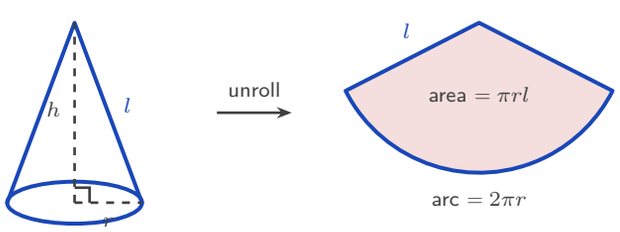
KEY · IB

Sphere, radius r	$V = \frac{4}{3}\pi r^3, \quad A = 4\pi r^2$
Cone, radius r, height h	$V = \frac{1}{3}\pi r^2 h, \quad$ curved $A = \pi r l$
Pyramid, base area B, height h	$V = \frac{1}{3}Bh$
Cylinder, radius r, height h	$V = \pi r^2 h, \quad$ curved $A = 2\pi r h$

The cylinder is listed in the *Prior learning* table at the front of the booklet rather than beside the other three.

Each formula has a short reason behind it.

- **Cylinder.** Cut the curved side along one vertical line and unroll it into a rectangle of height h and width $2\pi r$, the base circumference. Its area is $2\pi r h$.
- **Cone.** The same cut unrolls the curved side into a sector. Its radius is the **slant height** l, measured from the rim to the tip, and its arc is the base circumference $2\pi r$. A sector's area is half its arc times its radius, so the curved area is $\frac{1}{2}(2\pi r)(l) = \pi r l$.
- **The $\frac{1}{3}$.** A cube cuts into three identical square pyramids. Each one stands on a face of the cube, and all three tips meet at the opposite corner. Each is therefore a third of the cube, and the same fraction holds for every pyramid and every cone: $V = \frac{1}{3}Bh$.
- **Sphere.** Put a sphere inside the smallest cylinder that holds it, of radius r and height $2r$. Archimedes proved the sphere is two thirds of that cylinder. The cylinder holds $\pi r^2(2r) = 2\pi r^3$, so the sphere holds $\frac{4}{3}\pi r^3$. Its surface, $4\pi r^2$, is the area of four **great circles** (a circle round the sphere at its widest).



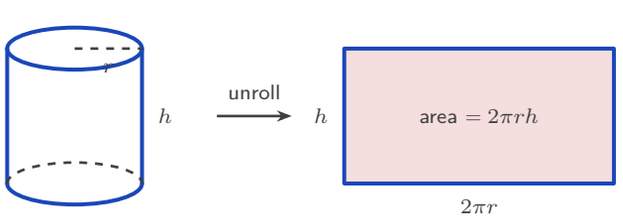


Figure 7.2 Cut the curved surface of a cone along one slant line and it opens into a sector of radius l whose arc is the base circumference. Cut a cylinder the same way and it opens into a rectangle. Each curved-area formula is the area of the flat shape it opens into.

Two mistakes are common. The cone's curved surface uses the slant height l, not the vertical height h; Pythagoras links them by $l = \sqrt{r^2 + h^2}$. And "total surface area" means every face: a solid hemisphere has a curved part $2\pi r^2$ *and* a flat base πr^2, giving $3\pi r^2$, while the same hemisphere as an open bowl has only $2\pi r^2$. For a solid made of two parts, count only the exposed surface: the faces where the parts join are hidden inside and do not count.

Worked through. For a solid made of a cone of base radius 8 and height 6 sitting on top of a hemisphere of radius 8, the exposed surface is the cone's curved part plus the hemisphere's curved part. The two flat circles where they join are hidden inside, so neither counts. The cone has slant height

$$l = \sqrt{8^2 + 6^2} = 10,$$

so its curved area is $\pi r l = 80\pi$. The hemisphere's curved area is $2\pi r^2 = 128\pi$, and the total surface area is

$$80\pi + 128\pi = 208\pi \approx 653.$$

Deciding which faces are exposed is the harder step; the formulas themselves are routine.

7.2.1 Similar solids

Every formula above is a constant times a length squared, for an area, or a length cubed, for a volume. That fixes what happens when a solid is enlarged.

Two solids are **similar** when they have the same shape and all corresponding lengths are in the same ratio. That ratio is the **linear scale factor** k. Multiplying every length by k multiplies $4\pi r^2$ by k^2 and $\frac{4}{3}\pi r^3$ by k^3.

KEY **Scale factors.** If two similar figures have lengths in the ratio k, then

$$\frac{\text{length}_2}{\text{length}_1} = k, \qquad \frac{\text{area}_2}{\text{area}_1} = k^2, \qquad \frac{\text{volume}_2}{\text{volume}_1} = k^3.$$

One ratio of lengths settles every area and every volume.

A cone cut parallel to its base is the case to recognise. The small cone above the cut has the same apex and the same slant, so it is similar to the whole, and one ratio of heights gives its radius, its surface and its volume at once. The piece left below the cut is not similar to either cone, so no scale factor applies to it. Find its volume as the difference between the whole cone and the small cone.

EXAMPLE 7.2

A cone has base radius 9 cm and height 12 cm. A cut parallel to the base is made 4 cm below the apex. Find the radius of the circular cut and the volume of the small cone above it.
The small cone is similar to the whole cone, and the two heights are known, so

$$k = \frac{4}{12} = \frac{1}{3}.$$

The radius is a length, so it scales by k:

$$r = \tfrac{1}{3}(9) = 3 \text{ cm}.$$

Similar triangles say the same thing directly: the vertical axis and the radius of the cut form a triangle similar to the one made by the axis and the base radius, so $\dfrac{r}{4} = \dfrac{9}{12}$.
The whole cone holds

$$V = \tfrac{1}{3}\pi(9)^2(12) = 324\pi \text{ cm}^3,$$

and volume scales by k^3, so the small cone holds

$$\left(\tfrac{1}{3}\right)^3 (324\pi) = \tfrac{1}{27}(324\pi) = 12\pi \text{ cm}^3.$$

Computing it directly gives $\frac{1}{3}\pi(3)^2(4) = 12\pi$, which agrees. The piece left below the cut has volume $324\pi - 12\pi = 312\pi$ cm^3, the difference of the two cones.

YOUR TURN 7.2 Volumes and Surface Areas of Solids

Volume and surface area

8. EXACT Find the volume of each solid. Leave your answers in terms of π where appropriate.

(a) sphere of radius 3
(b) cone of radius 3, height 4
(c) pyramid with base area 20, height 9
(d) cylinder of radius 4, height 10

9. EXACT Find the surface area asked for in each case.

(a) the surface area of a sphere of radius 5
(b) the curved surface area of a cone of radius 5, slant height 13
(c) the total surface area of a solid hemisphere of radius 4
(d) the curved surface area of a cylinder of radius 4, height 10
(e) the surface area of an open hemispherical bowl of radius 6 (the outside surface only)

10. EXACT A cone has radius 7 and height 24.

(a) Find its slant height.
(b) Hence find its curved surface area.

Composite solids

11. EXACT A solid is a cylinder of radius 3 and height 10, with a hemisphere of radius 3 fixed to its top face.

(a) Explain which surfaces are exposed.
(b) Find the total surface area, in terms of π.
(c) Find the total volume, in terms of π.

Finding a missing dimension

12. EXACT Work backwards from the given measurement.

(a) A sphere has volume 288π. Find its radius.
(b) A cone of radius 6 has volume 96π. Find its height, then its curved surface area.
(c) A cylinder of height 10 has curved surface area 120π. Find its radius, then its volume.

Similar solids

13. Two similar cones have heights 6 cm and 9 cm. The smaller cone has volume 40 cm^3 and total surface area 24 cm^2.

(a) Find the volume of the larger cone.
(b) Find the total surface area of the larger cone.

14. Two similar jugs hold 250 cm^3 and 2000 cm^3.

(a) Find the ratio of their heights.
(b) The smaller jug has surface area 150 cm^2. Find the surface area of the larger jug.

15. EXACT A cone has base radius 6 cm and height 12 cm. It is cut parallel to its base, 8 cm below the apex, and the small cone above the cut is removed. The piece left below the cut is called a *frustum*.

(a) Find the radius of the circular cut.
(b) Find the volume of the small cone as a fraction of the volume of the whole cone.
(c) Find the volume of the frustum, in terms of π.

Concept check

16. EXACT A cone has base radius 5 and slant height 13. A student writes: "$V = \frac{1}{3}\pi(5^2)(13) = \frac{325\pi}{3}$."

(a) Identify the error.
(b) Find the correct volume, in terms of π.

17. Two jugs are similar, and one is twice as tall as the other. A student says: "*the larger jug is twice as tall, so it holds twice as much.*" Explain the error, and state the correct factor for the capacity.

7.3 The Sine and Cosine Rules

Right-angled triangles are solved with $\sin$, $\cos$ and $\tan$. Name the sides opposite, adjacent and hypotenuse relative to an angle, and $\sin\theta = \frac{o}{h}$, $\cos\theta = \frac{a}{h}$, $\tan\theta = \frac{o}{a}$. Most triangles have no right angle. Two rules extend trigonometry to all of them, and between them they solve any triangle from three known parts, provided at least one of them is a side.

Both rules use one labelling. A side takes the lower-case letter of the angle *opposite* it, so side a faces angle A.

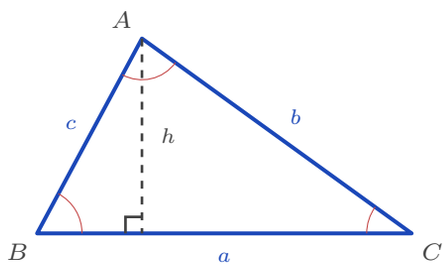


Figure 7.3 Each side is named after the angle opposite it: side a faces angle A, and so on. The dashed altitude h splits the triangle into two right triangles, which is where the sine rule comes from.

7.3.1 The sine rule

Drop a perpendicular of height h from A onto side a, as in the figure above. It cuts the triangle into two right triangles, and in each one h is the side opposite a known angle. On the left, $h = c\sin B$. On the right, $h = b\sin C$. Both expressions measure the same segment, so $c\sin B = b\sin C$, which rearranges to $\frac{b}{\sin B} = \frac{c}{\sin C}$. Dropping a perpendicular from a different corner gives the third ratio.

KEY · IB **Sine rule.** In any triangle,

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}.$$

The sine rule pairs a side with the angle facing it. Use it whenever you know one complete pair, plus one more side or angle. The standard shorthand writes S for a known side and A for a known angle, in the order they occur round the triangle. These are the cases AAS, ASA and SSA.

7.3.2 The ambiguous case

The sine rule needs one extra check when you use it to find an *angle*.

Suppose you know two sides and an angle that is *not* between them. This is the SSA case. The sine rule gives an equation of the form $\sin B = k$. If $k > 1$ there is no solution and no triangle; if $k = 1$ there is one solution, $B = 90^\circ$. If $0 < k < 1$, the equation has two solutions between 0° and 180°: the angle B and its supplement $180^\circ - B$. An angle and its supplement have the same sine, so the equation cannot separate them.

Your calculator returns only the first. Test the second yourself. Add it to the angle you were given: if the total is still below 180°, a positive third angle remains, so a second triangle exists and both answers are correct.

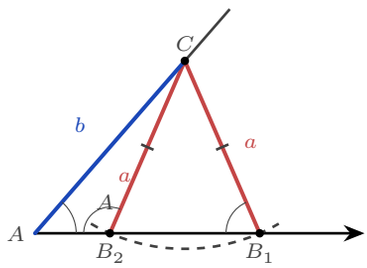


Figure 7.4 Two sides and a non-included angle. Swinging the fixed length a from C meets the base at B_2 and at B_1, so two different triangles share the same a, b and A. The angles at B_2 and B_1 are supplementary, which is why the sine rule returns both.

EXAMPLE 7.3

In triangle ABC, $a = 9$, $b = 11$, and $A = 40^\circ$. Find angle B.
By the sine rule, $\dfrac{\sin B}{11} = \dfrac{\sin 40^\circ}{9}$, so

$$\sin B = \frac{11 \sin 40^\circ}{9} \approx 0.786.$$

The calculator gives $B \approx 51.8^\circ$. The supplement is $180^\circ - 51.8^\circ = 128.2^\circ$, so test it: $A + B = 40^\circ + 128.2^\circ = 168.2^\circ$, which is below 180°. A third angle of 11.8° remains, so that triangle exists too.
Both values of B are correct. There are two triangles, and a complete answer reports both.

CAUTION This check is needed only when the sine rule gives an *angle*. If you are finding a *side*, or using the cosine rule, no supplement arises. The cosine rule returns an angle between 0° and 180° directly, because cos is negative for obtuse angles and positive for acute ones, so it distinguishes the two.

7.3.3 The cosine rule

When the known angle sits *between* the two known sides, no side is paired with the angle opposite it, so the sine rule cannot start. The cosine rule covers this case. Read it as Pythagoras' theorem with a correction term for the angle not being 90°.

Coordinates prove it directly. Put the corner C at the origin with side a along the x-axis, so B sits at $(a, 0)$. The point A lies a distance b from the origin, at angle C to the positive x-axis. By the definitions of cosine and sine, its coordinates are $(b\cos C,\ b\sin C)$. The distance formula now measures the third side:

$$c^2 = (a - b\cos C)^2 + (b\sin C)^2 = a^2 - 2ab\cos C + b^2\cos^2 C + b^2\sin^2 C,$$

and $\cos^2 C + \sin^2 C = 1$, so the last two terms add to b^2.

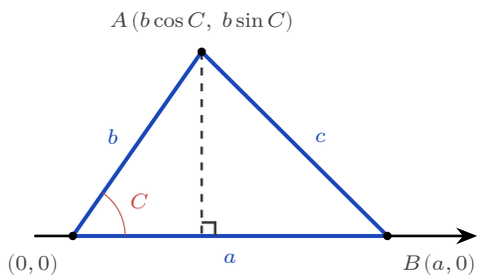


Figure 7.5 Placing C at the origin with a along the axis fixes all three corners. Only A needs trigonometry, and the distance from A to B is then the cosine rule.

KEY · IB **Cosine rule.** In any triangle,

$$c^2 = a^2 + b^2 - 2ab\cos C, \qquad \text{equivalently} \qquad \cos C = \frac{a^2 + b^2 - c^2}{2ab}.$$

The letters rotate. Written about a different corner the same rule reads $a^2 = b^2 + c^2 - 2bc\cos A$, and the pattern is the same in each version: the side opposite the named angle stands alone on the left.

When $C = 90°$, $\cos C = 0$ and the rule becomes $c^2 = a^2 + b^2$, which is Pythagoras' theorem. The first form finds a third side from two sides and the angle between them (SAS). The rearranged form finds an angle from all three sides (SSS).

7.3.4 The area of a triangle

The familiar area $\frac{1}{2}$ base $\times$ height becomes a trigonometric formula as soon as the height is written in terms of a side. Take b as the base. The perpendicular height from B down to it is the opposite side of the right triangle at C, so $h = a\sin C$ and

$$\text{Area} = \tfrac{1}{2}\, b\,(a\sin C).$$

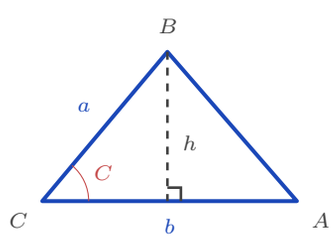


Figure 7.6 The height of the triangle on base b is the opposite side of the right triangle at C, so it is $a\sin C$. Substituting it into $\frac{1}{2}\times$ base $\times$ height gives the formula.

KEY · IB **Area of a triangle.** Area $= \frac{1}{2}ab\sin C$, using two sides and the angle between them.

This formula needs exactly what the cosine rule needs, so one sketch usually answers both questions at once.

Worked through. A triangular field has $AB = 90$ m, $AC = 70$ m, and the angle at A is $110°$. The known angle lies between the two known sides, so the cosine rule finds the third side BC. Rotate the letters to put A in the named position:

$$BC^2 = 90^2 + 70^2 - 2(90)(70)\cos 110° \approx 17\,309 \implies BC \approx 132 \text{ m}.$$

The area of the field needs the same three pieces of information, so no further measurement is required:

$$\text{Area} = \tfrac{1}{2}(90)(70)\sin 110° \approx 2960 \text{ m}^2.$$

Two sides and the angle between them fix a triangle completely. Once you hold that combination, every other part of the triangle can be found.

What you are given	What to use
Two angles and any side (AAS, ASA)	sine rule
Two sides and an angle opposite one (SSA)	sine rule, plus the ambiguous case
Two sides and the angle between them (SAS)	cosine rule, for the third side
All three sides (SSS)	cosine rule rearranged, for an angle
The same (SAS), when the area is wanted	area $= \frac{1}{2}ab\sin C$

Read the table by counting what you know. Three parts, at least one of them a side, determine the triangle, except that SSA may allow two. The arrangement of the three parts chooses the rule.

YOUR TURN 7.3 The Sine and Cosine Rules

The sine rule

18. Use the sine rule to find the required side or angle.

(a) $a = 7$, $A = 40°$, $B = 60°$: find b
(b) $a = 9$, $A = 35°$, $B = 55°$: find b
(c) $A = 45°$, $B = 60°$, $a = 10$: find C and b
(d) $a = 12$, $b = 10$, $A = 65°$: find B, and explain why there is only one possible value

The cosine rule

19. Use the cosine rule.

(a) $b = 5$, $c = 8$, $A = 60°$: find a
(b) $a = 6$, $b = 9$, $C = 110°$: find c
(c) $a = 6$, $b = 7$, $c = 8$: find angle A
(d) sides 5, 6, 7: find the largest angle

The area of a triangle

20. Find the area of each triangle.

(a) sides 5 and 8, included angle $30°$
(b) sides 9 and 6, included angle $125°$
(c) sides 7 and 9, included angle $40°$
(d) sides 5, 7 and 8: find the angle opposite the side 7 first

21. In triangle PQR, $PQ = 11$ cm, $PR = 7$ cm and $Q\hat{P}R = 52°$.

(a) Find the length QR.
(b) Find the area of the triangle.

The ambiguous case

22. For each triangle ABC, use the sine rule to find $\sin B$. State how many triangles fit the data, and give every possible value of B.

(a) $a = 6$, $b = 10$, $A = 45°$
(b) $a = 4$, $b = 8$, $A = 30°$
(c) $a = 7$, $b = 10$, $A = 35°$

23. In triangle ABC, $a = 7$, $b = 9$ and $A = 45°$.

(a) Find the two possible values of angle B.
(b) Show that both give a valid triangle.
(c) Explain why this situation cannot arise when the given angle is included between the two given sides.
(d) For each triangle, find the length of side c.

Concept check

24. To find a when $b = 5$, $c = 8$ and $A = 60°$, a student writes

$$a^2 = 5^2 + 8^2 - 2(5)(8)\cos 60° = 9\cos 60° = 4.5, \quad \text{so } a \approx 2.12.$$

Identify the error and find the correct value of a.

7.4 Elevation, Depression and Bearings

Real questions do not arrive as labelled triangles. They arrive as heights, distances and directions: how far away is the boat, what heading takes the ship home. Two conventions turn that language into a triangle you can solve.

An **angle of elevation** is measured upward from the horizontal to a line of sight. An **angle of depression** is measured downward from it. The observer's horizontal and the ground are parallel lines. So the angle of depression from a cliff top down to a boat equals the angle of elevation from the boat up to the cliff top. They are alternate angles, and either one can be used in the triangle.

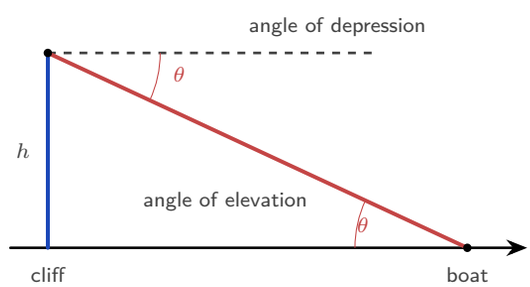


Figure 7.7 The dashed line at the top is the observer's horizontal, and the ground is parallel to it. The line of sight crosses both, so the two marked angles are alternate angles and are equal.

When the distance to the object is unknown, one sighting is not enough to find a height: two unknowns, one equation. A second sighting from a known distance nearer supplies the missing equation. The height is the same from both positions, so writing it twice and equating the two expressions eliminates the distance.

EXAMPLE 7.4

From a point on level ground the angle of elevation of the top of a tower is $25°$. From a point 30 m nearer the tower, on the same straight line, the angle of elevation is $40°$. Find the height of the tower.

Let h be the height and let d be the distance from the *nearer* point to the foot of the tower. The far point is then $d + 30$ from the foot. Each sighting gives a right triangle with h opposite

the angle and the ground distance adjacent to it, so

$$h = d\tan 40^\circ \qquad \text{and} \qquad h = (d+30)\tan 25^\circ.$$

Both measure the same tower, so set them equal and collect the terms in d:

$$d\tan 40^\circ = (d+30)\tan 25^\circ \implies d\,(\tan 40^\circ - \tan 25^\circ) = 30\tan 25^\circ,$$

$$d = \frac{30\tan 25^\circ}{\tan 40^\circ - \tan 25^\circ} \approx \frac{13.99}{0.3728} \approx 37.5 \text{ m}.$$

The height follows from either equation:

$$h = 37.5\tan 40^\circ \approx 31.5 \text{ m}.$$

Check it against the other sighting: the far point is 67.5 m away, and $\arctan\left(\frac{31.5}{67.5}\right) \approx 25^\circ$, as given. Measuring d from the nearer point is what keeps the second distance as $d+30$; taking d from the far point instead makes it $d-30$, and mixing the two gives an impossible answer.

A **bearing** gives a direction as an angle measured clockwise from north. It runs from 000° up to, but not including, 360°, and is always written with three figures, so due east is 090° and not 90°. Three figures keep the reading unambiguous on a chart. Due south is 180°, due west is 270°, and north-west is 315°.

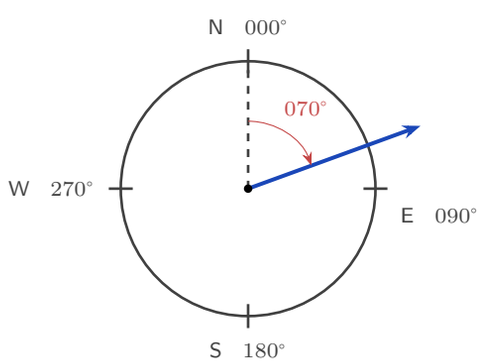


Figure 7.8 A bearing is measured clockwise from north, always with three figures. The arrow here is on a bearing of 070°.

The **back bearing** is the direction of the return journey, and it differs from the outward bearing by 180°. Add 180° when the bearing is below 180°, and subtract 180° when it is 180° or more. Either way the answer stays between 000° and 360°.

EXAMPLE 7.5

A ship sails 60 km from port P on a bearing of 050° to a point Q, then 80 km on a bearing of 140° to R. Find the distance PR.

The two bearings differ by $140^\circ - 50^\circ = 90^\circ$, so the ship turns through a right angle at Q: the

angle PQR is $90°$. Then PR is a hypotenuse, and Pythagoras gives

$$PR = \sqrt{60^2 + 80^2} = \sqrt{10000} = 100 \text{ km}.$$

Sketching the bearings from a north line at P and at Q is what shows the right angle. The arithmetic then follows from the picture.

Most journeys do not turn so conveniently. When the corner angle is not $90°$, read the two distances and that angle off the same sketch. Find the missing distance with the cosine rule, then find the direction back with the sine rule. Many bearings questions have this shape, and the steps below handle them in the same order.

KEY **Solving a bearings problem.**

1. Sketch the journey, and draw a north line at **every** point where a bearing is given.
2. Mark each bearing as an angle clockwise from its own north line.
3. Find the interior angle of the triangle at the corner: the difference between the back bearing of the first leg and the bearing of the second leg. If that difference is more than $180°$, subtract it from $360°$.
4. Use the cosine rule for the missing distance, then the sine rule for the missing angle.
5. Turn that angle back into a bearing: add it to the bearing of the first leg if the route turns right (clockwise) at the corner, and subtract it if the route turns left. The sketch shows which.

Do not skip step 1: without a north line drawn at the corner, step 3 is easy to get wrong.

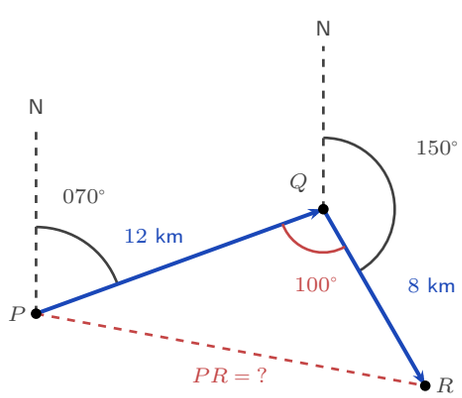


Figure 7.9 The two bearings are measured clockwise from a north line at P and at Q. Together they fix the interior angle $P\hat{Q}R = 100°$, which is what the cosine rule needs.

Worked through. For the journey in the figure, a hiker walks 12 km from P on a bearing of $070°$ to Q, then 8 km on a bearing of $150°$ to R. To find the distance PR and the bearing of R from P, first construct the angle at the corner. The bearing back from Q to P is $070° + 180° = 250°$, and the onward bearing is $150°$, so the interior angle is $P\hat{Q}R = 250° - 150° = 100°$. Subtracting the second bearing from the back bearing is the

standard method, and the sketch agrees: the angle at the corner is obtuse. There is no right angle, so the cosine rule finds the closing side:

$$PR^2 = 12^2 + 8^2 - 2(12)(8)\cos 100^\circ \approx 241.3 \implies PR \approx 15.5 \text{ km}.$$

For the bearing of R from P, find the angle $Q\hat{P}R$ by the sine rule:

$$\frac{\sin(Q\hat{P}R)}{8} = \frac{\sin 100^\circ}{15.54} \implies \sin(Q\hat{P}R) \approx 0.507 \implies Q\hat{P}R \approx 30.5^\circ.$$

The route turns right at Q, so add. The bearing of R from P is $070^\circ + 30.47^\circ \approx 100.47^\circ$, which is 100° to the nearest degree. Every step came from the labelled diagram: the bearings give the corner angle, the cosine rule the closing side, and the sine rule the final direction.

YOUR TURN 7.4 Elevation, Depression and Bearings

Elevation and depression

25. Find the angle of elevation in each case.

(a) a 50 m tower, from 120 m away on level ground

(b) a 45 m tower, from 90 m away on level ground

26. In each case, find the horizontal distance asked for.

(a) From the top of an 80 m cliff, the angle of depression to a boat is 25°. Find the distance from the foot of the cliff to the boat.

(b) From the top of a 30 m building, the angle of depression to a car is 40°. Find the distance from the foot of the building to the car.

(c) From the top of a 60 m cliff, the angles of depression to two buoys are 20° and 35°. The buoys and the foot of the cliff lie on one straight line, with both buoys on the same side of the cliff. Find the distance between the buoys.

Two sightings

27. From a point A on level ground, the angle of elevation of the top of a building is 30°. From a point B, 50 m nearer the building on the same straight line, the angle of elevation is 45°. Find the height of the building. *(Hint: let d be the distance from B to the foot of the building, and write the height twice)*

28. From a boat, the angle of elevation of the top of a lighthouse is 12°. The boat sails 200 m directly towards the lighthouse, and the angle of elevation is now 20°.

(a) Find the height of the lighthouse.
(b) Find the distance from the boat to the foot of the lighthouse at the second sighting.

Bearings

29. Answer each question about bearings.

(a) Write the bearing of due west as a three-figure bearing.
(b) A ship travels on a bearing of 120°. Find its back bearing.
(c) B is on a bearing of 065° from A. Find the bearing of A from B.
(d) Find the back bearing of 200°, and explain why you subtract here rather than add.

Journeys with two legs

30. A ship sails 15 km from P on a bearing of 040° to Q, then 20 km on a bearing of 110° to R.

(a) Show that angle $P\hat{Q}R = 110^\circ$.
(b) Find the distance PR.
(c) Find the bearing of R from P.

31. A hiker walks 25 km from P on a bearing of 080° to Q, then 18 km on a bearing of 330° to R.

(a) Find the angle $P\hat{Q}R$.
(b) Find the distance PR.
(c) Find the bearing of R from P.
(d) Find the bearing of P from R.

Concept check

32. From the top of a 50 m cliff, the angle of depression to a boat is 20°. A student places the 20° between the cliff face and the line of sight, and writes $d = 50\tan 20^\circ \approx 18.2$ m.

(a) Identify the error.
(b) Find the correct distance from the foot of the cliff to the boat.

7.5 Radians, Arc Length, and Sector Area

The degree is a human invention. The Babylonians cut the circle into 360 parts, and we inherited the choice. Nothing about a circle requires 360. The circle does suggest a unit of its own: take the radius, bend it round the circumference, and measure the angle it makes at the centre. That angle is one **radian**. Because the circumference is $2\pi r$, exactly 2π radii fit around it, so a full turn is 2π radians and one radian is about $57.3°$.

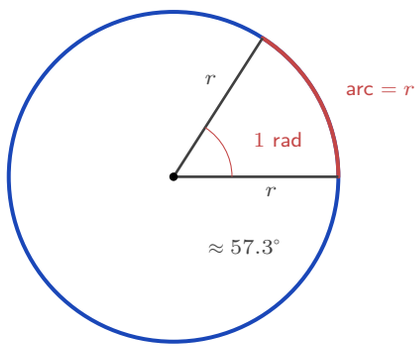


Figure 7.10 One radian is the angle at the centre for which the arc is exactly as long as the radius. A full turn takes 2π of them, because the circumference is 2π radii long.

KEY 2π radians $= 360°$, so π radians $= 180°$. Convert by multiplying: degrees to radians $\times \dfrac{\pi}{180}$; radians to degrees $\times \dfrac{180}{\pi}$.

The angles that appear most often should be known without conversion.

Degrees	$30°$	$45°$	$60°$	$90°$	$180°$	$360°$
Radians	$\dfrac{\pi}{6}$	$\dfrac{\pi}{4}$	$\dfrac{\pi}{3}$	$\dfrac{\pi}{2}$	π	2π

The reason for the new unit appears at once: arc length and sector area become simple fractions of the whole circle. An arc is the same fraction of the circumference as its angle θ is of a full turn, so $\dfrac{l}{2\pi r} = \dfrac{\theta}{2\pi}$, and the two 2π's cancel.

KEY · IB For a sector of radius r and angle θ **in radians**,

$$\text{arc length } l = r\theta, \qquad \text{sector area } A = \tfrac{1}{2}r^2\theta.$$

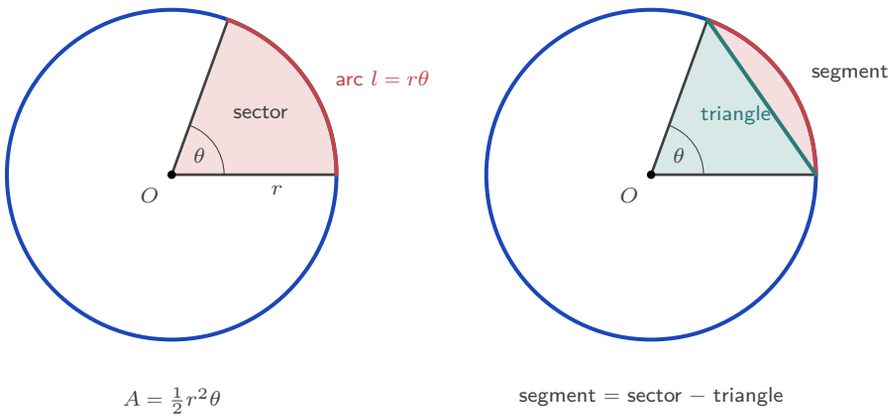


Figure 7.11 **Left:** the *sector* is the pie slice between two radii and the arc. **Right:** the chord cuts that same sector into a triangle and a *segment*, so subtracting the triangle's area from the sector's leaves the segment.

The area works the same way. The sector is the fraction $\dfrac{\theta}{2\pi}$ of the whole disc πr^2, so $A = \dfrac{\theta}{2\pi} \cdot \pi r^2 = \frac{1}{2}r^2\theta$. A **segment** is the region between a chord and its arc. It is a sector with the triangle removed, and that triangle has two sides r with the angle θ between them, so §7.3 gives its area as $\frac{1}{2}r^2 \sin\theta$:

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

CAUTION Every formula in this section needs θ in **radians**. Substituting an angle in degrees into $l = r\theta$ or $A = \frac{1}{2}r^2\theta$ gives an answer that is wrong by a factor of $\frac{180}{\pi}$ (for the segment formula the error is not a fixed factor), and the working looks correct, so the error is easy to miss. Check the mode on your calculator before you start.

Worked through. For a sector of radius 10 cm and angle $\frac{2\pi}{5}$ radians, the formulas give the arc length and area directly:

$$l = r\theta = 10 \cdot \frac{2\pi}{5} = 4\pi \approx 12.6 \text{ cm}, \qquad A = \tfrac{1}{2}r^2\theta = \tfrac{1}{2}(100)\frac{2\pi}{5} = 20\pi \approx 62.8 \text{ cm}^2.$$

For the segment cut off by the chord, subtract the triangle from the sector:

$$\tfrac{1}{2}r^2(\theta - \sin\theta) = 50\left(\frac{2\pi}{5} - \sin\frac{2\pi}{5}\right) \approx 50(1.2566 - 0.9511) \approx 15.3 \text{ cm}^2.$$

Each formula took the angle in radians, exactly as given. The segment is far smaller than the sector, because most of the sector is triangle.

A sector's boundary is not the arc alone. Walk round the edge and you travel out along one radius, round the arc, and back along the other radius. The perimeter of a sector is therefore the arc plus two radii, and it has to be assembled from $l = r\theta$ rather than looked up, because the formula booklet does not print it.

KEY **Perimeter of a sector** of radius r and angle θ in radians:

$$P = 2r + r\theta = r(2 + \theta).$$

Two straight radii and one curved arc. This one is not in the formula booklet.

EXAMPLE 7.6

A sector has radius 7 cm and angle 1.5 radians. Find its perimeter.

The arc is

$$l = r\theta = 7(1.5) = 10.5 \text{ cm},$$

and the two radii add 14 cm, so

$$P = 2(7) + 10.5 = 24.5 \text{ cm}.$$

A quick sanity check: the two radii alone measure 14 cm, and an angle of 1.5 radians is just under a quarter turn, so the arc should be 1.5 radii long, which 10.5 cm is.

Given the perimeter and the area together, neither formula can be used on its own, because each contains both r and θ. Treat them as simultaneous equations, and eliminate θ. The resulting quadratic usually gives two values of r. Find θ for each, and reject any value with $\theta > 2\pi$, since no sector has an angle of more than a full turn.

EXAMPLE 7.7

A sector has perimeter 36 cm and area 45 cm^2. Find its radius and its angle.

Write down both conditions:

$$2r + r\theta = 36, \qquad \tfrac{1}{2}r^2\theta = 45.$$

The second gives $r^2\theta = 90$, so $\theta = \dfrac{90}{r^2}$. Substituting that into the first,

$$2r + r \cdot \frac{90}{r^2} = 36 \implies 2r + \frac{90}{r} = 36.$$

Multiply through by r and collect:

$$2r^2 - 36r + 90 = 0 \implies r^2 - 18r + 45 = 0 \implies (r-3)(r-15) = 0,$$

so $r = 3$ or $r = 15$. Test each against $\theta = \dfrac{90}{r^2}$. With $r = 3$ the angle is $\theta = 10$ radians, which is more than a full turn of $2\pi \approx 6.28$, so no such sector exists. With $r = 15$ the angle is $\theta = 0.4$ radians.

So $r = 15$ cm and $\theta = 0.4$ radians. Check: the arc is $15(0.4) = 6$ cm, giving a perimeter of $30 + 6 = 36$ cm, and the area is $\frac{1}{2}(225)(0.4) = 45$ cm^2. ✓ The quadratic will usually offer two radii, and rejecting the one whose angle exceeds 2π is part of the answer, not an afterthought.

EXAMPLE 7.8

The opener measured the Earth from a hilltop. There is an older method, and it uses this section's formula. Eratosthenes measured the sun's rays at 7.2° to the vertical in Alexandria, while in Syene they were vertical. He took the arc between the towns to be 5000 stadia. Find the Earth's radius and circumference in stadia.

Convert the angle first:

$$7.2^\circ \times \frac{\pi}{180} = \frac{\pi}{25} \text{ radians.}$$

The two towns and the Earth's centre make a sector whose arc is the distance between them, so $l = r\theta$ can be used in reverse:

$$5000 = r \cdot \frac{\pi}{25} \implies r = \frac{125\,000}{\pi} \approx 39\,800 \text{ stadia,}$$

and the circumference is $2\pi r = 250\,000$ stadia. Since 7.2° is $\frac{1}{50}$ of a full turn, this is Eratosthenes' own argument: the circumference is 50 times the arc. Taking a stadion as roughly 158 m gives a circumference near $39\,500$ km, within a couple of percent of the true $40\,075$ km.

YOUR TURN 7.5 Radians, Arc Length, and Sector Area

Degrees and radians

33. EXACT Convert to radians, in terms of π.

(a) 60° (b) 135°

(c) 210° (d) 90°

34. Convert to degrees.

(a) $\dfrac{3\pi}{4}$ (b) $\dfrac{2\pi}{3}$

(c) 1 radian, to three significant figures

Arc length and sector area

35. Find the arc length and the area of each sector.

(a) radius 8, angle 1.5 radians (b) radius 6, angle $\dfrac{\pi}{3}$ radians

(c) radius 9, angle 2.4 radians (d) radius 4, angle 90°

36. In each case the angle is unknown. Find it from the information given.

(a) A sector has radius 10 and arc length 25. Find its angle in radians.

(b) A sector has radius 6 and area 54. Find its angle in radians, then find its arc length.

Segments and perimeters

37. A chord subtends an angle of 2 radians at the centre of a circle of radius 8.

(a) Find the area of the sector.

(b) Find the area of the triangle formed by the chord and the two radii.

(c) Find the area of the minor segment.

(d) Find the perimeter of the sector.

Perimeter and area together

38. In each case, find the radius and the angle of every sector that fits the information.

(a) The perimeter is 40 cm and the area is 64 cm^2.

(b) The perimeter is 20 cm and the area is 24 cm^2.

Concept check

39. EXACT A student finds the arc length of a sector of radius 6 and angle 60° as $s = r\theta = 6 \times 60 = 360$. Identify the error and find the correct arc length.

Chapter 7 · Shapes & Measures

Reflections

TOK The formulas $l = r\theta$ and $A = \frac{1}{2}r^2\theta$ are this simple only when θ is in radians, and later you will meet another case: the derivative of $\sin x$ is exactly $\cos x$, again only in radians. The degree was chosen by people, the Babylonians. The radian is fixed by the circle: it is the angle whose arc equals its radius. Does a unit that the circle chooses make the mathematics truer, or only simpler?

IM Trigonometry was built by many people over many centuries. The Greek astronomer Hipparchus made the first table of chords around 150 BCE. In India, Aryabhata replaced the chord with the half-chord and called it the *jyā*, which is our sine. The word passed into Arabic, was misread there as *jaib*, meaning "bay", and was translated into Latin as *sinus*. Islamic astronomers, al-Bīrūnī among them, extended the sine and cosine rules, partly to find the direction of Mecca from any city.

RW Fixing an unknown position from measured angles is called **triangulation**. Surveyors still use it, and once carried chains of triangles across whole countries. Astronomers use *parallax*: as the Earth moves round the Sun, a near star seems to shift against the far ones, and that small angle gives its distance. Your phone does the same job with distances instead of angles. A GPS receiver times the signal from each satellite, turns each time into a distance, and finds the one point lying at all of those distances at once: three satellites fix a point on the Earth's surface, and a fourth corrects the receiver's clock. This is **trilateration**, the sphere equation of §7.1 solved several times over.

EXT The ambiguous case is really a statement about congruence. Give a triangle by two angles and a side (ASA or AAS), by three sides (SSS), or by two sides and the angle between them (SAS), and only one triangle fits. Those are the congruence rules. Two sides and an angle that is *not* between them (SSA) is not one of them, so it can leave the triangle undecided. The sine rule's two answers are the two triangles that SSA allows.

End-of-Chapter Problems
Chapter 7: Shapes & Measures

1. A cuboid has dimensions $6 \times 8 \times 10$.

(a) Find the length of the space diagonal. [2]
(b) Find the angle between the space diagonal and the base. [3]

2. Consider the points $A(2, 0, 0)$, $B(0, 4, 0)$ and $C(0, 0, 4)$.

(a) Find the lengths AB, AC and BC. [3]
(b) Find angle BAC. [3]

3. EXACT A cone has radius 5 and height 12.

(a) Find the slant height. [1]
(b) Find the volume. [2]
(c) Find the total surface area. [2]

4. EXACT A solid consists of a cylinder of radius 3 and height 10 with a hemisphere of radius 3 on top. Find its total surface area. [5]

5. In triangle ABC, $a = 8$, $b = 11$, $c = 15$.

(a) Find the largest angle. [3]
(b) Find the area of the triangle. [2]

6. In triangle PQR, $p = 10$, $q = 7$ and $R = 95°$.

(a) Find r. [2]
(b) Find the area of the triangle. [2]

7. A triangle has two sides of length 9 and 12 with an included angle of $40°$.

(a) Find the length of the third side. [2]
(b) Find the area of the triangle. [2]

8. A boat sails 30 km from P on a bearing of $100°$ to Q, then 20 km on a bearing of $040°$ to R.

(a) Show that angle $PQR = 120°$. [3]
(b) Find the distance PR. [3]

(c) Find the bearing of R from P. [4]

9. A sector has radius 12 cm and angle 1.2 radians.

(a) Find the arc length. [1]
(b) Find the area of the sector. [2]
(c) Find the perimeter of the sector. [2]

10. A sector has area 30 cm^2 and radius 6 cm.

(a) Find the angle of the sector in radians. [2]
(b) Find the arc length. [2]

11. A chord subtends an angle of 1.5 radians at the centre of a circle of radius 10 cm.

(a) Find the area of the sector. [2]
(b) Find the area of the triangle formed by the chord and the two radii. [2]
(c) Find the area of the minor segment. [2]

12. In triangle ABC, $A = 50°$, $a = 6$ and $b = 7$.

(a) Find the two possible values of angle B. [3]
(b) Find the two possible areas of the triangle. [3]

13. EXACT A regular hexagon is inscribed in a circle of radius 8. Find its area. [4]

14. EXACT A cone has volume 100π and radius 5.

(a) Find the height. [2]
(b) Find the slant height. [2]
(c) Find the curved surface area. [1]

15. From a point on level ground, the angle of elevation of the top of a tower is $30°$. Moving 40 m closer, the angle of elevation becomes $45°$. Find the height of the tower. [5]

16. A square-based pyramid has base side 10 and height 12.

(a) Find the volume. [2]
(b) Find the slant height of a triangular face. [2]
(c) Find the total surface area. [2]

17. The area of a triangle is 24, with $a = 8$ and $b = 10$. Find the two possible values of the included angle C. [3]

18. Two ships leave a port at the same time. One sails on a bearing of 040° at 12 km/h, the other on a bearing of 100° at 16 km/h. Find the distance between them after 2 hours. [5]

19. Points $A(0,0,0)$ and $B(4,4,7)$ are given.

(a) Find AB. [2]
(b) Find the midpoint of $[AB]$. [1]
(c) Show that B lies on the sphere $x^2+y^2+z^2=81$. [1]

20. A pendulum of length 0.8 m swings through an angle of 0.5 radians. Find the length of the arc traced by the bob. [2]

21. A windscreen wiper arm OA of length 50 cm turns about O through an angle of 1.8 radians. The rubber blade covers the part of the arm between 20 cm and 50 cm from O.

(a) Find the length of the arc traced by the tip A. [2]
(b) Find the area of glass wiped by the blade. [4]
(c) Find the perimeter of the wiped region. [3]

22. Two observers stand 200 m apart on level ground, with a balloon in the vertical plane between them. From one the angle of elevation to the balloon is 40°; from the other it is 55°. Find the height of the balloon. [6]

23. In triangle ABC, $AB=12$, $BC=9$ and angle $BAC=35^\circ$. Find the two possible values of angle BCA. [4]

24. Radians and arc length.

(a) Convert $\dfrac{5\pi}{6}$ to degrees. [1]
(b) Convert $\dfrac{7\pi}{4}$ to degrees. [1]
(c) Find the arc length of a sector of radius 9 and angle $\dfrac{5\pi}{6}$. [2]

25. A sector of a circle of radius 12 and angle 1.5 radians is rolled into a cone, so that the radius becomes the slant height and the arc becomes the base circumference. Find the radius and height of the cone. [6]

26. A sector of a circle has perimeter 20 and area 16. Find the radius r and angle θ. [6]

27. A chord of a circle of radius 10 has length 12. Find the area of the minor segment it cuts off. [6]

28. The area of a triangle is 84, and two of its sides have lengths 13 and 14. Find both

possible lengths of the third side. [6]

29. A yacht leaves harbour H and sails 25 km on a bearing of $048°$ to a buoy A. It then sails 18 km on a bearing of $130°$ to a second buoy B.

(a) Show that angle $HAB = 98°$. [3]
(b) Find the distance HB. [4]
(c) Find the bearing of B from H. [4]
(d) Find the area of triangle HAB. [3]
(e) The yacht returns directly from B to H. Find the total distance it has sailed. [2]

30. A rectangular box $ABCDEFGH$ has a horizontal base $ABCD$ with $AB = 8$ cm and $BC = 6$ cm, and vertical edges of length 5 cm, so that G is the vertex above C.

(a) Find the length of the base diagonal AC. [2]
(b) Find the exact length of the space diagonal AG. [3]
(c) Find the angle that AG makes with the base $ABCD$. [3]
(d) Find the angle $G\hat{A}B$ between the diagonal AG and the edge AB. [4]
(e) Find the total surface area of the box. [3]

31. EXACT A container is a cone of base radius 6 cm and height 15 cm, held with its point downwards.

(a) Find the volume of the cone when full, in terms of π. [2]
(b) Water is poured in to a depth of 10 cm. Using similar triangles, find the radius of the circular water surface. [3]
(c) Find the volume of water, in terms of π. [3]
(d) Find the fraction of the cone's volume that is filled, and explain why it is not $\frac{2}{3}$ even though the water reaches $\frac{2}{3}$ of the height. [4]
(e) Find the depth of water when the cone is exactly half full, giving your answer to three significant figures. [3]

Challenge Questions

32. GDC **How many triangles fit the data?** In triangle ABC you are given the side b, the angle A, and the side a opposite A. This is the SSA case. This investigation finds exactly how many triangles fit, for every possible value of a, first from a diagram and then from a quadratic equation.

In parts (a) to (e), take A to be acute, and let h be the distance from C to the line AB.

(a) Show that $h = b\sin A$. [2]

(b) Take $b = 10$ and $A = 40°$. Find the number of triangles when $a = 5$, when $a = 8$, and when $a = 12$. [3]

(c) Explain why no triangle exists when $a < b\sin A$. *(Hint: what is the shortest distance from C to the line AB?)* [2]

(d) Explain why exactly one triangle exists when $a = b\sin A$, and state what is special about it. [2]

(e) Show that two triangles exist exactly when $b\sin A < a < b$, and that only one exists when $a > b$. Explain what happens at $a = b$. [4]

(f) Now let the side $c = AB$ be the unknown. Use the cosine rule to show that

$$c^2 - (2b\cos A)\,c + (b^2 - a^2) = 0,$$

and use this equation to find both values of c when $b = 10$, $A = 40°$ and $a = 8$. [4]

(g) Each positive root c of this quadratic gives one triangle. Use its discriminant, and the sum and product of its roots, to prove the counts of parts (c) to (e) again, including the case $a = b$. Then use the same method to find how many triangles exist when A is obtuse, for $a \le b$ and for $a > b$. [5]

(h) Take $b = 10$ and $A = 40°$ again. Find both values of c when $a = 6.44$ and when $a = 6.5$. Comment on what this means for a surveyor whose measurement of a carries a small error when a is close to $b\sin A$, and suggest how this could be investigated further. [3]

33. GDC **The best cone from a sector.** A sector of radius R and angle θ radians is cut from a sheet of paper. The two straight edges are joined to make a cone. This investigation finds the angle that gives the largest cone.

Write $t = \dfrac{\theta}{2\pi}$, so that $0 < t < 1$.

(a) The arc of the sector becomes the circumference of the base. Show that the base radius is $r = Rt$. [2]

(b) Show that the height of the cone is $h = R\sqrt{1 - t^2}$. [3]

(c) Hence show that the volume is $V = \frac{1}{3}\pi R^3 t^2\sqrt{1 - t^2}$. [2]

(d) Taking $R = 1$, use technology to find the value of t that makes V largest, and hence find θ to the nearest degree. [3]

(e) Let $u = t^2$. Show that

$$\frac{4}{27} - u^2(1 - u) = \left(u - \frac{2}{3}\right)^2\left(u + \frac{1}{3}\right),$$

and hence verify that V is largest when $t^2 = \frac{2}{3}$. Find the exact values of $\dfrac{r}{R}$ and $\dfrac{h}{r}$ at this maximum. *(Hint: V is largest where V^2 is largest, and V^2 is a multiple of $u^2(1 - u)$)* [5]

(f) Now the whole disc of radius R is cut into two sectors, a fraction t and a fraction $1 - t$ of the disc, and each is made into a cone. Write down the total volume of the two cones as a function of t, and use technology to find the value of t that makes it largest.

Compare the result with part (d), and suggest how this could be investigated further. [4]

34. How reliable was al-Bīrūnī's measurement? The opener described his method. From a hill of height h he measured the angle of dip θ to the horizon, and the radius of the Earth follows from one right-angled triangle. This investigation asks how much trust the answer deserves.

(a) The line of sight touches the Earth, so it meets the radius there at a right angle. Show that $\cos\theta = \dfrac{R}{R+h}$, and hence that $R = \dfrac{h\cos\theta}{1-\cos\theta}$. [3]

(b) Al-Bīrūnī recorded $h = 652$ cubits and $\theta = 34'$. Find R in cubits. *(Hint: $34' = \frac{34}{60}$ of a degree)* [3]

(c) Repeat the calculation for $\theta = 33'$ and for $\theta = 35'$. Find the percentage change in R in each case. [4]

(d) For small θ the approximation $\cos\theta \approx 1 - \frac{1}{2}\theta^2$ holds, with θ in radians. Use it to show that $R \approx \dfrac{2h}{\theta^2}$. [3]

(e) Hence explain why an error of about 3% in θ produces an error of about 6% in R. Comment on what this means for how confidently his result can be quoted. [4]

(f) Suppose his instrument could read any angle to within about δ radians, however large the angle. Use part (d) to show that the percentage error in R is then roughly proportional to $\dfrac{1}{\sqrt{h}}$, and state what this suggests about the choice of hill. Suggest one further effect that an extended investigation of the method should include. [3]

35. GDC A line's angles with three perpendicular faces. A line leaves the corner O of a cuboid. It makes some angle with each of the three faces that meet at O. This investigation finds the rule that links the three angles.

The cuboid has edges of lengths p, q and r along the x-, y- and z-axes, with O at the origin, and $G(p, q, r)$ is the opposite corner. Let α, β and γ be the angles that the diagonal OG makes with the faces $x = 0$, $y = 0$ and $z = 0$ respectively.

(a) Take $p = 2$, $q = 3$ and $r = 6$. Find the length OG, and find α, β and γ in degrees, correct to one decimal place. [4]

(b) Find $\sin^2\alpha + \sin^2\beta + \sin^2\gamma$ for the cuboid in part (a) and for the cuboid with $p = 1$, $q = 2$, $r = 2$. State a conjecture. [2]

(c) Prove your conjecture for every cuboid. *(Hint: the perpendicular from G to the face $x = 0$ has length p)* [3]

(d) Explain why any line through O that enters the region $x > 0$, $y > 0$, $z > 0$ is the diagonal of some cuboid with a corner at O. State what the rule becomes for a line that lies in the face $z = 0$, and what it then says about $\alpha + \beta$. [3]

(e) Show that no line can make an angle of 30° with each of the three faces. Find the exact value of $\sin\alpha$ when the three angles are equal, and name the cuboid whose diagonal does this. [3]

(f) Find $\alpha + \beta + \gamma$ for the three lines met in parts (a), (b) and (e), and for a line in the face

$z = 0$. Conjecture the least and greatest possible values of $\alpha + \beta + \gamma$, and suggest how this investigation could be extended. [3]

GEOMETRY & TRIGONOMETRY

Chapter 8

Trigonometry

8

SYLLABUS 3.5–3.11

Your voice is a wave. So is the light from your screen, the signal carrying this page to your device, the current in the wall socket, the tide in Tokyo Bay, and the beating of your own heart.

They do not look alike. A heartbeat is a sharp spike, a spoken vowel is jagged, a tide is a slow rise and fall over twelve hours. In 1822 Joseph Fourier claimed that a repeating signal, however jagged, can be built up from pure sine waves; Peter Dirichlet later proved it for signals that are not too irregular. Choose enough of them, with the right heights and the right timings, and the shape appears. Phone audio compression uses this idea many times a second.

So all of them can be built, at least approximately, from one curve. Where does it come from?

Watch one seat on a turning Ferris wheel. As the wheel turns steadily, the seat rises, slows near the top, falls, slows near the bottom, and rises again. Plot its height against time and the sine curve appears. The seat is a point moving round a circle, and its height is the vertical coordinate of that point. The circle and the wave are the same motion seen two ways: one going round, one going along.

That is why this chapter starts with a circle. In Ch. 7 sine and cosine were defined as ratios in a right-angled triangle, which only makes sense for acute angles, even though the sine and cosine rules already used obtuse ones. On the circle they become functions of *any* angle, and they repeat forever.

By the end of this chapter you will be able to read exact values of sine, cosine and tangent from the unit circle. You will use identities to rewrite one expression as another, sketch and transform the graphs of these functions, and work with the reciprocal and inverse functions. You will also solve trigonometric equations, where a single value of the sine or cosine is shared by two different angles, so answers usually have to be looked for in pairs.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

8.1 The Unit Circle

How do you find $\sin 120^\circ$? Or $\cos 360^\circ$, or $\tan(-30^\circ)$?

You already have a definition, and it is SOHCAHTOA. In a right triangle,

$$\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}}, \qquad \cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}}, \qquad \tan\theta = \frac{\text{opposite}}{\text{adjacent}}.$$

Now try to draw the triangle. For 120° there is none: the three angles add to 180°, one of them is already the right angle, so the other two must each be under 90°. For 360° there is none either, and -30° is worse, because a triangle has no negative angles. Under SOHCAHTOA these three expressions have no meaning at all.

Yet your calculator answers all three without complaint, and a turn of 120° is an ordinary thing to make. So the definition has to widen. Whatever replaces it must agree with SOHCAHTOA wherever SOHCAHTOA works, and keep working where it does not.

The unit circle does exactly that. Take the circle of radius 1 centred at the origin, and a point P on it at angle θ, measured anticlockwise from the positive x-axis.

For an acute θ, drop a perpendicular from P to the x-axis. This makes a right triangle whose hypotenuse is the radius, so the hypotenuse is 1. The old ratios then read

$$\sin\theta = \frac{\text{opposite}}{1} = \text{opposite}, \qquad \cos\theta = \frac{\text{adjacent}}{1} = \text{adjacent},$$

and those two sides are exactly the coordinates of P. So for acute angles the new definition agrees with the old one.

The new form never mentions a triangle. A point can sit anywhere on the circle, at any angle, and it always has coordinates. That is what extends the definition to every real θ.

DEF On the unit circle, for a point P at angle θ:

$$\cos\theta = x\text{-coordinate of } P, \qquad \sin\theta = y\text{-coordinate of } P.$$

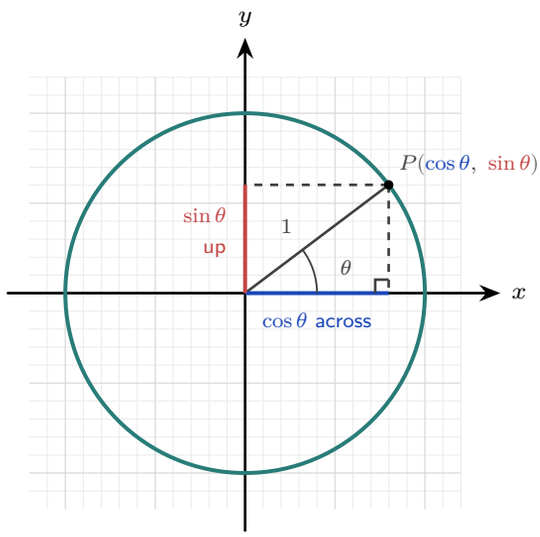


Figure 8.1 On a circle of radius 1 the hypotenuse is 1, so the two shorter sides *are* the coordinates of P. The cobalt side, measured across, is $\cos\theta$; the scarlet side, measured up, is $\sin\theta$.

The tangent is the gradient of the line OP, so $\tan\theta = \dfrac{\text{rise}}{\text{run}} = \dfrac{\sin\theta}{\cos\theta}$. It is undefined wherever $\cos\theta = 0$.

KEY · IB $\tan\theta = \dfrac{\sin\theta}{\cos\theta}$.

That word *gradient* is not accidental. The line OP passes through the origin, and its gradient is $\tan\theta$, so its equation is $y = x\tan\theta$. This ties the tangent to the straight lines of Ch. 4: the tangent converts an angle into a gradient, and $\arctan$ converts back. A line of gradient 1 makes an angle of $\frac{\pi}{4}$ with the positive x-axis, because $\tan\frac{\pi}{4} = 1$.

KEY A line through the origin making an angle θ with the positive x-axis has equation

$$y = x\tan\theta.$$

Read forwards it gives the gradient from the angle; read backwards, $\theta = \arctan(\text{gradient})$ gives the angle from the gradient.

8.1.1 Exact values

A few angles give exact coordinates, and they come from the 30–60–90 and 45–45–90 triangles.

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\sin\theta$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\cos\theta$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
$\tan\theta$	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	undefined

Only three numbers appear in the shaded row, and they are all that has to be memorised:

$$\frac{1}{2}, \qquad \frac{\sqrt{2}}{2}, \qquad \frac{\sqrt{3}}{2}.$$

The cosine row is those same three in reverse order. That is not a coincidence. Since $\cos\theta = \sin(\frac{\pi}{2} - \theta)$, reading the sine row backwards reads the cosine row forwards. The tangent row needs nothing new at all: $\tan\theta = \dfrac{\sin\theta}{\cos\theta}$, so divide one row by the other. At $\theta = \frac{\pi}{2}$ the cosine is zero, which is why the tangent has no value there.

The first-quadrant values repeat into the other three quadrants with signs changed. Collecting them all onto one circle gives the diagram to learn.

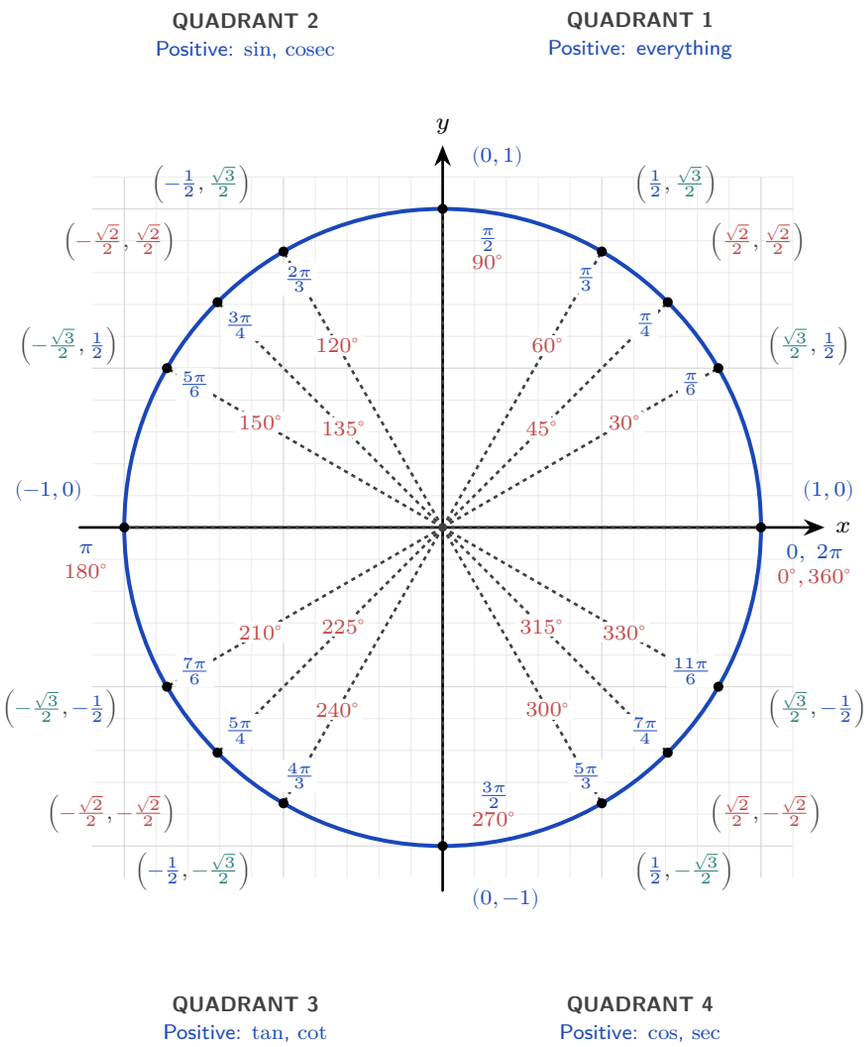


Figure 8.2 The unit circle with every special angle. Each spoke carries the angle twice: in cobalt on the outer ring in radians, and in scarlet on the inner ring in degrees. Outside the circle is the point $(\cos\theta,\ \sin\theta)$, with each coordinate coloured by its size: cobalt for $\frac{1}{2}$, scarlet for $\frac{\sqrt{2}}{2}$, teal for $\frac{\sqrt{3}}{2}$. To read a value, find the angle and take the matching coordinate. The corner labels say which functions are positive in each quadrant; $\sec$, cosec and $\cot$ are the reciprocal functions of §8.4.

The colours make the point. Only three numbers appear anywhere on the circle, and every coordinate on it is one of those three with a sign attached. Add the four points on the axes and the diagram is complete.

8.1.2 Symmetry and the sign of each function

Since $\cos\theta$ and $\sin\theta$ are the coordinates of P, their signs are just the signs of x and y in each quadrant. Nothing needs to be memorised separately: in the second quadrant x is negative and y is positive, so $\cos\theta < 0$ and $\sin\theta > 0$, and $\tan\theta = \frac{\sin}{\cos}$ is negative.

The unit-circle figure above records the answer at each corner. Reading the four quadrants anticlockwise from the first, the functions that stay positive are all three, then sine alone, then tangent alone, then cosine alone.

The mnemonic **All Students Take Calculus** is the usual way to remember that order. Use it as a check, not as the reason. The reason is always the sign of the coordinates, and if you can picture the point you never need the letters at all.

Each symmetry of the circle now gives an identity, and each one is a statement about where a reflected point lands.

Reflect in the y-axis. The point at angle $\pi - \theta$ is the mirror image, across the y-axis, of the point at angle θ. Reflecting in the y-axis negates the x-coordinate and leaves y alone, so the two points are $(-\cos\theta,\ \sin\theta)$ and $(\cos\theta,\ \sin\theta)$. Comparing coordinates:

$$\sin(\pi - \theta) = \sin\theta, \qquad \cos(\pi - \theta) = -\cos\theta.$$

Reflect in the x-axis. Measuring $-\theta$ means turning clockwise instead of anticlockwise, which lands on the mirror image of P across the x-axis. That reflection negates y and leaves x alone, so

$$\cos(-\theta) = \cos\theta, \qquad \sin(-\theta) = -\sin\theta.$$

Turn half a circle. Adding π sends P to the diametrically opposite point, which negates both coordinates:

$$\sin(\theta + \pi) = -\sin\theta, \qquad \cos(\theta + \pi) = -\cos\theta.$$

Turn a full circle. Adding 2π to the angle sends P once round and back to exactly the same point, so both coordinates are unchanged:

$$\sin(\theta + 2\pi) = \sin\theta, \qquad \cos(\theta + 2\pi) = \cos\theta.$$

This is why both functions are **periodic** with period 2π. Every identity of this kind says the same thing: move the point, see which coordinate changes sign.

You have met the first of these relations before. Two angles, θ and $\pi - \theta$, share the same sine. That shared value is what produced the ambiguous case of the sine rule in Ch. 7. A triangle built from a sine value can sometimes close in two ways, one acute and one obtuse.

An exact value at any special angle now comes out of the circle in two moves: find the acute angle the arm makes with the x-axis, which fixes the size, then look at the quadrant, which fixes the sign.

Worked through. To find $\sin\frac{5\pi}{3}$ and $\cos\frac{5\pi}{3}$ exactly, write the angle as $2\pi - \frac{\pi}{3}$. A full turn returns to the positive x-axis, so turning back through $\frac{\pi}{3}$ leaves the point in the fourth quadrant, where the x-coordinate is positive and the y-coordinate negative. The arm makes an acute angle of $\frac{\pi}{3}$ with the axis, so both values have the size they have at $\frac{\pi}{3}$:

$$\sin\frac{5\pi}{3} = -\sin\frac{\pi}{3} = -\frac{\sqrt{3}}{2}, \qquad \cos\frac{5\pi}{3} = \cos\frac{\pi}{3} = \frac{1}{2}.$$

Dividing them gives $\tan\frac{5\pi}{3} = -\sqrt{3}$, negative, because only the sine carries a minus sign.

That is the fourth quadrant's entry in the mnemonic, recovered from the coordinates rather than remembered.

YOUR TURN 8.1 The Unit Circle

Exact values

1. Write down the exact value of each.

(a) $\sin\dfrac{\pi}{6}$ (b) $\cos\dfrac{\pi}{4}$

(c) $\tan\dfrac{\pi}{3}$ (d) $\tan\dfrac{\pi}{4}$

2. Use the symmetry of the circle to find each exact value.

(a) $\sin\dfrac{5\pi}{6}$ (b) $\cos\dfrac{2\pi}{3}$

(c) $\sin\dfrac{4\pi}{3}$ (d) $\cos\dfrac{7\pi}{4}$

Signs in each quadrant

3. State the quadrant or quadrants in which each condition holds.

(a) $\sin\theta > 0$ (b) $\cos\theta < 0$

(c) $\tan\theta > 0$ (d) $\sin\theta < 0$ and $\cos\theta > 0$

The unit circle and gradients

4. The point P on the unit circle has angle θ.

(a) State the coordinates of P in terms of θ.

(b) Explain why $\cos^2\theta + \sin^2\theta = 1$ for every θ.

(c) Explain why $\tan\theta$ is undefined at $\theta = \dfrac{\pi}{2}$.

5. Lines through the origin, using $y = x\tan\theta$. Give each angle in radians.

(a) Write the equation of the line through the origin making an angle of $\dfrac{\pi}{3}$ with the positive x-axis.

(b) Find the angle a line of gradient 1 makes with the positive x-axis.

(c) Find the anticlockwise angle θ from the positive x-axis to the line $y = -x$, with $0 \le \theta < \pi$.

Concept check

6. A student writes $\cos\dfrac{2\pi}{3} = -\cos\dfrac{\pi}{6} = -\dfrac{\sqrt{3}}{2}$, taking the reference angle from the nearest axis. Identify the error and find the correct value.

8.2 Trigonometric Identities

An **identity** is an equation true for every angle. The first one is Pythagoras' theorem read off the unit circle: the point $(\cos\theta, \sin\theta)$ is distance 1 from the origin, so $\cos^2\theta + \sin^2\theta = 1$.

KEY · IB $\cos^2\theta + \sin^2\theta = 1.$

8.2.1 Double and compound angles

The functions of a sum of two angles can be expanded, and the double-angle formulas follow from the sum formulas. These **compound-angle** identities are usually just quoted, but the proof is short and uses only the unit circle and the cosine rule from Ch. 7.

Put two points on the unit circle, P at angle A and Q at angle B. Their coordinates are $P(\cos A,\ \sin A)$ and $Q(\cos B,\ \sin B)$, and the angle $\angle POQ$ at the centre is $A - B$.

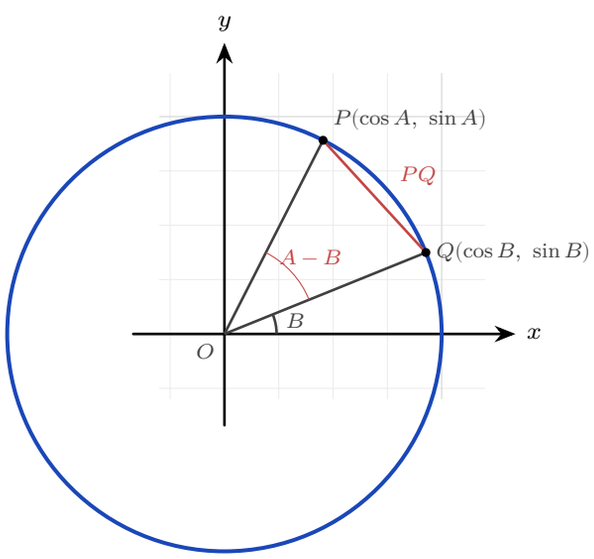


Figure 8.3 Two points on the unit circle, at angles A and B. The angle between the radii is $A - B$, and the chord PQ can be measured either by the distance formula or by the cosine rule.

Now find the distance PQ in two different ways.

By the distance formula:

$$PQ^2 = (\cos A - \cos B)^2 + (\sin A - \sin B)^2.$$

Expanding gives $\cos^2 A - 2\cos A\cos B + \cos^2 B + \sin^2 A - 2\sin A\sin B + \sin^2 B$. Both Pythagorean pairs collapse to 1, so

$$PQ^2 = 2 - 2(\cos A\cos B + \sin A\sin B).$$

By the cosine rule in triangle OPQ, where both sides from the centre have length 1 and the angle between them is $A - B$:

$$PQ^2 = 1^2 + 1^2 - 2(1)(1)\cos(A - B) = 2 - 2\cos(A - B).$$

The two expressions are the same length, so

$$\cos(A - B) = \cos A \cos B + \sin A \sin B.$$

Every other compound-angle formula follows from this one. Replacing B by $-B$, and using $\cos(-B) = \cos B$ with $\sin(-B) = -\sin B$, gives the formula for $\cos(A + B)$.

A sine is a cosine a quarter turn away: $\sin\theta = \cos(\frac{\pi}{2} - \theta)$. That single fact is all the sine version needs. Write the sum in that form, then group the result as a *difference* of two angles:

$$\sin(A + B) = \cos(\tfrac{\pi}{2} - A - B) = \cos((\tfrac{\pi}{2} - A) - B)\,.$$

That is exactly the shape just proved. Expanding it, and using $\cos(\frac{\pi}{2} - A) = \sin A$ together with $\sin(\frac{\pi}{2} - A) = \cos A$, gives

$$\sin(A + B) = \sin A \cos B + \cos A \sin B.$$

The tangent versions come from dividing the sine formula by the cosine formula.

KEY · IB HL **Compound angles:**

$$\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B, \qquad \cos(A \pm B) = \cos A \cos B \mp \sin A \sin B,$$

$$\tan(A \pm B) = \frac{\tan A \pm \tan B}{1 \mp \tan A \tan B}.$$

Setting $B = A$ reduces each of these to a **double-angle** identity. Working through the derivation once shows there is nothing new to learn here.

Put $B = A$ in the sine formula:

$$\sin 2A = \sin(A + A) = \sin A \cos A + \cos A \sin A = 2\sin A \cos A.$$

Do the same in the cosine formula:

$$\cos 2A = \cos(A + A) = \cos A \cos A - \sin A \sin A = \cos^2 A - \sin^2 A.$$

That is the first form. The other two come from $\sin^2 A + \cos^2 A = 1$, used to remove one function or the other. Replacing $\sin^2 A$ by $1 - \cos^2 A$ gives

$$\cos 2A = \cos^2 A - (1 - \cos^2 A) = 2\cos^2 A - 1,$$

and replacing $\cos^2 A$ by $1 - \sin^2 A$ instead gives

$$\cos 2A = (1 - \sin^2 A) - \sin^2 A = 1 - 2\sin^2 A.$$

The tangent version follows in the same way, by putting $B = A$ in the tangent formula.

The sine and cosine versions sit in the SL half of the booklet; the tangent version is HL.

KEY · IB **Double angles:**

$$\sin 2\theta = 2\sin\theta\cos\theta, \qquad \cos 2\theta = \cos^2\theta - \sin^2\theta = 2\cos^2\theta - 1 = 1 - 2\sin^2\theta.$$

KEY · IB HL $\tan 2\theta = \dfrac{2\tan\theta}{1 - \tan^2\theta}.$

The three forms of $\cos 2\theta$ are one identity rewritten by $\sin^2 + \cos^2 = 1$. Choose whichever leaves you with only sines or only cosines. That choice is what makes the identity so useful when solving equations.

The compound-angle formulas also give exact values beyond the basic table. Write the angle as a sum or difference of two special angles, such as $165^\circ = 120^\circ + 45^\circ$, and expand. The two angles need not be acute; carry the sign of each ratio from its quadrant.

Worked through. To find $\sin 165^\circ$ exactly, write $165^\circ = 120^\circ + 45^\circ$ and use the compound-angle formula:

$$\sin 165^\circ = \sin 120^\circ \cos 45^\circ + \cos 120^\circ \sin 45^\circ = \frac{\sqrt{3}}{2}\cdot\frac{\sqrt{2}}{2} + \left(-\frac{1}{2}\right)\cdot\frac{\sqrt{2}}{2} = \frac{\sqrt{6} - \sqrt{2}}{4}.$$

Here $\cos 120^\circ = -\frac{1}{2}$ carries a minus sign from the second quadrant, and that minus sign is what turns the sum into a difference.

Identities also let you move between ratios *without ever finding the angle*. The angle stays unknown from start to finish, and the Pythagorean identity supplies everything needed. The identity gives the square of the unknown ratio, so it leaves a choice of sign. Let the quadrant of the angle fix that sign before you use the value.

EXAMPLE 8.1

Given that $\cos x = \dfrac{3}{4}$ and x is acute, find the exact value of $\sin 2x$ without finding x.

From $\cos^2 x + \sin^2 x = 1$, $\sin^2 x = 1 - \frac{9}{16} = \frac{7}{16}$, and since x is acute, $\sin x = \frac{\sqrt{7}}{4}$ (positive).

Then

$$\sin 2x = 2\sin x\cos x = 2\cdot\frac{\sqrt{7}}{4}\cdot\frac{3}{4} = \frac{3\sqrt{7}}{8}.$$

The word "acute" fixed the sign of $\sin x$. If x had instead been in the fourth quadrant, the same

working would give $\sin 2x = -\frac{3\sqrt{7}}{8}$: always let the quadrant choose the sign before doubling.

Proving an identity uses the same substitutions: start from one side, substitute known identities, and simplify until the other side appears.

EXAMPLE 8.2

Prove that $\cos(A-B) - \cos(A+B) = 2\sin A \sin B$.
Work on the left side, expanding each cosine by the compound-angle formula:

$$\cos(A-B) = \cos A\cos B + \sin A\sin B, \qquad \cos(A+B) = \cos A\cos B - \sin A\sin B.$$

The two expansions differ only in the sign of the second term, so subtracting removes the $\cos A\cos B$ pair and doubles what is left:

$$\cos(A-B) - \cos(A+B) = 2\sin A\sin B. \qquad \blacksquare$$

Expanding the more complicated side first is the usual choice. Once both expansions are written down, the terms that have to cancel are visible before any algebra is done.

8.2.2 One wave from two

This subsection is beyond the syllabus. It is included because the method is short and it appears in physics.

A sum such as $8\sin x + 15\cos x$ mixes a sine and a cosine of the same angle. Both have period 2π, so the sum repeats every 2π as well, and it turns out to be a single sine wave, stretched and shifted. Reading the compound-angle formula backwards shows why.

Expanding $R\sin(x+\alpha)$ gives

$$R\sin(x+\alpha) = R\cos\alpha\,\sin x + R\sin\alpha\,\cos x,$$

which already has the shape $a\sin x + b\cos x$, with $a = R\cos\alpha$ and $b = R\sin\alpha$. Squaring those two and adding removes α, since $\cos^2\alpha + \sin^2\alpha = 1$, and leaves $a^2 + b^2 = R^2$. Dividing them removes R instead and leaves $\tan\alpha = \frac{b}{a}$. So any such sum can be written as one wave, and the two constants can be recovered from a and b alone.

KEY **The wave form.** For constants a and b not both zero,

$$a\sin x + b\cos x = R\sin(x+\alpha), \qquad R = \sqrt{a^2+b^2}, \qquad \tan\alpha = \frac{b}{a},$$

with $R > 0$ and α taken in the quadrant where $\cos\alpha$ has the sign of a and $\sin\alpha$ the sign of b. The same R and α give the cosine form

$$a\cos x + b\sin x = R\cos(x-\alpha).$$

Since $\sin$ and $\cos$ run between -1 and 1, the expression has maximum R and minimum $-R$.

This form is not printed in the formula booklet, so rebuild it from the expansion of $R\sin(x+\alpha)$ whenever it is needed. Its value is that the variable now appears once instead of twice, which turns a maximum, a minimum or an equation into the ordinary work of §8.6.

EXAMPLE 8.3

Write $8\sin x + 15\cos x$ in the form $R\sin(x+\alpha)$ with $R>0$ and $0<\alpha<\frac{\pi}{2}$. State the maximum and minimum values, and solve $8\sin x + 15\cos x = 6$ for $0 \le x \le 2\pi$.

Here $a = 8$ and $b = 15$. Both are positive, so α is in the first quadrant:

$$R = \sqrt{8^2+15^2} = \sqrt{289} = 17, \qquad \tan\alpha = \frac{15}{8} \implies \alpha = \arctan\frac{15}{8} = 1.0808,$$

giving $8\sin x + 15\cos x = 17\sin(x+1.0808)$.

The sine factor runs between -1 and 1, so the maximum is 17 and the minimum is -17.

For the equation, $17\sin(x+\alpha) = 6$, so $\sin(x+\alpha) = \frac{6}{17}$. Put $u = x+\alpha$. As x runs over $[0,\ 2\pi]$, u runs over $[1.0808,\ 7.3640]$. The reference angle is $\arcsin\frac{6}{17} = 0.3607$, and $\frac{6}{17}$ is positive, so sine takes it in the first and second quadrants:

$$u = 0.3607 \qquad \text{and} \qquad u = \pi - 0.3607 = 2.7809.$$

The first sits below 1.0808, so add a turn: $0.3607 + 2\pi = 6.6439$, which is inside. The second is already inside, and a further turn from it would pass 7.3640. So $u = 2.7809$ or $u = 6.6439$, and subtracting α,

$$x = 1.70 \qquad \text{or} \qquad x = 5.56 \quad (3\text{ s.f.}).$$

Transforming the interval along with the substitution is what saved the second answer. Had u been left to run over $[0,\ 2\pi]$, the pair 0.3607 and 2.7809 would have given one negative x and lost 5.56 altogether.

YOUR TURN 8.2 Trigonometric Identities

Recognising the identities

7. Simplify each expression.

(a) $\sin^2\theta + \cos^2\theta$

(b) $2\sin\theta\cos\theta$

(c) $\cos^2\theta - \sin^2\theta$

(d) $1 - 2\sin^2\theta$

Values without the angle

8. Given $\sin\theta = \dfrac{3}{5}$ with θ acute, find each exact value.

(a) $\cos\theta$

(b) $\tan\theta$

(c) $\sin 2\theta$

(d) $\cos 2\theta$

9. Given $\cos\theta = -\dfrac{3}{5}$ with θ in the second quadrant, find each exact value.

(a) $\sin\theta$

(b) $\tan\theta$

(c) $\sin 2\theta$

(d) $\cos 2\theta$

Compound-angle values

10. Use a compound-angle formula to find each exact value.

(a) $\cos 15^\circ$, from $45^\circ - 30^\circ$

(b) $\cos 105^\circ$, from $60^\circ + 45^\circ$

(c) $\cos 165^\circ$, from $120^\circ + 45^\circ$

Tangent formulas

11. Use the tangent compound-angle and double-angle formulas.

(a) Find the exact value of $\tan 75^\circ$, from $45^\circ + 30^\circ$.

(b) Given $\tan A = \dfrac{1}{2}$ and $\tan B = \dfrac{1}{3}$, with A and B acute, find $\tan(A+B)$, and hence $A + B$.

(c) Given $\tan\theta = \dfrac{3}{4}$, find the exact value of $\tan 2\theta$.

(d) Given $\tan\theta = -2$, find the exact value of $\tan 2\theta$.

Proving identities

12. Prove each identity.

(a) $\dfrac{\sin 2\theta}{1+\cos 2\theta} = \tan\theta$

(b) $\dfrac{1-\cos 2\theta}{\sin 2\theta} = \tan\theta$

(c) $(\sin\theta + \cos\theta)^2 = 1 + \sin 2\theta$

(d) $\dfrac{\cos 2\theta}{\cos\theta + \sin\theta} = \cos\theta - \sin\theta$

(e) $\tan\theta \sin 2\theta = 2\sin^2\theta$

(f) $\dfrac{\sin 3\theta}{\sin\theta} - \dfrac{\cos 3\theta}{\cos\theta} = 2$

Concept check

13. A student writes $\sin(A+B) = \sin A + \sin B$.

(a) Give a counter-example.

(b) State the correct expansion of $\sin(A+B)$.

8.3 Graphs of Trigonometric Functions

Each graph plots one coordinate of P against the angle. Send the point P round the circle and plot its y-coordinate against the angle: that plot *is* $y = \sin x$. Plot its x-coordinate instead and you get $y = \cos x$.

Every feature of the sine graph can be read off the circle before you plot a single point.

- The y-coordinate of P never leaves -1 to 1, so the curve stays between those two heights. Its **amplitude**, the distance from the middle to the top, is 1.
- P returns to its starting place after one full turn, so the curve repeats every 2π. Its **period**, the width of one complete cycle, is 2π.
- $y = 0$ when P sits on the x-axis, at $\theta = 0$, π and 2π, so the curve crosses the axis at those three angles.
- y is largest at the top of the circle, at $\theta = \frac{\pi}{2}$, so the maximum is 1 and it occurs at $x = \frac{\pi}{2}$.
- $y > 0$ everywhere in the upper half of the circle, so the curve is above the axis for $0 < x < \pi$.

Read the x-coordinate instead and the cosine graph appears the same way. It starts at 1, because P starts on the positive x-axis, and falls to 0 at $\frac{\pi}{2}$, where P is at the top and its x-coordinate is zero.

So both curves have amplitude 1 and period 2π. They differ only in where a cycle begins.

The x-coordinate at angle θ is the y-coordinate a quarter turn later, so $\cos\theta = \sin(\theta + \frac{\pi}{2})$, and the cosine curve is the sine curve shifted left by $\frac{\pi}{2}$.

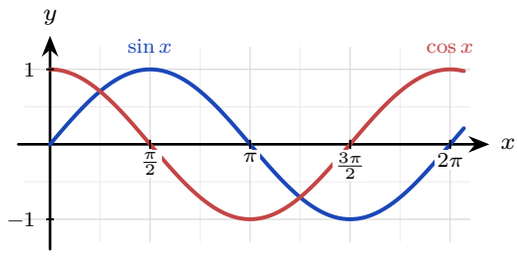


Figure 8.4 One turn of the circle unrolled. Both waves have amplitude 1 and period 2π; the cosine is the sine shifted left by $\frac{\pi}{2}$.

The tangent behaves differently, and the circle explains why. Since $\tan\theta = \frac{y}{x}$, it is undefined wherever $x = 0$, which is at the top and the bottom of the circle. That gives vertical asymptotes at $x = \frac{\pi}{2} + n\pi$.

Its period is π, not 2π. Half a turn sends P to the point diametrically opposite, whose coordinates are $(-x, -y)$. The ratio $\frac{-y}{-x}$ equals $\frac{y}{x}$, so the tangent repeats after only half a revolution.

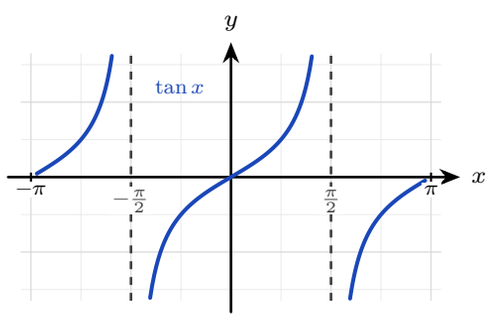


Figure 8.5 The tangent repeats every π, not every 2π, with a vertical asymptote wherever $\cos x = 0$.

That is the basic behaviour, collected in one place.

	$\sin x$	$\cos x$	$\tan x$
Domain	all x	all x	$x \neq \frac{\pi}{2} + n\pi$
Range	$-1 \le y \le 1$	$-1 \le y \le 1$	all y
Period	2π	2π	π
Amplitude	1	1	none
Zeros	$n\pi$	$\frac{\pi}{2} + n\pi$	$n\pi$

Here n stands for any integer, and the last row gives the values of x at which the curve crosses the axis. The tangent has no amplitude because it has no largest or smallest value.

Amplitude measures half the rise and fall, and the tangent rises without limit.

8.3.1 Transforming the wave

The general sinusoid stretches and slides the basic sine, and each parameter has a plain meaning.

KEY For $y = a\sin(b(x+c)) + d$:

$$\text{amplitude} = |a|, \qquad \text{period} = \frac{2\pi}{|b|},$$

$$\text{horizontal shift} = -c, \qquad \text{vertical shift (midline)} = d.$$

A question can also run the other way, giving you the graph's largest and smallest values and asking for a and d. Since the maximum is $d + |a|$ and the minimum is $d - |a|$, adding and subtracting recovers both constants.

- **Amplitude.** From the equation it is $|a|$. From a graph it is $\dfrac{\text{max} - \text{min}}{2}$, half the total rise and fall.
- **Midline.** From the equation it is the line $y = d$. From a graph it is $y = \dfrac{\text{max} + \text{min}}{2}$, the level halfway between the top and the bottom.
- **Maximum and minimum.** They are $d + |a|$ and $d - |a|$: the midline, plus and minus the amplitude.
- **Period.** From the equation it is $\dfrac{2\pi}{|b|}$. From a graph it is the horizontal distance from one repeat to the next.

Read the first two as sentences and they stop being formulas to memorise. Consider the wave $y = 5\sin(2x) - 3$, whose maximum is 2 and whose minimum is -8. The amplitude is $\frac{2-(-8)}{2} = 5$ and the midline is $\frac{2+(-8)}{2} = -3$, which is what the equation says.

Worked through. In $y = 2\sin(3x) + 1$ the amplitude is 2, so the wave rises and falls 2 either side of its midline. The period is $\frac{2\pi}{3}$, so it completes a full cycle three times as fast as $\sin x$. The midline is $y = 1$. The graph therefore runs from $1 - 2 = -1$ up to $1 + 2 = 3$, giving range $-1 \le y \le 3$.

8.3.2 Symmetry of the graphs HL

Each reflection symmetry of the circle from Section 8.1 becomes a visible symmetry of the graphs. The cosine curve is mirror-symmetric in the y-axis, so $\cos(-x) = \cos x$: cosine is an **even** function. The sine and tangent curves map onto themselves under a half-turn about the origin, so $\sin(-x) = -\sin x$ and $\tan(-x) = -\tan x$: both are **odd**, in the sense of Ch. 5. The same circle symmetries give the supplementary-angle relations.

KEY HL $\sin(\pi-\theta)=\sin\theta,\qquad \cos(\pi-\theta)=-\cos\theta,\qquad \tan(\pi-\theta)=-\tan\theta.$

You can see all three on the graphs directly: the sine curve is symmetric about the vertical line $x=\frac{\pi}{2}$, while cosine and tangent flip sign across it. In an exam these relations remove the need for a calculator: $\cos\frac{3\pi}{4}=-\cos\frac{\pi}{4}=-\frac{\sqrt{2}}{2}$.

YOUR TURN 8.3 Graphs of Trigonometric Functions

Amplitude, period and midline

14. State the amplitude, the period and the midline of each graph.

(a) $y=3\sin x$

(b) $y=\sin 2x$

(c) $y=4\sin(3x)-2$

(d) $y=5\sin(4x)+3$

15. State the range of each function.

(a) $y=2\sin x+1$

(b) $y=3\sin(2x)-1$

(c) $y=2\cos\left(\frac{x}{2}\right)-1$

Horizontal shift

16. State the amplitude, the period, the horizontal shift and the midline of each graph. *(Hint: write the argument in the form $b(x+c)$ first)*

(a) $y=3\sin\left(x-\frac{\pi}{4}\right)+1$

(b) $y=2\sin\left(2x+\frac{\pi}{3}\right)$

(c) $y=5\sin(3x-\pi)-2$

From the maximum and minimum

17. Find the constants from the information given.

(a) The graph of $y=a\sin(bx)+d$, with $a>0$ and $b>0$, has maximum value 7, minimum value -1 and period π. Find a, b and d.

(b) The depth of water in a harbour is modelled by $h(t)=a\sin(bt)+d$ metres, t hours after midnight, with $a>0$ and $b>0$. The greatest depth is 12.5 m, the least depth is 2.5 m, and the pattern repeats every 12 hours. Find a, b and d.

The tangent graph

18. Answer each question about $y = \tan x$.

(a) State its period.

(b) State the equations of two of its asymptotes.

(c) Explain why the asymptotes occur where they do.

Symmetry of the graphs

19. An angle θ has $\sin\theta = 0.6$ and $\cos\theta = 0.8$. Use the symmetries of the graphs to write down each value.

(a) $\sin(\pi - \theta)$

(b) $\cos(\pi - \theta)$

(c) $\tan(\pi - \theta)$

(d) $\sin(-\theta)$

Modelling

20. A wheel of diameter 40 m has its lowest point 2 m above the ground and turns once every 20 seconds. A seat starts at the bottom. Its height in metres after t seconds is modelled by $h(t) = a\sin(b(t + c)) + d$, with $a > 0$ and $b > 0$.

(a) State the amplitude a and the midline value d.

(b) Find the value of b.

(c) The seat is at its lowest point when $t = 0$. Show that $c = -5$ is a suitable value, and write down $h(t)$.

(d) Find the height after 8 seconds.

Concept check

21. A student says: "*the graph of* $y = \sin\left(2x - \dfrac{\pi}{3}\right)$ *is the graph of* $y = \sin 2x$ *translated* $\dfrac{\pi}{3}$ *to the right.*" Identify the error and state the correct translation.

22. A sinusoidal graph has maximum value 7 and minimum value -1. A student says: "*the amplitude is* 8." Identify the error, and find the correct amplitude and the equation of the midline.

8.4 Reciprocal Trigonometric Functions HL

Divide 1 by a sine, a cosine or a tangent and three more functions appear. Each has a name of its own, a graph with vertical asymptotes, and an identity that makes certain equations solvable.

KEY · IB $\sec\theta = \dfrac{1}{\cos\theta}, \qquad \operatorname{cosec}\theta = \dfrac{1}{\sin\theta}.$

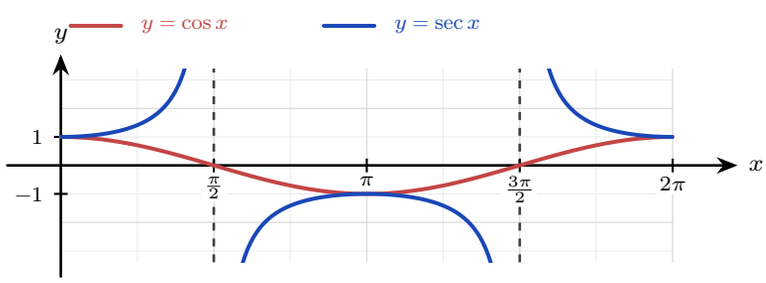
The third is the **cotangent**, $\cot\theta = \dfrac{1}{\tan\theta} = \dfrac{\cos\theta}{\sin\theta}$. Its definition is not printed in the formula booklet, so learn it.

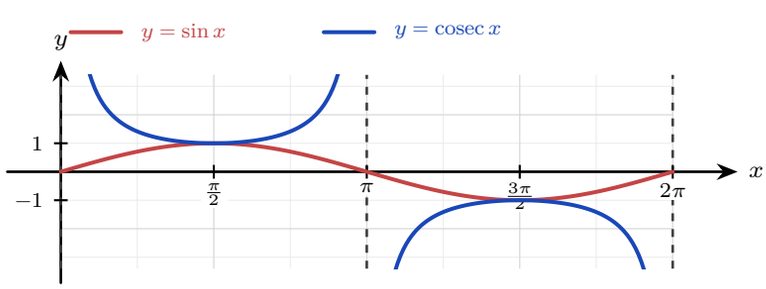
The pairing looks wrong at first: *sec*ant goes with *cos*ine, and *cosec*ant with *sin*e. The *co* means complement: $\operatorname{cosec}\theta = \sec(\frac{\pi}{2} - \theta)$, just as $\cos\theta = \sin(\frac{\pi}{2} - \theta)$. As a quick check, use the third letter. In **sec** the third letter is c, for cosine. In **cosec** it is s, for sine. In **cot** it is t, for tangent.

Two facts follow from the reciprocal alone, and between them they fix every graph.

First, a reciprocal has an asymptote wherever the original function is zero, because you cannot divide by zero. So $\sec$ has asymptotes where $\cos\theta = 0$, at $\theta = \frac{\pi}{2} + n\pi$, while cosec and $\cot$ have them where $\sin\theta = 0$, at $\theta = n\pi$.

Second, $\sin$ and $\cos$ never exceed 1 in size, so their reciprocals are never smaller than 1 in size. Neither $\sec\theta$ nor $\operatorname{cosec}\theta$ ever takes a value strictly between -1 and 1. State that gap whenever you give the range.





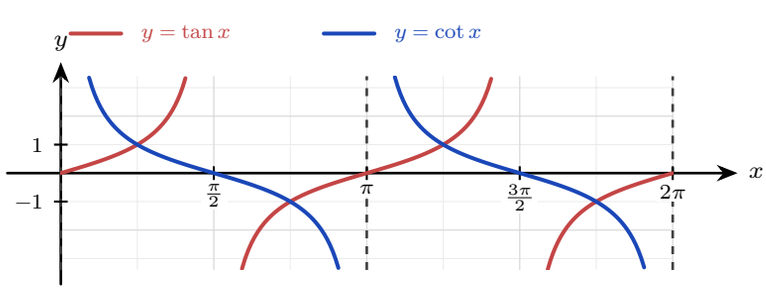


Figure 8.6 In each panel the parent function is scarlet and its reciprocal cobalt. The reciprocal has an asymptote, marked by a dashed line, wherever its parent crosses the axis, and the two curves meet wherever the parent reaches ± 1. Secant and cosecant never enter the band between -1 and 1; cotangent has no such gap.

Taking n to be any integer, the three read as follows.

Function	Domain	Range	Period
$\sec\theta = \dfrac{1}{\cos\theta}$	all θ except $\frac{\pi}{2} + n\pi$	$y \le -1$ or $y \ge 1$	2π
$\operatorname{cosec}\theta = \dfrac{1}{\sin\theta}$	all θ except $n\pi$	$y \le -1$ or $y \ge 1$	2π
$\cot\theta = \dfrac{\cos\theta}{\sin\theta}$	all θ except $n\pi$	all real y	π

Every excluded value in the domain column is a zero of the function underneath, and every one of them is an asymptote on the graph.

8.4.1 The two Pythagorean variants

The identity $\cos^2\theta + \sin^2\theta = 1$ has two companions, one for each reciprocal pair, and both come from dividing it through.

Dividing $\cos^2\theta + \sin^2\theta = 1$ by $\cos^2\theta$ gives

$$1 + \tan^2\theta = \sec^2\theta,$$

since $\frac{\sin\theta}{\cos\theta} = \tan\theta$ and $\frac{1}{\cos\theta} = \sec\theta$. Dividing the same identity by $\sin^2\theta$ instead gives the companion result. Both are in the formula booklet; the derivation shows where they come from.

KEY · IB $1 + \tan^2\theta = \sec^2\theta, \qquad 1 + \cot^2\theta = \operatorname{cosec}^2\theta.$

These identities exist to turn mixed equations into single-function ones, exactly as the double angles did.

Worked through. The equation $\tan^2 x + \sec x = 1$, for $0 \le x \le 2\pi$, mixes $\tan$ and $\sec$; the identity $\tan^2 x = \sec^2 x - 1$ rewrites it in $\sec x$ alone:

$$\sec^2 x - 1 + \sec x = 1 \implies \sec^2 x + \sec x - 2 = 0 \implies (\sec x + 2)(\sec x - 1) = 0.$$

So $\sec x = -2$ or $\sec x = 1$, which is $\cos x = -\frac{1}{2}$ or $\cos x = 1$. From the circle,

$$x = \frac{2\pi}{3},\ \frac{4\pi}{3} \quad \text{or} \quad x = 0,\ 2\pi.$$

Converting back to $\cos$ at the end is a sensible final step, because the unit circle is read in

sine and cosine, not in secant.

YOUR TURN 8.4 Reciprocal Trigonometric Functions HL

Exact values

23. Evaluate each exactly.

(a) $\sec \frac{\pi}{3}$

(b) $\operatorname{cosec} \frac{\pi}{2}$

(c) $\cot \frac{\pi}{4}$

(d) $\sec \frac{2\pi}{3}$

(e) $\cot \frac{2\pi}{3}$

(f) $\operatorname{cosec} \frac{7\pi}{6}$

Graphs of the reciprocal functions

24. Use the graphs of the reciprocal functions.

(a) State the range of $y = \sec x$.

(b) State the equations of the asymptotes of $y = \operatorname{cosec} x$ for $0 \leq x \leq 2\pi$.

(c) State the domain and the range of $y = \cot x$.

(d) Find the equations of the asymptotes of $y = \sec 2x$ for $0 \leq x \leq \pi$.

Simplifying

25. Simplify each expression.

(a) $\sec^2 \theta - \tan^2 \theta$

(b) $\cot \theta \sin \theta$

(c) $\dfrac{1}{\cot \theta}$

(d) $(\sec \theta - 1)(\sec \theta + 1)$

(e) $\dfrac{\sec \theta}{\operatorname{cosec} \theta}$

(f) $\sin \theta \operatorname{cosec} \theta + \cos \theta \sec \theta$

Proving identities

26. Prove each identity.

(a) $\cot \theta \sec \theta = \operatorname{cosec} \theta$

(b) $\dfrac{\operatorname{cosec} \theta}{\cot \theta + \tan \theta} = \cos \theta$

(c) $\sec^2 \theta + \operatorname{cosec}^2 \theta = \sec^2 \theta \operatorname{cosec}^2 \theta$

(d) $\dfrac{1 + \cot^2 \theta}{1 + \tan^2 \theta} = \cot^2 \theta$

Solving equations

27. Solve for $0 \leq x \leq 2\pi$.

(a) $\sec x = 2$

(b) $2 \tan^2 x + \sec^2 x = 4$

(c) $\cot^2 x + \operatorname{cosec} x = 1$

Concept check

28. A student writes: "$\cos^{-1} 0 = \dfrac{1}{\cos 0} = 1$." Explain the difference between $\cos^{-1} x$ and $\sec x$, and evaluate both at $x = 0$.

8.5 Inverse Trigonometric Functions HL

Everything so far has gone one way: give an angle, get a ratio. The inverse functions run the other way, recovering the angle from the ratio.

You have already used them without naming them. When you solve $\sin x = \frac{1}{2}$ and press the $\sin^{-1}$ key, that key is the inverse function.

Two notations, one function. The inverse of sine is written either $\arcsin x$ or $\sin^{-1} x$. They mean exactly the same thing. These notes use $\arcsin$, because the other form invites a serious confusion:

$$\sin^{-1} x \text{ means the } \textit{inverse}, \qquad \text{but} \qquad (\sin x)^{-1} = \frac{1}{\sin x} = \operatorname{cosec} x.$$

The -1 is not an exponent here. Writing $\arcsin$, $\arccos$ and $\arctan$ removes the ambiguity, and your calculator's $\sin^{-1}$ key is the same function under the other name.

8.5.1 Why a restriction is needed

There is an immediate difficulty. A function has an inverse only if it is one-to-one (Ch. 5), and $\sin$, $\cos$ and $\tan$ are not. Each takes every value in its range infinitely often, so the horizontal line $y = \frac{1}{2}$ meets $y = \sin x$ infinitely many times.

That is a real problem, not a technicality. If you ask "which angle has sine $\frac{1}{2}$?", there is no single answer: $\frac{\pi}{6}$, $\frac{5\pi}{6}$, $\frac{13\pi}{6}$ and infinitely many others all qualify. A function must return one output, so it has to choose.

The fix is the one from Ch. 5: cut the domain down until the function is one-to-one, then invert what is left. Which piece to keep is a choice, and mathematicians have agreed on one for each function. Those agreed pieces are shaded in the diagrams below, and they are what fix the ranges in the table.

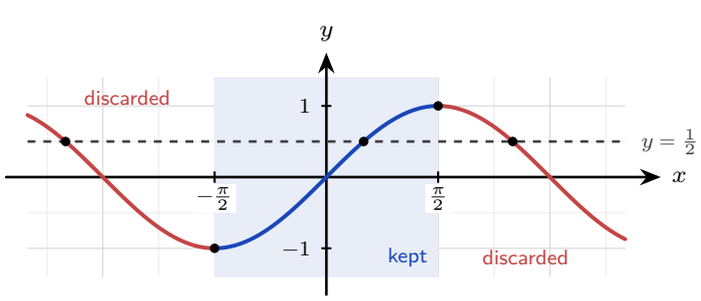


Figure 8.7 The dashed line $y = \frac{1}{2}$ meets $y = \sin x$ again and again, which is why "the angle whose sine is $\frac{1}{2}$" has no single answer. Keeping only the shaded piece from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ leaves one intersection. On that piece the curve rises steadily, so it is one-to-one and every output comes from exactly one input.

The same is done for the other two: cosine keeps $0 \le x \le \pi$, where it falls steadily, and tangent keeps $-\frac{\pi}{2} < x < \frac{\pi}{2}$, one branch between two asymptotes. Those kept pieces become the *ranges* of the inverses, because domain and range swap when a function is inverted.

Function	Domain	Range
$\arcsin x$	$-1 \le x \le 1$	$-\dfrac{\pi}{2} \le y \le \dfrac{\pi}{2}$
$\arccos x$	$-1 \le x \le 1$	$0 \le y \le \pi$
$\arctan x$	all real x	$-\dfrac{\pi}{2} < y < \dfrac{\pi}{2}$

Each inverse is a reflection of its restricted parent in the line $y = x$ (Ch. 5), so the graphs are the familiar curves turned on their sides. Arcsine climbs steeply from $(-1, -\frac{\pi}{2})$ to $(1, \frac{\pi}{2})$. Arccosine falls from $(-1, \pi)$ to $(1, 0)$. Arctangent flattens towards the horizontal asymptotes $y = \pm\frac{\pi}{2}$, which are the vertical asymptotes of tangent after the reflection.

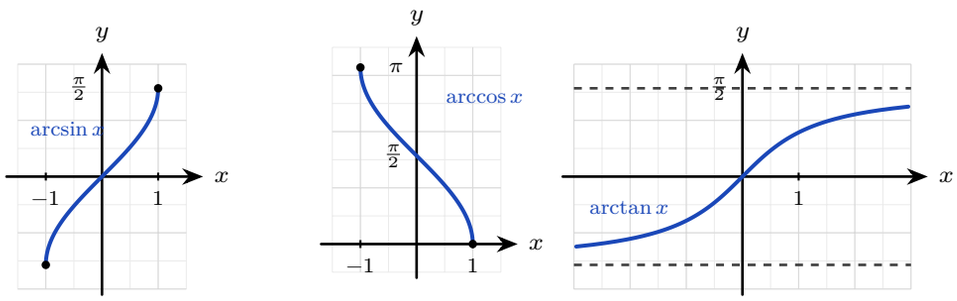


Figure 8.8 Each inverse is its restricted parent reflected in $y = x$. The marked endpoints show the range each one is confined to.

Worked through. To find $\arcsin\left(-\frac{\sqrt{3}}{2}\right)$, look for the angle in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ whose sine is $-\frac{\sqrt{3}}{2}$: it is $-\frac{\pi}{3}$. Not $\frac{4\pi}{3}$ and not $\frac{5\pi}{3}$: those solve the equation but lie outside arcsine's range. In the same way, $\arccos\left(-\frac{\sqrt{3}}{2}\right)$ is the angle in $[0, \pi]$ whose cosine is $-\frac{\sqrt{3}}{2}$, which is $\pi - \frac{\pi}{6} = \frac{5\pi}{6}$. Arccosine handles negatives by reaching into the second quadrant. Finally, $\sin\frac{5\pi}{4} = -\frac{\sqrt{2}}{2}$, and $\arcsin\left(-\frac{\sqrt{2}}{2}\right) = -\frac{\pi}{4}$, so $\arcsin\left(\sin\frac{5\pi}{4}\right)$ returns $-\frac{\pi}{4}$, not $\frac{5\pi}{4}$: arcsine can only answer with angles from its own range.

CAUTION $\arcsin(\sin x)$ does not always give back x. Every answer from arcsin must lie between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$, and 2 does not. So $\arcsin(\sin 2) = \pi - 2$, an angle with the same sine as 2 but inside the range. The restriction is working, not failing.

8.5.2 Compositions the other way round

A composition can also run with the inverse on the inside, as in $\cos(\arcsin k)$. The angle is never found here, and it never needs to be. Name it: let $\theta = \arcsin k$, so $\sin\theta = k$. A right triangle with opposite k and hypotenuse 1 has exactly that sine, and Pythagoras gives its adjacent side as $\sqrt{1-k^2}$, so

$$\cos(\arcsin k) = \sqrt{1-k^2}.$$

The positive root is the right one because $\arcsin$ answers in $[-\frac{\pi}{2},\ \frac{\pi}{2}]$, and cosine is never negative there. The same triangle serves $\arccos$ and $\arctan$: place the known ratio on the two sides it names, find the third by Pythagoras, then read off whatever the question asks for.

EXAMPLE 8.4

Find the exact value of (a) $\cos\left(\arcsin\dfrac{20}{29}\right)$, (b) $\sin\left(\arccos\dfrac{9}{41}+\arctan\dfrac{20}{21}\right)$.

(a) Let $\theta = \arcsin\frac{20}{29}$, so $\sin\theta = \frac{20}{29}$: opposite 20, hypotenuse 29. The adjacent side is $\sqrt{29^2-20^2} = \sqrt{441} = 21$, so

$$\cos\left(\arcsin\frac{20}{29}\right) = \frac{21}{29},$$

positive because θ lies in $[-\frac{\pi}{2},\ \frac{\pi}{2}]$.

(b) Name both angles: $A = \arccos\frac{9}{41}$ and $B = \arctan\frac{20}{21}$. A 9–40–41 triangle gives $\cos A = \frac{9}{41}$ and $\sin A = \frac{40}{41}$, taken positive because $\arccos$ answers in $[0,\ \pi]$, where sine is never negative. A 20–21–29 triangle, the same one as in part (a), gives $\sin B = \frac{20}{29}$ and $\cos B = \frac{21}{29}$, both positive because $\tan B > 0$ puts B in the first quadrant. The sum is now a compound angle:

$$\sin(A+B) = \sin A\cos B + \cos A\sin B = \frac{40}{41}\cdot\frac{21}{29} + \frac{9}{41}\cdot\frac{20}{29} = \frac{840+180}{1189} = \frac{1020}{1189}.$$

Neither A nor B was ever evaluated. The range of each inverse function is what fixes the sign of every side, and that is the only place in the working where a decision is made.

YOUR TURN 8.5 Inverse Trigonometric Functions HL

Exact values

29. Evaluate each exactly.

(a) $\arcsin\dfrac{1}{2}$ (b) $\arccos 0$

(c) $\arctan(-1)$ (d) $\arccos\left(-\dfrac{1}{2}\right)$

Domains, ranges and graphs

30. Use the graphs of the inverse functions.

(a) Sketch the graph of $y = \arccos x$, and label the coordinates of its endpoints.

(b) Write down the equations of the horizontal asymptotes of $y = \arctan x$.

(c) Find the domain of $y = \arcsin(x - 2)$.

(d) Find the range of $y = 2\arctan x + \pi$.

The inverse on the outside

31. An inverse undoes its function only on the piece that was kept.

(a) State the range of $\arcsin x$, of $\arccos x$ and of $\arctan x$.

(b) Evaluate $\arcsin\left(\sin\dfrac{3\pi}{4}\right)$, and explain why the answer is not $\dfrac{3\pi}{4}$.

(c) Evaluate $\arccos\left(\cos\dfrac{4\pi}{3}\right)$.

(d) Evaluate $\arctan\left(\tan\dfrac{5\pi}{6}\right)$.

(e) State every value of x for which $\arcsin(\sin x) = x$.

32. Find the exact value, or write **undefined**.

(a) $\arcsin\left(\sin\dfrac{2\pi}{3}\right)$

(b) $\arccos(\cos 4)$

(c) $\arctan\left(\tan\left(-\dfrac{3\pi}{4}\right)\right)$

(d) $\arcsin(\sin 3)$

(e) $\arcsin\left(\tan\dfrac{\pi}{3}\right)$

(f) $\arctan\left(2\sin\dfrac{\pi}{3}\right)$

The inverse on the inside

33. Find the exact value, or write **undefined**.

(a) $\tan(\arctan 12)$

(b) $\cos\left(\arccos\dfrac{2\pi}{3}\right)$

(c) $\sin(\arcsin\pi)$

(d) $\sin\left(\arctan\dfrac{3}{4}\right)$

(e) $\cos\left(\arcsin\dfrac{7}{25}\right)$

(f) $\cos\left(\arctan\dfrac{1}{2}\right)$

(g) $\cos(\arcsin 0.6)$

Sums of inverse angles

34. Use a compound-angle formula to find each exact value.

(a) $\sin\left(\arccos\dfrac{3}{5} + \arctan\dfrac{5}{12}\right)$

(b) $\cos\left(\arctan 3 + \arcsin\dfrac{1}{3}\right)$

Concept check

35. A student says: "$\arcsin\left(\sin\dfrac{5\pi}{6}\right)=\dfrac{5\pi}{6}$, *because* $\arcsin$ *undoes* $\sin$." Explain why this is false, and find the correct value.

8.6 Solving Trigonometric Equations

An equation like $\sin x=\frac{1}{2}$ has not one solution but infinitely many, because the sine repeats every 2π. Your calculator returns only one of them, the **principal value**. Everything else you must find yourself, and the circle is what tells you how.

Start by counting the solutions in a single turn. Solving $\sin x=k$ means finding the points on the circle whose y-coordinate is k, so draw the horizontal line at height k and see where it meets the circle. Three things can happen.

- If $-1<k<1$ the line cuts the circle in **two** places, so two angles in each turn have sine k.
- If $k=1$ or $k=-1$ the line touches the circle at **one** point, the top or the bottom. The two angles have merged into one.
- If $|k|>1$ the line misses the circle, and there is **no** solution at all.

The same three cases hold for $\cos x=k$, read with a vertical line at $x=k$ instead. Tangent behaves differently: it takes every real value, so $\tan x=k$ has solutions whatever k is.

Two words are needed next. You will use both constantly.

DEF The **reference angle** $\hat{\theta}$ is the acute angle between the arm OP and the x-axis. It is positive, it is less than $\frac{\pi}{2}$, and it ignores signs entirely: you get it from the size of k alone, as $\arcsin|k|$, $\arccos|k|$ or $\arctan|k|$. It is defined whenever the arm lies off the axes.

The **principal value** is the single angle your calculator returns. It is the one solution lying inside the range of the matching inverse function:

$$\arcsin k\in\left[-\tfrac{\pi}{2},\,\tfrac{\pi}{2}\right],\qquad \arccos k\in[0,\,\pi],\qquad \arctan k\in\left(-\tfrac{\pi}{2},\,\tfrac{\pi}{2}\right).$$

The two are easy to confuse. A principal value can be negative, and for a cosine it can be obtuse; a reference angle is never either. Take $\cos x=-\frac{1}{2}$. The calculator returns the principal value $\frac{2\pi}{3}$, while the reference angle is $\frac{\pi}{3}$, the acute angle the arm makes with the negative x-axis. Work from the reference angle and the calculator's choice of quadrant stops mattering.

The method. Every solution is built from the reference angle and the quadrant it sits in. The sign of k tells you which two quadrants to use, by the rule of §8.1.2, and each quadrant has one construction.

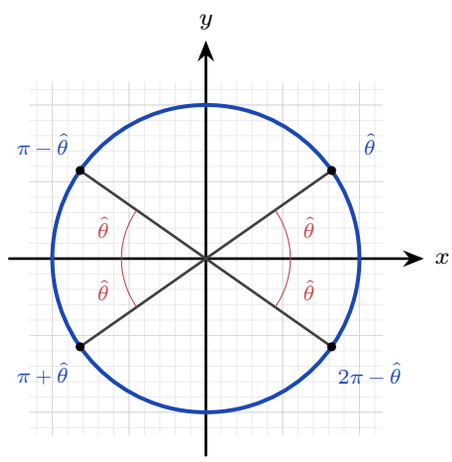


Figure 8.9 One reference angle $\hat{\theta}$, measured from the x-axis, gives an angle in each quadrant. The scarlet arcs are all the same size; only the direction they are measured from changes.

KEY **Solving a trigonometric equation.**

1. Rearrange until a single trigonometric function is alone on one side.
2. If the argument is not plain x, substitute u for it and **transform the interval as well**.
3. Find the reference angle from the *size* of k: $\hat{\theta} = \arcsin|k|$, $\arccos|k|$ or $\arctan|k|$.
4. Use the sign of k to choose the two quadrants, then build the angle in each: $\hat{\theta}$, $\pi - \hat{\theta}$, $\pi + \hat{\theta}$ or $2\pi - \hat{\theta}$.
5. Add or subtract whole turns of 2π until you have every solution the interval asks for.
6. If you substituted, convert back to x at the very end.

There is a shorter route once you already have one solution, which you do whenever a solution exists (for sine and cosine, when $|k| \leq 1$): the calculator has just given you one. Call it θ. Its partner in the same turn is fixed by the function alone, whatever quadrant θ landed in.

KEY **The partner solution.** If θ is one solution, then the other in the same turn is

$$\sin x = k: \quad \pi - \theta, \qquad \cos x = k: \quad 2\pi - \theta, \qquad \tan x = k: \quad \pi + \theta.$$

Then add or subtract whole turns of 2π (for the tangent, half turns of π) to reach every solution the interval asks for.

The interval decides how the partner is written, not what it is. On $-\pi \leq x \leq \pi$ the cosine pair is θ and $-\theta$, because $2\pi - \theta$ is one whole turn away from $-\theta$. That is why a cosine equation on a symmetric interval always answers $\pm$ the same number, and it is worth

recognising on sight: $\cos x = \frac{1}{2}$ on $-\pi \le x \le \pi$ gives $x = \pm\frac{\pi}{3}$ and nothing else.

Both routes give the same angles. The reference angle is the one to fall back on when the signs get confusing, or when the equation has been rearranged so far that you no longer trust which quadrant you are in.

Two checks belong in every solution. Your two angles should lie in the quadrants where that function has the sign of k; if only one does, you have built the wrong pair. And when $k = \pm 1$ the two constructions land on the same point of the circle, so report that angle once in each turn. Keep both ends when the interval includes them: $\cos x = 1$ on $0 \le x \le 2\pi$ gives $x = 0$ and $x = 2\pi$.

Take care at steps 2 and 5: a solution missed there does not reappear later.

Worked through. To solve $\cos x = -\frac{1}{2}$ for $0 \le x \le 2\pi$, take the reference angle from the size of k alone:

$$\hat{\theta} = \arccos\left|-\tfrac{1}{2}\right| = \arccos\tfrac{1}{2} = \frac{\pi}{3}.$$

Cosine is negative in the second and third quadrants, so build one angle in each:

$$\pi - \frac{\pi}{3} = \frac{2\pi}{3}, \qquad \pi + \frac{\pi}{3} = \frac{4\pi}{3}.$$

Both lie in the interval, so $x = \frac{2\pi}{3}$ or $x = \frac{4\pi}{3}$. Notice that the calculator alone would have given only $\frac{2\pi}{3}$. The reference angle gives both, and the sign of k says where they live. The same steps work when k is not an exact value: only the reference angle now comes from the calculator.

TIP On a calculator paper, always check your solution count graphically: plot both sides and count the intersections in the interval. If your algebra found three solutions and the screen shows four, look for the one you missed.

When the argument is not plain x but $2x$ or $3(x-4)$, substitute for the whole argument and, crucially, transform the interval along with it. A doubled argument sweeps the circle twice as fast, so it collects twice as many solutions.

EXAMPLE 8.5

Solve $\cos 2x = \dfrac{\sqrt{3}}{2}$ for $0 \le x \le 2\pi$.

Let $u = 2x$. As x runs over $[0, 2\pi]$, u runs over $[0, 4\pi]$: two full turns.

The reference angle is $\hat{\theta} = \arccos\frac{\sqrt{3}}{2} = \frac{\pi}{6}$. Cosine is positive, so the first turn gives

$$u = \frac{\pi}{6} \quad \text{and} \quad u = 2\pi - \frac{\pi}{6} = \frac{11\pi}{6}.$$

The interval covers a second turn, so add 2π to each:

$$\frac{\pi}{6} + 2\pi = \frac{13\pi}{6}, \qquad \frac{11\pi}{6} + 2\pi = \frac{23\pi}{6}.$$

A third turn would pass 4π, so the list is complete:

$$u = \frac{\pi}{6},\ \frac{11\pi}{6},\ \frac{13\pi}{6},\ \frac{23\pi}{6} \implies x = \frac{\pi}{12},\ \frac{11\pi}{12},\ \frac{13\pi}{12},\ \frac{23\pi}{12}.$$

Four solutions, not two. Note where the doubling acts: the step in u is still 2π, because $\cos u$ has period 2π whatever u came from. What doubling changes is the *interval*, and that is why two extra solutions appear. In x the period has become π.

8.6.1 When the interval is not 0 to 2π

The interval is part of the question, and it is not always $0 \le x \le 2\pi$. Another interval you will meet is $-\pi \le x \le \pi$. That is still one whole turn, but measured differently: half a turn clockwise from the positive x-axis, and half a turn anticlockwise.

The method does not change. Build the two angles from the reference angle exactly as before, then shift either one by a whole turn if it falls outside the interval: subtract 2π from an angle above π, add 2π to an angle below $-\pi$. A whole turn returns the point to exactly where it was, so the value of the function is untouched.

EXAMPLE 8.6

Solve $\cos x = -\dfrac{1}{2}$ for $-\pi \le x \le \pi$.

The reference angle is $\hat{\theta} = \arccos \frac{1}{2} = \frac{\pi}{3}$. Cosine is negative, so the two quadrants are the second and the third:

$$\pi - \frac{\pi}{3} = \frac{2\pi}{3}, \qquad \pi + \frac{\pi}{3} = \frac{4\pi}{3}.$$

The first already lies in $[-\pi,\ \pi]$. The second does not, because $\frac{4\pi}{3} > \pi$, so subtract one turn:

$$\frac{4\pi}{3} - 2\pi = -\frac{2\pi}{3}.$$

So $x = \frac{2\pi}{3}$ or $x = -\frac{2\pi}{3}$.

The two answers form a $\pm$ pair, and that is not a coincidence. Cosine is even, so on any interval symmetric about 0 its answers are $\pm\alpha$ for a single angle α. Sine is odd, and its pair on the same interval is

$$\hat{\theta} \text{ and } \pi - \hat{\theta} \quad \text{when } k > 0, \qquad -\hat{\theta} \text{ and } -\pi + \hat{\theta} \quad \text{when } k < 0.$$

8.6.2 Using identities

An equation mixing $\sin 2x$ with $\sin x$, or $\cos 2x$ with $\cos x$, cannot be solved as it stands. An identity rewrites everything in terms of a single function, turning the equation into an algebraic one, usually a quadratic.

EXAMPLE 8.7

Solve $\cos 2x + 3\cos x + 2 = 0$ for $0 \le x \le 2\pi$.

Choose the form of $\cos 2x$ that gives only cosines, $\cos 2x = 2\cos^2 x - 1$:

$$2\cos^2 x - 1 + 3\cos x + 2 = 0 \implies 2\cos^2 x + 3\cos x + 1 = 0.$$

This factors as $(2\cos x + 1)(\cos x + 1) = 0$, so $\cos x = -\frac{1}{2}$ or $\cos x = -1$. From the circle,

$$x = \frac{2\pi}{3},\ \frac{4\pi}{3} \quad \text{or} \quad x = \pi.$$

Picking the form of the double-angle identity that matches the other terms is the step that turns these problems into quadratics.

YOUR TURN 8.6 Solving Trigonometric Equations

Exact solutions

36. Solve for $0 \le x \le 2\pi$.

(a) $\sin x = \dfrac{1}{2}$

(b) $\cos x = 0$

(c) $\tan x = 1$

(d) $\cos x = 1$

37. Solve for $0 \le x \le 2\pi$.

(a) $\sin x = -\dfrac{1}{2}$

(b) $2\cos x = \sqrt{3}$

(c) $\sqrt{2}\sin x = -1$

(d) $\tan x = -\sqrt{3}$

Other intervals

38. Solve for $-\pi \le x \le \pi$. Give every answer inside that interval.

(a) $\cos x = \dfrac{\sqrt{3}}{2}$

(b) $\sin x = -\dfrac{1}{2}$

(c) $\tan x = -1$

(d) $2\sin x = \sqrt{2}$

Answers that are not exact

39. GDC Solve for $0 \leq x \leq 2\pi$, giving your answers to three significant figures. If there is no solution, say why.

(a) $\cos x = 0.3$
(b) $3\sin x = -2$
(c) $4\cos x + 5 = 0$

40. GDC Solve $\sin x = -0.8$ for $0 \leq x \leq 2\pi$, giving your answers to three significant figures.

Multiple angles

41. Solve for $0 \leq x \leq 2\pi$. Transform the interval before solving.

(a) $\sin 2x = 0$
(b) $\cos 2x = \dfrac{1}{2}$
(c) $\tan 2x = 1$

Shifted arguments

42. Solve for $0 \leq x \leq 2\pi$. Transform the interval before solving.

(a) $\sin\left(x + \dfrac{\pi}{3}\right) = \dfrac{1}{2}$
(b) $\cos\left(x - \dfrac{\pi}{4}\right) = -\dfrac{\sqrt{2}}{2}$
(c) $2\sin\left(2x - \dfrac{\pi}{3}\right) = \sqrt{3}$

Factorising

43. Factorise, then solve for $0 \leq x \leq 2\pi$.

(a) $2\sin^2 x - \sin x = 0$
(b) $2\cos^2 x + \cos x - 1 = 0$
(c) $2\sin^2 x + 3\sin x + 1 = 0$
(d) $\tan^2 x - \tan x = 0$

Using an identity first

44. Use an identity to rewrite each equation, then solve for $0 \leq x \leq 2\pi$.

(a) $\cos 2x + \sin x = 0$
(b) $\sin 2x = \cos x$
(c) Explain why dividing both sides of the equation in (b) by $\cos x$ would lose solutions.
(d) $2\sin^2 x = 3\cos x$

Concept check

45. To solve $\sin 2x = \frac{1}{2}$ for $0 \le x \le 2\pi$, a student writes: "$2x = \frac{\pi}{6}$ *or* $\frac{5\pi}{6}$, *so* $x = \frac{\pi}{12}$ *or* $\frac{5\pi}{12}$." Identify the error and give the full solution set.

Chapter 8 · Trigonometry

Reflections

TOK A sine wave repeats forever, identical in every cycle, and nothing in the physical world does that exactly. Yet sums of sine waves describe voices, tides and heartbeats closely enough to compress, predict and diagnose them. Is the sinusoid a human tool, or an 'alphabet' of nature, a basic set of shapes from which the repeating patterns of the physical world are built?

IM Why is a full turn $360°$? The number comes from Babylonian astronomy, which counted in base 60, and 360 divides exactly by 2, 3, 4, 5, 6, 8, 9, 10 and 12. During the French Revolution, the reformers who introduced the metre also proposed a decimal angle, the *grade*, with 400 in a full turn and 100 in a right angle. Most of the world kept degrees, but the grade, now called the *gon*, is still used by surveyors in parts of continental Europe, and many calculators still offer a GRAD mode. Which unit is right depends on who is measuring and why, and calculus, as Chapter 13 shows, uses neither: it uses radians.

RW Because a sine wave is what steady rotation looks like from the side, it models many oscillations in science and engineering. Alternating current rises and falls as a sine while the generator spins. Sound is pressure oscillating in the same way, and its pitch is the frequency of the wave. Light and radio are sinusoids too, and the signals your phone sends and receives are built from them. Fourier analysis, decomposing a signal into its sine components, is used to compress music and images and to read earthquake and heart signals.

EXT In radians, $\sin x$ and x almost agree for small x: $\sin 0.1 = 0.09983\ldots$ On the unit circle x is the arc and $\sin x$ is the height, and over a tiny arc there is almost no difference between them. Physicists use this as the **small-angle approximation**: for $|x|$ small and in radians,

$$\sin x \approx x, \qquad \tan x \approx x, \qquad \cos x \approx 1 - \tfrac{1}{2}x^2.$$

At $x = 0.1$ these give 0.1, 0.1 and 0.995, against true values 0.09983, 0.10033 and 0.99500. The pendulum equation can be solved in elementary functions only when $\sin\theta$ is replaced by θ, which is why the simple period formula holds for small swings and not for large ones. All three fail in degrees, and that is one reason calculus uses radians. Chapter 13 turns the first into $\lim_{x\to 0} \dfrac{\sin x}{x} = 1$, the limit that gives $\dfrac{d}{dx}\sin x = \cos x$.

End-of-Chapter Problems
Chapter 8: Trigonometry

1. Use the unit circle to write down the exact value of each.

(a) $\sin\dfrac{2\pi}{3}$ [1] (b) $\cos\dfrac{2\pi}{3}$ [1]

(c) $\tan\dfrac{2\pi}{3}$ [2] (d) $\cos\dfrac{5\pi}{4}$ [2]

(e) $\sin\dfrac{11\pi}{6}$ [2] (f) $\tan\dfrac{3\pi}{4}$ [2]

2. The point P lies on the unit circle at angle θ.

(a) State the coordinates of P when $\theta = \dfrac{5\pi}{6}$. [2]

(b) P has coordinates $\left(-\dfrac{\sqrt{2}}{2}, -\dfrac{\sqrt{2}}{2}\right)$. Find θ for $0 \le \theta \le 2\pi$. [2]

(c) P lies in the third quadrant. State the sign of $\sin\theta$, of $\cos\theta$ and of $\tan\theta$. [3]

(d) Given $\sin\theta = \dfrac{\sqrt{3}}{2}$, write down the two possible values of θ in $0 \le \theta \le 2\pi$, and explain why there are exactly two. [3]

3. Prove each identity. They are arranged in increasing order of difficulty.

(a) $\cos^2\theta - \sin^2\theta = 1 - 2\sin^2\theta$ [2]

(b) $\dfrac{\sin 2\theta}{\sin\theta} - \dfrac{\cos 2\theta}{\cos\theta} = \sec\theta$ [3]

(c) $\tan\theta + \cot\theta = 2\operatorname{cosec} 2\theta$ [3]

(d) $\cos^4\theta - \sin^4\theta = \cos 2\theta$ [3]

(e) $\dfrac{1-\cos 2\theta}{1+\cos 2\theta} = \tan^2\theta$ [4]

(f) $\dfrac{1}{\sec\theta - \tan\theta} = \sec\theta + \tan\theta$ [4]

(g) $\dfrac{\sin\theta}{1+\cos\theta} + \dfrac{1+\cos\theta}{\sin\theta} = 2\operatorname{cosec}\theta$ [5]

(h) $\dfrac{1+\sin 2\theta}{\cos 2\theta} = \dfrac{\cos\theta + \sin\theta}{\cos\theta - \sin\theta}$ [5]

4. Given $\sin\theta = \dfrac{5}{13}$ with θ in the second quadrant, find the exact value of each.

(a) $\cos\theta$ [2] (b) $\tan\theta$ [1]

(c) $\sin 2\theta$ [2] (d) $\cos 2\theta$ [2]

5. Given $\tan\theta = 2$ with θ acute, find the exact value of each.

(a) $\sin\theta$ and $\cos\theta$ [2] (b) $\sin 2\theta$ [2]
(c) $\cos 2\theta$ [2]

6. Use a compound-angle formula to find each exact value.

(a) $\tan 75^\circ$ [4] (b) $\sin 15^\circ$ [3]
(c) $\sin 105^\circ$ [3] (d) $\tan 105^\circ$ [4]

7. Prove each result.

(a) $\sin(A+B)+\sin(A-B)=2\sin A\cos B$ [4]
(b) $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ [5]
(c) $\sin 3\theta = 3\sin\theta - 4\sin^3\theta$ [5]

8. Solve for $0 \le x \le 2\pi$.

(a) $2\cos x + 1 = 0$ [2] (b) $\sqrt{3}\tan x = 1$ [3]
(c) $2\cos^2 x = 1$ [4] (d) $\tan^2 x = 3$ [4]

9. Solve for $0 \le x \le 2\pi$. Each becomes a quadratic.

(a) $2\sin^2 x + \sin x - 1 = 0$ [5] (b) $4\sin^2 x - 3 = 0$ [4]
(c) $\cos 2x = \cos x$ [5] (d) $\sin 2x = \sin x$ [5]

10. Solve for $0 \le x \le 2\pi$.

(a) $2\sin x\cos x = \dfrac{1}{2}$ [4]
(b) $\cos 2x + 3\sin x = 2$ [6]
(c) $\sin x + \cos x = 1$ [6]

11. Solve each equation for $-\pi \le x \le \pi$. Build every answer from the reference angle, then shift by a whole turn where the interval requires it.

(a) $\sin x = \dfrac{\sqrt{3}}{2}$ [3] (b) $2\cos x = -\sqrt{3}$ [3]
(c) $\sqrt{3}\tan x = -1$ [3] (d) $2\cos^2 x - 1 = 0$ [4]
(e) $\sin 2x = \dfrac{\sqrt{3}}{2}$ [5] (f) $\tan\left(x + \dfrac{\pi}{4}\right) = \sqrt{3}$ [5]

12. Given the identity $1 + \tan^2\theta = \sec^2\theta$, solve $\sec^2 x = 4$ for $0 \le x \le \dfrac{\pi}{2}$. [4]

13. Evaluate, giving exact values.

(a) $\arctan 1$ [1] (b) $\arcsin\left(-\dfrac{1}{2}\right)$ [1]

(c) $\arccos(-1)$ [1] (d) $\arcsin\left(\sin\dfrac{5\pi}{6}\right)$ [3]

14. Consider $f(x) = 3\sin(2x) + 1$.

(a) State the amplitude and the period. [2]
(b) State the maximum and minimum values. [2]

15. The graph of $y = a\sin(bx) + d$ has maximum 7, minimum 1, and period π. Find a, b and d. [4]

16. Sketch $y = \cos x - 1$ for $0 \le x \le 2\pi$, stating the maximum, the minimum, and the intercepts. [4]

17. The depth of water in a harbour is $d(t) = 5 + 3\sin\left(\dfrac{\pi t}{6}\right)$ metres, t hours after midnight.

(a) State the maximum and minimum depths. [2]
(b) Find the period, and say what it represents. [2]

18. A seat on a Ferris wheel has height $h(t) = 10 - 8\cos\left(\dfrac{\pi t}{15}\right)$ metres, t seconds after the start.

(a) State the minimum and maximum heights. [2]
(b) Find the time for one full revolution. [2]

19. GDC Express $3\sin x + 4\cos x$ in the form $R\sin(x + \alpha)$ with $R > 0$.

(a) Find R and α. [3]
(b) Hence state the maximum value. [1]

20. GDC Find the maximum value M of $5\sin x + 12\cos x$, and the value of x in $[0, 2\pi]$ at which it occurs. [5]

21. Use the symmetry of the unit circle to answer each part.

(a) Given $\sin 0.6 = 0.565$, write down $\sin(\pi - 0.6)$. [2]
(b) Given $\cos 1.2 = 0.362$, write down $\cos(-1.2)$ and $\cos(2\pi - 1.2)$. [3]
(c) Explain why $\sin(\theta + 2\pi) = \sin\theta$ for every θ. [2]
(d) Explain why $\tan(\theta + \pi) = \tan\theta$, but $\sin(\theta + \pi) \neq \sin\theta$. [4]

22. Solve each equation on the interval given. Take care: the same equation has different answers on different intervals.

(a) $\cos x = -\dfrac{\sqrt{2}}{2}$ for $0 \le x \le 2\pi$, and then for $-\pi \le x \le \pi$. [4]

(b) $\sin x = -\dfrac{\sqrt{3}}{2}$ for $-\pi \le x \le \pi$. [3]

(c) $\tan 2x = \sqrt{3}$ for $-\pi \le x \le \pi$. [5]

(d) Explain why an equation of the form $\cos x = k$, with $|k| < 1$, always has two solutions on $-\pi \le x \le \pi$, and why they are negatives of each other. [3]

23. GDC **Daylight in Sapporo.** The number of hours of daylight is modelled by

$$L(t) = 12.2 + 3.4\sin\left(\frac{2\pi(t-81)}{365}\right),$$

where t is the day of the year, with $t = 1$ on 1 January.

(a) State the maximum and minimum daylight, and find the value of t at which each occurs. [4]

(b) State the period of L, and say what it represents. [2]

(c) Find the daylight on 1 January, to three significant figures. [2]

(d) Find the two days of the year on which the daylight is exactly 12.2 hours, and explain what is special about them. [4]

(e) Find the values of t for which $L(t) \ge 15$, and state how many days of the year have at least 15 hours of daylight. [6]

24. One function, two forms. Let $f(x) = 2\sin x + \cos 2x$ for $0 \le x \le 2\pi$.

(a) Show that $f(x) = -2\sin^2 x + 2\sin x + 1$. [3]

(b) Let $u = \sin x$. Write f as a quadratic in u, and state the interval of values u can take. [3]

(c) Hence show that the maximum value of f is $\dfrac{3}{2}$. [4]

(d) Find the values of x in $[0, 2\pi]$ at which that maximum occurs. [3]

(e) Solve $f(x) = 1$ for $0 \le x \le 2\pi$. [5]

(f) State the range of f. [3]

25. Find the exact value where there is one, and write **undefined** where there is not.

(a) $\arcsin\left(\sin\dfrac{7\pi}{6}\right)$ [2]

(b) $\arccos\left(\cos\dfrac{7\pi}{4}\right)$ [2]

(c) $\arctan\left(\tan\dfrac{2\pi}{3}\right)$ [2]

(d) $\sin\left(\arccos\dfrac{5}{13}\right)$ [2]

(e) $\tan\left(\arcsin\dfrac{8}{17}\right)$ [3]

(f) $\cos(\arctan 2)$ [2]

(g) $\arccos(\sin\pi)$ [2]

(h) $\sin(\arcsin 2)$ [2]

(i) $\cos\left(\arcsin\dfrac{3}{5} + \arctan\dfrac{4}{3}\right)$ [4]

(j) $\tan\left(2\arctan\dfrac{1}{3}\right)$ [3]

26. Ten identities. Prove each one. They draw on the double-angle, compound-angle and reciprocal results of this chapter, and they are arranged in increasing order of difficulty.

(a) $\sin(A+B)-\sin(A-B)=2\cos A\sin B$ [3]

(b) $\dfrac{\sin 2\theta}{1-\cos 2\theta}=\cot\theta$ [3]

(c) $\operatorname{cosec} 2\theta=\frac{1}{2}\sec\theta\operatorname{cosec}\theta$ [3]

(d) $\dfrac{\sec\theta+1}{\tan\theta}=\dfrac{\tan\theta}{\sec\theta-1}$ [4]

(e) $\tan\left(\theta+\dfrac{\pi}{4}\right)=\dfrac{1+\tan\theta}{1-\tan\theta}$ [4]

(f) $\cos(A+B)\cos(A-B)=\cos^2 A-\sin^2 B$ [5]

(g) $\dfrac{1+\tan^2\theta}{1-\tan^2\theta}=\sec 2\theta$ [5]

(h) $\cot\theta-\tan\theta=2\cot 2\theta$ [5]

(i) $\dfrac{\cos 2\theta}{1+\sin 2\theta}=\dfrac{1-\tan\theta}{1+\tan\theta}$ [6]

(j) $\operatorname{cosec}\theta-\cot\theta=\tan\dfrac{\theta}{2}$ [6]

Challenge Questions

27. GDC **When is a sum of waves periodic?** A function f is *periodic* with period $T>0$ if $f(x+T)=f(x)$ for every x. Every sine wave is periodic, but a sum of two sine waves need not be. This investigation finds when the sum repeats, what it looks like when the two periods are close, and what happens when it never repeats.

(a) State the periods of $\sin 2x$ and $\sin 3x$. Show that $f(x)=\sin 2x+\sin 3x$ satisfies $f(x+2\pi)=f(x)$, and by evaluating $f(\frac{\pi}{2})$ and $f(\frac{3\pi}{2})$ show that π is not a period of f. [3]

(b) Let p and q be positive numbers with $\dfrac{p}{q}=\dfrac{m}{n}$, where m and n are positive integers. Show that $\sin px+\sin qx$ is periodic, with period $\dfrac{2\pi m}{p}$, and check this against part (a). [3]

(c) By writing $A=u+v$ and $B=u-v$ and expanding $\cos(u+v)+\cos(u-v)$, show that

$$\cos A+\cos B=2\cos\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right).$$

[3]

(d) Two tuning forks sound at 440 Hz and 442 Hz together. The combined air pressure is modelled by $y=\cos(880\pi t)+\cos(884\pi t)$, with t in seconds. Write y as a product, and use it to describe the graph of y. Hence find how many times each second the sound is loudest, and state the general rule for forks of frequencies f_1 and f_2 Hz that are close together. [4]

(e) Let $g(x)=\cos x+\cos(\sqrt{2}\,x)$. Show that $g(x)=2$ only when $x=0$, and deduce that g is not periodic. You may assume that $\sqrt{2}$ is irrational. [5]

(f) Use technology to find the greatest value of $g(x)$ for $5 \le x \le 100$, and the value of x at which it occurs. Explain why this x is close to a multiple of 2π, by relating it to a fraction close to $\sqrt{2}$. Suggest how this could be investigated further. [3]

28. $\cos n\theta$ **as a polynomial in** $\cos\theta$**.** You have already met $\cos 2\theta = 2\cos^2\theta - 1$ and, in the end-of-chapter problems, $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$. Each is a polynomial in $\cos\theta$ alone. This investigation shows that this always happens, and finds the pattern.

Write $c = \cos\theta$ throughout.

(a) Use the compound-angle formulas to show that

$$\cos((n+1)\theta) = 2\cos\theta\,\cos(n\theta) - \cos((n-1)\theta).$$

(Hint: expand $\cos(n\theta+\theta)$ *and* $\cos(n\theta-\theta)$*, then add)* [5]

(b) Starting from $\cos 0 = 1$ and $\cos\theta = c$, use the recurrence to obtain $\cos 2\theta$ and $\cos 3\theta$ in terms of c. [3]

(c) Use it again to find $\cos 4\theta$ and $\cos 5\theta$ in terms of c. [4]

(d) For $n \ge 1$, state the degree of the polynomial for $\cos n\theta$, and the coefficient of its highest power. Justify both from the recurrence. [4]

(e) Explain why no such polynomial in $\cos\theta$ alone exists for $\sin n\theta$, for any $n \ge 1$. *(Hint: compare* $\sin(-\theta)$ *with* $\sin\theta$*)* [3]

(f) Find the three roots of $4c^3 - 3c = 0$ by solving $\cos 3\theta = 0$ for $0 \le \theta \le \pi$. Generalise: state the n roots of the polynomial for $\cos n\theta$, explain why there are no others, and suggest how this could be investigated further. [3]

29. GDC **The window of safe passage.** At a harbour mouth the depth of water is modelled by

$$d(t) = 6 + 2\sin\left(\frac{\pi t}{6}\right) \text{ metres},$$

where t is the number of hours after midnight. A ferry can enter only while the depth is at least the figure its hull requires. This investigation finds the window of time available, and then finds it in general.

(a) State the greatest and least depths, the period of the tide, and the time of the first high tide. [4]

(b) The ferry needs at least 7 m of water. Solve $d(t) \ge 7$ on $0 \le t \le 12$, and state the window as a pair of clock times. [5]

(c) Show that the window is exactly 4 hours long, and explain why the same window opens once in every tidal cycle. [3]

(d) Now take a general tide $d(t) = m + a\sin\left(\dfrac{2\pi t}{T}\right)$ and a required depth k with $m < k < m + a$. Show that the time available in each cycle is

$$\frac{T}{\pi}\arccos\left(\frac{k-m}{a}\right).$$

(Hint: solve for the two ends in terms of $\arcsin$*, subtract, then use* $\frac{\pi}{2} - \arcsin c = \arccos c$*)* [6]

(e) A larger ferry needs 7.5 m. Use the formula to find its window, to the nearest minute. Then explain, from the shape of the arccosine graph, why the last half-metre of clearance costs far more time than the first. [4]

(f) Real tides are closer to a sum of waves. The two largest have periods of about 12.42 hours (from the Moon) and 12.00 hours (from the Sun), so a better model is

$$d(t) = 6 + 1.5\sin\left(\frac{2\pi t}{12.42}\right) + 0.5\sin\left(\frac{2\pi t}{12}\right).$$

Explain why the range between high and low water grows and shrinks, state the largest and smallest ranges, and find, in days, how long the range takes to return to its largest value. Suggest how the window of part (d) could be studied for this model. [3]

30. GDC **Machin's formula for π.** In 1706 John Machin used the compound-angle formula for the tangent to write π in terms of two small arctangents, and so found 100 decimal places of π by hand. This investigation proves his formula and shows why small arguments help.

(a) Use the formula for $\tan(A+B)$ to show that $\arctan\frac{1}{2} + \arctan\frac{1}{3} = \frac{\pi}{4}$. Explain why the sum on the left must lie between 0 and $\frac{\pi}{2}$. [3]

(b) Let $\alpha = \arctan\frac{1}{5}$. Show that $\tan 2\alpha = \frac{5}{12}$ and $\tan 4\alpha = \frac{120}{119}$. [3]

(c) Show that $\tan\left(4\alpha - \frac{\pi}{4}\right) = \frac{1}{239}$. [2]

(d) Show that $0 < 4\alpha - \frac{\pi}{4} < \frac{\pi}{2}$, and hence prove Machin's formula

$$\frac{\pi}{4} = 4\arctan\frac{1}{5} - \arctan\frac{1}{239}.$$

(Hint: use part (b) to show first that $2\alpha < \frac{\pi}{4}$*)* [4]

(e) For small x, $\arctan x \approx x - \frac{x^3}{3} + \frac{x^5}{5}$ (you will meet this series in Ch. 16). Use these three terms for $\arctan\frac{1}{5}$, and only the first term for $\arctan\frac{1}{239}$, to estimate π, and compare your estimate with the value of π. Explain why Machin's formula gives π far faster than the formula of part (a) would. [4]

(f) Show that for every positive integer n,

$$\arctan\frac{1}{n} - \arctan\frac{1}{n+1} = \arctan\frac{1}{n^2+n+1},$$

and check that $n = 1$ gives part (a). Suggest how this could be used to investigate other formulas of Machin's kind. [3]

GEOMETRY & TRIGONOMETRY

CHAPTER 9

Vectors

9

SYLLABUS 3.12–3.18 HL ONLY

A camera tripod has three legs, not four. The choice is deliberate. Part of the reason is geometry.

Any three points lie in one flat surface, whatever their heights. All three feet of a tripod therefore reach the ground together, on any floor, however uneven. A fourth point need not lie in that surface. When it does not, the stand rocks between two pairs of legs. This is why a stool with four legs can wobble on an uneven floor, where a stool with three legs does not.

Ordinary arithmetic cannot express that fact. The fact is not about how large anything is. It is about position, direction and flatness. To state it, and to prove it, we need a language in which a direction is something we can calculate with.

A **vector** is a quantity that has both a size and a direction. Speed is a number, but velocity is a vector. Distance is a number, but displacement is a vector. A pilot who flies due north through a wind from the west does not travel due north, and does not travel at the aircraft's airspeed. The motion and the wind point in different directions, and both must be counted.

Once a quantity has a direction, addition changes. Three plus three need not give six. If the two vectors point the same way, the answer is six. If they point in opposite directions, the answer is zero. Any value between those two is possible, and the angle between them fixes which one.

By the end of this chapter you will add and scale vectors, use the scalar product to find angles and to test for perpendicularity, and use the vector product to find areas and normals. You will write the equation of a line and of a plane, decide how two lines or three planes are arranged, and find shortest distances. A vector holds a size and a direction in a single object, which makes it the natural tool for describing position and movement in three dimensions. Physics, computer graphics and navigation all rely on it, from the pilot's crosswind to the satellites overhead.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

9.1 Vectors: Magnitude and Direction

A vector has exactly two properties: a **magnitude**, which is its length, and a **direction**. Where it sits in space does not matter. Two arrows of the same length pointing the same way are the same vector.

An arrow drawn to represent a vector is called a **directed line segment**. "Directed" because it has an arrowhead, and "line segment" because it is a piece of a line with two ends. Examination questions use this phrase; it means an arrow and nothing more.

Vectors must be kept apart from ordinary numbers, which in this context are called **scalars**. Three notations are in common use, and they all mean the same thing. Printed books, including this one, use a bold letter $\mathbf{a}$. By hand you cannot write bold type, so you write either a letter with an arrow above it, $\vec{a}$, or an underlined letter, $\underline{a}$. An arrow is also used for a vector named by its two endpoints: $\overrightarrow{AB}$ means the vector from the point A to the point B.

The usual representation is the **column vector**. It lists how far to travel along each axis to get from the tail to the head.

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = a_1\hat{\mathbf{i}} + a_2\hat{\mathbf{j}} + a_3\hat{\mathbf{k}}.$$

Here $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are the **base vectors**: arrows of length 1 along the x-, y- and z-axes. The hat is there to record that each one has length 1, a point returned to in the next subsection. Many books print them without hats, as $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$; the two notations are equivalent, and the formula booklet uses only the column form. The column form and the base-vector form describe the same vector, so use whichever is convenient.

9.1.1 The algebra of arrows

Vectors are added **tip to tail**. Draw the second vector starting where the first one ends. The sum is the arrow from the start of the first to the end of the second. In components it is simple: everything happens coordinate by coordinate.

KEY

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix}, \qquad k\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} ka_1 \\ ka_2 \\ ka_3 \end{pmatrix}.$$

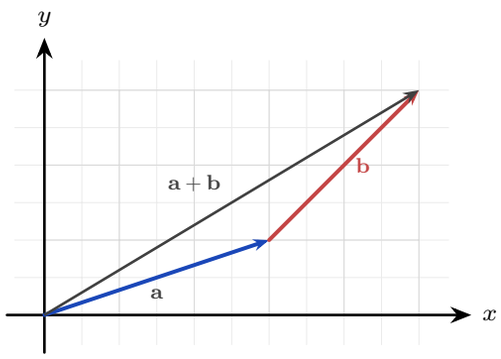


Figure 9.1 Adding tip to tail. Draw $\mathbf{b}$ starting where $\mathbf{a}$ ends; the sum runs from the start of $\mathbf{a}$ to the end of $\mathbf{b}$.

Scalar multiplication by k scales the length by $|k|$ and reverses the direction when k is negative. This gives a clean test: two nonzero vectors are **parallel** exactly when one is a scalar multiple of the other, $\mathbf{a} = k\mathbf{b}$.

Two particular vectors have names of their own.

The **negative** of $\mathbf{v}$, written $-\mathbf{v}$, is the vector of the same length pointing the opposite way. It is the case $k = -1$ of scalar multiplication. Subtraction is then just addition of the negative:

$$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b}).$$

The **zero vector** $\mathbf{0}$ has magnitude 0 and no direction at all. It is what you get from $\mathbf{v} + (-\mathbf{v})$, or from $0\mathbf{v}$. It is the only vector whose direction is undefined, and it is written in bold to keep it apart from the number 0.

Worked through. Take $\mathbf{a} = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 2 \\ -4 \\ 1 \end{pmatrix}$. To find $2\mathbf{a} - \mathbf{b}$, scale first, then subtract, working one coordinate at a time:

$$2\mathbf{a} - \mathbf{b} = \begin{pmatrix} 6 \\ 2 \\ -4 \end{pmatrix} - \begin{pmatrix} 2 \\ -4 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ -5 \end{pmatrix}.$$

Is it parallel to $\mathbf{c} = \begin{pmatrix} -8 \\ -12 \\ 10 \end{pmatrix}$, or to $\mathbf{d} = \begin{pmatrix} 12 \\ 18 \\ -14 \end{pmatrix}$? Parallel means one vector is a scalar multiple of the other, so look for a single k that works in all three coordinates. For $\mathbf{c}$, the first coordinate forces $k = -2$, and that value holds throughout:

$$-2\begin{pmatrix} 4 \\ 6 \\ -5 \end{pmatrix} = \begin{pmatrix} -8 \\ -12 \\ 10 \end{pmatrix} = \mathbf{c},$$

so $\mathbf{c}$ is parallel to $2\mathbf{a}-\mathbf{b}$, pointing the opposite way because k is negative. For $\mathbf{d}$, the first two coordinates both give $k=3$, but $3\times(-5)=-15$, and the third coordinate of $\mathbf{d}$ is -14. No single k works, so $\mathbf{d}$ is not parallel to $2\mathbf{a}-\mathbf{b}$. Test all three coordinates before answering. If two coordinates give the same k and the third does not, the vectors are not parallel.

9.1.2 Magnitude and unit vectors

Three words are used for the same quantity. The **magnitude** of a vector, its **length**, and its **modulus** all mean the size of the arrow, and all are written $|\mathbf{v}|$. These notes use whichever reads more naturally in the sentence. The value comes straight from Pythagoras, extended into three dimensions.

KEY · IB The magnitude of $\mathbf{v}=\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}$ is $|\mathbf{v}|=\sqrt{v_1^2+v_2^2+v_3^2}$.

Dividing a vector by its own length leaves the direction untouched but sets the length to 1. The result is a **unit vector**, written $\hat{\mathbf{v}}$, the pure direction stripped of any magnitude, ready to be scaled to any length you like.

KEY $\hat{\mathbf{v}}=\dfrac{\mathbf{v}}{|\mathbf{v}|}$ is the unit vector in the direction of $\mathbf{v}$.

The base vectors are unit vectors, which is what their hats record: $|\hat{\mathbf{i}}|=|\hat{\mathbf{j}}|=|\hat{\mathbf{k}}|=1$, one unit along each axis. So writing $\mathbf{a}=a_1\hat{\mathbf{i}}+a_2\hat{\mathbf{j}}+a_3\hat{\mathbf{k}}$ says that $\mathbf{a}$ is built by taking a_1 steps of length 1 along x, then a_2 along y, then a_3 along z.

EXAMPLE 9.1

Find a vector of magnitude 14 in the direction of $\mathbf{v}=\begin{pmatrix}2\\3\\6\end{pmatrix}$.

First the length: $|\mathbf{v}|=\sqrt{2^2+3^2+6^2}=\sqrt{49}=7$. The unit vector is $\hat{\mathbf{v}}=\frac{1}{7}\mathbf{v}$, and scaling it to length 14 means multiplying by 14:

$$14\hat{\mathbf{v}}=\frac{14}{7}\begin{pmatrix}2\\3\\6\end{pmatrix}=\begin{pmatrix}4\\6\\12\end{pmatrix}.$$

Reduce to a unit vector first, then rescale. Those two steps answer every question of the form "find a vector of magnitude m in the direction of".

YOUR TURN 9.1 Vectors: Magnitude and Direction

Magnitude and combining vectors

1. Find the magnitude of each vector.

(a) $\begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}$ (b) $\begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$

(c) $\begin{pmatrix} 6 \\ -2 \\ 3 \end{pmatrix}$ (d) $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

2. Given $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix}$, evaluate each of the following.

(a) $\mathbf{a} + \mathbf{b}$ (b) $\mathbf{a} - \mathbf{b}$

(c) $\mathbf{a} + 2\mathbf{b}$ (d) $3\mathbf{a} - \mathbf{b}$

Base vectors

3. Work with the base vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$.

(a) Write $3\hat{\mathbf{i}} - 2\hat{\mathbf{j}} + 5\hat{\mathbf{k}}$ as a column vector.

(b) Write $\begin{pmatrix} 4 \\ 0 \\ -1 \end{pmatrix}$ in terms of $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$.

(c) Given $\mathbf{a} = 2\hat{\mathbf{i}} - \hat{\mathbf{j}} + 2\hat{\mathbf{k}}$ and $\mathbf{b} = \hat{\mathbf{i}} + 3\hat{\mathbf{k}}$, find $2\mathbf{a} - \mathbf{b}$ in terms of $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$, and find its magnitude.

Unit vectors and scaling

4. Unit vectors and scaling.

(a) Find the unit vector in the direction of $\begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix}$.

(b) Find the unit vector in the direction of $\begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$.

(c) Find a vector of magnitude 6 in the direction of $\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$.

(d) Explain why $\dfrac{\mathbf{v}}{|\mathbf{v}|}$ has magnitude 1 for every non-zero vector $\mathbf{v}$.

Parallel vectors

5. Decide whether each pair of vectors is parallel, and justify your answer.

(a) $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$ (b) $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix}$

(c) $\begin{pmatrix} 3 \\ -6 \\ 9 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 2 \\ -3 \end{pmatrix}$

(d) Find p and q so that $\begin{pmatrix} 2 \\ p \\ 6 \end{pmatrix}$ and $\begin{pmatrix} -3 \\ 6 \\ q \end{pmatrix}$ are parallel.

Concept check

6. A student says: "$|\mathbf{a} + \mathbf{b}| = |\mathbf{a}| + |\mathbf{b}|$."

(a) Give a counter-example.

(b) State when the claim is true.

9.2 Position Vectors and Geometry

Fix the origin O. The **position vector** of a point A is the arrow $\overrightarrow{OA}$ from the origin to A; its components are simply the coordinates of A. Position vectors connect the algebra of vectors to the fixed points of a geometry problem.

The arrow from one point to another is found by subtraction. Much of the chapter uses this fact.

> KEY $\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$, the position vector of the head minus that of the tail.

The distance between A and B is then the magnitude of that displacement, $|\overrightarrow{AB}| = |\mathbf{b} - \mathbf{a}|$, and the midpoint of AB has position vector $\frac{1}{2}(\mathbf{a} + \mathbf{b})$.

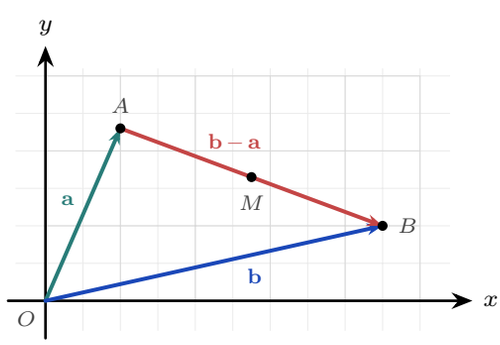


Figure 9.2 Head minus tail. The two position vectors run from the origin; the arrow between the points closes the triangle, so $\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$. The midpoint M sits at the average of the two position vectors, $\frac{1}{2}(\mathbf{a} + \mathbf{b})$.

Worked through. For the points $A(2, -3, 1)$ and $B(5, 1, 6)$,

$$\overrightarrow{AB} = \mathbf{b} - \mathbf{a} = \begin{pmatrix} 5 \\ 1 \\ 6 \end{pmatrix} - \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}, \qquad AB = \sqrt{3^2 + 4^2 + 5^2} = \sqrt{50} = 5\sqrt{2}.$$

Head minus tail gives the arrow; the magnitude of that arrow is the distance. Every distance between two points reduces to those two steps.

Three points A, B, C are **collinear** (on one straight line) when $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are parallel, that is, one is a scalar multiple of the other. This is the vector version of "same direction from a shared point", and the equation of a line in the next section is built from it.

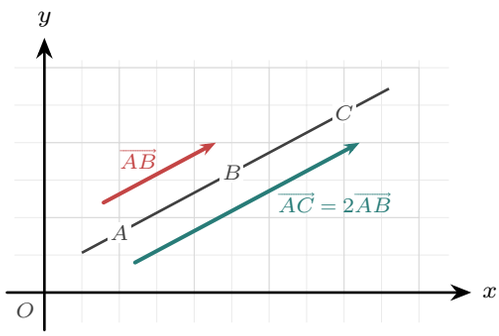


Figure 9.3 Collinear points. Both displacements start at A, and one is a scalar multiple of the other, so the three points lie on a single line. The two arrows are drawn slightly to either side of that line so they can be told apart.

9.2.1 Proving geometry with vectors

Vectors give a way of proving geometric facts by algebra, with no diagram-chasing and no coordinates. The method has three steps.

KEY **Proving a geometric property with vectors.**

1. Choose one point as the origin, and name the position vectors of the others with single letters.
2. Write every point in the figure in terms of those letters.
3. Show the required relation by algebra.

Choosing the origin well is what keeps the algebra short. Put it at a vertex that many edges meet.

EXAMPLE 9.2

Prove that the diagonals of a parallelogram bisect each other.

Let the parallelogram be $OABC$, with the origin at O. Write $\overrightarrow{OA} = \mathbf{a}$ and $\overrightarrow{OC} = \mathbf{c}$.

Because $OABC$ is a parallelogram, $\overrightarrow{CB} = \overrightarrow{OA} = \mathbf{a}$, so the fourth vertex is

$$\overrightarrow{OB} = \mathbf{a} + \mathbf{c}.$$

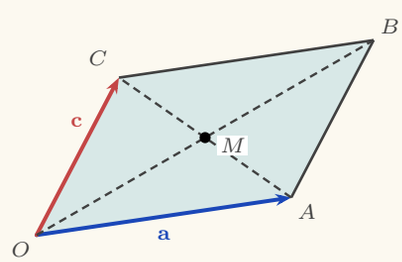


Figure 9.4 The parallelogram $OABC$. The two dashed diagonals meet at the point M, and the proof below shows that M is the midpoint of both.

The two diagonals are OB and AC. Find the midpoint of each.

Midpoint of OB: $\frac{1}{2}(\mathbf{0}+\mathbf{a}+\mathbf{c}) = \frac{1}{2}(\mathbf{a}+\mathbf{c})$.
Midpoint of AC: $\frac{1}{2}(\mathbf{a}+\mathbf{c})$.
The two midpoints are the same point. A single point lying halfway along both diagonals is exactly what "bisect each other" means. ■
Notice that no coordinates were used, and the result holds for every parallelogram, in a plane or in space.

EXAMPLE 9.3

M and N are the midpoints of the sides OA and OB of a triangle. Prove that MN is parallel to AB and half its length.
Take the origin at O, with $\overrightarrow{OA} = \mathbf{a}$ and $\overrightarrow{OB} = \mathbf{b}$. The midpoints are then

$$\overrightarrow{OM} = \tfrac{1}{2}\mathbf{a}, \qquad \overrightarrow{ON} = \tfrac{1}{2}\mathbf{b}.$$

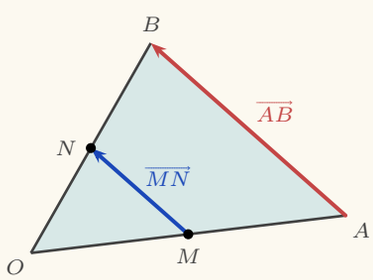


Figure 9.5 The midpoint theorem. Joining the midpoints of two sides gives a segment parallel to the third side and half as long.

Now find the two displacement vectors, each as head minus tail:

$$\overrightarrow{MN} = \tfrac{1}{2}\mathbf{b} - \tfrac{1}{2}\mathbf{a} = \tfrac{1}{2}(\mathbf{b}-\mathbf{a}), \qquad \overrightarrow{AB} = \mathbf{b}-\mathbf{a}.$$

So $\overrightarrow{MN} = \frac{1}{2}\overrightarrow{AB}$.
One vector is a scalar multiple of the other, so the two segments are parallel, and the scalar is $\frac{1}{2}$, so MN is half the length of AB. ■
That is the entire proof. This result is known as the midpoint theorem.

YOUR TURN 9.2 Position Vectors and Geometry

Vectors between points

7. Points $A(1, 2, -1)$ and $B(4, 0, 3)$ are given.

(a) Find $\overrightarrow{AB}$.
(b) Find the distance AB.
(c) Find the midpoint of AB.
(d) Find $\overrightarrow{BA}$.

8. Further position-vector work.

(a) A is $(1, 1, 1)$ and $\overrightarrow{AB} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$. Find the coordinates of B.
(b) B is the midpoint of AC, with $A(2, 1, 3)$ and $B(4, 5, 3)$. Find C.
(c) Show that $A(1, 1, 1)$, $B(2, 3, 5)$ and $C(4, 7, 13)$ are collinear.

Proving geometry with vectors

9. $OABC$ is any quadrilateral. Take the origin at O and write $\overrightarrow{OA} = \mathbf{a}$, $\overrightarrow{OB} = \mathbf{b}$, $\overrightarrow{OC} = \mathbf{c}$. The points P, Q, R and S are the midpoints of $[OA]$, $[AB]$, $[BC]$ and $[CO]$.

(a) Write the position vectors of P, Q, R and S in terms of $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$.
(b) Find $\overrightarrow{PQ}$ and $\overrightarrow{SR}$.
(c) Hence prove that $PQRS$ is a parallelogram.

10. In triangle OAB, take the origin at O and write $\overrightarrow{OA} = \mathbf{a}$, $\overrightarrow{OB} = \mathbf{b}$. M is the midpoint of $[AB]$, N is the midpoint of $[OA]$, and G is the point on $[OM]$ with $OG = \frac{2}{3}OM$.

(a) Show that $\overrightarrow{OG} = \frac{1}{3}(\mathbf{a} + \mathbf{b})$.
(b) Find $\overrightarrow{BG}$ and $\overrightarrow{BN}$, and hence prove that B, G and N are collinear.

Concept check

11. For $A(1, 2, -1)$ and $B(4, 0, 3)$, a student writes $\overrightarrow{AB} = \mathbf{a} - \mathbf{b} = \begin{pmatrix} -3 \\ 2 \\ -4 \end{pmatrix}$. Identify the error and find $\overrightarrow{AB}$.

9.3 The Scalar Product

There are two ways to multiply vectors, and they answer two different questions. The first, the **scalar product** (or dot product), takes two vectors and returns a number, and that number carries the angle between them. It has two forms, and they always agree.

KEY · IB **Scalar product.** For $\mathbf{a}, \mathbf{b}$ with angle θ between them,

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 \qquad \text{and} \qquad \mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \cos\theta.$$

In the second form, the $\cos\theta$ is not decoration. Ordinary multiplication assumes both quantities are measured along the same line: 3×4 makes sense because both numbers sit on one number line. Two vectors usually point different ways, so multiplying their lengths would measure nothing in particular.

The repair is to bring them onto the same line first. Keep $\mathbf{a}$ as it is, and use only the part of $\mathbf{b}$ that lies along $\mathbf{a}$: the **projection** of $\mathbf{b}$ onto $\mathbf{a}$, found by dropping a perpendicular from the tip of $\mathbf{b}$, which has length $|\mathbf{b}| \cos\theta$. Multiplying those two gives

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

the length of $\mathbf{a}$ times how much of $\mathbf{b}$ goes in $\mathbf{a}$'s direction. Doing it the other way round, projecting $\mathbf{a}$ onto $\mathbf{b}$, gives $|\mathbf{b}| \times |\mathbf{a}| \cos\theta$, the same number. That is why the order does not matter.

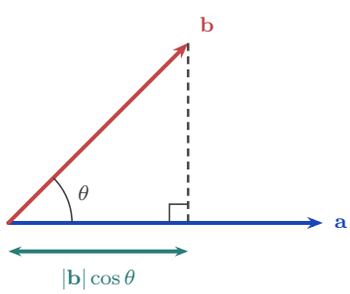


Figure 9.6 The projection of $\mathbf{b}$ onto $\mathbf{a}$. Dropping the tip of $\mathbf{b}$ perpendicularly onto $\mathbf{a}$ leaves a length of $|\mathbf{b}| \cos\theta$, and the scalar product multiplies that by $|\mathbf{a}|$.

The scalar product then obeys the rules you would expect of a multiplication.

- **Commutative:** $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$. The order does not matter.
- **Distributive:** $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$. Brackets expand as usual.
- **Scalars move outside:** $(k\mathbf{a}) \cdot \mathbf{b} = k(\mathbf{a} \cdot \mathbf{b})$. A constant factor can be taken out.
- **Dotted with itself:** $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$, since $\theta = 0$ and $\cos 0 = 1$. The result is the squared length, and several proofs in this chapter use it.

The base vectors show both extremes at once. Each has length 1 and lies along its own axis, so dotting one with itself gives $1 \times 1 \times \cos 0 = 1$. Different axes are at right angles, so dotting two different base vectors gives 0.

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

This is exactly why the component form works. Expanding $(a_1\hat{\mathbf{i}} + a_2\hat{\mathbf{j}} + a_3\hat{\mathbf{k}}) \cdot (b_1\hat{\mathbf{i}} + b_2\hat{\mathbf{j}} + b_3\hat{\mathbf{k}})$ produces nine terms, and every term pairing two *different* base vectors is zero, because of that zero dot product. Only the three matching terms survive, leaving $a_1b_1 + a_2b_2 + a_3b_3$.

The component form is easy to compute; the cosine form carries the geometry. Setting the two equal is what lets you find the angle between any two directions, and it is used throughout the chapter.

KEY · IB $\cos\theta = \dfrac{a_1b_1 + a_2b_2 + a_3b_3}{|\mathbf{a}|\,|\mathbf{b}|}$.

Worked through. To find the angle between $\mathbf{a} = \begin{pmatrix} 3 \\ 2 \\ 6 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 2 \\ 6 \\ 3 \end{pmatrix}$, compute the three ingredients: $\mathbf{a} \cdot \mathbf{b} = 6 + 12 + 18 = 36$, $|\mathbf{a}| = \sqrt{9 + 4 + 36} = 7$, $|\mathbf{b}| = 7$. Then

$$\cos\theta = \frac{36}{49} \implies \theta = \arccos\frac{36}{49} \approx 42.7^\circ.$$

Dot product over product of lengths: the recipe never changes, whether the vectors come from a triangle, a pair of lines, or a pair of planes.

A special case used often is a right angle. Since $\cos 90^\circ = 0$, the scalar product of perpendicular vectors is zero, and the converse holds too. This gives a test for perpendicularity that needs no angles at all, just a multiply-and-add.

KEY $\mathbf{a} \perp \mathbf{b} \iff \mathbf{a} \cdot \mathbf{b} = 0$ (for nonzero $\mathbf{a}, \mathbf{b}$).

Parallel vectors sit at the opposite extreme. There θ is 0 or π, so $\cos\theta$ is $+1$ or -1 and the scalar product reaches the largest size it can. Taking the modulus covers both directions at once.

KEY For nonzero $\mathbf{a}, \mathbf{b}$: $\mathbf{a} \parallel \mathbf{b} \iff |\mathbf{a} \cdot \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|$.

So the scalar product recognises both extremes. It is zero when the vectors are perpendicular, and its size is as large as possible when they are parallel. Every angle in between falls somewhere on that scale.

EXAMPLE 9.4

Find the angle between $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$.

The dot product is $\mathbf{a} \cdot \mathbf{b} = (1)(2) + (2)(0) + (2)(-1) = 0$. Because it vanishes, the vectors are perpendicular: $\theta = 90^\circ$. Neither magnitude was needed: a zero dot product settles the angle at once.

YOUR TURN 9.3 The Scalar Product

Calculating the scalar product

12. Evaluate each scalar product.

(a) $\begin{pmatrix}1\\2\\3\end{pmatrix}\cdot\begin{pmatrix}4\\-1\\2\end{pmatrix}$ (b) $\begin{pmatrix}1\\0\\-1\end{pmatrix}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}$

(c) $\begin{pmatrix}2\\1\\-1\end{pmatrix}\cdot\begin{pmatrix}1\\3\\1\end{pmatrix}$ (d) $\begin{pmatrix}3\\4\end{pmatrix}\cdot\begin{pmatrix}-2\\5\end{pmatrix}$

13. Use the definition $\mathbf{a}\cdot\mathbf{b}=|\mathbf{a}|\,|\mathbf{b}|\cos\theta$.

(a) Given $|\mathbf{a}|=3$, $|\mathbf{b}|=4$ and an angle of 60° between them, find $\mathbf{a}\cdot\mathbf{b}$.

(b) Given $|\mathbf{a}|=2$, $|\mathbf{b}|=5$ and an angle of 120° between them, find $\mathbf{a}\cdot\mathbf{b}$.

(c) Explain why $\mathbf{a}\cdot\mathbf{a}=|\mathbf{a}|^2$ for every vector $\mathbf{a}$.

(d) Given $|\mathbf{a}|=3$, $|\mathbf{b}|=5$ and $\mathbf{a}\cdot\mathbf{b}=-15$, state what this tells you about the directions of $\mathbf{a}$ and $\mathbf{b}$, and justify your answer.

Angles between vectors

14. GDC Find the angle between each pair of vectors.

(a) $\begin{pmatrix}1\\0\\0\end{pmatrix}$ and $\begin{pmatrix}1\\1\\0\end{pmatrix}$ (b) $\begin{pmatrix}2\\1\\2\end{pmatrix}$ and $\begin{pmatrix}1\\2\\2\end{pmatrix}$

(c) $\begin{pmatrix}1\\1\\0\end{pmatrix}$ and $\begin{pmatrix}1\\0\\1\end{pmatrix}$ (d) $\begin{pmatrix}3\\0\\4\end{pmatrix}$ and $\begin{pmatrix}-2\\1\\-2\end{pmatrix}$

Perpendicular vectors

15. Use the scalar product to find k.

(a) Find k so that $\begin{pmatrix}2\\k\\1\end{pmatrix}\cdot\begin{pmatrix}3\\-1\\2\end{pmatrix}=0$.

(b) Find the possible values of k for which $\begin{pmatrix}k\\1\\2\end{pmatrix}$ is perpendicular to $\begin{pmatrix}k\\-5\\2\end{pmatrix}$.

Proofs with the scalar product

16. A rhombus has sides $\mathbf{a}$ and $\mathbf{b}$ drawn from one vertex, so $|\mathbf{a}|=|\mathbf{b}|$.

(a) Write the two diagonals in terms of $\mathbf{a}$ and $\mathbf{b}$.

(b) Expand $(\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}-\mathbf{b})$.

(c) Hence prove that the diagonals of a rhombus are perpendicular.

(d) Verify your result for $\mathbf{a}=\begin{pmatrix}2\\1\\2\end{pmatrix}$ and $\mathbf{b}=\begin{pmatrix}1\\2\\2\end{pmatrix}$.

Concept check

17. For $A(3,1,0)$, $B(1,1,0)$ and $C(2,2,0)$, a student finds angle $A\hat{B}C$ using $\overrightarrow{AB}=\begin{pmatrix}-2\\0\\0\end{pmatrix}$ and $\overrightarrow{BC}=\begin{pmatrix}1\\1\\0\end{pmatrix}$, and obtains 135°. Identify the error and find the correct angle.

9.4 The Vector Equation of a Line

A line is fixed by two pieces of information: one point it passes through, and a direction it runs along. Start at a known point $\mathbf{a}$ and move any distance along a direction $\mathbf{b}$. As the **parameter** λ takes every real value, this reaches every point on the line.

KEY · IB **Vector equation of a line:** $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, where $\mathbf{a}$ is the position vector of a point on the line and $\mathbf{b}$ is a direction vector.

This differs from the equation of a line met earlier, and the difference matters. An equation such as $y = 2x + 1$ is a condition: a point either satisfies it or does not. The vector equation produces points: each value of λ gives one point of the line. That is why the vector form works in three dimensions, where a single equation in x, y and z describes a plane, not a line.

Writing $\mathbf{r} = (x, y, z)$ and reading off each coordinate gives the **parametric form**; eliminating λ between them gives the **Cartesian form**.

KEY · IB **Parametric:** $x = x_0 + \lambda l, \quad y = y_0 + \lambda m, \quad z = z_0 + \lambda n.$

Cartesian: $\dfrac{x - x_0}{l} = \dfrac{y - y_0}{m} = \dfrac{z - z_0}{n}$, where $\mathbf{b} = \begin{pmatrix} l \\ m \\ n \end{pmatrix}$.

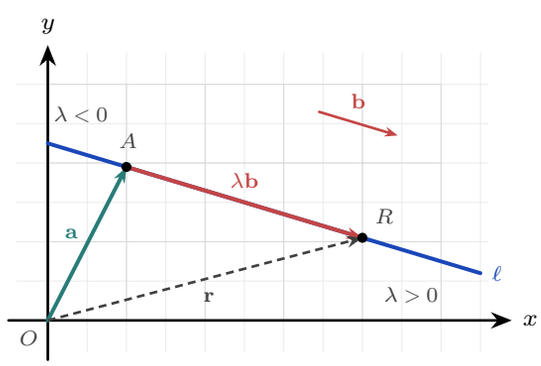


Figure 9.7 A line built from a point and a direction. The vector $\mathbf{a}$ reaches the known point A; travelling $\lambda\mathbf{b}$ from there lands on a general point R, whose position vector $\mathbf{r}$ closes the triangle. Negative λ runs back the other way, so the whole line is covered. A vector can be drawn anywhere, so the copy of $\mathbf{b}$ drawn above is the same vector.

The direction vector carries the geometry of how two lines relate. The **angle between two lines** is the angle between their direction vectors, found by the scalar product (take the acute value). Two lines are **parallel** when their directions are scalar multiples. And in three dimensions there is a possibility with no two-dimensional analogue: lines that are neither parallel nor crossing.

Two lines in space fall into exactly one of four cases. Two questions settle which: are the directions parallel, and do the lines share a point?

- **Coincident.** Directions parallel, every point shared. The same line written two ways.
- **Parallel.** Directions parallel, no point shared. Different lines running the same way.
- **Intersecting.** Directions not parallel, exactly one point shared. They cross once.
- **Skew.** Directions not parallel, no point shared. They pass without meeting, like an overpass above a road.

Only the last case is new. Two lines in a plane must either be parallel or cross, but in three dimensions they can miss each other entirely.

KEY **Classifying two lines.**

1. Compare the directions. If they are scalar multiples, check whether a point of one lies on the other: yes means **coincident**, no means **parallel**.
2. If the directions differ, set the two equations equal and solve for the parameters. Two components fix them; the third decides.
3. A consistent solution means the lines **intersect**, and substituting back gives the point of intersection. An inconsistent system means they are **skew**.

Worked through. Do the lines $\mathbf{r}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $\mathbf{r}_2 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$ intersect?

The directions are not multiples of each other, so the lines are not parallel. Setting components equal:

$$2 + \lambda = 1 + 2\mu, \qquad 1 + \lambda = 3 - \mu, \qquad \lambda = 2 + \mu.$$

From the first two, $\lambda - 2\mu = -1$ and $\lambda + \mu = 2$, so $\mu = 1$, $\lambda = 1$. Check in the third: $\lambda = 1$ and $2 + 1 = 3$. Since $1 \neq 3$, the system is inconsistent, so the lines are **skew**. Two equations fix the parameters; the third decides whether the lines really meet.

9.4.1 Lines in motion: kinematics

So far the parameter has had no meaning attached to it. Give it one and the line equation describes a moving object. Read λ as *time* t, and $\mathbf{r} = \mathbf{a} + t\mathbf{b}$ says: start at $\mathbf{a}$, and move with constant **velocity** $\mathbf{b}$. The direction of $\mathbf{b}$ is the heading; its magnitude $|\mathbf{b}|$ is the **speed**. This is the pilot from the opening page: the velocity is the tip-to-tail sum of the aircraft's velocity through the air and the wind velocity.

EXAMPLE 9.5

An aircraft passes the point $(20, 10)$ (coordinates in km) at $t = 0$ and flies with constant

velocity $\begin{pmatrix}240\\100\end{pmatrix}$ km h^{-1}. Find its speed, its position at $t=2$ hours, and whether its path passes through the airport at $(500, 210)$.

The speed is the magnitude of the velocity: $\sqrt{240^2+100^2}=260$ km h^{-1}. The position at time t is

$$\mathbf{r}=\begin{pmatrix}20\\10\end{pmatrix}+t\begin{pmatrix}240\\100\end{pmatrix}, \qquad \text{so at } t=2: \quad \mathbf{r}=\begin{pmatrix}500\\210\end{pmatrix}.$$

That *is* the airport, reached at exactly $t=2$. To test any point, solve the first coordinate for t, then check the others. One value of t must satisfy every equation at once. This is the same test used for intersecting lines.

For two objects moving at different constant velocities, the distance between them usually falls, reaches a least value, and then rises again. The moment of **closest approach** is found from the vector between them. Subtracting the position vectors gives $\mathbf{r}_B-\mathbf{r}_A$, whose components are linear in t, so the squared distance

$$d^2=|\mathbf{r}_B-\mathbf{r}_A|^2$$

is a quadratic in t. Completing the square gives its least value and the time at which that value occurs.

Work with d^2 rather than with d. Squaring clears the square root, and the two are least at the same instant, since squaring preserves order among non-negative numbers.

EXAMPLE 9.6

Two birds fly in straight lines at constant velocity. Taking t in seconds and distances in metres, their position vectors are

$$\mathbf{r}_A=\begin{pmatrix}8\\5\\1\end{pmatrix}+t\begin{pmatrix}1\\-1\\3\end{pmatrix}, \qquad \mathbf{r}_B=\begin{pmatrix}2\\3\\3\end{pmatrix}+t\begin{pmatrix}3\\1\\2\end{pmatrix}.$$

Find the time at which they are closest, and the least distance between them.

Subtract, component by component:

$$\mathbf{r}_B-\mathbf{r}_A=\begin{pmatrix}2-8\\3-5\\3-1\end{pmatrix}+t\begin{pmatrix}3-1\\1-(-1)\\2-3\end{pmatrix}=\begin{pmatrix}2t-6\\2t-2\\2-t\end{pmatrix}.$$

Its squared length is

$$d^2=(2t-6)^2+(2t-2)^2+(2-t)^2=(4t^2-24t+36)+(4t^2-8t+4)+(t^2-4t+4)=9t^2-36t+44.$$

Complete the square:

$$d^2 = 9\left(t^2 - 4t\right) + 44 = 9(t-2)^2 - 36 + 44 = 9(t-2)^2 + 8.$$

The squared bracket is never negative and is zero only at $t = 2$, so the least value of d^2 is 8, reached at $t = 2$ seconds. The least distance is

$$d = \sqrt{8} = 2\sqrt{2} \approx 2.83 \text{ m}.$$

Check either side: at $t = 1$ and at $t = 3$ the formula gives $d^2 = 17$ both times, larger than 8 and equal to each other, which is what a parabola symmetric about $t = 2$ requires.

YOUR TURN 9.4 The Vector Equation of a Line

Writing the equation of a line

18. Writing the equation of a line.

(a) Write a vector equation of the line through $A(1, 2, 3)$ with direction $\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$.

(b) Find a direction vector of the line through $A(1, 0, 2)$ and $B(4, 3, 5)$.

(c) Write a vector equation of that line.

19. Forms of the equation.

(a) Write $\mathbf{r} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$ in parametric form.

(b) Write $\mathbf{r} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$ in Cartesian form.

(c) Does $(3, 3, 2)$ lie on $\mathbf{r} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$? Justify your answer.

Classifying two lines

20. Decide whether each pair of lines is coincident, parallel, intersecting or skew. Where the lines intersect, find the point of intersection.

(a) $\mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ and $\mathbf{r} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} -4 \\ 2 \\ -6 \end{pmatrix}$

(b) $\mathbf{r} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ and $\mathbf{r} = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ -2 \\ 4 \end{pmatrix}$

(c) $\mathbf{r} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ and $\mathbf{r} = \begin{pmatrix} 0 \\ 5 \\ -3 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$

(d) $\mathbf{r} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ and $\mathbf{r} = \begin{pmatrix} 0 \\ 5 \\ -2 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$

Angle between two lines

21. GDC Find the acute angle between each pair of lines.

(a) $\mathbf{r} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ and $\mathbf{r} = \begin{pmatrix} 0 \\ 5 \\ -3 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$

(b) $\mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and the x-axis, $\mathbf{r} = \mu \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

Lines in motion

22. A boat leaves the point $(2, -1)$, in km, at $t = 0$ and moves with constant velocity $\binom{6}{8}$ km h^{-1}.

(a) Find the speed of the boat.
(b) Write down its position vector at time t.
(c) Find its position after 30 minutes.
(d) A lighthouse stands at $(20, 15)$. Show that the boat does not pass through it.

23. Two drones fly in straight lines at constant velocity. With t in seconds and distances in metres, their position vectors are

$$\mathbf{r}_A = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + t \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \qquad \mathbf{r}_B = \begin{pmatrix} 7 \\ 0 \\ 5 \end{pmatrix} + t \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}.$$

(a) Show that $\mathbf{r}_B - \mathbf{r}_A = \begin{pmatrix} 6-3t \\ t-2 \\ 5 \end{pmatrix}$.
(b) Show that the square of the distance between the drones is $d^2 = 10t^2 - 40t + 65$.
(c) Find the time at which the drones are closest, and their least distance apart.

Concept check

24. A student says: "*the direction vectors of the two lines are not parallel, so the lines intersect.*" State when this is false, and describe the check that decides.

9.5 The Vector Product

The second way to multiply vectors returns not a number but a *vector*. The **vector product** (or cross product) $\mathbf{a} \times \mathbf{b}$ points perpendicular to *both* of the vectors it is built from, and its length measures the area they span. Where the dot product was about alignment, the cross product is about the plane the two vectors sweep out.

KEY · IB **Vector product.** For $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$,

$$\mathbf{a} \times \mathbf{b} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}, \qquad |\mathbf{a} \times \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}| \sin\theta,$$

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

The magnitude and the direction can be combined into one statement,

$$\mathbf{a} \times \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \sin\theta\, \hat{\mathbf{n}},$$

where $\hat{\mathbf{n}}$ is the unit vector perpendicular to both $\mathbf{a}$ and $\mathbf{b}$. Taking magnitudes gives the second formula in the box, since $|\hat{\mathbf{n}}| = 1$.

The direction is fixed by the **right-hand screw rule**: turn a screw from $\mathbf{a}$ towards $\mathbf{b}$ and it advances along $\mathbf{a} \times \mathbf{b}$. Four properties are used constantly.

- **Perpendicular to both.** The result is at right angles to the plane of $\mathbf{a}$ and $\mathbf{b}$. This is what makes it the tool for building normal vectors in the next section.
- **Not commutative.** It is **anti-commutative**: $\mathbf{b} \times \mathbf{a} = -(\mathbf{a} \times \mathbf{b})$. The order matters, and reversing it reverses the answer.
- **Distributive, and scalars come out.** $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$ and $(k\mathbf{v}) \times \mathbf{w} = k(\mathbf{v} \times \mathbf{w})$.
- **Zero for parallel vectors.** If $\mathbf{a}$ and $\mathbf{b}$ are parallel then $\sin\theta = 0$, so $\mathbf{a} \times \mathbf{b} = \mathbf{0}$. In particular $\mathbf{v} \times \mathbf{v} = \mathbf{0}$, the counterpart of $\mathbf{v} \cdot \mathbf{v} = |\mathbf{v}|^2$.

The base vectors again show the pattern. Each is perpendicular to the other two, so crossing two different ones gives a unit vector along the third axis, and the screw rule fixes which way it points. Crossing one with itself gives $\mathbf{0}$, since the angle is 0 and $\sin 0 = 0$.

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

Read the first three in the cyclic order $\hat{\mathbf{i}} \to \hat{\mathbf{j}} \to \hat{\mathbf{k}} \to \hat{\mathbf{i}}$: going forwards gives a plus, and going backwards gives a minus, so $\hat{\mathbf{j}} \times \hat{\mathbf{i}} = -\hat{\mathbf{k}}$.

9.5.1 One array, two products

Both products come out of the same nine numbers. Multiply every component of $\mathbf{a}$ by every component of $\mathbf{b}$ and lay the results in a square.

	b_1	b_2	b_3
a_1	a_1b_1	a_1b_2	a_1b_3
a_2	a_2b_1	a_2b_2	a_2b_3
a_3	a_3b_1	a_3b_2	a_3b_3

The **diagonal** is the scalar product. Those are the three terms where a base vector meets itself, and $\hat{\mathbf{i}} \cdot \hat{\mathbf{i}} = 1$ keeps them; everything off the diagonal is removed by $\hat{\mathbf{i}} \cdot \hat{\mathbf{j}} = 0$. Adding the shaded entries gives $a_1b_1 + a_2b_2 + a_3b_3$.

The **off-diagonal** entries are the vector product, taken in opposite pairs. Here the diagonal

vanishes, since $\hat{\mathbf{i}} \times \hat{\mathbf{i}} = \mathbf{0}$, and each pair of mirror-image entries subtracts to give one component:

$$\underbrace{a_2b_3 - a_3b_2}_{\text{the pair missing } 1}, \qquad \underbrace{a_3b_1 - a_1b_3}_{\text{the pair missing } 2}, \qquad \underbrace{a_1b_2 - a_2b_1}_{\text{the pair missing } 3}.$$

So the component of $\mathbf{a} \times \mathbf{b}$ in a given direction is built from the two rows and columns that leave that direction out. That the result is perpendicular to both vectors is checked by the scalar product: $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b})$ expands to six terms that cancel in pairs. The two products are the same array read two ways: down the diagonal, or across it.

You need not write the array out. For each component, cover its own row in both vectors and multiply the remaining four numbers crosswise. The subtraction then runs one way for the outer two components and the other way for the middle one:

$$\mathbf{a} \times \mathbf{b} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}.$$

Only the middle component reverses. Covering row 1 leaves rows 2 and 3, and covering row 3 leaves rows 1 and 2; both pairs are in the cyclic order $1 \to 2 \to 3 \to 1$. Covering row 2 leaves rows 1 and 3, which run against that order, so the two products swap places before subtracting. Taking all three the same way is a common error here, and the vector it produces is usually not perpendicular to both $\mathbf{a}$ and $\mathbf{b}$, so a dot-product check will usually catch it. A cross product built correctly answers "find a vector perpendicular to both" immediately; a *unit* vector perpendicular to both needs one further step.

Worked through. For $\mathbf{a} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$, each component of $\mathbf{a} \times \mathbf{b}$ leaves out its own row, and the middle one subtracts the other way round:

$$\mathbf{a} \times \mathbf{b} = \begin{pmatrix} (1)(1) - (0)(2) \\ (0)(2) - (2)(1) \\ (2)(2) - (1)(2) \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}.$$

The middle entry is $a_3b_1 - a_1b_3$, not $a_1b_3 - a_3b_1$; reading it in the same direction as the other two would give $+2$ there instead of -2. Check it against both vectors: $(1)(2) + (-2)(1) + (2)(0) = 0$ and $(1)(2) + (-2)(2) + (2)(1) = 0$. The rejected version fails this test, since $(1)(2) + (2)(1) + (2)(0) = 4$. Perpendicular is not the same as unit. For a unit vector perpendicular to both, note that the magnitude is $\sqrt{1 + 4 + 4} = 3$, so divide by it:

$$\hat{\mathbf{n}} = \frac{1}{3} \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}.$$

The vector $-\hat{\mathbf{n}} = \frac{1}{3}\begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix}$ is equally correct. It is $\mathbf{b} \times \mathbf{a}$ divided by its magnitude, and it runs the opposite way along the same perpendicular line, so unless the question fixes a direction either answer is complete.

9.5.2 Area of a parallelogram and a triangle

The magnitude has a clean geometric meaning: it is the area of the parallelogram with $\mathbf{a}$ and $\mathbf{b}$ as sides. To see why, read the two factors of $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}| \sin\theta$ separately. Take $|\mathbf{a}|$ as the base. Then $|\mathbf{b}| \sin\theta$ is the perpendicular height, because $\mathbf{b}$ is the slanted side and $\sin\theta$ keeps only the part of it at right angles to the base. So the magnitude is base $\times$ height, which is the area.

KEY · IB Area of a parallelogram $= |\mathbf{a} \times \mathbf{b}|$, where $\mathbf{a}$ and $\mathbf{b}$ are two adjacent sides.

A triangle is half of such a parallelogram, so its area is $\frac{1}{2}|\mathbf{a} \times \mathbf{b}|$; remember the half, because the booklet prints only the parallelogram.

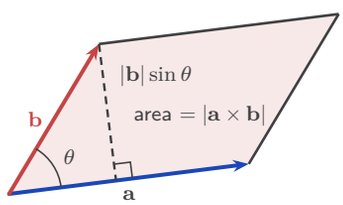


Figure 9.8 The parallelogram spanned by $\mathbf{a}$ and $\mathbf{b}$. Its area is $|\mathbf{a} \times \mathbf{b}|$, and a triangle on the same two sides has half that area.

EXAMPLE 9.7

Find the area of the triangle with vertices $A(0,0,0)$, $B(3,1,0)$, $C(1,0,2)$.

Take $\overrightarrow{AB} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$ and $\overrightarrow{AC} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$. Their cross product is

$$\overrightarrow{AB} \times \overrightarrow{AC} = \begin{pmatrix} (1)(2) - (0)(0) \\ (0)(1) - (3)(2) \\ (3)(0) - (1)(1) \end{pmatrix} = \begin{pmatrix} 2 \\ -6 \\ -1 \end{pmatrix}.$$

Its magnitude is $\sqrt{4 + 36 + 1} = \sqrt{41}$, so the triangle's area is $\frac{1}{2}\sqrt{41}$. The cross product avoids trigonometry here: one calculation and one square root give the area.

YOUR TURN 9.5 The Vector Product

Calculating the vector product

25. Evaluate each vector product.

(a) $\begin{pmatrix}1\\0\\0\end{pmatrix} \times \begin{pmatrix}0\\1\\0\end{pmatrix}$

(b) $\begin{pmatrix}2\\0\\0\end{pmatrix} \times \begin{pmatrix}0\\3\\0\end{pmatrix}$

(c) $\begin{pmatrix}1\\2\\0\end{pmatrix} \times \begin{pmatrix}0\\1\\2\end{pmatrix}$

(d) $\begin{pmatrix}1\\2\\3\end{pmatrix} \times \begin{pmatrix}4\\5\\6\end{pmatrix}$

26. Use $\hat{\mathbf{i}} \times \hat{\mathbf{j}} = \hat{\mathbf{k}}$, $\hat{\mathbf{j}} \times \hat{\mathbf{k}} = \hat{\mathbf{i}}$ and $\hat{\mathbf{k}} \times \hat{\mathbf{i}} = \hat{\mathbf{j}}$ to find each product without writing column vectors.

(a) $\hat{\mathbf{j}} \times \hat{\mathbf{k}}$

(b) $\hat{\mathbf{k}} \times \hat{\mathbf{j}}$

(c) $\hat{\mathbf{i}} \times (\hat{\mathbf{i}} + \hat{\mathbf{j}})$

(d) $(2\hat{\mathbf{i}}) \times (3\hat{\mathbf{k}})$

Perpendicular to both

27. Let $\mathbf{a} = \begin{pmatrix}2\\1\\-1\end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix}1\\3\\2\end{pmatrix}$.

(a) Find $\mathbf{a} \times \mathbf{b}$.

(b) Verify that $\mathbf{a} \times \mathbf{b}$ is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$.

(c) Find a unit vector perpendicular to both.

(d) Write down $\mathbf{b} \times \mathbf{a}$, without further calculation.

Parallel vectors

28. Use the vector product to decide whether each pair of vectors is parallel.

(a) $\begin{pmatrix}2\\-1\\3\end{pmatrix}$ and $\begin{pmatrix}-4\\2\\-6\end{pmatrix}$

(b) $\begin{pmatrix}1\\2\\3\end{pmatrix}$ and $\begin{pmatrix}2\\4\\5\end{pmatrix}$

Areas

29. Areas and magnitudes.

(a) Given $|\mathbf{a}| = 3$, $|\mathbf{b}| = 5$ and an angle of $30°$, find $|\mathbf{a} \times \mathbf{b}|$.

(b) Find the area of the parallelogram with sides $\begin{pmatrix}1\\0\\0\end{pmatrix}$ and $\begin{pmatrix}0\\2\\0\end{pmatrix}$.

(c) Find the area of the triangle with vertices $A(1,0,1)$, $B(3,2,1)$, $C(2,-1,3)$.

(d) Explain why $\mathbf{v} \times \mathbf{v} = \mathbf{0}$ for every vector $\mathbf{v}$.

Concept check

30. A student calculates $\begin{pmatrix}1\\2\\1\end{pmatrix} \times \begin{pmatrix}3\\0\\2\end{pmatrix} = \begin{pmatrix}4\\-1\\-6\end{pmatrix}$.

(a) Use a scalar product to show that the answer is wrong.

(b) Find the correct vector product, and state the likely error.

9.6 Planes

A plane is a flat, endless sheet, and like a line it can be built from a point and some directions. A line needed only one direction, because a line is one-dimensional: one number, the parameter, is enough to say where you are on it. A plane is two-dimensional, so it needs two directions and two numbers. Anchor at $\mathbf{a}$, then travel λ lots of $\mathbf{b}$ and μ lots of $\mathbf{c}$. The pair (λ, μ) acts as a coordinate system inside the sheet, and every point of the plane has exactly one such pair. The two directions must not be parallel: if they were, every combination of them would still point along a single line, and the points would all lie on that line.

KEY · IB **Vector equation of a plane:** $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b} + \mu\mathbf{c}$.

Two parameters make this form awkward to use. It is hard to tell whether a point lies on the plane, or whether two such equations describe the same plane. A better description uses not the directions the plane contains but the single direction it does *not*: the **normal vector** $\mathbf{n}$, perpendicular to the whole sheet. Every displacement within the plane is perpendicular to $\mathbf{n}$, and the scalar product states exactly that.

KEY · IB **Normal (scalar-product) form:** $\mathbf{r} \cdot \mathbf{n} = \mathbf{a} \cdot \mathbf{n}$,

Cartesian form: $n_1x + n_2y + n_3z = d$, where $\mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}$ and $d = \mathbf{a} \cdot \mathbf{n}$.

The picture behind this is simple. Take any point R on the plane. The displacement $\mathbf{r} - \mathbf{a}$ from the known point A to R lies flat in the sheet, so it is perpendicular to $\mathbf{n}$ and its scalar product with $\mathbf{n}$ is zero. Expanding $(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0$ gives the form above.

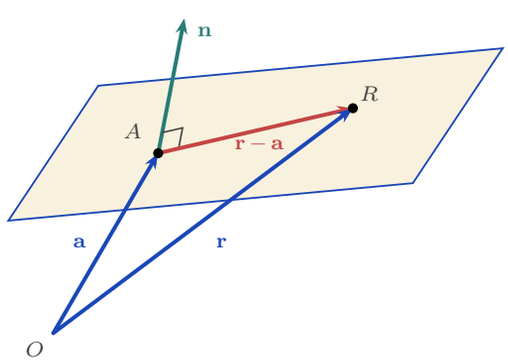


Figure 9.9 A plane held by one point and one perpendicular direction. Every displacement inside the sheet, such as $\mathbf{r} - \mathbf{a}$, is at right angles to the normal $\mathbf{n}$, so $(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0$ for every point R of the plane, and for no other point.

A Cartesian plane equation can now be read directly: the coefficients of x, y, z *are* the normal vector. To find a plane through three points, use the cross product of two edge

vectors to build $\mathbf{n}$, then fix d from any one point.

Worked through. For the plane with normal $\mathbf{n} = \begin{pmatrix} 3 \\ -2 \\ 2 \end{pmatrix}$ through $A(4, 8, -1)$, the Cartesian equation is $3x - 2y + 2z = d$. Substitute A to find d:

$$d = 3(4) - 2(8) + 2(-1) = 12 - 16 - 2 = -6, \qquad \text{so} \qquad 3x - 2y + 2z = -6.$$

The normal supplies the coefficients and a single point supplies the constant. That is the method.

When no normal is given, the cross product produces one.

EXAMPLE 9.8

Find the Cartesian equation of the plane through $A(3, 0, 0)$, $B(0, 4, 0)$, $C(0, 0, 2)$.

Two edge vectors span the plane: $\overrightarrow{AB} = \begin{pmatrix} -3 \\ 4 \\ 0 \end{pmatrix}$ and $\overrightarrow{AC} = \begin{pmatrix} -3 \\ 0 \\ 2 \end{pmatrix}$. Their cross product is normal to both, hence to the plane:

$$\mathbf{n} = \overrightarrow{AB} \times \overrightarrow{AC} = \begin{pmatrix} (4)(2) - (0)(0) \\ (0)(-3) - (-3)(2) \\ (-3)(0) - (4)(-3) \end{pmatrix} = \begin{pmatrix} 8 \\ 6 \\ 12 \end{pmatrix} = 2 \begin{pmatrix} 4 \\ 3 \\ 6 \end{pmatrix}.$$

Any scalar multiple of a normal is a normal, so use the cleaner $(4, 3, 6)$. Through A: $d = 4(3) = 12$, giving $4x + 3y + 6z = 12$. Check with B: $3(4) = 12$; and C: $6(2) = 12$. Both lie on it. Scaling the normal down before computing d keeps every later number small.

That example settles the stool from the opening page. Any three points not in a straight line give two edge vectors; their cross product is a normal, and a normal with a point fixes a plane. So three points always have exactly one plane through them, and three feet always reach the floor. A fourth point comes with no such guarantee, so a fourth foot need not touch the floor when the other three do.

9.6.1 Angles with a plane

Two planes meet at an angle, and that angle is the angle between their normals, computed by the ordinary scalar product. To see why, look along the line where the planes meet, so that each plane is seen edge-on and appears as a line. Turning one arm of an angle through $90°$, and then the other arm through $90°$, leaves the angle unchanged. The normals are exactly the two arms turned in this way, so they carry the same angle.

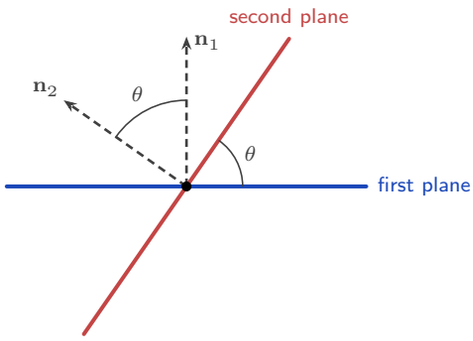


Figure 9.10 Two planes seen end-on, looking straight along the line where they meet. Each plane shows as a line. The angle between the planes and the angle between their normals are the same angle.

Taking the absolute value of the dot product guarantees the acute answer.

KEY Angle between two planes (normals $\mathbf{n}_1, \mathbf{n}_2$): $\cos\theta = \dfrac{|\mathbf{n}_1 \cdot \mathbf{n}_2|}{|\mathbf{n}_1|\,|\mathbf{n}_2|}$.

EXAMPLE 9.9

Find the acute angle between the planes $2x + 2y - z = 4$ and $x - 2y + 2z = 1$.

Read the normals straight off the coefficients: $\mathbf{n}_1 = (2, 2, -1)$ and $\mathbf{n}_2 = (1, -2, 2)$. Then $\mathbf{n}_1 \cdot \mathbf{n}_2 = 2 - 4 - 2 = -4$, and both magnitudes are 3, so

$$\cos\theta = \frac{|-4|}{3 \times 3} = \frac{4}{9} \implies \theta \approx 63.6^\circ.$$

The dot product came out negative, which means the angle between the normals is obtuse. Taking the absolute value gives the acute angle instead, which is what the question asks for.

A line and a plane need one extra step. The angle we want is between the line and its projection *in* the plane: the line formed by the feet of perpendiculars dropped from the line to the plane. But the scalar product gives the angle φ between the line and the *normal*. These two angles add to 90°, so subtract from 90° to get the one we want.

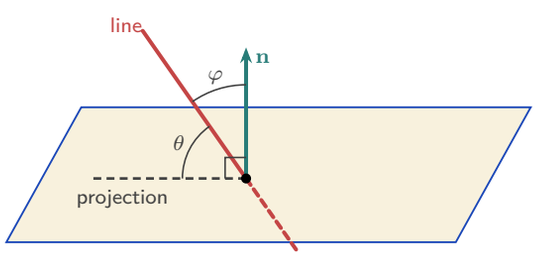


Figure 9.11 The angle between a line and a plane is measured to the line's projection in the plane, marked θ, not to the normal. Since the normal stands at 90° to the projection, the two angles at the point of entry satisfy $\theta + \varphi = 90^\circ$.

KEY Angle between a line (direction $\mathbf{d}$) and a plane (normal $\mathbf{n}$): $\sin\theta = \dfrac{|\mathbf{d}\cdot\mathbf{n}|}{|\mathbf{d}|\,|\mathbf{n}|}$.

Using $\sin$ instead of $\cos$ is what builds the "90° minus" into the formula. The line's angle to the plane is the complement of its angle to the normal.

EXAMPLE 9.10

Find the angle between the line $\mathbf{r} = \begin{pmatrix}1\\0\\-2\end{pmatrix} + \lambda\begin{pmatrix}2\\1\\2\end{pmatrix}$ and the plane $2x + 2y + z = 4$.

The direction is $\mathbf{d} = \begin{pmatrix}2\\1\\2\end{pmatrix}$, and the normal is read straight off the coefficients, $\mathbf{n} = \begin{pmatrix}2\\2\\1\end{pmatrix}$. Then $\mathbf{d}\cdot\mathbf{n} = 4 + 2 + 2 = 8$, with $|\mathbf{d}| = 3$ and $|\mathbf{n}| = 3$, so

$$\sin\theta = \frac{|8|}{3\times 3} = \frac{8}{9} \implies \theta = \arcsin\frac{8}{9} \approx 62.7^\circ.$$

The same fraction fed into a cosine would give 27.3°, and that is the angle to the *normal*, not to the plane. The two add to 90°, so a solution that reaches for $\arccos$ out of habit lands on the complement of the answer.

9.6.2 Intersections: line and plane, plane and plane

To find where a line meets a plane, put the line's parametric coordinates into the plane's Cartesian equation. That gives one equation in λ. Solve it, then substitute back to get the point. If the λ terms cancel, the line is parallel to the plane. The line then misses the plane when the constants disagree, and lies entirely in it when they agree.

EXAMPLE 9.11

Find where the line $\mathbf{r} = \begin{pmatrix}1\\2\\0\end{pmatrix} + \lambda\begin{pmatrix}2\\-1\\3\end{pmatrix}$ meets the plane $3x + y - z = 9$.

Parametrically, $x = 1 + 2\lambda$, $y = 2 - \lambda$, $z = 3\lambda$. Substituting into the plane,

$$3(1 + 2\lambda) + (2 - \lambda) - 3\lambda = 9 \implies 2\lambda + 5 = 9 \implies \lambda = 2,$$

so the point is $(5, 0, 6)$. Check: $3(5) + 0 - 6 = 9$.

Two non-parallel planes do not meet at a point; they meet along a whole **line**. That line runs within both planes, so its direction is perpendicular to both normals, and the cross product builds it directly. For a point on the line, set one coordinate to zero and solve the two plane equations for the other two. If that system has no solution, the line never crosses that coordinate plane; set a different coordinate to zero instead.

EXAMPLE 9.12

Find a vector equation of the line of intersection of the planes $x + y + 2z = 9$ and $3x - y + z = 7$.

The direction is perpendicular to both normals:

$$\mathbf{d} = \mathbf{n}_1 \times \mathbf{n}_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ -4 \end{pmatrix}.$$

For a point on the line, pick a coordinate and set it to zero. With $z = 0$: $x + y = 9$ and $3x - y = 7$, so $x = 4$, $y = 5$. Hence

$$\mathbf{r} = \begin{pmatrix} 4 \\ 5 \\ 0 \end{pmatrix} + t \begin{pmatrix} 3 \\ 5 \\ -4 \end{pmatrix}.$$

Check the direction: $(3, 5, -4) \cdot (1, 1, 2) = 0$ and $(3, 5, -4) \cdot (3, -1, 1) = 0$. Perpendicular to both normals, as the geometry demands.

9.6.3 Three planes

Three planes form a system of three linear equations, the same systems solved by row reduction in Ch. 4. Nothing new has to be calculated here. What is new is that every algebraic outcome now has a shape, and the shape is often easier to remember than the algebra.

There are five arrangements of three distinct planes. The first two have solutions; the last three do not, and they are told apart by how many of the normals are parallel.

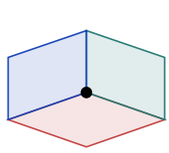
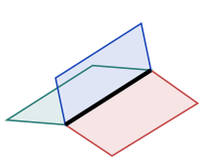
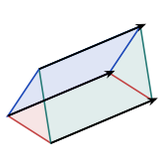
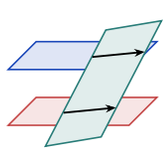
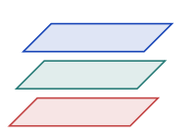
Unique solution	Infinitely many solutions	No solutions (inconsistent system)		
		No normals parallel	Two normals parallel	Three normals parallel
				
Three planes meet at a point	Three planes meet along a line	Three planes form a prism	One plane cutting two parallel planes	Three parallel planes

Figure 9.12 The five arrangements of three planes. Only the first has a single solution; the second has a whole line of them. The last three have none, and the number of parallel normals says which one you are looking at.

Three further cases arise when two or more of the equations are multiples of one another, so that they describe the same plane.

- All three describe the same plane: every point of that plane is a solution.
- Two describe the same plane and the third crosses it: the solutions form the line where they meet.
- Two describe the same plane and the third is parallel to it: there is no solution.

Row reduction tells you which case you are in without any picture. A row of zeros with a zero on the right means infinitely many solutions, so the planes share a line (or a plane, if two rows vanish). A row of zeros with a non-zero on the right means an inconsistent one, so there is no common point. The pictures only tell you what that answer looks like.

Three planes that share a common line are called a **sheaf**. To test for one, take the line where two of the planes meet and substitute its parametric coordinates into the third plane's equation. If the parameter cancels, the third plane is parallel to the line: when the constants agree it contains the line and the planes form a sheaf; when they disagree, and no two of the planes are parallel, the three planes form a prism.

EXAMPLE 9.13

The planes $x + y + 2z = 9$ and $3x - y + z = 7$ from the previous example are joined by a third, $4x + 3z = c$. Describe the configuration of the three planes for $c = 16$ and for $c \neq 16$.

Adding the first two equations gives $4x + 3z = 16$, so every point on their line of intersection satisfies $4x + 3z = 16$. Indeed, check the line $\mathbf{r} = (4, 5, 0) + t(3, 5, -4)$ in the third plane: $4(4 + 3t) + 3(-4t) = 16 + 12t - 12t = 16$, for every t.

So when $c = 16$ the third plane contains the entire line, and all three planes share it: they form a sheaf. When $c \neq 16$ the third plane is parallel to that line but misses it. The three planes then bound a triangular prism: each pair meets in a line, the three lines are parallel, and no point lies on all three. In the language of Ch. 4, the system has infinitely many solutions when $c = 16$ and is inconsistent otherwise; the geometry is what those words look like.

YOUR TURN 9.6 Planes

Equations of a plane

31. Equations of planes.

(a) Write the Cartesian equation of the plane with normal $\begin{pmatrix}1\\2\\3\end{pmatrix}$ through the origin.

(b) State a normal vector to the plane $2x - y + 3z = 5$.

(c) Find the Cartesian equation of the plane with normal $\begin{pmatrix}1\\1\\1\end{pmatrix}$ through $(2, 0, 1)$.

(d) Does $(1, 1, 1)$ lie on the plane $x + y + z = 3$?

32. Forms of the equation of a plane.

(a) Write the equation of the plane through $(1, -2, 3)$ with normal $\begin{pmatrix}2\\1\\-1\end{pmatrix}$ in the form $\mathbf{r} \cdot \mathbf{n} = \mathbf{a} \cdot \mathbf{n}$, and in Cartesian form.

(b) A plane has vector equation $\mathbf{r} = \begin{pmatrix}1\\0\\1\end{pmatrix} + \lambda\begin{pmatrix}1\\2\\0\end{pmatrix} + \mu\begin{pmatrix}0\\1\\3\end{pmatrix}$. Find a normal vector and the Cartesian equation of the plane.

(c) Does the point $(2, 1, -2)$ lie on the plane in part (b)? Justify your answer.

33. Planes from points.

(a) Find the Cartesian equation of the plane through $A(1, 0, 2)$, $B(3, 1, 0)$ and $C(0, 2, 1)$.

(b) Find the Cartesian equation of the plane that contains the line $\mathbf{r} = \begin{pmatrix}1\\2\\0\end{pmatrix} + \lambda\begin{pmatrix}1\\0\\1\end{pmatrix}$ and the point $(2, 1, 3)$.

Intersections

34. Intersections.

(a) Find where the line $\mathbf{r} = \begin{pmatrix}1\\0\\2\end{pmatrix} + \lambda\begin{pmatrix}2\\1\\-1\end{pmatrix}$ meets the plane $3x - 2y + 2z = 11$.

(b) Find a direction vector of the line of intersection of $x + y + z = 6$ and $2x - y + z = 3$.

(c) Find a point on both planes in part (b), and hence a vector equation of their line of intersection.

Angles with a plane

35. GDC Find each acute angle.

(a) The angle between the planes $x + 2y - z = 3$ and $2x - y + z = 1$.

(b) The angle between the planes $x + y = 3$ and $y + z = 1$.

(c) The angle between the line $\mathbf{r} = \begin{pmatrix}1\\0\\2\end{pmatrix} + \lambda\begin{pmatrix}2\\1\\-1\end{pmatrix}$ and the plane $3x - 2y + 2z = 11$.

(d) The angle between the line $\mathbf{r} = \lambda\begin{pmatrix}1\\1\\1\end{pmatrix}$ and the plane $z = 0$.

Three planes

36. For each system of three planes, decide whether they meet at a single point, in a line, or not at all, and describe the arrangement (for example, a sheaf or a triangular prism). Give the point or the line where there is one.

(a) $x + y + z = 6, \quad x - y + z = 2, \quad 2x + y - z = 1$

(b) $x + y + z = 6, \quad x - y + z = 2, \quad 2x + 2z = 8$

(c) $x + y + z = 6, \quad 2x + 2y + 2z = 10, \quad x - y + z = 2$

(d) $x + y + z = 6, \quad x - y + z = 2, \quad 2x + 2z = 3$

Concept check

37. GDC To find the angle between the line $\mathbf{r} = \lambda \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ and the plane $z = 0$, a student uses the direction vector and the normal $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ in $\cos\theta = \dfrac{|\mathbf{d} \cdot \mathbf{n}|}{|\mathbf{d}|\,|\mathbf{n}|}$, and gets $\theta \approx 48.2°$. Identify the error and find the correct angle.

9.7 Distances

Two objects that never meet are still some distance apart, and asking for that distance always means the same thing: the *shortest* one. There is a single idea behind every case in this section, and it needs no new formula.

Shortest distance is not itself listed in the syllabus, which asks for intersections and angles. Every step below is built from tools that are listed, and exam questions can reach the same answers by asking for the foot of a perpendicular. Treat this section as the natural extension of what comes before, not as separate material.

KEY **The shortest link is perpendicular.** Of all the segments joining two objects, the shortest is the one at right angles to what it joins. Perpendicular means a scalar product of zero, so the geometry turns straight into an equation.

That gives a method rather than a formula to memorise. Nothing here is printed in the formula booklet, so working it out is the point.

KEY **Finding a shortest distance.**

1. Write the unknown point on each object in terms of its parameter.
2. Set the scalar product of the joining vector with each direction it must be perpendicular to equal to zero.
3. Solve for the parameter or parameters, which locates the feet of the perpendicular.
4. Find the length of the joining vector.

9.7.1 From a point to a line

A general point on the line is $F = \mathbf{a} + \lambda\mathbf{b}$. As λ varies, F slides along the line and the length PF changes. It is shortest exactly when $\overrightarrow{PF}$ is perpendicular to the line, so $\overrightarrow{PF} \cdot \mathbf{b} = 0$. That is one equation in one unknown.

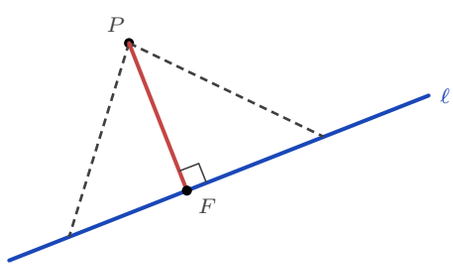


Figure 9.13 Every dashed segment joins P to the line. The shortest of them meets the line at right angles, at the foot of the perpendicular F.

Worked through. To find the distance from $P(0, 1, 3)$ to the line $\mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$, write a general point on the line as $F = (1 + \lambda,\ 2\lambda,\ -1 + 2\lambda)$, so

$$\overrightarrow{PF} = \begin{pmatrix} 1 + \lambda \\ 2\lambda - 1 \\ 2\lambda - 4 \end{pmatrix}.$$

Now make it perpendicular to the direction:

$$\overrightarrow{PF} \cdot \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} = (1 + \lambda) + 2(2\lambda - 1) + 2(2\lambda - 4) = 9\lambda - 9 = 0,$$

so $\lambda = 1$ and the foot is $F(2, 2, 1)$. Then $\overrightarrow{PF} = (2, 1, -2)$ and the distance is $\sqrt{4 + 1 + 4} = 3$. Notice what the method gives you for free: not just the distance but the point on the line that is closest to P, which many questions ask for as well.

9.7.2 From a point to a plane

For a plane the perpendicular direction is already known: it is the normal $\mathbf{n}$. So no scalar product needs to be set to zero. Instead, walk from P along $\mathbf{n}$ until you land on the plane. The general point on that walk is $\mathbf{p} + t\mathbf{n}$, and putting it into $\mathbf{r} \cdot \mathbf{n} = d$ gives one equation in t. The distance travelled is then $|t|\,|\mathbf{n}|$.

EXAMPLE 9.14

Find the distance from $P(1, 1, 1)$ to the plane $2x - y + 2z = 12$.
The normal is $\mathbf{n} = (2, -1, 2)$, so a point on the walk from P is

$$\mathbf{p} + t\mathbf{n} = (1 + 2t,\ 1 - t,\ 1 + 2t).$$

This lies on the plane when

$$2(1 + 2t) - (1 - t) + 2(1 + 2t) = 12 \implies 9t + 3 = 12 \implies t = 1.$$

The foot is $(3, 0, 3)$, and the distance is $|t|\,|\mathbf{n}| = 1 \times 3 = 3$. Check the foot on the plane: $6 - 0 + 6 = 12$. ✓

9.7.3 Between two skew lines

Skew lines never meet, yet there is still a shortest gap between them, and the same idea finds it. Take a general point on each, P_1 on the first and P_2 on the second. The joining vector $\overrightarrow{P_1P_2}$ is shortest when it is perpendicular to *both* directions at once. That is two scalar products set to zero, so two equations in the two unknowns λ and μ.

EXAMPLE 9.15

Find the shortest distance between $l_1 : \mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $l_2 : \mathbf{r} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.
General points are $P_1 = (\lambda,\ \lambda,\ 1 + \lambda)$ and $P_2 = (1 + \mu,\ 1,\ \mu)$, so

$$\overrightarrow{P_1P_2} = \begin{pmatrix} 1 + \mu - \lambda \\ 1 - \lambda \\ \mu - 1 - \lambda \end{pmatrix}.$$

Perpendicular to both directions gives two equations:

$$\overrightarrow{P_1P_2} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 1 - 3\lambda + 2\mu = 0, \qquad \overrightarrow{P_1P_2} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 2\mu - 2\lambda = 0.$$

The second gives $\mu = \lambda$; the first then gives $1 - \lambda = 0$, so $\lambda = \mu = 1$. The feet are $P_1(1, 1, 2)$

and $P_2(2,1,1)$, and

$$\overrightarrow{P_1P_2} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \qquad \text{distance} = \sqrt{2}.$$

Two ordinary simultaneous equations, and the answer comes with both endpoints of the shortest segment attached.

9.7.4 The remaining cases

Nothing else needs a new method. Every other pairing either meets, and so has distance zero, or is parallel, and then any convenient point reduces it to one of the three cases above.

The two objects	What to do
Two parallel lines	Take any point on one line, then use point to line
Two intersecting lines	They meet, so the distance is 0
Two skew lines	Two scalar products set to zero, as above
A line parallel to a plane	Take any point on the line, then use point to plane
A line meeting a plane	They meet, so the distance is 0
Two parallel planes	Take any point on one plane, then use point to plane
Two non-parallel planes	They meet along a line, so the distance is 0

The first step in any of these questions is therefore not calculation but classification. Decide whether the objects meet, and if they do not, decide which of the three methods the pair reduces to.

EXAMPLE 9.16

Find the distance between the parallel planes $2x - y + 2z = 12$ and $2x - y + 2z = 3$.
Both have normal $(2, -1, 2)$, so they are parallel and never meet. Take any point on the second plane; setting $y = z = 0$ gives $x = \frac{3}{2}$, so $A(\frac{3}{2}, 0, 0)$ lies on it. Now walk from A along the normal to the first plane:

$$2\left(\tfrac{3}{2} + 2t\right) - (-t) + 2(2t) = 12 \implies 9t + 3 = 12 \implies t = 1,$$

so the distance is $|t|\,|\mathbf{n}| = 3$. Any other starting point on the second plane gives the same answer, which is what parallel means.

YOUR TURN 9.7 Distances

From a point to a line

38. Distance from a point to a line. Throughout, $l : \mathbf{r} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$.

(a) Find the foot of the perpendicular from $P(6, 2, 1)$ to l, and hence the distance from P to l.

(b) Find the distance from $Q(3, 3, 0)$ to l.

(c) The line m has equation $\mathbf{r} = \begin{pmatrix} 6 \\ 2 \\ 1 \end{pmatrix} + t \begin{pmatrix} -4 \\ 2 \\ -4 \end{pmatrix}$. Explain why m is parallel to l, and write down the distance between them.

(d) Check your answer to part (c) by repeating the method with the point on m where $t = 1$.

From a point to a plane

39. Distance from a point to a plane. Throughout, $\Pi : x + 2y + 2z = 14$.

(a) Find the foot of the perpendicular from $P(0, -1, -1)$ to Π, and hence the distance from P to Π.

(b) Find the distance from the origin to Π.

(c) Find the distance between Π and the parallel plane $x + 2y + 2z = 5$.

(d) Show that the line $\mathbf{r} = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$ is parallel to Π, and write down the distance between the line and the plane.

Between two lines

40. Distance between two lines. Throughout, $l_1 : \mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ and $l_2 : \mathbf{r} = \begin{pmatrix} 4 \\ -2 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$.

(a) Show that l_1 and l_2 are skew.

(b) Write down general points P_1 on l_1 and P_2 on l_2, and find the values of λ and μ for which $\overrightarrow{P_1P_2}$ is perpendicular to both direction vectors.

(c) Hence state the two feet of the common perpendicular and the shortest distance between l_1 and l_2.

(d) The lines $n_1 : \mathbf{r} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ and $n_2 : \mathbf{r} = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ are not parallel. Apply the same method to them and explain what your answer shows.

Concept check

41. A student says: "*the distance from the origin to the plane* $x + 2y + 2z = 14$ *is* 14." Identify the error and find the correct distance.

Chapter 9 · Vectors

Reflections

TOK The same plane can be a drawing, a vector equation, or the single equation $2x - y + 3z = 7$. The symbols are easier to calculate with. The picture is easier to believe.

If the two disagree in your head, which one do you trust, and what happens in four dimensions, where the picture is no longer available to you?

IM The products of this chapter were chosen, not found. On 16 October 1843 the Irish mathematician William Rowan Hamilton, walking beside a canal in Dublin, saw how to multiply numbers of the form $a + bi + cj + dk$, and carved the rule $i^2 = j^2 = k^2 = ijk = -1$ into the stone of Brougham Bridge. His *quaternions* multiply in a way that depends on the order, as the cross product does. About forty years later the American Josiah Willard Gibbs and the English engineer Oliver Heaviside cut quaternions down to the two operations physicists actually used, the dot product and the cross product, and their vector notation replaced Hamilton's. What should decide which of two definitions survives: which is more elegant, or which is easier to use?

RW Three-dimensional games and animated films rely on this chapter. The dot product decides how brightly a surface is lit, by comparing the direction of the light with the direction the surface faces: a value near 1 points at the light and is bright, a value near 0 faces edge-on and stays dark. The cross product builds the normals that say which way each surface points in the first place. At the same time the physics engine adds force and velocity vectors tip to tail, many times a second, so a thrown object curves the way your eye expects.

EXT The cross product, as defined here, works only in three dimensions.

In two dimensions there is no room for a vector perpendicular to two others. In four dimensions or more there is too much room: no single perpendicular direction is picked out, because a whole family of them works. Three dimensions is the only case where two vectors determine exactly one perpendicular line. That is why $\mathbf{a} \times \mathbf{b}$ can be a vector at all. (A product with most of the same properties also exists in seven dimensions, and in no others.)

End-of-Chapter Problems

1. Given $\mathbf{a} = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}$:

(a) find $\mathbf{a} + 2\mathbf{b}$; [2]

(b) find $|\mathbf{a}|$; [1]

(c) find the unit vector in the direction of $\mathbf{a}$. [2]

2. For $A(1, -1, 2)$ and $B(3, 2, 4)$:

(a) find $\overrightarrow{AB}$; [1]

(b) find $|\overrightarrow{AB}|$; [2]

(c) find the midpoint of AB. [2]

3. The vectors $\mathbf{a} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ k \end{pmatrix}$ are perpendicular. Find k. [3]

4. Find the angle between $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$. [4]

5. Consider the line $l : \mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$.

(a) Find the point on l when $\lambda = 2$. [2]

(b) Determine whether $P(5, 2, 3)$ lies on l. [2]

6. GDC Find, in degrees, the acute angle between two lines with direction vectors $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}$. [4]

7. Given $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$, find $\mathbf{a} \times \mathbf{b}$ and $|\mathbf{a} \times \mathbf{b}|$. [4]

8. Find the area of the triangle with vertices $A(1, 0, 0)$, $B(0, 1, 0)$, $C(0, 0, 1)$. [4]

9. A plane Π has normal $\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ and passes through $A(1, 2, -1)$.

(a) Find the Cartesian equation of Π. [3]

(b) Determine whether $B(2, 1, 0)$ lies on Π. [1]

10. Find the Cartesian equation of the plane through $A(1, 0, 0)$, $B(0, 1, 0)$ and $C(0, 0, 1)$. [4]

11. Given $\mathbf{a} = \begin{pmatrix} 4 \\ -3 \\ 0 \end{pmatrix}$:

(a) find $|\mathbf{a}|$; [1]
(b) find the unit vector $\hat{\mathbf{a}}$; [2]
(c) find the vector of magnitude 10 in the direction of $\mathbf{a}$. [2]

12. The point $M(3, -1, 4)$ is the midpoint of AB, where A is $(1, 2, 5)$. Find the coordinates of B. [3]

13. The line $l : \mathbf{r} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ meets the plane $z = 4$. Find the point of intersection. [4]

14. GDC Find the angle between the vectors $\begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix}$. [4]

15. Show that the points $A(1, 2, 3)$, $B(2, 4, 5)$ and $C(4, 8, 9)$ are collinear. [4]

16. GDC Given $|\mathbf{a}| = 5$, $|\mathbf{b}| = 3$ and $\mathbf{a} \cdot \mathbf{b} = 9$, find the angle between $\mathbf{a}$ and $\mathbf{b}$. [4]

17. The lines $l_1 : \mathbf{r} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ and $l_2 : \mathbf{r} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ intersect. Find the point of intersection. [5]

18. The plane Π has equation $2x + y - 2z = 6$.

(a) State a normal vector to Π. [1]
(b) Find where the line $\mathbf{r} = \lambda \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$ meets Π. [4]

19. Given $\mathbf{a} = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$:

(a) find $\mathbf{a} \cdot \mathbf{b}$; [2]
(b) hence state the angle between $\mathbf{a}$ and $\mathbf{b}$. [2]

20. Find the area of the triangle with vertices $A(0, 0, 0)$, $B(3, 0, 0)$, $C(0, 4, 0)$ using the vector product. [4]

21. Find a vector equation of the line through $A(2, -1, 3)$ and $B(4, 1, 7)$. [4]

22. GDC Find the acute angle between the planes $x + y + z = 1$ and $x - y + z = 3$. [5]

23. Given $\mathbf{a} = \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 5 \end{pmatrix}$, find $\mathbf{a} \times \mathbf{b}$ and $|\mathbf{a} \times \mathbf{b}|$. [4]

24. Find the value of t for which $\begin{pmatrix} 1 \\ t \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ -1 \\ t \end{pmatrix}$ are perpendicular. [4]

25. At time $t = 0$ seconds a drone is at the point $(2, -1, 10)$ (coordinates in metres) and moves with constant velocity $\begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}$ m s^{-1}.

(a) Write down a vector equation for the drone's position at time t. [2]
(b) Find the drone's speed. [2]
(c) Determine whether the drone passes through the point $(17, 19, 10)$, and if so, at what time. [3]

26. Find a vector equation of the line of intersection of the planes $x + y - z = 2$ and $2x - y + z = 1$. [6]

27. Consider the planes $\Pi_1 : x + y + z = 3$, $\Pi_2 : x - y + z = 1$ and $\Pi_3 : x + z = c$, where $c \in \mathbb{R}$.

(a) Find a vector equation of the line of intersection of Π_1 and Π_2. [4]
(b) Show that when $c = 2$ the three planes intersect in this line. [2]
(c) Describe geometrically the configuration of the three planes when $c \neq 2$. [2]

28. Show that the lines $l_1 : \mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ and $l_2 : \mathbf{r} = \begin{pmatrix} 1 \\ 5 \\ -1 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$ intersect, and find the point of intersection. [6]

29. Find the shortest distance from the point $P(1, 2, 3)$ to the plane $2x - y + 2z = 0$, by finding the foot of the perpendicular from P. [6]

30. Find a unit vector perpendicular to both $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$. [5]

31. The points $A(1, 0, 0)$, $B(0, 2, 0)$ and $C(0, 0, 3)$ are given.

(a) Find the Cartesian equation of the plane ABC. [3]
(b) Find the area of triangle ABC. [3]

32. GDC Find the angle that the line $l : \mathbf{r} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$ makes with the plane $x + 2y + 2z = 5$. [6]

33. The points $A(2, 0, 1)$, $B(1, 3, 2)$ and $C(4, 1, 0)$ are given, and O is the origin.

(a) Find $\overrightarrow{AB}$ and $\overrightarrow{AC}$. [2]
(b) Find $\overrightarrow{AB} \times \overrightarrow{AC}$. [3]
(c) Find the exact area of triangle ABC. [2]
(d) Find the Cartesian equation of the plane through A, B and C. [4]
(e) Find the exact shortest distance from O to that plane. [3]
(f) Hence find the volume of the tetrahedron $OABC$. [2]

34. Consider the lines

$$l_1 : \mathbf{r} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \qquad l_2 : \mathbf{r} = \begin{pmatrix} 4 \\ 1 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

(a) Show that l_1 and l_2 are not parallel. [2]
(b) Determine whether the lines intersect, and if so find the point of intersection. [5]
(c) Find the acute angle between the lines. [3]
(d) Find the Cartesian equation of the plane containing both lines. [4]
(e) Explain why such a plane exists only because of your answer to part (b). [2]

35. GDC Two aircraft fly at the same altitude. At time t hours their position vectors, in km, are

$$\mathbf{r}_A = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} + t \begin{pmatrix} 300 \\ 400 \\ 0 \end{pmatrix}, \qquad \mathbf{r}_B = \begin{pmatrix} 600 \\ 0 \\ 2 \end{pmatrix} + t \begin{pmatrix} 0 \\ 500 \\ 0 \end{pmatrix}.$$

(a) Find the speed of each aircraft. [3]
(b) Find $\overrightarrow{AB}$ at time t. [3]
(c) Show that the distance d between them satisfies $d^2 = 100\,000t^2 - 360\,000t + 360\,000$. [4]
(d) Find the time at which the aircraft are closest, and the least distance between them. [4]

36. Consider the planes

$$\Pi_1 : x + y + z = 1, \qquad \Pi_2 : x + 2y + 3z = 4, \qquad \Pi_3 : x + y + kz = m,$$

where k and m are real constants.

(a) Find a vector equation of the line of intersection of Π_1 and Π_2. [4]
(b) Show that when $k = 2$ the three planes meet at a single point, and find that point in terms of m. [4]
(c) Find the values of k and m for which the three planes meet in a line. [3]
(d) Find the values of k and m for which the three planes have no common point, and describe that arrangement geometrically. [3]

37. The points $A(1, 2, 1)$, $B(3, 1, 2)$ and $C(2, 4, 3)$ are given.

(a) Find $\overrightarrow{AB}$ and $\overrightarrow{AC}$. [2]
(b) Find $\overrightarrow{AB} \times \overrightarrow{AC}$. [3]
(c) Hence find the area of triangle ABC. [3]
(d) Find the Cartesian equation of the plane through A, B and C. [3]
(e) Explain why the vector product gives a normal to that plane. [2]

38. GDC Two drones move with constant velocity. At time t seconds their positions, in metres, are

$$\mathbf{r}_A = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + t \begin{pmatrix} 4 \\ -3 \\ 12 \end{pmatrix}, \qquad \mathbf{r}_B = \begin{pmatrix} 3 \\ -13 \\ 6 \end{pmatrix} + t \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix}.$$

(a) Find the speed of each drone. [3]
(b) Show that the two flight paths cross, and find the point where they cross. [5]
(c) Explain why the drones nevertheless do not collide. [3]
(d) Find the distance between the drones at $t = 2$. [3]

(e) Find the acute angle between the two flight paths. [3]

Challenge Questions

39. Closest approach of two moving objects. Two drones fly in straight lines at constant velocity. Taking $t \geq 0$ in seconds and distances in metres, their positions are

$$\mathbf{r}_A = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix} + t \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}, \qquad \mathbf{r}_B = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} + t \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}.$$

Section 9.4 finds a closest approach by completing the square for one pair of objects. This investigation turns the method into a formula, links it to the vector product, and compares it with the distance between the two flight paths.

In general write $\mathbf{r}_A = \mathbf{a}_A + t\mathbf{v}_A$ and $\mathbf{r}_B = \mathbf{a}_B + t\mathbf{v}_B$, with $\Delta\mathbf{a} = \mathbf{a}_B - \mathbf{a}_A$ and $\Delta\mathbf{v} = \mathbf{v}_B - \mathbf{v}_A \neq \mathbf{0}$, and let d be the distance between the two objects at time t.

(a) For the drones, show that $d^2 = 4(t-2)^2 + 4$. State the time of closest approach and the least distance. [4]

(b) In general, show that $d^2 = |\Delta\mathbf{v}|^2 t^2 + 2(\Delta\mathbf{a} \cdot \Delta\mathbf{v})\, t + |\Delta\mathbf{a}|^2$. By completing the square, show that d^2 is least at

$$t^* = -\frac{\Delta\mathbf{a} \cdot \Delta\mathbf{v}}{|\Delta\mathbf{v}|^2}, \qquad \text{with} \qquad d^2_{\min} = |\Delta\mathbf{a}|^2 - \frac{(\Delta\mathbf{a} \cdot \Delta\mathbf{v})^2}{|\Delta\mathbf{v}|^2},$$

and state what it means for the two objects when $t^* < 0$. [5]

(c) Use $|\mathbf{u} \times \mathbf{v}| = |\mathbf{u}||\mathbf{v}| \sin\theta$ and $\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}||\mathbf{v}| \cos\theta$ to show that $|\mathbf{u} \times \mathbf{v}|^2 = |\mathbf{u}|^2|\mathbf{v}|^2 - (\mathbf{u} \cdot \mathbf{v})^2$. Hence show that $d_{\min} = \dfrac{|\Delta\mathbf{a} \times \Delta\mathbf{v}|}{|\Delta\mathbf{v}|}$, and check this against part (a). [4]

(d) The two flight paths are lines. Use the method of Section 9.7 to find the two points where the common perpendicular of the paths meets them, and show that the shortest distance between the paths is $\sqrt{2}$ metres. [4]

(e) Explain why the least distance between two moving objects can never be less than the distance between their paths. Find the times at which A and B pass the two points found in part (d), and use them to explain why the drones never come as close as $\sqrt{2}$ metres. State when the two distances are equal. [3]

(f) Suppose drone B starts s seconds late, so that $\mathbf{r}_B = \mathbf{a}_B + (t-s)\mathbf{v}_B$, where a negative s means an early start. Investigate how the least distance between the drones depends on s, and suggest how this model of a near miss could be taken further. [3]

40. How far apart are two skew lines? Two skew lines never meet, yet there is still a shortest distance between them. This investigation finds it, and then deals with the case the

formula cannot handle.

Let $l_1 : \mathbf{r} = \mathbf{a}_1 + \lambda\mathbf{d}_1$ and $l_2 : \mathbf{r} = \mathbf{a}_2 + \mu\mathbf{d}_2$ be skew.

(a) Show that $l_1 : \mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $l_2 : \mathbf{r} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ are skew. [4]

(b) The shortest join between the two lines is perpendicular to both. Explain why its direction must be parallel to $\mathbf{n} = \mathbf{d}_1 \times \mathbf{d}_2$. [3]

(c) Hence explain why the shortest distance is the length of the projection of $\mathbf{a}_2 - \mathbf{a}_1$ onto $\mathbf{n}$, that is

$$\text{distance} = \frac{|(\mathbf{a}_2 - \mathbf{a}_1) \cdot \mathbf{n}|}{|\mathbf{n}|}.$$

(Hint: let the shortest join run from P_1 on l_1 to P_2 on l_2, and write $\mathbf{a}_2 - \mathbf{a}_1$ in terms of $\overrightarrow{P_1P_2}$, $\mathbf{d}_1$ and $\mathbf{d}_2$) [4]

(d) Use the formula on the two lines in part (a), giving an exact answer. [5]

(e) What does the formula give when the two lines intersect? Explain why your answer is what it should be. [3]

(f) Now let $l_3 : \mathbf{r} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$, which is parallel to l_1. Explain why the formula of part (c) cannot be used for l_1 and l_3. Show that the distance between two parallel lines is $\dfrac{|(\mathbf{a}_3 - \mathbf{a}_1) \times \mathbf{d}_1|}{|\mathbf{d}_1|}$, and find the exact distance between l_1 and l_3. [4]

(g) The lines here are infinite. Suggest how the method would have to change to find the shortest distance between two line segments, such as two rods or two edges of a solid, and how this could be investigated further. [2]

41. The triple product and volume. Combining both products gives $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$, a single number called the *scalar triple product*. It measures volume, and it decides whether three planes meet in a single point.

In parts (a) to (e), take $\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$.

(a) Find $\mathbf{v} \times \mathbf{w}$, then $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$. [3]

(b) The parallelepiped with edges $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ has base area $|\mathbf{v} \times \mathbf{w}|$. Explain why its volume is $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$. *(Hint: volume = base area × perpendicular height)* [3]

(c) Evaluate $\mathbf{v} \cdot (\mathbf{w} \times \mathbf{u})$, $\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})$ and $\mathbf{u} \cdot (\mathbf{w} \times \mathbf{v})$. Make a conjecture about what happens to the triple product when the three vectors are moved round in order, and when two of them are swapped. Prove both parts of your conjecture for any three vectors. [5]

(d) Three vectors are *coplanar* when they all lie in one plane. Prove that three vectors are coplanar if and only if their triple product is zero, proving both directions. Test $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 5 \\ 2 \end{pmatrix}$. [5]

(e) A tetrahedron on the same three edges has one sixth of the parallelepiped's volume. Find the volume of the tetrahedron with vertices at the origin and at the three points $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ above. [2]

(f) Three planes have normals $\mathbf{n}_1$, $\mathbf{n}_2$, $\mathbf{n}_3$. Explain why they meet in exactly one point when $\mathbf{n}_1 \cdot (\mathbf{n}_2 \times \mathbf{n}_3) \neq 0$, and why they cannot meet in exactly one point when it is zero. Use this to show that the planes $x + 2y = 3$, $y + 2z = 1$ and $2x + 5y + 2z = k$ never meet in

a single point, and find the value of k for which they share a line. [5]

(g) The sign of the triple product has been set aside so far. Explain what the sign of $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ tells you about the three vectors, and suggest how this could be investigated further. [2]

42. Varignon's theorem in three dimensions. Varignon's theorem states that joining the midpoints of the sides of a quadrilateral in a plane, in order, gives a parallelogram. This investigation asks whether the four corners need to lie in a plane at all.

Let A, B, C, D have position vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, and let P, Q, R, S be the midpoints of AB, BC, CD and DA.

(a) Take $A(0,0,0)$, $B(4,0,0)$, $C(6,4,2)$ and $D(0,2,6)$. Show that these four points do not lie in one plane. *(Hint: use the triple product of Challenge Question 41)* [2]

(b) For the points in part (a), find P, Q, R and S, and show that $PQRS$ is a parallelogram with four equal sides. [3]

(c) Prove that for any four points $\overrightarrow{PQ} = \overrightarrow{SR} = \frac{1}{2}(\mathbf{c} - \mathbf{a})$ and $\overrightarrow{QR} = \overrightarrow{PS} = \frac{1}{2}(\mathbf{d} - \mathbf{b})$. Deduce that $PQRS$ is always a parallelogram, lying in one plane even when A, B, C, D do not. [3]

(d) Show that the diagonals PR and QS bisect each other at the point G with position vector $\frac{1}{4}(\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d})$, and that G is also the midpoint of the segment joining the midpoints of AC and BD. Find G for the points in part (a). [3]

(e) Compare the lengths of AC and BD in part (a). Conjecture a condition on AC and BD for $PQRS$ to be a rhombus, and another for it to be a rectangle, and prove both. Hence state what shape $PQRS$ is when $ABCD$ are the vertices of a regular tetrahedron, justifying your answer. [5]

(f) Show that the area of $PQRS$ is $\frac{1}{4}|(\mathbf{c} - \mathbf{a}) \times (\mathbf{d} - \mathbf{b})|$, and find it for the points in part (a). Compare it with the area of $ABCD$ when $ABCD$ is a convex quadrilateral in a plane, and suggest how this could be investigated further. [3]

STATISTICS AND PROBABILITY

Chapter 10

Statistics

10

SYLLABUS 4.1 4.2 4.3 4.4 4.10 4.14

Statistics begins with a practical difficulty: a list of numbers is usually too long to read and too varied to summarise by eye.

In the 1880s Francis Galton met exactly this problem. He recorded the heights of 928 adult children and of the 205 sets of parents who raised them, in order to answer a single question: how well does a parent's height predict a child's? Tall parents did tend to have tall children. But the children of very tall parents were, on average, slightly shorter than their parents, and the children of very short parents slightly taller. Extreme parents tended to have less extreme children. Galton called this "regression toward mediocrity". The judgement in that phrase has since been dropped; the name *regression* has not.

Before any of this could be seen, the ordinary work had to be done first. One number says very little, and nine hundred say too much at once. The data had to be summarised, their variation measured, and one list set against another.

Those are the tasks of this chapter. By the end of it you will be able to collect data by a defensible method, describe a single list by its centre and its spread, display a distribution as a histogram, a cumulative frequency curve or a box plot, compare two lists, measure how closely two variables move together, and fit the line that best predicts one from the other.

Each of these tools has a range in which it can be trusted and a range in which it cannot, and knowing where the boundary lies is part of the subject. Correlation does not establish cause. A fitted line says nothing reliable beyond the data it was fitted to. Galton's drift is the standing example: it is not a fact about families, and no biology is needed to explain it. The reason appears at the end of the chapter.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

10.1 Sampling and Data

You almost never have all the data. Galton could not measure every family in England, a factory cannot test every light bulb it makes, and a biologist cannot weigh every fish in a lake. So you measure a part and reason about the whole. How far that reasoning can be trusted depends on how the part was chosen.

The complete group you want to know about is the **population**. The smaller group you actually measure is the **sample**. A number describing the population is a **parameter**; the matching number computed from a sample is a **statistic**. You use the statistic to estimate the parameter you cannot see.

DEF A **population** is the entire set of individuals under study. A **sample** is a subset of the population from which data are actually collected.
A **parameter** describes a population; a **statistic** describes a sample.

10.1.1 Types of data

Data come in two broad kinds. **Qualitative** data describe a quality or category, such as eye colour or country of birth. **Quantitative** data are numerical. Quantitative data split again. **Discrete** data take separate values you can count, such as the number of goals in a match. **Continuous** data can take any value in a range, such as height or time, limited only by the precision of your instrument.

The distinction matters because it decides how you may display and summarise the data. You can find the mean number of goals, but "mean eye colour" is meaningless.

10.1.2 Sampling techniques

A good sample resembles the population on a smaller scale. The techniques below are different ways of achieving that. Each one balances fairness against the effort required.

DEF **Simple random:** every member of the population has an equal chance of selection.
Systematic: order the population and select every k-th member from a random start.
Stratified: divide the population into groups (strata), then sample from each group in proportion to its size.
Quota: like stratified, but members within each group are chosen non-randomly until each quota is filled.
Convenience: select whoever is easiest to reach.

Stratified sampling is useful when a population has clear subgroups that may differ in what is being measured. If a school is 60% junior and 40% senior students, a stratified sample of

50 takes 30 juniors and 20 seniors, preserving the structure.

EXAMPLE 10.1

A school has 800 juniors and 400 seniors. A stratified sample of 60 students is taken. How many should be seniors?

Seniors are $\frac{400}{1200} = \frac{1}{3}$ of the school. The sample must keep that proportion:

$$\frac{1}{3} \times 60 = 20 \text{ seniors} \quad (\text{and } 40 \text{ juniors}).$$

Stratified sampling guarantees both groups are represented in their true ratio, which a simple random sample only achieves on average.

10.1.3 Bias and reliability

A sample is **biased** when its method systematically favours some outcomes over others. The statistic then misses the parameter, no matter how large the sample grows. A survey about reading habits conducted in a library will overstate how much people read. Bias is a fault in the method, not in the sample size; collecting more biased data only makes you more confident in the wrong answer.

Reliability also depends on the records themselves. Real data sets contain missing entries and recording errors. Before any analysis, ask how the data were collected, what might be missing, and whether any value could have been recorded wrongly.

CAUTION A large sample is not automatically a good one. The 1936 *Literary Digest* poll collected 2.4 million replies and still predicted the wrong US election winner. Its mailing list came largely from telephone directories and car registrations, whose owners were wealthier than the average voter, and the people who chose to reply differed again from those who did not. Size cannot repair a biased method.

10.1.4 Outliers

An **outlier** is a value far from the rest of the data. It may be a genuine extreme value, or it may be a recording error. Either way it can distort a summary badly. We will give it a precise definition once we have quartiles, in §10.4. For now, treat any suspiciously distant value as something to investigate, not something to delete immediately.

YOUR TURN 10.1 Sampling and Data

Types of data and key terms

1. State whether each variable is qualitative, discrete or continuous.

(a) the number of goals scored in a match

(b) the mass of a parcel

(c) the favourite sport of a student

(d) the time taken to run 100 m

(e) the number of students in a class

2. A health authority wants to know the mean height of all the 12-year-olds in a city. It measures 200 of them, chosen at random, and their mean height is 151.3 cm.

(a) Identify the population and the sample.

(b) State whether 151.3 cm is a parameter or a statistic, and give a reason.

Naming a sampling method

3. Name the sampling method used in each case.

(a) The names of all 600 members of a club are put in a hat and 30 are drawn.

(b) A shop surveys every 20th customer through the door, starting from a customer chosen at random among the first 20.

(c) A reporter asks the first 25 people who leave a station.

(d) A researcher is told to interview 20 men and 20 women, choosing whoever is willing to stop. Name this method and give one weakness of it.

4. A systematic sample of 40 is taken from a list of 2000 people by choosing every kth person, starting from a person chosen at random among the first k.

(a) Find k.

(b) Explain why this is not a simple random sample.

Stratified samples

5. A college has 1500 students, 600 of them male. A stratified sample of 50 is taken.

(a) Find the number of males in the sample.

(b) Find the number of females in the sample.

(c) Explain one advantage of stratified sampling over simple random sampling here.

6. Find the size of each stratum in the sample.

(a) A town of $15\,000$ people has 4000 residents over 65, 6500 aged 18 to 65 and the rest under 18. A stratified sample of 150 is taken.

(b) A school has 412 juniors, 356 middle-school students and 232 seniors. A stratified sample of 60 is taken. Round each stratum to the nearest whole number and check that the total is still 60.

Bias and reliability

7. Each survey below is biased.

(a) A researcher surveys people at a gym about how much they exercise. Explain why this is biased.

(b) A magazine prints a questionnaire on how many books its readers read each year, and 8000 readers post it back. The editor plans to collect $80\,000$ replies in the same way to make the result more accurate. Explain whether this would help.

8. A club's spreadsheet records the ages of its 120 members. Among the entries are 34, 41, 0, 29, 347 and one blank cell.

(a) Identify the entries that should be investigated before a mean age is calculated, giving a reason for each.

(b) A member is recorded as 88. Explain why this value should not be deleted just because it is far from the others.

Concept check

9. A college has 1500 students, 900 of them female. A stratified sample of 50 is to be taken by gender. A student says: "*there are two strata, so the sample should contain* 25 *females and* 25 *males.*" Identify the error, and find the correct numbers.

10.2 Measures of Central Tendency

Once you have a list of numbers, the usual first question is this: what is a typical value? A single number that stands in for the whole list is a measure of **central tendency**. There are three, and they answer slightly different questions.

The **mean** is the balance point: add everything and share it equally. The **median** is the middle value when the data are ordered. The **mode** is the most frequent value.

10.2.1 The mean

For n raw values, the mean is the total divided by the count. When values repeat, you weight each value by how often it occurs, which is just a shortcut for the same sum.

KEY · IB

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{n}, \qquad \text{where } n = \sum_{i=1}^{k} f_i$$

Here x_i are the distinct values, f_i their frequencies. A table listing each value alongside its frequency is called a **frequency distribution**, and one is worked through below. It is the standard compact way to present discrete data. With no repeats, every $f_i = 1$ and this reduces to $\bar{x} = \frac{\sum x_i}{n}$, the familiar form.

Worked through. In 20 football matches a team scored the numbers of goals below. To find the mean, multiply each value by its frequency, then divide by the total count.

Goals x	**0**	**1**	**2**	**3**	**4**
Frequency f	4	7	5	3	1

$$\bar{x} = \frac{(0)(4) + (1)(7) + (2)(5) + (3)(3) + (4)(1)}{4 + 7 + 5 + 3 + 1} = \frac{30}{20} = 1.5.$$

The mean need not be a value the data can actually take. No team scores 1.5 goals, but 1.5 is the fair share.

10.2.2 Combining two groups

Two sets of data are often summarised separately and then have to be described together. Averaging the two means will not do it, unless the groups happen to be the same size, because the larger group holds more values and must count for more.

Go back to totals. A group of n values with mean $\bar{x}$ has total $n\bar{x}$, so each group's total can be recovered from its summary. Add the totals, add the counts, and divide.

KEY A group of n_1 values with mean $\bar{x}_1$ and a group of n_2 values with mean $\bar{x}_2$ combine to give

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}.$$

The weights need not be counts. A final grade made of 40% coursework and 60% examination is the same calculation with 0.4 and 0.6 in place of n_1 and n_2, over a denominator of $0.4 + 0.6 = 1$.

EXAMPLE 10.2

Twelve students in one group scored a mean of 54 on a test, and the other eight scored a mean

of 69. Find the mean for all twenty. Five more students then sit the same test, and the mean of all twenty-five becomes 58. Find the mean of those five.

Recover each total, then combine:

$$\bar{x} = \frac{12(54) + 8(69)}{12 + 8} = \frac{648 + 552}{20} = \frac{1200}{20} = 60.$$

Run the same identity backwards for the missing group. All twenty-five values total $25 \times 58 = 1450$, and the first twenty account for 1200, so the five total $1450 - 1200 = 250$ and have mean $\frac{250}{5} = 50$.

A combined mean always lies between the two group means and, when the groups differ in size, nearer the mean of the larger group, so 60 falling between 54 and 69 and leaning towards the twelve is the check. Averaging the two means directly gives 61.5, which leans the wrong way.

10.2.3 Median and mode

The median splits ordered data into two equal halves. For n values, it sits at position $\frac{n+1}{2}$. If n is odd this is a single middle value; if n is even it is the average of the two middle values.

EXAMPLE 10.3

Find the median and mode of $4, 7, 2, 9, 4, 11, 6$.

Order first: $2, 4, 4, 6, 7, 9, 11$. With $n = 7$, the median is at position $\frac{7+1}{2} = 4$, the fourth value: 6.

The mode is the most frequent value: 4 appears twice, everything else once, so the mode is 4.

10.2.4 Grouped data

Continuous data are often recorded in classes, such as "time between 10 and 20 minutes." You no longer know the individual values, only the class each fell into. To estimate the mean, assume every value sits at its class **midpoint**, then apply the frequency formula. The result is an estimate, because the assumption is rarely exact.

The same loss of detail means that the mode and the median are reported as classes, not values. The **modal class** is the class with the highest frequency. The **median class** is the class containing the middle value: add the frequencies in order until the running total reaches the median's position, $\frac{n+1}{2}$. The two classes need not be the same.

EXAMPLE 10.4

Fifty phone calls to a help line had the durations below. Estimate the mean duration.

Duration (min)	$0 \le t < 4$	$4 \le t < 8$	$8 \le t < 12$	$12 \le t < 16$
Frequency	8	18	16	8

Use the midpoint of each class: $2, 6, 10, 14$.

$$\bar{x} = \frac{(2)(8) + (6)(18) + (10)(16) + (14)(8)}{50} = \frac{16 + 108 + 160 + 112}{50} = \frac{396}{50} = 7.92 \text{ min.}$$

The modal class is $4 \le t < 8$, the class with the highest frequency. With grouped data you report a modal class, not a single modal value.

The middle is located by a running cumulative count rather than by a formula. With $n = 50$ the median lies between the 25th and 26th values. The first class holds values 1 to 8; the second carries the running total to $8 + 18 = 26$, so it holds values 9 to 26. Both the 25th and the 26th fall there, and the median class is $4 \le t < 8$.

Again you name a class, not a value: the individual durations are gone, so the most that can be said is which class the middle value fell in. The median class and the modal class coincide here, which is common but not automatic, since one is located by a running total and the other by the largest single frequency.

10.2.5 Which measure to use

The mean and median agree when the data are symmetric (and, for a single-peaked distribution, the mode as well). They disagree when the data are skewed, and the disagreement is itself information.

The mean uses every value, so one very large value pulls it towards that value. The median depends only on position, so the same value moves it hardly at all. This is why incomes are often reported by median, not mean: a few very large incomes can pull the mean well above what most people earn.

The same difference makes the two measures a quick test of shape. A long tail on the right holds a few unusually large values, and those values pull the mean above the median. A long tail on the left does the reverse. Skew is named after the side the tail is on, not the side the peak is on, so check which side the tail is on before naming the skew.

Comparison	Shape	Usually report
$\bar{x} >$ median	positively skewed, tail to the right	median
$\bar{x} \approx$ median	roughly symmetric	mean
$\bar{x} <$ median	negatively skewed, tail to the left	median

For the single-peaked distributions in this course the comparison is a dependable guide, and it is fast: it needs two numbers, not a graph. It is a guide rather than a proof, since data

sets can be built that break it, so check it against a diagram whenever one is available.

CAUTION The mean is sensitive to outliers; the median is resistant to them. When a distribution is strongly skewed or contains a suspect extreme value, the median usually describes the centre better.

YOUR TURN 10.2 Measures of Central Tendency

Mean, median and mode of a list

10. For the data $5,\ 8,\ 8,\ 10,\ 14,\ 15$, find each of the following.

(a) the mean
(b) the median
(c) the mode
(d) the range

11. Find the required measure in each case.

(a) the median of $3, 7, 9, 12, 15, 20$
(b) the mode of $2, 4, 4, 5, 7, 7, 7, 9$
(c) the mean of $3, 7, 7, 8, 10, 11, 12, 14$
(d) the median of $12, 15, 15, 18, 20, 22$

Frequency tables and grouped data

12. The table shows a frequency distribution.

Value x	1	2	3	4	5
Frequency f	3	5	8	4	2

(a) Find n.
(b) Find $\sum fx$.
(c) Find the mean.
(d) State the mode.

13. The grouped data below record a measurement x.

Class	$0 \le x < 10$	$10 \le x < 20$	$20 \le x < 30$	$30 \le x < 40$
Frequency	4	12	9	10

(a) Write down the four class midpoints.
(b) Estimate the mean.
(c) State the modal class.
(d) Find the median class.
(e) Explain why the mean is only an estimate.

Working backwards and combining groups

14. Work backwards from a given mean.

(a) The numbers $4, 7, x, 12, 15$ have a mean of 10. Find x.

(b) The mean of six numbers is 14. A seventh number is added and the mean becomes 15. Find the seventh number.

15. Combine the information to find a single mean.

(a) One class of 18 students has a mean height of 162 cm and another class of 12 students has a mean height of 170 cm. Find the mean height of all 30 students.

(b) A final grade is made up of 40% coursework and 60% examination. A student scores 72 in the coursework and 58 in the examination. Find the final grade.

16. The 30 members of a running club have a mean 5 km time of 26.4 minutes. Six new members join, and the mean time of all 36 members becomes 27.0 minutes. Find the mean time of the six new members.

Choosing a measure and skew

17. The prices of eight houses sold in a village, in thousands of euros, were

$$180,\ 195,\ 205,\ 210,\ 220,\ 235,\ 240,\ 1150.$$

(a) Find the mean and the median.

(b) State which of the two better describes a typical price, and give a reason.

(c) Use your answers to part (a) to describe the skew of the prices.

18. Each distribution below has a single peak. State whether it is positively skewed, negatively skewed or roughly symmetric.

(a) mean 48, median 55

(b) mean 12.1, median 12.0

(c) mean 3.9, median 2.5

Concept check

19. A student is asked for the median of $7,\ 3,\ 12,\ 9,\ 5$ and writes: "*the middle value is* 12*, so the median is* 12." Identify the error, and find the median.

20. A group of 10 students has a mean mark of 60, and a group of 30 students has a mean mark of 80. A student says: "*the mean of all* 40 *students is* $\frac{60+80}{2} = 70$."

(a) Explain why this is wrong.

(b) Find the correct mean.

10.3 Measures of Dispersion

The centre alone does not describe a data set. Consider two classes that both average 60% on a test. In the first, every student scored between 58 and 62. In the second, the scores were spread evenly from 30 to 90. The mean is the same; the two classes are not. To capture the difference you need a measure of **spread**.

10.3.1 Range and interquartile range

The simplest measure is the **range**: the largest value minus the smallest. It is easy to find but unreliable. It depends only on the two most extreme values, which are exactly the ones most likely to be outliers.

A better approach is to ignore the extremes and measure the spread of the middle. Just as the median cuts the data in half, the **quartiles** cut it into quarters. Q_1 is the value one quarter of the way through the ordered data, Q_2 is the median, and Q_3 is three quarters of the way. The spread of the central half is the **interquartile range**.

By hand, Q_1 is the median of the values below the median and Q_3 is the median of the values above it; when n is odd, the median itself is left out of both halves. Textbooks and calculators place quartiles in slightly different ways, so in this course report the value your GDC gives.

KEY · IB

$$\text{IQR} = Q_3 - Q_1$$

Because it discards the top and bottom quarters, the IQR is little affected by outliers, which is why it is normally reported alongside the median.

Worked through. The data $3, 5, 6, 8, 8, 9, 11, 13, 15$ are already ordered, with $n = 9$. Their range is $15 - 3 = 12$. The median (Q_2) is the 5th value, 8. The lower quartile is the median of the four values below it, $3, 5, 6, 8$:

$$Q_1 = \frac{5 + 6}{2} = 5.5.$$

The upper quartile is the median of the four values above, $9, 11, 13, 15$:

$$Q_3 = \frac{11 + 13}{2} = 12.$$

So $\text{IQR} = 12 - 5.5 = 6.5$.

10.3.2 Variance and standard deviation

The IQR describes the middle half and ignores the rest, so two data sets with very different extremes can share an IQR. A measure built from every value avoids that. The natural quantity to start from is the **deviation** of each value from the mean, $x_i - \bar{x}$, which says how far that value sits from the centre and on which side.

Averaging the deviations themselves does not work, and it fails for every data set, not just awkward ones. Adding them gives

$$\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} x_i - n\bar{x} = n\bar{x} - n\bar{x} = 0,$$

because $\sum x_i = n\bar{x}$ is what the mean means. Values above the centre and values below it cancel exactly. The total is 0 for tightly packed data and 0 for widely scattered data, so it separates nothing.

The fix is to remove the signs before averaging, and squaring is the standard way to do it. Each $(x_i - \bar{x})^2$ is positive or zero, squares are far easier to handle in algebra than absolute values, and squaring gives a large deviation more weight than a small one, which is usually what you want from a measure of spread. The mean of the squared deviations is the **variance**, written σ^2.

Squaring changes the units: for times in minutes the variance is in minutes squared, which is hard to picture as a spread. Taking the square root undoes that and returns to the original units. The result is the **standard deviation** σ, and it is the one normally reported.

KEY · IB HL

$$\sigma^2 = \frac{\sum_{i=1}^{k} f_i(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{k} f_i x_i^2}{n} - \bar{x}^2, \qquad \sigma = \sqrt{\sigma^2}$$

The second form is the same quantity rearranged, and it is quicker by hand because it needs only the totals $\sum f_i x_i^2$ and $\sum f_i x_i$, not a separate subtraction for every value. The booklet writes μ where this chapter writes $\bar{x}$; for a set of data they are the same number.

The standard deviation σ is, roughly, the typical distance of a value from the mean. A small σ means the values lie close to the centre. A large σ means they are widely spread.

CAUTION In this course you find σ with your **GDC** (graphic display calculator), not by hand. Both forms appear in the HL formula booklet, so you are never asked to recall them; they are there so that you understand what the calculator computes. Know what σ *means*.

EXAMPLE 10.5

Return to the two classes, both with mean 60. Class A scored $58, 59, 60, 61, 62$; Class B scored $30, 45, 60, 75, 90$. Compare their standard deviations.
Both means are 60. For Class A the deviations are $-2, -1, 0, 1, 2$:

$$\sigma_A = \sqrt{\frac{4+1+0+1+4}{5}} = \sqrt{2} \approx 1.41.$$

For Class B the deviations are $-30, -15, 0, 15, 30$:

$$\sigma_B = \sqrt{\frac{900+225+0+225+900}{5}} = \sqrt{450} \approx 21.2.$$

The means are identical, but the standard deviations differ by a factor of 15. The single number σ captures exactly the difference that the mean alone does not show.

10.3.3 The effect of changing the data

Suppose every value in a data set is changed in the same way, either by adding a constant or by multiplying by a constant. Temperatures convert from Celsius to Fahrenheit; prices rise by a fixed tax. You can predict the new mean and standard deviation without recomputing anything.

Adding a constant c to every value shifts the whole distribution sideways. The centre moves by c, but the spread does not change. Multiplying every value by k stretches the distribution away from zero, scaling both the centre and the spread.

KEY If $y_i = kx_i + c$ for every value, then

$$\bar{y} = k\bar{x} + c, \qquad \sigma_y = |k|\,\sigma_x, \qquad \sigma_y^2 = k^2\sigma_x^2.$$

The constant c does not affect the spread. Moving every value by the same amount changes no distance between them. Only the multiplier k rescales σ.

EXAMPLE 10.6

A set of test marks has mean 54 and standard deviation 8. Every mark is scaled by multiplying by 1.5 and then adding 5. Find the new mean and standard deviation.
Here $k = 1.5$ and $c = 5$:

$$\bar{y} = 1.5(54) + 5 = 86, \qquad \sigma_y = 1.5 \times 8 = 12.$$

The $+5$ moves the mean but does not change the spread. Only the $\times 1.5$ affects σ.

The same transformation carries the median and the quartiles across. For $k > 0$ the map $x \mapsto kx + c$ is increasing, so it preserves order, and whichever value sat in the middle before

still sits in the middle afterwards. The new median is therefore $k \times (\text{old median}) + c$, and Q_1 and Q_3 obey the identical rule. When $k < 0$ the order reverses, and the new Q_1 is $k \times (\text{old } Q_3) + c$.

Questions often run this backwards: the target mean and standard deviation are given, and the transformation has to be found. Take the standard deviation equation first, because c does not appear in it.

EXAMPLE 10.7

A set of examination marks has mean 52 and standard deviation 16. The marks are rescaled by $y = ax + b$ with $a > 0$, so that the scaled marks have mean 65 and standard deviation 12. Find a and b. The original median was 50; write down the scaled median.

Only a appears in the standard deviation, so start there. With $a > 0$, $\sigma_y = a\sigma_x$:

$$12 = 16a \implies a = 0.75.$$

Now the mean equation has one unknown left:

$$65 = 0.75(52) + b = 39 + b \implies b = 26,$$

so $y = 0.75x + 26$. Since $a > 0$ the rescaling preserves order, and the median maps by that same rule:

$$\text{median}_y = 0.75(50) + 26 = 63.5.$$

Beginning with the mean equation instead leaves two unknowns in one equation and stalls, so starting with the standard deviation is the quicker route.

YOUR TURN 10.3 Measures of Dispersion

Quartiles and interquartile range

21. For the data $2,\ 4,\ 5,\ 7,\ 8,\ 10,\ 12,\ 14,\ 16,\ 20$, find each of the following.

(a) Q_1 (b) Q_3

(c) the interquartile range

22. Find the median and the interquartile range of $3,\ 8,\ 9,\ 12,\ 15,\ 18,\ 21,\ 25,\ 30$.

Variance and standard deviation

23. GDC Use your GDC for each part.

(a) Find the mean and standard deviation of $5,\ 8,\ 9,\ 11,\ 13,\ 14$.

(b) Find the mean and standard deviation of the values $10, 20, 30, 40$ with frequencies $2, 5, 8, 3$.

(c) Find the variance and the standard deviation of $6,\ 9,\ 9,\ 11,\ 14,\ 17$.

24. Find the variance and the standard deviation without a GDC.

(a) The data $2,\ 4,\ 6,\ 8,\ 10$. Give σ in exact form.

(b) A set of 20 values with $\sum x = 240$ and $\sum x^2 = 3100$.

Changing the data

25. Each data set is transformed. Find the new mean and the new standard deviation.

(a) Mean 50, standard deviation 6; every value is multiplied by 3.

(b) Mean 30, standard deviation 5; the value 10 is added to every data point.

(c) Mean 60, standard deviation 12; each value is transformed by $y = 0.5x + 20$.

(d) Mean 40, standard deviation 9; each value is transformed by $y = 2x - 5$.

(e) Mean 20, standard deviation 3; each value is transformed by $y = 50 - 2x$. Find also the new variance.

26. A set of marks has mean 45, standard deviation 10, $Q_1 = 38$ and $Q_3 = 52$. The marks are rescaled by $y = ax + b$ with $a > 0$, so that the new marks have mean 60 and standard deviation 8.

(a) Find a and b.

(b) Find the new quartiles and the new interquartile range.

27. A data set has mean 8.5, standard deviation 3, median 8, $Q_1 = 5$ and $Q_3 = 12$. Every value x is replaced by $y = 30 - 2x$.

(a) Find the new mean and the new standard deviation.

(b) Find the new median, Q_1 and Q_3.

Explaining spread

28. Explain each of the following.

(a) Why adding a constant to every value leaves the standard deviation unchanged.

(b) Why the interquartile range is preferred to the range when the data may contain outliers.

Concept check

29. A data set has mean 20 and standard deviation 3. Each value x is replaced by $y = 50 - 2x$. A student writes: "*new standard deviation* $= 50 - 2(3) = 44$." Identify the two errors, and find the correct new standard deviation.

10.4 Visualising Distributions

Numbers summarise a data set; a picture shows its shape. A mean and standard deviation tell you where the data sit and how far they spread. They do not tell you whether the distribution is symmetric, or leans to one side, or is concentrated at one end. Three kinds of display are used in this course.

10.4.1 Histograms

A **histogram** groups continuous data into classes and draws a bar for each, with height equal to the frequency. Unlike a bar chart for categories, the bars touch, because the horizontal axis is a continuous number line. In this course every class has the same width.

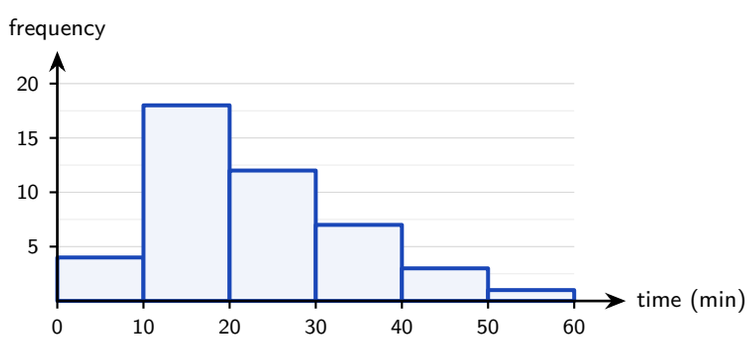


Figure 10.1 A histogram of 45 waiting times in classes of equal width. The bars touch because the horizontal axis is a continuous number line. The long tail to the right is positive skew.

The shape is what a histogram is for. It may be a symmetric peak, a long tail to the right (**positive skew**), as above, or a long tail to the left (**negative skew**).

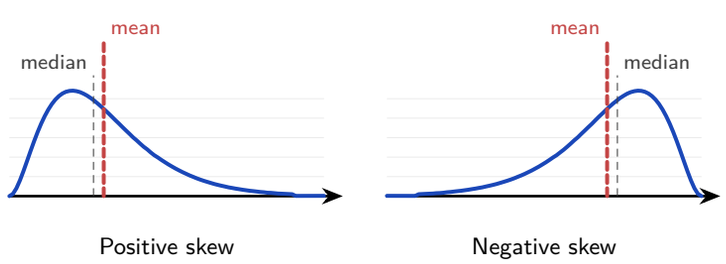


Figure 10.2 The long tail pulls the mean towards it, and the median hardly moves. With a tail to the right the mean sits above the median; with a tail to the left it sits below. Comparing the two numbers therefore

tells you which way a distribution leans.

10.4.2 Cumulative frequency

A **cumulative frequency** curve answers "how many values fall below this point?" You plot the running total of frequencies against the *upper* boundary of each class, then join the points with a smooth curve. The result is an S-shape, and the median and quartiles can be read from it directly. Enter from the vertical axis at one quarter, one half and three quarters of the total. Move across to the curve, then drop to the horizontal axis to read Q_1, Q_2 and Q_3.

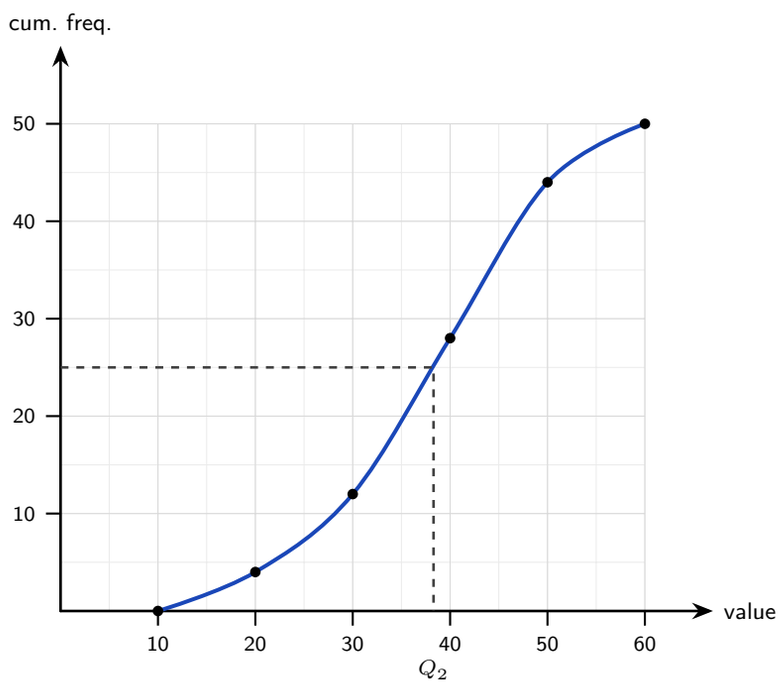


Figure 10.3 A cumulative frequency curve for 50 values. Each point is plotted at the *upper* boundary of its class. The dashed line reads the median: enter at half the total frequency, cross to the curve, drop to the value axis.

The dashed line shows the median. Enter at half the total frequency, here 25 of 50, cross to the curve, then drop to the value axis.

The same read-off gives any **percentile**: the p-th percentile is the value below which $p\%$ of the data lie. Enter the vertical axis at $\frac{p}{100} \times n$ and read across and down exactly as for the median. The median is the 50th percentile, and the quartiles Q_1 and Q_3 are the 25th and 75th. For the curve above, the 90th percentile is read off at a cumulative frequency of $0.9 \times 50 = 45$, giving a value of roughly 52.

EXAMPLE 10.8

The times taken by 80 competitors to finish a course are recorded below.

Time (min)	$0 \le t < 10$	$10 \le t < 20$	$20 \le t < 30$	$30 \le t < 40$
Frequency	8	24	28	20

Tabulate the cumulative frequencies, then estimate Q_1, the median, Q_3 and the 90th percentile from the curve.

Add the frequencies as you go, and record each running total against the *upper* boundary of its class. Nobody has finished before $t = 0$, so the table opens at $(0, 0)$.

Time less than	0	10	20	30	40
Cumulative frequency	0	8	32	60	80

Plot those five points and join them with a smooth increasing curve. Every reading then enters from the vertical axis, crosses to the curve, and drops to the time axis.

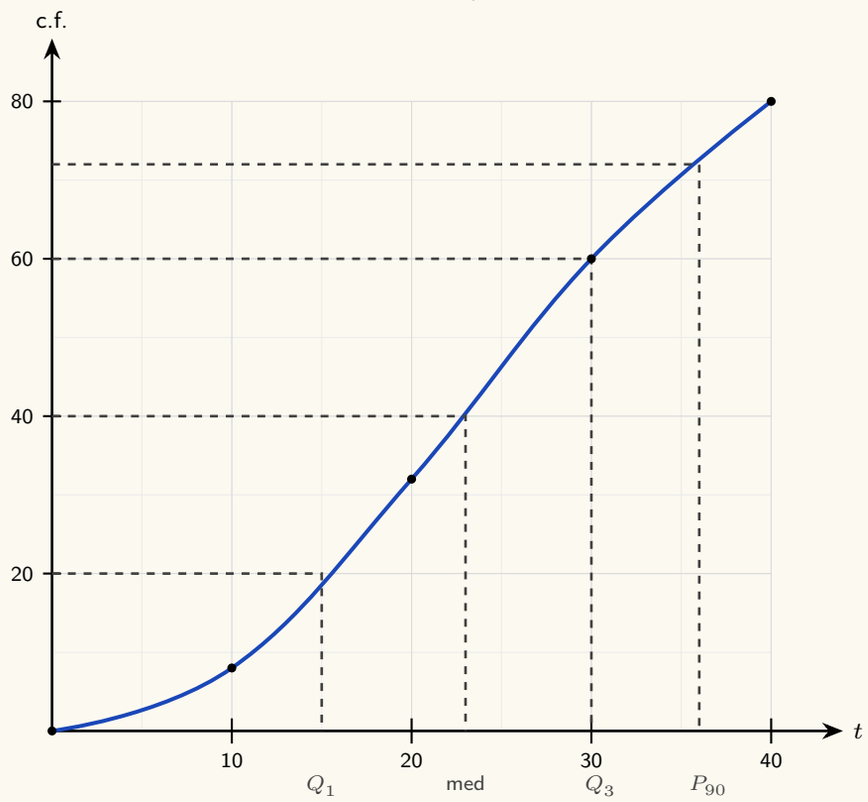


Figure 10.4 The cumulative frequency curve for the 80 finishing times, plotted at the upper class boundaries. Each read-off enters at a height on the c.f. axis, crosses to the curve, and drops to the time axis.

With $n = 80$, enter at $\frac{1}{4}(80) = 20$ for Q_1. That level lies between the plotted points $(10, 8)$ and $(20, 32)$, halfway between them, and the curve is met at about 15 minutes. The median needs $\frac{1}{2}(80) = 40$, a little above $(20, 32)$ and well below $(30, 60)$, giving about 23 minutes. For Q_3, enter at $\frac{3}{4}(80) = 60$, which is the plotted point $(30, 60)$ itself, so $Q_3 = 30$ minutes exactly. Hence

$$\text{IQR} \approx 30 - 15 = 15 \text{ minutes}.$$

The 90th percentile is read the same way, entering at $0.9 \times 80 = 72$. That falls between $(30, 60)$ and $(40, 80)$, giving about 36 minutes, so nine competitors in ten finished within 36 minutes. The upper boundary is the part to get right. A cumulative frequency counts everything that has happened *by* a value, so the 32 belongs at $t = 20$ and not at the midpoint 15; plotting at midpoints moves every quartile you later read off half a class width too low.

10.4.3 The outlier rule

With quartiles available, "outlier" can be defined precisely. A value is treated as an outlier if it lies more than 1.5 interquartile ranges beyond the nearer quartile.

KEY A value x is an **outlier** if

$$x < Q_1 - 1.5\,(\text{IQR}) \qquad \text{or} \qquad x > Q_3 + 1.5\,(\text{IQR}).$$

Worked through. To decide whether a value of 70 is an outlier in a data set with $Q_1 = 24$ and $Q_3 = 40$, first find $\text{IQR} = 40 - 24 = 16$. The upper boundary is

$$Q_3 + 1.5(\text{IQR}) = 40 + 1.5(16) = 40 + 24 = 64.$$

Since $70 > 64$, the value is an outlier. On a box plot it would be marked with a cross, with the whisker stopping at the largest value that is not an outlier.

10.4.4 Box-and-whisker plots

A **box-and-whisker plot** draws the **five-number summary**: minimum, Q_1, median, Q_3, maximum. The box spans the interquartile range, and a line inside it marks the median. The whiskers reach to the smallest and largest values that are *not* outliers. Any outlier is marked separately with a cross. In one small picture you see the centre, the spread, and the skew.

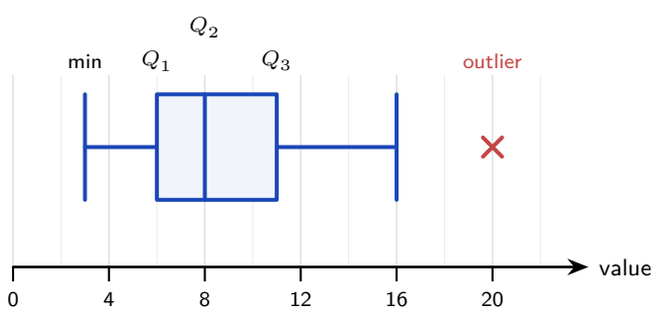


Figure 10.5 A box-and-whisker plot of the five-number summary. The box spans the interquartile range and the inner line marks the median. The right whisker stops at 16, the largest value that is not an outlier; the value 20 is marked with a cross.

A box positioned to the left, with a long right whisker, as above, indicates positive skew.

Here the value 20 is an outlier, by the outlier rule. The whisker therefore stops at 16, the largest value that is not an outlier, and 20 is marked with a cross. Comparing two box plots side by side is a quick way to contrast two distributions.

A box plot also gives a first hint about the normal distribution of Ch. 12. If the box and whiskers are roughly symmetric about the median, the data *may* be normally distributed. Clear skew suggests that they are not.

EXAMPLE 10.9

The masses, in kilograms, of eleven parcels are

$$8,\ 11,\ 13,\ 14,\ 16,\ 16,\ 18,\ 21,\ 22,\ 25,\ 41.$$

Find the five-number summary, determine whether 41 is an outlier, and state where each whisker of the box plot stops.

The values are already ordered, with $n = 11$. The median is the 6th value, 16. The lower quartile is the median of the five values below it, $8, 11, 13, 14, 16$, so $Q_1 = 13$; the upper quartile is the median of the five above, $18, 21, 22, 25, 41$, so $Q_3 = 22$. The five-number summary is

$$8, \quad 13, \quad 16, \quad 22, \quad 41.$$

Now apply the $1.5 \times$ IQR outlier rule. Here IQR $= 22 - 13 = 9$, so the boundaries are

$$Q_1 - 1.5(9) = 13 - 13.5 = -0.5, \qquad Q_3 + 1.5(9) = 22 + 13.5 = 35.5.$$

Since $41 > 35.5$, the heaviest parcel is an outlier, and it is marked with a cross. No value falls below -0.5, so there is no outlier at the low end.

The left whisker therefore reaches the minimum, 8. The right whisker stops at 25, the largest mass that is not an outlier.

The maximum in the summary is still 41: the whisker stops at 25, the data do not. Reading a five-number summary back off a plot therefore means taking the maximum from the cross, not from the end of the whisker.

10.4.5 Comparing two distributions

One box plot describes a data set. Two drawn on the same scale compare them, and examination questions often ask for that comparison. Four things should be read off: centre, spread, shape and outliers. A good answer mentions the centre, the spread, and at least one shape or outlier feature, always in context.

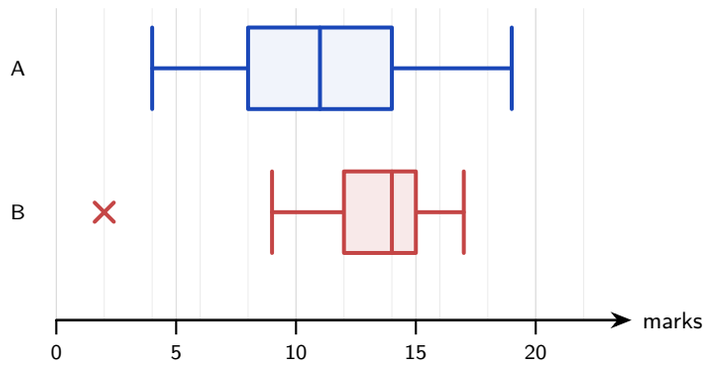


Figure 10.6 Two classes' marks on the same scale. Class B has the higher median and the smaller interquartile range, so it did better and more consistently. Class A is roughly symmetric; class B is negatively skewed and has one low outlier, marked with a cross.

Written out, that comparison reads: class B has the higher median (14 against 11), so B performed better on average. B also has the smaller interquartile range (3 against 6), so B's marks are more consistent. A is roughly symmetric while B is negatively skewed, and B contains one low outlier that A does not. Notice that every claim names a statistic and then says what it means for the classes. A comparison that only says "B is better" gives no evidence for the claim.

YOUR TURN 10.4 Visualising Distributions

Five-number summaries and outliers

30. Consider the data $4,\ 6,\ 7,\ 9,\ 10,\ 12,\ 13,\ 15,\ 18$.

(a) Find the median.

(b) Find Q_1 and Q_3.

(c) Write down the five-number summary.

(d) State the interquartile range.

31. Apply the $1.5 \times$ IQR outlier rule.

(a) A data set has $Q_1 = 30$ and $Q_3 = 50$. Determine whether 95 is an outlier.

(b) A data set has $Q_1 = 12$ and $Q_3 = 20$. Determine whether 2 is an outlier.

(c) A data set has $Q_1 = 9$ and $Q_3 = 19$. Find both outlier boundaries.

32. The data below are to be drawn as a box plot.

$$2,\ 15,\ 16,\ 18,\ 19,\ 21,\ 22,\ 24,\ 26,\ 27$$

(a) Find the five-number summary.

(b) Show that 2 is an outlier and that there is no outlier at the upper end.

(c) State where each whisker of the box plot ends.

Cumulative frequency and percentiles

33. The masses of 60 apples are recorded below.

Mass m (g)	Frequency
$100 \le m < 120$	6
$120 \le m < 140$	14
$140 \le m < 160$	22
$160 \le m < 180$	12
$180 \le m < 200$	6

(a) Write out the cumulative frequency table, using the upper class boundaries.
(b) Draw the cumulative frequency curve and use it to estimate the median.
(c) Estimate the interquartile range.
(d) Estimate the 90th percentile.
(e) Estimate the number of apples heavier than 170 g.

34. A cumulative frequency curve is drawn for 80 values.

(a) State the cumulative frequencies at which you read off Q_1, the median and Q_3.
(b) State the cumulative frequency at which you read off the 15th percentile.

Reading and comparing displays

35. Answer each part about the shape of a display.

(a) Explain why the bars of a histogram touch, while those of a bar chart do not.
(b) A histogram of waiting times has its tallest bar at the left and a long tail to the right. Name the skew and state whether the mean or the median is larger.
(c) A box plot has minimum 10, $Q_1 = 14$, median 16, $Q_3 = 25$ and maximum 40. Describe the skew, giving a reason.

36. The test marks of two classes, P and Q, have these five-number summaries (minimum, Q_1, median, Q_3, maximum).

Class	Min	Q_1	Median	Q_3	Max
P	32	48	55	61	78
Q	40	52	58	63	70

(a) Find the interquartile range of each class.
(b) Compare the marks of the two classes, referring to a measure of centre and a measure of spread.

Concept check

37. To draw a cumulative frequency curve for the apple masses in Question 33, a student plots the cumulative frequencies $6, 20, 42, 54, 60$ at the class midpoints $110, 130, 150, 170, 190$. Explain the error, and state its effect on the median read from the curve.

10.5 Correlation

Everything so far has described one variable. Galton's question was different: he had *two* measurements on each family, parent height and child height, and wanted to know whether they were linked. This is the study of **bivariate** data, meaning data in which two variables are recorded for each individual. Start by looking at the data.

10.5.1 Scatter diagrams

A **scatter diagram** plots each pair (x, y) as a point. The cloud of points has a shape, and the shape is the relationship. If y tends to rise as x rises, the correlation is **positive**. If y falls as x rises, it is **negative**. If there is no tilt, there is no linear correlation. The tighter the points cluster around a straight line, the **stronger** the correlation.

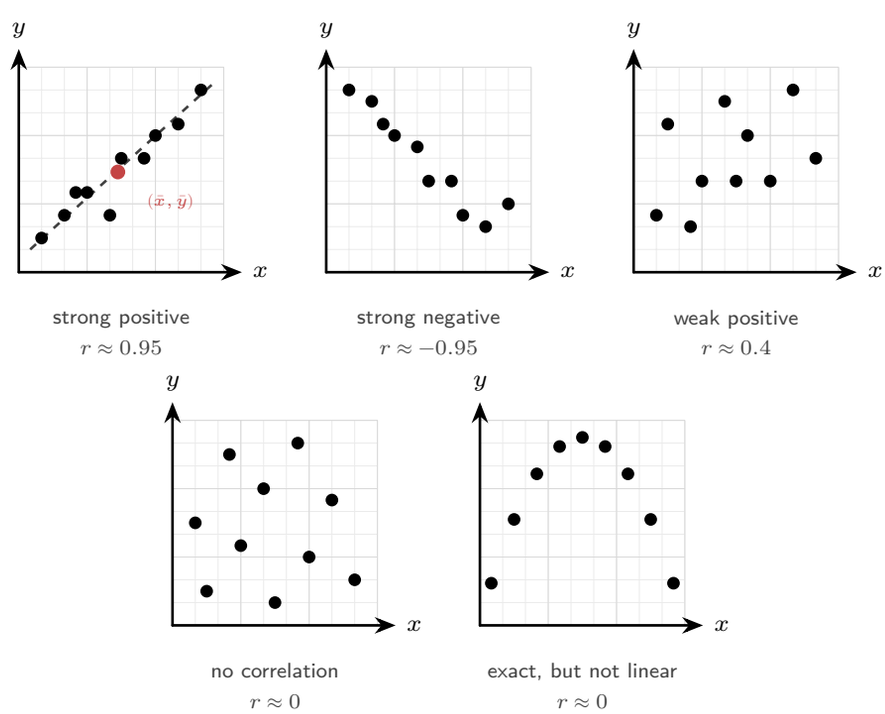


Figure 10.7 Five clouds and the value of r each gives. The sign of r is the direction of the tilt and the size of r is how close the points lie to a line. The last panel is the warning: the points lie exactly on a curve, so the relationship could not be stronger, yet r is near zero because the curve is not straight.

Read the panels from left to right. The first two tilt clearly, one up and one down, and the points sit close to a line: those are strong correlations, positive and negative. The third tilts upward but the points scatter widely, so the correlation is weak. The fourth has no tilt at all.

The last panel carries a warning. Its points lie exactly on a curve, so the two variables are perfectly related, yet r is near zero. That is because r only measures how close the points are to a *straight* line.

The first panel also shows a **line of best fit** drawn by eye: a straight line balancing the points above and below it. To position it, first compute the **mean point** $(\bar{x}, \bar{y})$, the mean of the x-values paired with the mean of the y-values, marked in scarlet. Draw your line through it. Every correctly drawn line of best fit passes through the mean point; in §10.6 the calculator makes this exact.

10.5.2 Pearson's correlation coefficient

You can judge correlation well by eye, but not precisely. **Pearson's product-moment correlation coefficient**, written r, turns the strength and direction into a single number between -1 and 1.

DEF Pearson's correlation coefficient r measures the strength and direction of the **linear** relationship between two variables. It satisfies $-1 \le r \le 1$.

The sign of r gives the direction, and its size gives the strength. The extremes $r = \pm 1$ mean the points lie exactly on a straight line; $r = 0$ means no linear relationship. A rough reading:

$\lvert r \rvert$ near	Interpretation
1	strong linear correlation
0.5	moderate linear correlation
0	weak or no linear correlation

You compute r with your GDC by entering the two lists. In this course you read and interpret r; you do not compute it from the formula.

A linear change of scale leaves r unchanged. Record the heights in metres rather than centimetres, or convert the temperatures from Celsius to Fahrenheit, and every point moves while r stays exactly where it was, because r measures how close the points lie to a straight line, and stretching or shifting an axis keeps collinear points collinear. Multiplying by a

negative number reverses one axis, which flips the sign of r and leaves its size alone.

Those words, strong and weak, are only a rough guide, because how impressive a given r is depends on how much data produced it. With five points a value of $r = 0.7$ is unremarkable, since five points can fall this close to a line by chance alone. With fifty points the same 0.7 would be hard to explain away. A **critical value** settles the question: it is the smallest $|r|$ that counts as evidence of real correlation for a sample of that size.

KEY Compare $|r|$ with the **critical value** for your sample size n. If $|r|$ is larger, the correlation is taken as evidence of a genuine linear relationship; if it is smaller, the data are consistent with no correlation at all. Critical values shrink as n grows.

EXT Critical values of r belong to *Applications and Interpretation*, not to Analysis and Approaches, so no AA examination will ask for one. They are set out here because they answer the obvious question of how large r has to be, and because they make clear that a value of r means nothing until you know how many points produced it.

EXAMPLE 10.10

A study of 10 pairs gives $r = 0.58$. The critical value for $n = 10$ is 0.632. What can be concluded?

Here $|r| = 0.58 < 0.632$, so the correlation does not reach the critical value. There is not enough evidence of a linear relationship in these data.

Notice what this does *not* say. It does not say the two variables are unrelated, only that ten points are too few to demonstrate it. The same $r = 0.58$ from 30 pairs, where the critical value is about 0.361, would support the opposite conclusion.

CAUTION r measures *linear* association only. Data lying perfectly on a parabola can have r near 0, because the relationship, though exact, is not straight. Always look at the scatter diagram before trusting r.

10.5.3 Correlation is not causation

A strong r says two variables move together. It does not say one causes the other. There are at least three quite different reasons a correlation can appear without the causal link it seems to suggest, and they must be told apart.

A third variable drives both. Ice-cream sales and drowning deaths rise and fall together across the year. Neither causes the other. Summer heat raises both, because hot weather sends people to the ice-cream shop and to the water. A variable like this, sitting behind both and explaining the link, is called a **confounding variable**.

The cause runs the other way. A variable that looks like the cause may in fact be the effect.

This reversal turns up in school data. Students who receive extra support often have lower grades than those who do not, and the figures can be read as though the support were doing harm. The direction is the other way round: low grades are why the support was offered in the first place. Judge the support by comparing those students with how they were doing before, not with everyone else.

Coincidence, found by searching. Between 2000 and 2009, cheese consumption per person in the United States tracked the number of people who died by becoming entangled in their bedsheets, with r above 0.94. Nobody thinks these are connected. If you compare thousands of pairs of unrelated variables, some pairs will match closely by chance alone, and a determined search will almost certainly find some.

Keep the third case in mind whenever a surprising correlation is reported. A high r is easy to find if you go looking through enough data.

None of this makes correlation useless. The link between smoking and lung cancer was found statistically, in the 1950s, long before anyone could say how a cigarette damages a cell. The correlation was what made the question urgent, and biology then had to supply the mechanism. That is the usual division of labour: mathematics can show that two things move together and how strongly, but the case for cause has to be built outside the mathematics.

CAUTION A high correlation never establishes cause. There may be a hidden third variable, the cause may run the other way, or the link may be coincidence. Causation must be argued from the context, never from r alone.

YOUR TURN 10.5 Correlation

Describing correlation

38. Describe the correlation indicated by each value of r.

(a) $r = -0.85$ (b) $r = 0.12$

(c) $r = 0.97$ (d) $r = -0.40$

39. GDC Find Pearson's correlation coefficient for each data set.

(a) $x = 1, 2, 3, 4, 5$ and $y = 2, 5, 4, 7, 9$

(b) $x = 2, 4, 6, 8, 10$ and $y = 5, 9, 10, 14, 17$

40. For the heights x (in cm) and masses y (in kg) of a group of adults, $r = 0.78$. State the value of r in each case, giving a reason.

(a) The heights are converted to metres.

(b) Every mass is increased by 5 kg.

(c) Each height is replaced by $200 - x$, its distance below 200 cm.

The mean point and a line of best fit

41. The table gives six pairs of values.

x	2	4	5	7	8	10
y	3	6	6	9	11	13

(a) Find the mean point $(\bar{x}, \bar{y})$.

(b) Draw a scatter diagram and a line of best fit by eye through the mean point.

(c) Use your line to estimate y when $x = 9$.

Correlation and cause

42. Explain each result about r.

(a) The seven points with $x = -3, -2, \ldots, 3$ lie exactly on the parabola $y = x^2$. Explain why $r = 0$ even though the relationship is exact.

(b) A study finds a strong positive correlation between shoe size and reading ability in children. Explain why shoe size does not cause reading ability.

43. Each finding below shows a strong correlation. State whether the most likely explanation is a confounding variable, reverse causation or coincidence, and explain.

(a) People who use a walking stick walk more slowly than people who do not.

(b) Over twelve years, the number of films a particular actor appeared in each year and the number of people who drowned in swimming pools that year give $r = 0.67$.

(c) In a town, monthly sales of hot chocolate and the monthly number of people with a cold rise and fall together.

Concept check

44. Data set A has $r = -0.92$ and data set B has $r = 0.55$. A student says: "*B shows the stronger correlation, because* $0.55 > -0.92$." Identify the error, and state which data set shows the stronger correlation.

45. For a data set, $r = 0.9$. A student says: "*this means* y *increases by* 0.9 *each time* x *increases by* 1." Explain why this is false.

10.6 Linear Regression

If the scatter diagram shows a clear linear trend, you can do more than measure it. You can draw the line that best fits the data and use it to predict. The question is what "best" should mean.

10.6.1 The least-squares line

For any candidate line, each data point sits some vertical distance above or below it. That distance is the **residual**, the error of the line at that point. The **least-squares regression line** is the line that makes the total of the squared residuals as small as possible. Squaring, again, stops positive and negative errors from cancelling, and it counts large errors much more heavily than small ones.

KEY The **regression line of y on x** has the form

$$y = ax + b,$$

chosen to minimise the sum of squared vertical residuals. It always passes through the mean point $(\bar{x}, \bar{y})$.

You find a and b from your GDC. You must also be able to say what they *mean* in context. The gradient a is the change in y for each one-unit increase in x. The intercept b is the predicted y when $x = 0$.

Worked through. Suppose a GDC gives the regression line of exam score y on hours studied x as $y = 6.2x + 41$, valid for students who studied between 2 and 9 hours. To predict the score of a student who studied 5 hours, substitute $x = 5$ into the line:

$$y = 6.2(5) + 41 = 31 + 41 = 72.$$

In context, the gradient says each additional hour of study raises the predicted score by 6.2 marks. The intercept 41 is the predicted score of a student who did not study at all. That reading is itself a mild extrapolation, because $x = 0$ lies below the data range.

Predicting at $x = 20$ is unsafe. The data only cover 2 to 9 hours, so 20 lies far outside the known range, and the linear trend may not continue there.

10.6.2 Interpolation and extrapolation

Predicting *inside* the range of the data is **interpolation**, and it is reliable when the correlation is strong. Predicting *outside* the range is **extrapolation**, and it is unreliable. You are assuming the pattern continues where you have no evidence. The further beyond the data you go, the less trustworthy the prediction.

CAUTION A regression line is only trustworthy for prediction within the range of the original data, and only when $|r|$ is reasonably large. Extrapolating far beyond the data, or fitting a line to weakly correlated points, produces predictions worth little.

10.6.3 The two regression lines

There is an asymmetry hidden in "least squares." The line of y on x minimises *vertical* residuals, because it treats y as the quantity to predict from x. If instead you want to predict x from y, you minimise *horizontal* residuals, giving a different line: the **regression line of x on y**.

KEY Use the regression line of y **on** x to predict y from a given x. Use the regression line of x **on** y to predict x from a given y. Both lines pass through $(\bar{x}, \bar{y})$, and they coincide only when $r = \pm 1$.

The rule is simple. Choose the line whose name begins with the quantity you want to predict. Choosing the wrong line gives a different prediction, as the second example below shows.

The two gradients are linked: the gradient of y on x times the gradient of x on y equals r^2, and r takes the sign the two gradients share. The further r is from ± 1, the further apart the two lines lie, and the more a wrong choice of line costs.

EXAMPLE 10.11

For a set of bivariate data, $\bar{x} = 10$ and $\bar{y} = 25$. The regression line of y on x is $y = 2x + 5$. Verify that the mean point lies on this line.

Substitute $x = \bar{x} = 10$:

$$y = 2(10) + 5 = 25 = \bar{y}. \checkmark$$

Every least-squares line of y on x passes through $(\bar{x}, \bar{y})$, so this is a quick check on a computed line.

EXAMPLE 10.12

For the same data the regression line of x on y is $x = 0.35y + 1.25$. Estimate x when $y = 30$, and compare with what rearranging the other line would have given.

The quantity wanted is x, so use the line of x on y:

$$x = 0.35(30) + 1.25 = 11.75.$$

Rearranging the y on x line instead would give $x = \dfrac{30 - 5}{2} = 12.5$, a different answer. Both lines pass through the mean point, but they are not the same line, and only one of them was fitted to predict x.

The two agree only when $r = \pm 1$. Here the product of the gradients is $2 \times 0.35 = 0.7$, which is r^2, so $r = \sqrt{0.7} \approx 0.84$.

YOUR TURN 10.6 Linear Regression

Finding and using the line

46. GDC Find the regression line of y on x for each data set.

(a) $x = 1, 2, 3, 4, 5$ and $y = 3, 5, 4, 8, 10$

(b) $x = 2, 3, 5, 7, 8$ and $y = 14, 12, 9, 6, 4$

47. Using a regression line.

(a) A regression line is $y = 2.5x + 4$. Predict y when $x = 6$.

(b) A regression line $y = -1.5x + 30$ was fitted using data with $2 \leq x \leq 15$. Predict y when $x = 8$.

(c) Explain why using that same line to predict y at $x = 40$ is unreliable.

The gradient and intercept in context

48. For car journeys of between 50 km and 400 km, the regression line of fuel used y (litres) on distance x (km) is $y = 0.064x + 1.8$.

(a) Interpret the value 0.064 in context.

(b) Interpret the value 1.8 in context, and comment on how far it can be relied on.

(c) Predict the fuel used on a journey of 250 km.

49. For used cars between 1 and 10 years old, the regression line of price y (thousands of euros) on age x (years) is $y = -1.9x + 24.5$.

(a) Interpret the value -1.9 in context.

(b) Predict the price of a car that is 6 years old.

(c) Use the line to predict the price of a 15-year-old car, and explain why the answer shows that the line should not be used there.

The mean point and the two lines

50. The mean point of a data set is $(8, 20)$.

(a) The regression line of y on x has gradient 1.5. Find its equation.

(b) You are given a value of y and want to predict the corresponding x. State which regression line you should use, and why.

51. For a set of data, the regression line of y on x is $y = 1.6x + 4$ and the regression line of x on y is $x = 0.5y - 0.5$.

(a) Estimate y when $x = 10$.

(b) Estimate x when $y = 12$.

(c) Find the value of r.

(d) Find the mean point.

Concept check

52. For a data set, the regression line of y on x is $y = 2x + 3$ and the regression line of x on y is $x = 0.4y - 0.2$. To estimate x when $y = 15$, a student rearranges the first line: "$x = \frac{15-3}{2} = 6$." Identify the error, and find the correct estimate.

Chapter 10 · Statistics

Reflections

TOK Much of what you are told about the world arrives as a number: your grade, the mass of an atom, how happy a country is. Mass really is a quantity, and 12.011 describes carbon fairly. Happiness is not a quantity in that sense: a score of 7.2 is the output of a questionnaire somebody designed, with questions somebody chose. So every statistic makes two claims, that the thing measured has a size at all and that this number captures it, and either can fail while the arithmetic stays perfect. Which of the numbers used to describe you measure something that was already there, and which create the thing they claim to measure?

IM The word "statistics" shares a root with "state." It began in the 18th century as the numerical bookkeeping of nations: population, taxes, soldiers, harvests. A census still decides how seats in a parliament and money for schools are shared out, so what a government chooses to count, and how it defines the categories it counts in, is a political decision before it is a mathematical one. Can a census ever be neutral?

RW Galton's "regression toward mediocrity" is now called regression to the mean, and it affects decisions in sport, business and medicine. A pilot who flies badly and is scolded usually does better next time. The reason is not the scolding. An extreme performance is, on average, followed by a less extreme one. When $r < 1$, part of an extreme value is chance that does not repeat, so the next value is on average closer to the mean. Mistaking this statistical drift for cause and effect has misled trainers, investors, and doctors for over a century.

The same error appears in school data. A class that scores unusually low one term tends to score closer to average the next, whatever was changed in between, so any change made after a bad term will appear to have worked. It is also why a drug trial needs a control group: patients often enrol when their illness is at its worst, and an illness measured at its worst tends to be measured better next time on its own.

EXT Your calculator returns r without showing its working. The formula behind it is

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \ \sum (y_i - \bar{y})^2}},$$

and it should be read rather than memorised.

Take one term of the numerator, $(x_i - \bar{x})(y_i - \bar{y})$. It asks where a single point sits relative to *both* means. Above average in both, or below average in both, and the product is

positive; above in one and below in the other, and it is negative. The numerator adds these products over all the points.

That total depends on the units, and the denominator divides them out, so r is a pure number. It also forces $-1 \leq r \leq 1$: the size of the numerator can never exceed the denominator, a fact called the Cauchy–Schwarz inequality, and equality holds exactly when every point lies on one straight line. That is why $r = \pm 1$ means perfect linearity, and nothing else does.

End-of-Chapter Problems
Chapter 10: Statistics

1. The times, in minutes, taken by 40 students to complete a puzzle are summarised below.

Time (min)	$0 \le t < 10$	$10 \le t < 20$	$20 \le t < 30$	$30 \le t < 40$	$40 \le t < 50$
Frequency	4	11	15	8	2

(a) Estimate the mean time. [3]
(b) Write down the modal class. [1]
(c) State the class containing the median. [2]

2. A data set is $12, 15, 18, 19, 21, 22, 24, 25, 27, 28, 30, 46$.

(a) Find the five-number summary. [3]
(b) Determine whether 46 is an outlier, showing your working. [3]
(c) Describe the shape of the distribution. [1]

3. GDC The table shows the number of hours x eight students revised and their test score y.

x	2	3	5	6	8	9	11	12
y	10	14	15	20	24	27	30	34

(a) Find Pearson's correlation coefficient and interpret it. [3]
(b) Find the equation of the regression line of y on x. [2]
(c) Interpret, in context, the gradient and the y-intercept of your line. [2]
(d) Use your line to estimate the score of a student who revised for 7 hours. [2]
(e) Explain whether your line should be used to estimate the score for 25 hours. [2]

4. A set of data has mean 45 and standard deviation 8.

(a) Each value is increased by 12. Write down the new mean and standard deviation. [2]
(b) Instead, each value is multiplied by 1.5. Write down the new mean and standard deviation. [2]
(c) Instead, each value is transformed by $y = 20 - 0.5x$. Write down the new mean and standard deviation. [2]

5. A region has populations of 4000, 10 000, and 6000 in three towns A, B, and C. A stratified sample of 100 households is required.

(a) Find the number of households to sample from each town. [3]

(b) The mean household sizes in the samples from A, B and C are 3.2, 2.5 and 2.9 people. Estimate the mean household size in the region. [2]

6. The cumulative frequencies of a data set of 60 values are given.

Value $\leq$	10	20	30	40	50
Cumulative frequency	4	16	34	52	60

(a) Draw a cumulative frequency curve. [3]

(b) Use it to estimate the median and the interquartile range. [4]

(c) Estimate the 70th percentile. [2]

7. In a test, 20 students in class A had a mean score of 72, and 30 students in class B had a mean score of 65. Find the mean score of all 50 students combined. [3]

8. A frequency distribution of the values $1, 2, 3, 4$ has frequencies $5, 8, k, 4$. The mean is 2.5. Find the value of k. [4]

9. Two machines fill bottles. A sample from each gives the same mean volume of 500 mL, but machine X has standard deviation 2 mL and machine Y has standard deviation 9 mL.

(a) Explain what this tells you about the two machines. [2]

(b) State which machine is more reliable, with a reason. [2]

10. GDC The table gives the area x (in m^2) and monthly rent y (in \$) of five apartments.

x	10	20	30	40	50
y	14	20	25	33	38

(a) Find the regression line of y on x. [2]

(b) Use it to estimate the rent of a 35 m^2 apartment. [2]

(c) Find the regression line of x on y, and use it to estimate the area of an apartment renting for \$30. [3]

11. GDC A data set is $4, 6, 7, 8, 9, 10, 55$.

(a) Find the mean and the median. [3]

(b) The value 55 is removed. Find the new mean and median. [2]

(c) Comment on which measure of centre is more affected by the outlier, and why. [2]

12. A student's final grade is 40% coursework and 60% final exam. The student scores 70 on coursework and 85 on the exam. Find the final grade. [2]

13. The masses, in grams, of seven eggs are $52, 55, 58, 60, 61, 63, 90$.

(a) Find Q_1, Q_3, and the interquartile range. [3]
(b) Show that 90 is an outlier. [2]
(c) State, with a reason, whether the mean or the median better represents a typical egg. [2]

14. For eight used cars aged between 2 and 10 years, x is the age in years and y the resale value in thousands of dollars. The data give $\sum x = 48$ and $\sum y = 152$, and the regression line of y on x has gradient -2.

(a) Find $\bar{x}$ and $\bar{y}$, and hence the equation of the regression line. [2]
(b) Interpret, in context, the gradient and the y-intercept of the line. [2]

15. The five-number summaries of test scores in two classes are given below.

	Min	Q_1	Median	Q_3	Max
Class A	32	45	55	62	90
Class B	42	50	58	66	74

(a) Write down the interquartile range for Class A. [1]
(b) Show that the score of 90 in Class A is an outlier. [2]
(c) Compare the two distributions, referring to centre and spread. [2]
(d) State, with a reason, whether the scores in Class B may be normally distributed. [2]

16. GDC The waiting times, in minutes, of a group of customers are summarised below. The estimated mean waiting time is 20 minutes.

Time (min)	$0 \le t < 10$	$10 \le t < 20$	$20 \le t < 30$	$30 \le t < 40$
Frequency	6	k	12	8

(a) Show that $k = 18$. [3]
(b) Write down the modal class. [1]
(c) State the class containing the median. [2]
(d) Use your GDC to estimate the standard deviation. [2]

17. A set of examination marks x has mean 62 and standard deviation 10. The marks are scaled linearly to $y = ax + b$, $a > 0$, so that the scaled marks have mean 70 and standard deviation 8.

(a) Find the values of a and b. [4]
(b) A student's original mark was 75. Find their scaled mark. [1]
(c) The median of the original marks was 61. Write down the median of the scaled marks. [1]

18. GDC The table shows the maximum temperature x (°C) and the number of ice creams

sold y on six days.

x (°C)	18	20	23	25	27	30
y	110	125	150	160	175	195

(a) Find Pearson's correlation coefficient r and comment on it. [2]
(b) Find the equation of the regression line of y on x. [2]
(c) Interpret, in context, the gradient of your line. [1]
(d) Estimate the number of ice creams sold on a day with maximum temperature $24\,°\text{C}$. [2]
(e) Explain why the line should not be used when the maximum temperature is $40\,°\text{C}$. [1]

19. For a bivariate data set, the regression line of y on x is $y = 2x + 1$ and the regression line of x on y is $x = 0.4y + 1$.

(a) Find $\bar{x}$ and $\bar{y}$. [4]
(b) Estimate the value of x when $y = 20$. [2]
(c) Explain why the regression line of y on x should not be used in part (b). [1]

20. A supermarket wants to survey its customers.

(a) State the sampling method used when every 20th customer entering the store is selected. [1]
(b) State the sampling method used when customers are grouped by age band and sampled in proportion to the size of each band. [1]
(c) State the sampling method used when the first 30 customers on Monday morning are selected. [1]
(d) Explain why the method in (c) is likely to produce a biased sample. [2]
(e) For the method in (a), about 1200 customers are expected on the survey day. State how many customers will be sampled, and explain why the first customer chosen should be selected at random from the first 20 to enter. [2]

21. A group of n students has a mean mark of 68. When 5 students with mean mark 60 join the group, the overall mean falls to 66. Find n. [4]

22.

(a) A set of points lies exactly on the line $y = 5 - 2x$. Write down the value of r. Every y-value is then multiplied by -3; write down the new value of r. [2]
(b) A study finds a strong positive correlation between the hours children spend watching television and their body mass. Explain why this does not show that television causes weight gain, and suggest a possible third variable that could explain both. [2]
(c) Heights are recorded in centimetres and r is calculated. The heights are converted to metres. State, with a reason, the effect on the value of r. [2]

23. GDC The table gives the monthly mean air temperature in Tokyo, in °C, taken from the Japan Meteorological Agency climatological normals for 1991–2020.

Month	J	F	M	A	M	J	J	A	S	O	N	D
Temp (°C)	5.4	6.1	9.4	14.3	18.8	21.9	25.7	26.9	23.3	18.0	12.5	7.7

(a) Find the five-number summary of the twelve values. [4]

(b) Show that the data contain no outliers. [3]

(c) The JMA publishes an annual normal of $15.8\,°\text{C}$. Calculate the mean of the twelve monthly values and compare it with the published figure. [3]

(d) Describe the shape of the distribution, and explain why these twelve values tell you very little about the temperature on any particular day in Tokyo. [3]

24. A data set of n values has mean $\bar{x}$. A new value equal to $\bar{x}$ is added to the set.

(a) Show that the mean is unchanged. [3]

(b) Show that the standard deviation decreases (or stays the same). State the condition for equality. [4]

25. Five positive integers have a mean of 8, a median of 7, and a mode of 4 (occurring exactly twice). Find all possible sets of five integers. [6]

26. The regression line of y on x for a data set is $y = 0.8x + 3$, and the regression line of x on y is $x = 1.05y - 2$.

(a) It can be shown that the gradient of the line of y on x is $r\dfrac{\sigma_y}{\sigma_x}$ and the gradient of the line of x on y is $r\dfrac{\sigma_x}{\sigma_y}$. Given that $\sigma_x = 5$, find σ_y. [4]

(b) It can be shown that the product of the two gradients equals r^2. Find r. [3]

27. A set of n data values has mean $\bar{x}$ and standard deviation σ. Prove algebraically that adding a constant c to every value leaves σ unchanged, starting from the definition $\sigma^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$. [5]

28. A sports club has 900 members, made up of 225 families of exactly four. The membership register lists each family together, always in the same order: father, mother, elder child, younger child. A sample of 45 members is to be taken.

(a) The club secretary takes a systematic sample. Find the sampling interval k, and describe how the 45 members are selected. [2]

(b) State one advantage of systematic sampling over simple random sampling for a register of this size. [2]

(c) Explain why this particular systematic sample is very likely to be badly unrepresentative, and state what the sample would consist of if the random start were 2. [3]

(d) The club has 360 junior members and 540 senior members. Find the number of each in a stratified sample of 45. [2]

(e) All 45 selected members are sent a questionnaire and 18 reply. Explain why the 18 replies may not represent the club, and suggest one thing the secretary could do about it. [3]

29. A city council wants to estimate the mean distance cycled per week by its $12\,000$ adult residents. Two methods are proposed.

Method A: leave questionnaires at the counter of the city's three bicycle shops and use the forms returned.

Method B: telephone every 200th name on the electoral register of all $12\,000$ residents.

(a) State the sampling method used in Method A, and explain why it will almost certainly overestimate the mean. [3]

(b) State the sampling method used in Method B, and find the size of the sample obtained. [2]

(c) The council instead takes a stratified sample of 300 residents by age band. Using the table below, find the number of residents sampled from each band. [3]

Age band	18–29	30–49	50–69	70+
Residents	3000	4800	2800	1400

(d) Explain why stratifying by age band is a sensible choice for this particular investigation. [2]

(e) The council suggests improving Method A by collecting ten times as many returned forms. Explain why this would not fix the problem identified in (a). [2]

30. A university has 2400 students, half of them men and half women. A researcher wants to know student opinion on the library's opening hours. She stands at the library entrance and interviews students as they arrive, stopping once she has 40 men and 40 women.

(a) Write down the fraction of the university's students that she interviews. [1]

(b) Explain why this is not a stratified sample, even though the sample matches the university's split exactly. [2]

(c) Explain why the choice of location makes the sample unsuitable for the question being asked. [2]

(d) Describe a better method for this investigation, and justify your choice. [3]

Challenge Questions

31. **How much can one value move a summary?** Consider the data set

$$3,\ 5,\ 7,\ 9,\ 11.$$

(a) Write down the mean and the median. [2]

(b) The largest value is replaced by k, where $k \geq 9$. Find the mean in terms of k, and write down the median. [3]

(c) Find the value of k for which the mean is 20. [2]

(d) Explain why the median is 7 for *every* $k \geq 9$. [2]

(e) Return to the original data. The smallest value, 3, is changed to 8. Find the new mean and the new median, and show that the new median is not one of the original five values. Explain why no limit can be placed on how far a change to a single value can move the mean. [3]

(f) A data set has an odd number of values, $n = 2m + 1$, ordered as $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$, so that its median is $x_{(m+1)}$. Exactly one value is changed, to any real number. Prove that the new median lies between $x_{(m)}$ and $x_{(m+2)}$ inclusive. Hence explain what is meant by calling the median *resistant* and the mean *not resistant*. *(Hint: count how many of the unchanged values are at most $x_{(m+2)}$)* [5]

(g) Prove that the smallest number of values that must be changed to make the median as large as you wish is $\lceil \frac{n}{2} \rceil$, the smallest integer greater than or equal to $\frac{n}{2}$. Treat odd and even n separately. [4]

(h) The fraction of a data set that must be changed before a summary can be made as large as you wish is called its *breakdown point*: it is $\frac{1}{n}$ for the mean and about $\frac{1}{2}$ for the median. Investigate the breakdown point of the interquartile range, or of a *trimmed mean* that discards the largest 10% and the smallest 10% of the values before averaging. Discuss what your findings suggest about choosing a summary for data that may contain recording errors. [3]

32. GDC **Combining groups and splitting the variance.** Group A has n_1 values with mean $\bar{x}_1$ and variance σ_1^2. Group B has n_2 values with mean $\bar{x}_2$ and variance σ_2^2. The two groups are combined into one set of $N = n_1 + n_2$ values, with mean $\bar{x}$ and variance σ^2.

(a) Using $\sum x^2 = n(\sigma^2 + \bar{x}^2)$ for each group, show that

$$\sigma^2 = \frac{n_1(\sigma_1^2 + \bar{x}_1^2) + n_2(\sigma_2^2 + \bar{x}_2^2)}{N} - \bar{x}^2.$$

[4]

(b) Group A has 12 values with mean 50 and standard deviation 4. Group B has 8 values with mean 60 and standard deviation 3. Find the mean and the standard deviation of the

combined set. [4]

(c) Define the *within-group* and *between-group* variances

$$V_{\text{w}} = \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{N}, \qquad V_{\text{b}} = \frac{n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2}{N}.$$

Show that $\sigma^2 = V_{\text{w}} + V_{\text{b}}$, and verify this with your answers to part (b). [5]

(d) Show that $V_{\text{b}} = \frac{n_1 n_2}{N^2}(\bar{x}_1 - \bar{x}_2)^2$. Hence show that $\sigma^2 \geq V_{\text{w}}$, state the condition for equality, and interpret this result. [3]

(e) The data are now split into k groups, where group i has n_i values, mean $\bar{x}_i$ and variance σ_i^2. State and prove the corresponding result for σ^2. [3]

(f) The ratio $\frac{V_{\text{b}}}{\sigma^2}$ is the fraction of the variance accounted for by which group a value belongs to. Choose a data set that falls naturally into groups, such as the heights of students grouped by year, and discuss what a ratio close to 0, and one close to 1, would tell you. Suggest how you might decide whether a ratio is large enough to matter. [3]

33. Which centre is closest? For a data set $x_1, x_2, \ldots, x_n$ and a real number a, define

$$S(a) = \sum_{i=1}^{n}(x_i - a)^2 \qquad \text{and} \qquad T(a) = \sum_{i=1}^{n}|x_i - a|.$$

Each one measures how far the data lie from a. This question finds the value of a that makes each one least.

(a) For the data $1, 2, 4, 9, 14$, show that $S(a) = 5a^2 - 60a + 298$. Hence find the value of a that minimises S, and the least value of S. State which summary of the data this value of a is. [3]

(b) For any data set with mean $\bar{x}$ and standard deviation σ, prove that

$$S(a) = \sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - a)^2.$$

Hence show that S is least when $a = \bar{x}$, and that its least value is $n\sigma^2$. [4]

(c) For the data in part (a), find $T(4)$ and $T(6)$. Explain why, on any interval between two consecutive data values, T is a linear function of a whose gradient is the number of data values less than a minus the number of data values greater than a. [4]

(d) Hence show that, when n is odd, T is least when a is the median. For the data $1, 2, 4, 9, 14, 20$, show that T takes its least value at every a with $4 \leq a \leq 9$, and find that least value. [4]

(e) In the data of part (a), the value 14 is replaced by 140. Find the new values of a that minimise S and T, and explain how your answer relates to Challenge Question 31. [2]

(f) The value of a that minimises the largest distance, $\max_i|x_i - a|$, is the *midrange* $\frac{1}{2}(x_{\min} + x_{\max})$. Investigate how the value of a that minimises $\sum|x_i - a|^p$ changes as p increases from 1, through 2, to large values, and explain what the choice of p says about how heavily large deviations are weighted. [3]

34. Whose quartile? In this course the lower quartile Q_1 is the median of the values below the median, with the median left out of both halves when n is odd. Other textbooks, calculators and spreadsheets use other rules. For ordered data $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$, compare four methods for Q_1.

- **Method A:** the median of the lower half, leaving the median out when n is odd.
- **Method B:** the median of the lower half, including the median in both halves when n is odd.
- **Method C:** the value at position $\frac{n+1}{4}$ in the ordered list.
- **Method D:** the value at position $1 + \frac{n-1}{4}$ in the ordered list.

In Methods C and D, a position that is not a whole number is read by linear interpolation: position 2.25 means $x_{(2)} + 0.25\left(x_{(3)} - x_{(2)}\right)$. The upper quartile Q_3 is found in the same way, from the upper half in Methods A and B, and at positions $\frac{3(n+1)}{4}$ and $1 + \frac{3(n-1)}{4}$ in Methods C and D.

(a) For the data $2, 4, 5, 7, 8, 10, 13$, find Q_1 and Q_3 by each method. [4]

(b) For the data $2, 4, 5, 7, 8, 10, 13, 21$, find Q_1, Q_3 and the interquartile range by each method. [4]

(c) Using the rule that an outlier is a value more than $1.5 \times$ IQR above Q_3, decide by each method whether 21 is an outlier in the data of part (b). Comment on your results. [3]

(d) Using parts (a) and (b), and further data sets of your own, conjecture which methods always give the same Q_1 when n is odd, and which when n is even. [2]

(e) Prove your conjecture for odd n, treating $n = 4k + 1$ and $n = 4k + 3$ separately. [4]

(f) Find out which method your GDC and one spreadsheet program use. Discuss whether the choice of method matters more for small or for large data sets, and suggest how you could decide which method gives the best estimate of the quartile of the population a sample was drawn from. [3]

STATISTICS AND PROBABILITY

Chapter 11

Probability

11

SYLLABUS 1.10 4.5 4.6 4.11 4.13

Probability puts a number on how likely an event is. That step lets you measure risk, compare it and act on it, so that a decision can rest on more than a feeling about what might happen.

Here is a decision that depends on such a number. A health service screens everyone in an age group for a rare disease. The test is a good one. Of every 100 people who have the disease, it correctly identifies 99. Of every 100 people who do not have it, it correctly clears 95. About one person in a hundred has the disease.

Your result comes back positive. How likely is it that you have the disease?

A common answer is about 99%, and in published studies many doctors given a question of this kind have answered in the same way. The correct answer is about 17%. The test is not faulty and none of the numbers are unusual. The error is in the reasoning. That 99% says how the test behaves *when the disease is present*, and the quantity you actually want is a different one: how likely the disease is *now that the result is positive*. Those two numbers are not the same, and this chapter builds the rule that connects them.

So probability gives clear answers, and it is easy to apply a clear answer to the wrong question. Two habits guard against that, and both are practised throughout the chapter. The first is to state precisely which event you are asking about. The second is to keep track of where a probability came from: some are calculated from the symmetry of a fair die, others only measured by counting what happened in 200 trials, and the two are found, and can be trusted, in different ways.

We begin with a simple experiment, a single roll of a die. From there the chapter builds the language of events, the rules for combining them, conditional probability and independence, tree diagrams, and finally Bayes' theorem. By the end of this chapter you will be able to find probabilities from Venn diagrams, tree diagrams and counting, test events for independence, find conditional probabilities, and use Bayes' theorem to reverse a conditional probability. When we return to the test in the last section, you will be able to find its answer in one step.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

11.1 The Language of Probability

Probability measures how likely something is, on a scale from 0 (impossible) to 1 (certain). To compute it we need to be precise about what "something" means.

Five words are used precisely from here on. The die is used to illustrate each one.

- **Experiment**: any process with an uncertain result. Rolling a die is an experiment.
- **Trial**: one repetition of the experiment. One roll is one trial.
- **Outcome**: the result of a single trial, such as a 3.
- **Sample space** U: the set of all possible outcomes. For one die, $U = \{1, 2, 3, 4, 5, 6\}$.
- **Event**: any subset of the sample space, that is, a collection of outcomes you are interested in. "An even number" is the event $\{2, 4, 6\}$.

When every outcome is equally likely, probability is just counting: the fraction of outcomes that belong to the event.

KEY · IB

$$P(A) = \frac{n(A)}{n(U)}$$

Here $n(A)$ is the number of outcomes in event A, and $n(U)$ the number in the whole sample space. Every probability obeys $0 \le P(A) \le 1$.

Worked through. When a fair die is rolled once, the sample space has $n(U) = 6$ equally likely outcomes, so the probabilities of an even number and of a number greater than 4 are

$$P(\text{even}) = \frac{n(\{2,4,6\})}{6} = \frac{3}{6} = \frac{1}{2}, \qquad P(>4) = \frac{n(\{5,6\})}{6} = \frac{2}{6} = \frac{1}{3}.$$

Counting works only because the die is fair. If the outcomes were not equally likely, this formula would not apply.

11.1.1 Relative frequency

When there is no symmetry to count with, the probability can be estimated by experiment. Repeat the trial many times and record how often the event occurs. The fraction of trials in which it happens is the **relative frequency** of the event, and it serves as an experimental estimate of the probability.

KEY After n trials,

$$P(A) \approx \frac{\text{number of trials in which } A \text{ occurred}}{n}$$

The estimate is rough when n is small, and it becomes more stable as n grows. This is the difference between the two kinds of probability you will meet: a *theoretical* probability comes from the structure of the experiment (a fair die, by symmetry), while an *experimental* one comes from data, and is what you fall back on when there is no symmetry to use.

EXAMPLE 11.1

A drawing pin is dropped 200 times and lands point-up 78 times. Estimate the probability that a dropped pin lands point-up.

There is no symmetry argument for a drawing pin, so use the relative frequency:

$$P(\text{point-up}) \approx \frac{78}{200} = 0.39.$$

A further 200 drops would almost certainly not give exactly 78 again, but the estimate becomes more accurate as the number of trials increases.

11.1.2 The complement

Sometimes the event you want is awkward to count, but its opposite is easy. The **complement** of A, written A', is the event that A does *not* happen. Since A either happens or it does not, their probabilities must fill the whole scale.

KEY · IB

$$P(A) + P(A') = 1 \qquad \Longrightarrow \qquad P(A') = 1 - P(A)$$

The complement is the route to take whenever the event itself is hard to count. In particular, when a question contains the phrase "at least one," the complement is usually the quicker route.

EXAMPLE 11.2

A fair coin is tossed three times. Find the probability of getting at least one head.

Counting every way to get one, two, or three heads is possible but slow. The complement of "at least one head" is "no heads at all," a single outcome: TTT.

$$P(\text{at least one head}) = 1 - P(\text{TTT}) = 1 - \frac{1}{8} = \frac{7}{8}.$$

One subtraction replaces a count of seven outcomes. The saving becomes far greater as the number of tosses increases.

11.1.3 Expected number of occurrences

If an event has probability p and you repeat the experiment n times, you expect it to happen about np times.

KEY If an event has probability p and the experiment is repeated n times, the **expected number of occurrences** is

$$\text{expected number} = np.$$

This is not a guarantee. It is a long-run average. If you repeated the whole set of n trials many times and averaged the counts, that average would be close to np.

EXAMPLE 11.3

A fair die is rolled 60 times. How many sixes would you expect?
Each roll gives a six with probability $\frac{1}{6}$, so over 60 rolls:

$$\text{expected sixes} = 60 \times \frac{1}{6} = 10.$$

You will rarely get exactly 10, but the average over many sets of 60 rolls will be close to 10.

YOUR TURN 11.1 The Language of Probability

Equally likely outcomes

1.

(a) A fair die is rolled. Find the probability that the number is odd.

(b) A card is drawn from a deck of 52. Find the probability that it is a heart.

(c) A bag holds 4 red, 6 green and 5 blue counters. Find the probability that a counter drawn at random is green.

(d) A letter is chosen at random from the word STATISTICS. Find the probability that it is an S.

(e) A spinner numbered 1 to 8 is spun. Find the probability that it lands on a multiple of 3.

2. List the sample space in each case.

(a) Two fair coins are tossed. List the sample space, and find the probability of exactly one head.

(b) Three fair coins are tossed. List the sample space, and find the probability of exactly two heads.

Relative frequency and expected number

3. A spinner has five sectors numbered 1 to 5, but it is not fair. In 400 spins it lands on 3 exactly 96 times.

(a) Estimate the probability of landing on 3.

(b) Using that estimate, find the expected number of 3s in 250 spins.

(c) Explain why the probability in question 2(a) can be found without any experiment, while the one in part (a) of this question cannot.

4.

(a) A fair die is rolled 90 times. Find the expected number of 5s.

(b) Records show that a train was late on 45 of the last 300 days. Estimate the probability that it is late on a given day, and the expected number of late days in a month of 20 working days.

The complement

5.

(a) Given $P(A) = 0.35$, find $P(A')$.

(b) A letter is chosen at random from the word PROBABILITY. Find the probability that it is not a B.

(c) A fair coin is tossed four times. Find the probability of at least one tail.

(d) A fair die is rolled twice. Find the probability that at least one of the rolls is not a six.

Concept check

6. A fair coin has landed heads five times in a row. A student says: "*the next toss is more likely to be tails, because the results have to even out.*" Explain why the student is wrong.

7. A student writes: "*the probability of a six on one roll is* $\frac{1}{6}$*, so the probability of at least one six in six rolls is* $6 \times \frac{1}{6} = 1$." Identify the error, and find the correct probability.

11.2 Counting the Outcomes HL

The formula $P(A) = \dfrac{n(A)}{n(U)}$ is only as useful as your ability to find the two numbers in it. For a die you list the outcomes. For a committee of four chosen from twelve people you cannot: there are 495 of them. The counting tools of Ch. 2, §2.8 are what make the formula work at that scale, and this section uses them to find probabilities.

11.2.1 Which tool, and when

Three questions decide which tool to use: *how many objects are you choosing from, how many are you taking, and does the order matter?* The last one decides between the two formulas below.

Order matters when the positions are different from one another: first, second and third place; president, secretary and treasurer; the order of speakers at a meeting. Order does not matter when you end up with a set: a committee, a hand of cards, six lottery numbers. Choosing Ali then Ben gives the same committee as choosing Ben then Ali, and a count that treats those as different has counted the committee twice.

KEY · IB HL Choosing r objects from n distinct objects:

$$\text{order matters: } {}^nP_r = \frac{n!}{(n-r)!}, \qquad \text{order does not: } {}^nC_r = \frac{n!}{r!\,(n-r)!}.$$

Every unordered selection corresponds to $r!$ *ordered ones, which is the only difference between the two.*

CAUTION Decide this from the wording of the question. Read the question for a word that fixes an order: *arrange*, *rank*, *in a row*, *first and second*. Without one, it is usually a combination.

11.2.2 Probability by counting

Once both counts use the same tool, the ratio is the probability. Count the sample space with one method, count the event with the same method, and divide.

To count the event, split the objects into the groups the question names, choose the required number from each group with ${}^{n}C_{r}$, and multiply the counts, because the event needs one choice from each group. Many counting probability questions have this shape.

Worked through. To find the probability that a committee of 4, chosen at random from 7 girls and 5 boys, contains exactly 3 girls, note first that a committee is a set, so order does not matter and both counts are combinations. The sample space is every committee of 4 from the 12 students:

$$n(U) = {}^{12}C_4 = 495.$$

For the event, read the requirement as an *and*: 3 of the 7 girls *and* 1 of the 5 boys. And multiplies Ch. 2, §2.8, so the two counts are multiplied:

$$n(A) = {}^{7}C_3 \times {}^{5}C_1 = 35 \times 5 = 175.$$

So

$$P(\text{exactly 3 girls}) = \frac{175}{495} = \frac{35}{99} \approx 0.354.$$

Notice the shape of the event count: one combination for each group the question splits the people into, multiplied together.

EXAMPLE 11.4

Five cards are dealt from a standard pack of 52. Find the probability that exactly two of them are kings.

A hand is a set, so combinations again:

$$n(U) = {}^{52}C_5 = 2\,598\,960.$$

Split the pack into the two groups the question cares about: the 4 kings and the 48 non-kings. Take 2 from the first *and* 3 from the second, so the counts multiply:

$$n(A) = {}^{4}C_2 \times {}^{48}C_3 = 6 \times 17\,296 = 103\,776.$$

$$P(\text{exactly two kings}) = \frac{103\,776}{2\,598\,960} \approx 0.0399.$$

The 48 in the second factor matters: taking ${}^{4}C_2 \times {}^{50}C_3$ instead would let the other two kings into the hand, so it would count hands with three or four kings, some of them more than once.

TIP Write down $n(U)$ before anything else. It fixes what an outcome *is* for the rest of the question, and both counts must use that same definition of an outcome.

11.2.3 At least one, again

The complement of §11.1 is at its most useful here, because "at least one" spreads over many cases while "none" is a single count.

EXAMPLE 11.5

A box holds 20 components, of which 3 are defective. Four are drawn at random without replacement. Find the probability that at least one is defective.

Counting the cases directly means one defective, or two, or three: three separate counts. The complement is one count:

$$P(\text{none defective}) = \frac{{}^{17}C_4}{{}^{20}C_4} = \frac{2380}{4845} = \frac{28}{57}.$$

Therefore

$$P(\text{at least one defective}) = 1 - \frac{28}{57} = \frac{29}{57} \approx 0.509.$$

Three bad components in twenty, and a sample of four is already more likely than not to contain one. So even a small sample has a good chance of finding a defective component.

11.2.4 When order does matter

EXAMPLE 11.6

Eight athletes run a final, and all orders of finishing are equally likely. Find the probability that a particular three of them take the first three places, in some order.

Here the positions are distinguishable, so the sample space is ordered:

$$n(U) = {}^{8}P_3 = 336$$

ordered ways to fill the three medal positions. The three named athletes can fill them in $3! = 6$ orders, so

$$P = \frac{6}{336} = \frac{1}{56}.$$

The same answer comes from counting unordered: ${}^{8}C_3 = 56$ possible medal-winning sets, one of which is the named three. Either tool works, provided the top and the bottom use the same one.

The counting principles, factorials, ${}^{n}P_r$ and ${}^{n}C_r$ are developed in Ch. 2, §2.8. The binomial distribution of Ch. 12 is this section's ${}^{n}C_r$ applied to repeated trials.

YOUR TURN 11.2 Counting the Outcomes HL

Choosing the counting tool

8. In each part decide whether order matters, find $n(U)$ and then the probability.

(a) Three of nine books are chosen at random to take on holiday. Find the probability that one particular book is among them.

(b) Gold, silver and bronze medals are won by three of nine finalists, and all finishing orders are equally likely. Find the probability that Lena wins gold and her sister Mia wins silver.

(c) Five people sit in a row of five seats in a random order. Find the probability that Ana and Ben sit next to each other.

Probability by counting groups

9. A team of 5 is chosen at random from 6 women and 4 men.

(a) Find the number of possible teams.

(b) Find the probability that the team contains exactly 3 women.

(c) Find the probability that the team is all women.

(d) Find the probability that the team contains at least one man.

10. A box contains 15 light bulbs, of which 4 are faulty. Three are chosen at random without replacement.

(a) Find the probability that none is faulty.

(b) Find the probability that at least one is faulty.

(c) Find the probability that exactly two are faulty.

When order matters

11. A four-digit code is made from four different digits chosen from $1, 2, \ldots, 9$, and every such code is equally likely.

(a) Find the probability that the digits of the code are in increasing order.

(b) Find the probability that the code starts with 1 and ends with 9.

12. Four of the 12 students in a club are chosen at random to be president, vice-president, secretary and treasurer. Find the probability that Maya and Leo are both given a post,

(a) counting ordered selections for both $n(U)$ and $n(A)$;

(b) counting unordered selections for both, and check that the answers agree.

Concept check

13. Three of nine books are chosen at random. To find the probability that one particular book is chosen, a student writes: "$n(A) = {}^8C_2 = 28$ *and* $n(U) = {}^9P_3 = 504$, *so the probability is* $\frac{28}{504} = \frac{1}{18}$." Identify the error, and find the correct probability.

11.3 Combining Events

Many questions involve more than one event. You want the probability that a student plays football *or* tennis, or that a card is a king *and* a heart. Two operations combine them: the **union** $A \cup B$ ("A or B or both") and the **intersection** $A \cap B$ ("A and B together").

The word *or* needs care, because everyday English and mathematics do not use it the same way. "Tea or coffee?" offers you one drink, not both. In mathematics $A \cup B$ always includes the case where both happen. This is called the **inclusive or**, and it is the reason the addition rule below needs a correction term. Questions that mean one but not both usually say so explicitly.

A **Venn diagram** makes the structure visible. The rectangle is the sample space, written U and called the **universal set**: everything under discussion lies inside it. Each circle is an event, and every point of the diagram lies in exactly one of four regions. Combining an event with a complement names each region precisely.

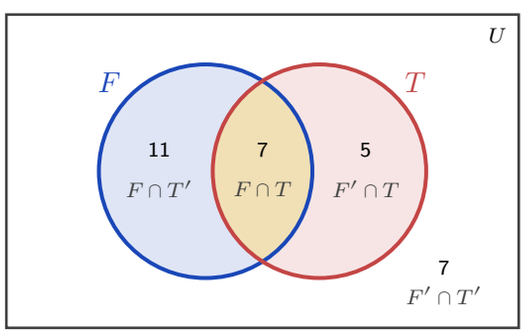


Figure 11.1 The four regions of a two-event Venn diagram, with the class of 30 filled in. Football only is $F \cap T'$, both is $F \cap T$, tennis only is $F' \cap T$, and neither is $F' \cap T'$, the region outside both circles. The four counts add to 30.

Fill a Venn diagram from the middle outwards. Put the overlap in first, then subtract it from each circle total to get the "only" regions, then use the grand total to find the outside. Working the other way round double counts the overlap, which is the same mistake the addition rule is built to correct.

11.3.1 The addition rule

If you add $P(A)$ and $P(B)$ to find $P(A \cup B)$, you count the overlap twice: once in A, once in B. Subtracting it once corrects the double count.

KEY · IB

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

When two events cannot happen together, they are **mutually exclusive**: their intersection

contains no outcomes at all, written $A \cap B = \emptyset$, where $\emptyset$ is the **empty set**. Then $P(A \cap B) = 0$ and the rule simplifies.

KEY · IB For mutually exclusive events:

$$P(A \cup B) = P(A) + P(B)$$

Worked through. In a class of 30, 18 students play football, 12 play tennis, and 7 play both. For a student chosen at random, the probability of playing football or tennis comes from the addition rule directly:

$$P(F \cup T) = \frac{18}{30} + \frac{12}{30} - \frac{7}{30} = \frac{23}{30}.$$

The probability of playing neither is the complement of "football or tennis":

$$P(\text{neither}) = 1 - \frac{23}{30} = \frac{7}{30}.$$

The Venn diagram above shows the same thing: 11 play only football, 7 play both, 5 play only tennis, and 7 play neither, totalling 30.

EXAMPLE 11.7

A single card is drawn from a standard deck of 52. Find the probability that it is a king or a heart.

There are 4 kings and 13 hearts, but the king of hearts is both, so it is counted in each:

$$P(\text{king} \cup \text{heart}) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}.$$

Forgetting to subtract the king of hearts gives $\frac{17}{52}$, which counts the king of hearts twice.

Three events need three circles, and the same instruction applies: start in the middle. The centre region belongs to all three events, and every other region is found by subtracting what has already been placed.

EXAMPLE 11.8

In a year group of 40 students, 20 study French, 18 study German and 15 study Spanish. Of these, 8 study French and German, 6 study German and Spanish, 7 study French and Spanish, and 3 study all three. A student is chosen at random. Find the probability that the student studies (a) exactly one of the three languages, (b) none of them.

Start with the 3 in the centre. The 8 who study French and German include those 3, so $8 - 3 = 5$ study those two and no other. In the same way $6 - 3 = 3$ study German and Spanish only, and $7 - 3 = 4$ study French and Spanish only.

Now the "only" regions. French has 20 altogether, of which $5+3+4=12$ are already placed, leaving 8 for French only. German gives $18-(5+3+3)=7$, and Spanish gives $15-(4+3+3)=5$.

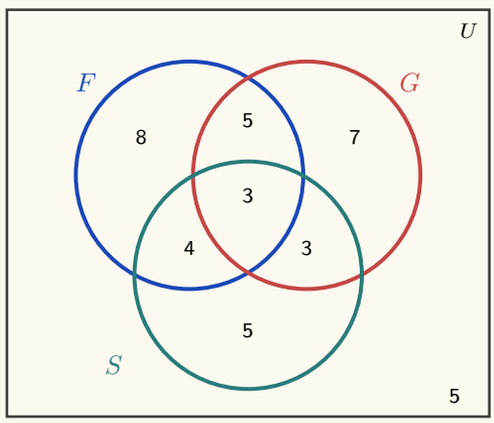


Figure 11.2 The three-circle diagram, filled from the centre outwards. The seven inner counts total 35, so 5 of the 40 students lie outside all three circles.

(a) Exactly one language means the three outer regions: $8+7+5=20$, so

$$P(\text{exactly one}) = \frac{20}{40} = \frac{1}{2}.$$

(b) The seven regions inside the circles total 35, so 5 students study none:

$$P(\text{none}) = \frac{5}{40} = \frac{1}{8}.$$

No formula was used. Once the diagram is filled, every probability is a count divided by 40.

11.3.2 Sample space diagrams and two-way tables

When an experiment has two parts, drawing the whole sample space is often faster than any formula. A **sample space diagram** plots one part on each axis. For two dice, the 36 equally likely outcomes form a grid, and any event is just a set of grid points to count.

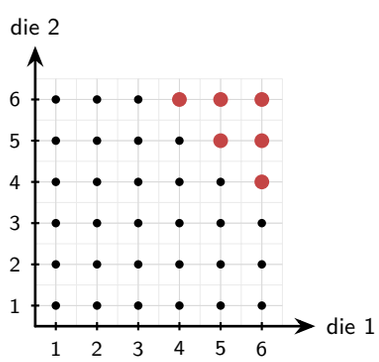


Figure 11.3 The 36 equally likely outcomes of rolling two dice. The six larger scarlet points are the outcomes whose total is at least 10.

The six larger scarlet points are the outcomes with a total of at least 10. There are six of

them, so

$$P(\text{total} \geq 10) = \frac{6}{36} = \frac{1}{6}.$$

No rule was needed, only a picture and a count.

A **two-way table** (a table of outcomes) does the same for grouped data: it splits a population by two characteristics at once, and every probability is a cell, row, or column total divided by the grand total.

EXAMPLE 11.9

The table shows the 100 students of a year group by age group and club membership.

	Club	No club	Total
Junior	14	26	40
Senior	21	39	60
Total	35	65	100

A student is chosen at random. Find the probability that the student (a) is a junior in a club, (b) is in a club or is a junior.

(a) Read the single cell where the two conditions meet:

$$P(J \cap C) = \frac{14}{100} = 0.14.$$

(b) Apply the addition rule with the row and column totals:

$$P(J \cup C) = \frac{40}{100} + \frac{35}{100} - \frac{14}{100} = \frac{61}{100} = 0.61.$$

The table already contains every count used in both calculations. We will return to this table in the next section, where it answers conditional questions just as directly.

YOUR TURN 11.3 Combining Events

The addition rule

14.

(a) Given $P(A) = 0.4$, $P(B) = 0.5$ and $P(A \cap B) = 0.2$, find $P(A \cup B)$.

(b) Events A and B are mutually exclusive, with $P(A) = 0.6$ and $P(B) = 0.3$. Find $P(A \cup B)$.

(c) Given $P(A \cup B) = 0.8$, $P(A) = 0.5$ and $P(B) = 0.4$, find $P(A \cap B)$.

(d) Given $P(A) = 0.3$, $P(B) = 0.4$ and $P(A \cup B) = 0.7$, determine whether A and B are mutually exclusive.

15.

(a) A card is drawn from a deck of 52. Find the probability that it is neither a heart nor a picture card (jack, queen or king).

(b) In a group of 40 people, 25 like coffee, 20 like tea and 10 like both. Find the probability that a person chosen at random likes coffee or tea.

(c) For the group in part (b), find the probability that the person likes neither.

Venn diagrams

16.

(a) Two events satisfy $n(A \text{ only}) = 8$, $n(B \text{ only}) = 5$, $n(A \cap B) = 4$ and $n(\text{neither}) = 3$. Find $P(A)$.

(b) Given $n(U) = 50$, $n(A) = 22$, $n(B) = 18$ and $n(A \cap B) = 7$, find $n(A \cap B')$ and $n(A' \cap B')$.

17. A survey of 50 households finds that 28 have a cat, 22 have a dog and 15 have a rabbit. Ten have a cat and a dog, 6 have a dog and a rabbit, 5 have a cat and a rabbit, and 3 have all three.

(a) Draw a Venn diagram showing the number of households in each region.

(b) A household is chosen at random. Find the probability that it has at least two of the three kinds of pet.

(c) Find the probability that it has none of the three.

Sample space diagrams and two-way tables

18. Two fair dice are rolled.

(a) Find the probability that the total is 7.
(b) Find the probability that the two numbers differ by exactly 1.
(c) Find the probability that the two dice show the same number.

19. The table shows 80 people classified by sex and by writing hand.

	Left	Right	Total
Male	7	33	40
Female	5	35	40
Total	12	68	80

One person is chosen at random.

(a) Find the probability that the person is left-handed.
(b) Find the probability that the person is female and right-handed.
(c) Find the probability that the person is female or left-handed.

Concept check

20. A card is drawn from a deck of 52. A student writes:
"$P(\textit{heart or king}) = \frac{13}{52} + \frac{4}{52} = \frac{17}{52}$." Identify the error, and find the correct probability.

11.4 Conditional Probability and Independence

Information changes probability. The chance that a random person has a particular illness is small, but the chance *given* that they tested positive is different. **Conditional probability** is the probability of one event once you know another has happened.

Knowing B has occurred **reduces the sample space** to B alone. Within that smaller set, you ask what fraction also lies in A.

KEY · IB

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B) \neq 0$$

Read $P(A \mid B)$ as "the probability of A given B." The denominator $P(B)$ is the new, restricted sample space; the numerator is the part of it where A also holds.

Worked through. Using the class of 30 from before, suppose a student chosen at random is known to play football. To find the probability that they also play tennis, restrict to the 18 football players. Of those, 7 also play tennis:

$$P(T \mid F) = \frac{P(T \cap F)}{P(F)} = \frac{7/30}{18/30} = \frac{7}{18}.$$

Note this differs from $P(F \mid T) = \frac{7}{12}$. Conditioning is not symmetric: the order matters.

11.4.1 The multiplication rule

Rearranging the conditional formula gives a rule for the probability that two events both happen.

KEY

$$P(A \cap B) = P(A)\,P(B \mid A)$$

In words: the chance of both is the chance of the first times the chance of the second given the first. This is exactly how you move along the branches of a tree diagram, as the next section shows.

EXAMPLE 11.10

A drawer holds nine pens, four of which are red. Two pens are taken at random, one after the other, without replacement. Find the probability that both are red.

Let A be "the first pen is red" and B "the second pen is red". Four of the nine pens are red, so $P(A) = \frac{4}{9}$.

For the second factor, assume A has happened. Eight pens remain and three of them are red, so $P(B \mid A) = \frac{3}{8}$. The multiplication rule combines the two:

$$P(A \cap B) = P(A)\,P(B \mid A) = \frac{4}{9} \times \frac{3}{8} = \frac{12}{72} = \frac{1}{6}.$$

The second factor is $\frac{3}{8}$ and not $\frac{4}{9}$ because the first pen is not put back: removing a red pen changes the number of red pens and the total together, so both parts of the fraction drop by one.

11.4.2 Independence

Two events are **independent** when knowing one tells you nothing about the other, so $P(A \mid B) = P(A)$. Substituting this into the multiplication rule gives a cleaner, symmetric test.

KEY · IB A and B are **independent** if and only if

$$P(A \cap B) = P(A)\,P(B).$$

Provided $0 < P(B) < 1$, this is equivalent to $P(A \mid B) = P(A) = P(A \mid B')$, which is not itself printed in the booklet. To test independence, compute both sides of either form and compare. The conditional form states directly what independence means: whether B happened or did not happen, the probability of A is unchanged.

EXAMPLE 11.11

A fair coin and a fair die are tossed together. Let H be "heads" and S be "rolling a six." Are H and S independent?

Here $P(H) = \frac{1}{2}$ and $P(S) = \frac{1}{6}$. The outcome "heads and a six" is one of the 12 equally likely combined outcomes:

$$P(H \cap S) = \frac{1}{12} = \frac{1}{2} \times \frac{1}{6} = P(H)\,P(S).$$

The two sides match, so the events are independent, which fits intuition: a coin cannot influence a die.

The two-way table of §11.3 passes the same test. There, $P(C \mid J) = \frac{14}{40} = 0.35$ and $P(C) = \frac{35}{100} = 0.35$: knowing a student is a junior does not change the chance they are in a club, so age group and club membership are independent.

Exact independence like this is unusual in real data. Two characteristics recorded on the same group of people are often linked at least a little, and then the test fails. The calculation is identical when it comes out that way, and so is the shape the answer must take: the two numbers first, then the verdict.

EXAMPLE 11.12

The table classifies the 200 students of a year group by whether they take music and whether they take drama.

	Drama	No drama	Total
Music	30	50	80
No music	20	100	120
Total	50	150	200

A student is chosen at random. Determine whether the events M (takes music) and D (takes drama) are independent.

Read the two single-event probabilities from the totals, and the intersection from the cell where the two conditions meet:

$$P(M) = \frac{80}{200} = 0.4, \qquad P(D) = \frac{50}{200} = 0.25, \qquad P(M \cap D) = \frac{30}{200} = 0.15.$$

Compare the product with the intersection:

$$P(M)\,P(D) = 0.4 \times 0.25 = 0.10 \neq 0.15 = P(M \cap D),$$

so M and D are **not** independent.

The conditional form says the same thing and adds the direction. Restricting to the 80 music students gives $P(D \mid M) = \frac{30}{80} = 0.375$, against $P(D) = 0.25$ for the year group as a whole, so knowing a student takes music raises the chance of drama from a quarter to three eighths.

Independence is an exact condition, so any difference between the two numbers, however small, means the events are not independent. Write both numbers down and then the verdict; a bare "not independent" with only one side computed shows nothing.

CAUTION Mutually exclusive and independent are different ideas. If two events with non-zero probability are mutually exclusive, then one happening prevents the other completely, so $P(A \mid B) = 0 \neq P(A)$ and the events cannot be independent. Test independence with $P(A \cap B) = P(A)\,P(B)$; do not assume it.

YOUR TURN 11.4 Conditional Probability and Independence

Conditional probability

21.

(a) Given $P(A \cap B) = 0.2$ and $P(B) = 0.5$, find $P(A \mid B)$.

(b) Given $P(A) = 0.6$ and $P(B \mid A) = 0.3$, find $P(A \cap B)$.

(c) Given $P(A) = 0.5$, $P(B) = 0.4$ and $P(A \cup B) = 0.7$, find $P(A \mid B)$.

(d) Two cards are drawn from a deck of 52 without replacement. Find the probability that both are kings.

(e) A card is drawn from a deck of 52. Given that it is a picture card, find the probability that it is a king.

(f) Two fair dice are rolled. Given that the first die shows an even number, find the probability that the total is 8.

22. The table classifies 120 patients by the treatment they received and whether their condition improved.

	Improved	Not improved	Total
New treatment	36	24	60
Standard treatment	30	30	60
Total	66	54	120

One patient is chosen at random. Let I be "the condition improved" and N be "received the new treatment".

(a) Find $P(I \mid N)$.
(b) Find $P(N \mid I)$.
(c) Find $P(I \mid N')$, and use it to determine whether I and N are independent.

Independence

23.

(a) Events A and B are independent, with $P(A) = 0.5$ and $P(B) = 0.4$. Find $P(A \cap B)$.
(b) Given $P(A) = 0.3$, $P(B) = 0.5$ and $P(A \cap B) = 0.2$, determine whether A and B are independent.
(c) Events A and B have $P(A) = 0.3$, $P(B) = 0.2$ and $P(A \cup B) = 0.5$. Show that A and B are mutually exclusive, and determine whether they are independent.

24. In a class of 25 students, 15 study Spanish, 10 study art and 6 study both.

(a) Given that a student studies Spanish, find the probability that they also study art.
(b) Given that a student studies art, find the probability that they also study Spanish.
(c) Determine whether studying Spanish and studying art are independent.

25. Events A and B have $P(A) = 0.4$, $P(B) = 0.5$ and $P(A \mid B) = 0.4$. Find $P(A \mid B')$, and state whether A and B are independent.

Concept check

26. Events A and B have $P(A) = 0.3$ and $P(B) = 0.4$, and are mutually exclusive. A student says: "*mutually exclusive events do not affect each other, so A and B are independent.*" Explain the difference between *mutually exclusive* and *independent*, and show that A and B are not independent.

11.5 Tree Diagrams

When events happen in sequence, a **tree diagram** organises every path. Each branch carries the probability of that step, and two rules then do the whole calculation: **multiply along the branches** of a single path (the multiplication rule), and **add across paths** that satisfy the condition (the paths are mutually exclusive).

When each draw is made **with replacement**, the object goes back before the next draw, every stage repeats the same probabilities, and the stages are independent. Care is needed with sequences **without replacement**, where the first outcome changes the probabilities on the second branches.

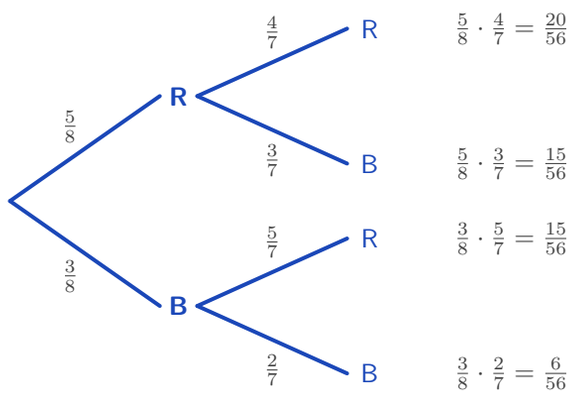


Figure 11.4 A tree diagram for drawing two marbles from 5 red and 3 blue without replacement. The probabilities on the second stage differ from the first, because one marble has already been removed. Multiply along a path; add across paths.

Worked through. For the bag in the figure, with 5 red and 3 blue marbles and two drawn without replacement, the second branches change because after drawing one red, only 4 reds remain out of 7 marbles. For both red, multiply along the top path:

$$P(\text{both red}) = \frac{5}{8} \times \frac{4}{7} = \frac{20}{56} = \frac{5}{14}.$$

"One of each" covers two paths, red-then-blue and blue-then-red; add them:

$$P(\text{one of each}) = \frac{5}{8} \cdot \frac{3}{7} + \frac{3}{8} \cdot \frac{5}{7} = \frac{15+15}{56} = \frac{30}{56} = \frac{15}{28}.$$

For "at least one blue," use the complement, which is "both red":

$$P(\text{at least one blue}) = 1 - \frac{5}{14} = \frac{9}{14}.$$

All three rules appear here: multiply along a path, add across alternatives, and take the complement for "at least one."

Putting the marble back changes every second-stage branch. The bag returns to 5 red and 3

blue, so both stages read $\frac{5}{8}$ and $\frac{3}{8}$, and the two draws are independent.

EXAMPLE 11.13

The same bag holds 5 red and 3 blue marbles, but now the first marble is replaced before the second is drawn. Find the probability that both are red, and compare with the answer without replacement.

Every branch keeps its original probability, so

$$P(\text{both red}) = \frac{5}{8} \times \frac{5}{8} = \frac{25}{64} \approx 0.391.$$

Without replacement the answer was $\frac{5}{14} \approx 0.357$, which is smaller. Removing a red marble leaves fewer reds for the second draw, so the second red becomes less likely. Replacing it removes that effect.

Independent repeated trials, such as draws with replacement, give a useful shortcut. When the same trial is repeated n times independently, "at least one success" has a complement that is a single product, because "no successes at all" means failure every time.

KEY For n independent trials in which the event has probability p each time,

$$P(\text{at least one}) = 1 - (1 - p)^n.$$

EXAMPLE 11.14

A machine part fails on any given day with probability 0.02, independently of other days. Find the probability that it fails at least once during a 30-day month.

Counting the ways to fail once, twice, three times and so on would need a separate calculation for each number of failures. The complement is a single product: the part survives all 30 days, each day with probability 0.98.

$$P(\text{at least one failure}) = 1 - (0.98)^{30} = 1 - 0.545 = 0.455.$$

A part that is safe on 98 days out of 100 still has close to an even chance of failing within a month. A small daily probability, repeated over many days, gives a much larger probability of at least one failure.

11.5.1 Three stages without replacement

A third draw adds a third stage and nothing else. Each branch still carries the probability of that step given everything before it, and without replacement the denominator falls by one at every draw.

What does change is the counting. An event such as "exactly two of the three" now corresponds to three paths, not one or two, because the odd draw out can come first, second or last. Identify those paths before computing anything. Each of these paths uses the same numerators and the same denominators in a different order, so all of them have the same probability. Evaluate one path and multiply by the number of orders.

EXAMPLE 11.15

A bag contains 7 green and 3 yellow counters. Three are drawn at random without replacement. Find the probability that (a) all three are green, (b) exactly two are green.

(a) One path, with the bag shrinking under it:

$$P(\text{GGG}) = \frac{7}{10} \times \frac{6}{9} \times \frac{5}{8} = \frac{210}{720} = \frac{7}{24}.$$

(b) Exactly two green means one yellow, and the yellow may be drawn first, second or third. The three paths are YGG, GYG and GGY:

$$P(\text{YGG}) = \frac{3}{10} \cdot \frac{7}{9} \cdot \frac{6}{8}, \qquad P(\text{GYG}) = \frac{7}{10} \cdot \frac{3}{9} \cdot \frac{6}{8}, \qquad P(\text{GGY}) = \frac{7}{10} \cdot \frac{6}{9} \cdot \frac{3}{8}.$$

Every one of them is $\frac{126}{720}$, so adding across the three paths gives

$$P(\text{exactly two green}) = 3 \times \frac{126}{720} = \frac{378}{720} = \frac{21}{40} = 0.525.$$

The three products agree because each is built from the same numerators, 7, 6 and 3, over the same denominators $10 \times 9 \times 8$; only the order of the factors differs. Count the orders first, because a missed path gives an answer that is too small.

YOUR TURN 11.5 Tree Diagrams

Two draws, with and without replacement

27.

(a) A bag holds 3 white and 2 black balls. Two are drawn with replacement. Find the probability that both are white.

(b) Repeat part (a) for two balls drawn without replacement.

(c) A box holds 6 working and 4 defective bulbs. Two are drawn without replacement. Find the probability that exactly one is defective.

Different probabilities on the second branches

28. The probability that it rains on a given morning is 0.3. If it rains, the probability that Sam is late for school is 0.6; if it does not rain, the probability is 0.1.

(a) Find the probability that Sam is late.
(b) Find the probability that it rains and Sam is not late.

Repeated independent trials

29.

(a) A machine works on any given day with probability 0.9, independently of other days. Find the probability that it works on two consecutive days.
(b) Ali and Bea take turns to roll a fair die, Ali first, and the first to roll a six wins. Find the probability that Ali wins with his second roll.
(c) Using the machine in part (a), find the probability that it fails on at least one of five consecutive days.

30. A component works with probability 0.95, independently of other components. A circuit contains four of them.

(a) Find the probability that all four components work.
(b) Find the probability that exactly one of the four components fails.

Three draws without replacement

31. A bag holds 4 red and 6 blue counters. Three are drawn without replacement.

(a) Find the probability that all three are blue.
(b) Find the probability that exactly two are red.

Concept check

32. A bag holds 3 red and 2 blue counters. Two are drawn without replacement. A student writes: "$P(\textit{one of each colour}) = \frac{3}{5} \times \frac{2}{4} = \frac{3}{10}$." Identify the error, and find the correct probability.

11.6 Bayes' Theorem HL

We can now return to the screening test from the start of the chapter. About one person in a hundred has the disease. The test correctly identifies 99% of those who have it, and wrongly reports a positive for 5% of those who do not. A positive result has just arrived.

The difficulty is that the test gives you $P(\text{positive} \mid \text{disease})$, but what you actually want is the reverse, $P(\text{disease} \mid \text{positive})$. Conditional probability is not symmetric, so these are

different numbers. **Bayes' theorem** is the tool that reverses the condition.

It is built from two facts you already have. The conditional formula gives $P(B \mid A) = \frac{P(A \cap B)}{P(A)}$. The multiplication rule rewrites the numerator as $P(A \cap B) = P(B)\,P(A \mid B)$, and the denominator can be split according to whether B happened or not: $P(A) = P(B)P(A \mid B) + P(B')P(A \mid B')$. Putting these together:

KEY · IB HL

$$P(B \mid A) = \frac{P(B)\,P(A \mid B)}{P(B)\,P(A \mid B) + P(B')\,P(A \mid B')}$$

Read the letters the way the booklet intends them: B is the thing you want to know about, and A is the evidence that has arrived. In the screening test, B is having the disease and A is the positive result.

The numerator is the single path through B; the denominator is every path that leads to A. You are asking what fraction of all the ways to observe A passed through B.

That description also gives the method. Written out in full the formula is long, but a tree diagram carries the same information in a form that is quicker to set up. Draw the first stage, B and B'. Draw the second stage, A given each of them. Multiply along every path. Then divide the path that passes through B by the total of all the paths that end in A.

TIP Do not memorise the Bayes formula as a string of symbols. Draw the tree, multiply along the paths, and divide the path you want by the total of the paths that match the evidence. It gives the same answer, it shows every step, and a correct method earns the same marks.

The probability of B before any evidence arrives, $P(B)$, is called the **base rate**. When the base rate is small, the few true positives can be outnumbered by false positives from the much larger group without the condition, so $P(B \mid A)$ can be small even for an accurate test.

Worked through. For the screening test, let D be "has the disease" and $+$ be "tests positive." We know $P(D) = 0.01$, $P(+ \mid D) = 0.99$, and $P(+ \mid D') = 0.05$. Draw the tree first, then multiply along every path.

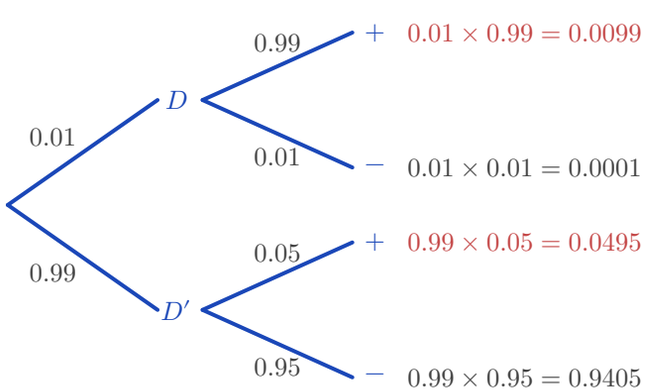


Figure 11.5 The screening test as a tree. Two paths end in a positive result, shown in scarlet. Bayes' theorem divides the scarlet path through D by the total of both scarlet paths.

Only the two scarlet paths match the evidence, so they are the only ones that count. The path through D divided by the total of both gives

$$P(D \mid +) = \frac{(0.01)(0.99)}{(0.01)(0.99) + (0.99)(0.05)} = \frac{0.0099}{0.0099 + 0.0495} = \frac{0.0099}{0.0594} \approx 0.167.$$

That is the formula and the tree saying the same thing: the numerator is one path, the denominator is every path that ends in a positive.

So a person who tests positive has only about a 17% chance of having the disease. The reason is the base rate. The disease is so rare that the few true positives are greatly outnumbered by the false positives, which come from the very large healthy majority. The test is good; the disease is simply very rare.

EXAMPLE 11.16

A city runs two taxi companies. 85% of its taxis are Green and 15% are Blue. A taxi is involved in a night-time accident and a witness says it was Blue. Tested under the same conditions, the witness identifies each colour correctly 80% of the time. Find the probability that the taxi really was Blue.

The evidence is the witness's statement, not the colour of the taxi, so put the colour on the first stage and the statement on the second.

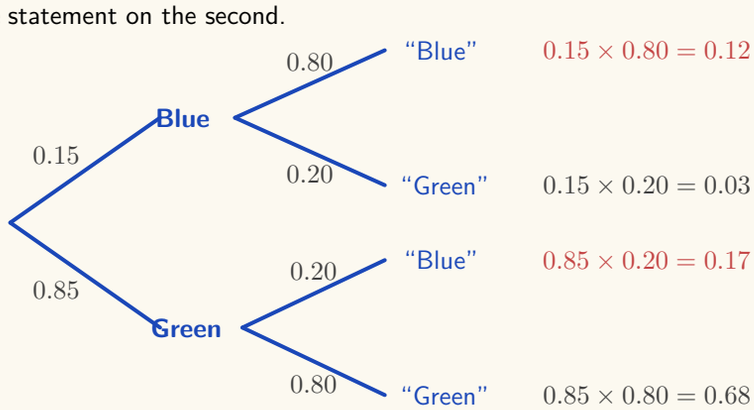


Figure 11.6 The witness problem. The two scarlet paths are the ones on which the witness says "Blue". Note that the larger of the two comes from a Green taxi being misreported.

The witness said "Blue", so only the two scarlet paths remain:

$$P(\text{Blue} \mid \text{"Blue"}) = \frac{0.12}{0.12 + 0.17} = \frac{0.12}{0.29} \approx 0.414.$$

A witness who is right 80% of the time reports "Blue", and the taxi is still more likely to have been Green. There are so many more Green taxis that the 20% of them misreported as Blue (0.17) outnumbers the 80% of Blue taxis reported correctly (0.12).

This is the same base-rate effect as the screening test, in a setting with no medicine in it. An

answer of 80% is tempting, because the witness's reliability is vivid and the proportion of Blue taxis is not. The tree keeps both in view.

11.6.1 Three alternatives

When the first stage has three possibilities instead of two, the denominator simply sums over all three paths. The structure is identical, and the booklet prints this version too.

KEY · IB HL

$$P(B_i \mid A) = \frac{P(B_i)\,P(A \mid B_i)}{P(B_1)\,P(A \mid B_1) + P(B_2)\,P(A \mid B_2) + P(B_3)\,P(A \mid B_3)}$$

Here B_1, B_2 and B_3 are the three possibilities at the first stage, and B_i is whichever one you are asking about. Nothing has changed except the number of terms underneath.

EXAMPLE 11.17

Three machines produce all the parts in a factory: machine A makes 50%, B makes 30%, C makes 20%. Their defect rates are 1%, 2%, and 3% respectively. A part is found to be defective. Find the probability it came from machine C.

Weight each machine's defect rate by its share of production, then take C's share of the total:

$$\begin{aligned} P(C \mid \text{def}) &= \frac{(0.20)(0.03)}{(0.50)(0.01) + (0.30)(0.02) + (0.20)(0.03)} \\ &= \frac{0.006}{0.005 + 0.006 + 0.006} = \frac{0.006}{0.017} \approx 0.353. \end{aligned}$$

Machine C makes only a fifth of the parts but, because it is the least reliable, accounts for over a third of the defects.

YOUR TURN 11.6 Bayes' Theorem HL

Two alternatives

33.

(a) The probability of rain on a given morning is 0.3. If it rains, the probability of being late is 0.6; if it does not, the probability is 0.1. A person is late. Find the probability that it rained.

(b) A disease affects 2% of a population. A test is positive for 95% of those who have it and for 10% of those who do not. Given a positive result, find the probability of having the disease.

(c) Factory A makes 60% of all items, of which 2% are defective; factory B makes the other 40%, of which 5% are defective. Given that an item is defective, find the probability it came from factory B.

34.

(a) In a mailbox, 30% of messages are spam. The word "free" appears in 90% of spam messages and in 5% of the others. A message contains the word "free". Find the probability that it is spam.

(b) A box contains two coins: one fair, and one with heads on both sides. A coin is chosen at random and tossed once, showing heads. Find the probability that it is the two-headed coin.

(c) Of the students in a year group, 70% revise for a test. Of those who revise, 80% pass; of those who do not, 30% pass. A student passes. Find the probability that the student revised.

Three alternatives

35. Three suppliers provide the components used by a factory. Supplier A provides 40%, supplier B 35% and supplier C 25%. Of their components, 1%, 3% and 4% respectively are defective.

(a) Find the probability that a component chosen at random is defective.

(b) Given that a component is defective, find the probability that it came from supplier B.

(c) Given that a component is defective, find the probability that it came from supplier A.

(d) Supplier A provides the most components but accounts for the fewest defective ones. Explain why.

Concept check

36. A disease affects 1% of a population. A test is positive for 95% of people who have the disease and for 5% of people who do not. A student says: "*the test is 95% accurate, so a person who tests positive has a 95% chance of having the disease.*"

(a) Explain the difference between $P(\text{positive} \mid \text{disease})$ and $P(\text{disease} \mid \text{positive})$.

(b) Find $P(\text{disease} \mid \text{positive})$, and explain why it is so much smaller than 0.95.

Chapter 11 · Probability

Reflections

TOK Bayes' theorem describes how a belief should be updated when new evidence arrives: start from a *prior*, the probability believed before the evidence, weigh the evidence, and arrive at a *posterior*, the probability believed afterwards. Some philosophers argue this is not just a formula but a model of knowledge itself. If that is right, the screening-test result of §11.6 is not special to medicine: evidence means little until you account for how rare the thing you are testing for is. Is Bayes' theorem a description of how people reason, or a rule for how they ought to?

TOK We say a fair die has probability $\frac{1}{6}$ because of its symmetry, and we say a drawing pin lands point-up with probability 0.39 because that is what 200 drops showed. The first number was never measured; the second was never proved.

Are these two the same kind of claim? And when a measured value disagrees with a calculated one, which of the two should be revised?

IM Probability theory is usually dated to 1654. That year the French writer Antoine Gombaud, the Chevalier de Méré, put several gambling problems to Blaise Pascal, among them how to divide the stakes fairly when a game is stopped before it ends. Pascal's letters with Pierre de Fermat on these problems set out many of the methods of probability theory, although Gerolamo Cardano had written about dice about a century earlier. A subject now central to physics, finance, and medicine grew out of questions about games of chance.

RW Ignoring the base rate has real costs. Courts have been misled when a "one-in-a-million" DNA match probability was presented as if it were the probability that the defendant was innocent, without asking how many people in the population would also match by chance. Mass screening programmes for rare cancers can produce more false alarms than true detections. Understanding $P(D \mid +)$ versus $P(+ \mid D)$ is not only an examination topic. It changes verdicts and medical decisions.

End-of-Chapter Problems
Chapter 11: Probability

1. Four cards are dealt at random from a standard pack of 52. Find the probability that the four cards are all from different suits. [5]

2. In a class of 50 students, 28 study Mathematics (M), 20 study Art (A), and 10 study neither.

(a) Find the number of students who study both Mathematics and Art. [3]
(b) Find the probability that a randomly chosen student studies only Art. [2]
(c) Find $P(A \mid M)$. [2]

3. The two-way table shows whether 100 people exercise regularly.

	Exercises	Does not	Total
Male	30	20	50
Female	24	26	50
Total	54	46	100

(a) Find the probability that a person is male and exercises. [2]
(b) Find the probability that a person exercises, given that they are female. [2]
(c) Determine whether the events "male" and "exercises" are independent. [3]

4. A bag contains 7 red and 3 green sweets. Two are drawn without replacement.

(a) Draw a tree diagram for the two draws. [3]
(b) Find the probability that both sweets are red. [2]
(c) Find the probability that both sweets are the same colour. [3]
(d) Find the probability that at least one sweet is green. [2]

5. A condition affects 5% of a population. A screening test is positive for 90% of those who have it, and for 15% of those who do not.

(a) Find the probability that a randomly chosen person tests positive. [3]
(b) Given a positive result, find the probability the person has the condition. [3]
(c) Comment on what your answer says about screening for rare conditions. [2]

6. Events A and B satisfy $P(A) = 0.5$ and $P(B) = 0.3$.

(a) Find $P(A \cup B)$ if A and B are independent. [3]
(b) Find $P(A \cup B)$ if A and B are mutually exclusive. [2]
(c) Explain why A and B cannot be both independent and mutually exclusive here. [2]

7. Two fair dice are rolled and the scores noted.

(a) Find the probability that the sum is a prime number. [2]
(b) Find the probability that at least one die shows a 6. [2]
(c) Given that at least one die shows a 4, find the probability that the sum is 7. [3]

8. Three suppliers provide components: X supplies 25%, Y supplies 45%, and Z supplies 30%. The probabilities that a component from X, Y and Z is defective are 0.04, 0.02 and p respectively. Overall, 3.4% of all components are defective.

(a) Find the value of p. [3]
(b) Given that a component is defective, find the probability that it came from X. [3]

9. Events A and B are independent. Given $P(A) = 0.4$ and $P(A \cup B) = 0.7$, find $P(B)$. [4]

10. A basketball player makes each free throw independently with probability 0.8. She takes 4 free throws.

(a) Find the probability that she makes exactly three of them. [2]
(b) Find the probability that she makes more free throws than she misses. [2]

11. At a school, 60% of students study French and 25% study both French and Spanish. Find the probability that a student studies Spanish, given that they study French. [3]

12. In a class of 40 students, the numbers studying two activities A and B are: only A is $2x$, both is x, only B is $x + 4$, and neither is 8.

(a) Find the value of x. [3]
(b) Find $P(A \cap B)$. [1]
(c) Find $P(A \cup B)$. [2]

13. A biased four-sided spinner numbered 1 to 4 was spun 250 times. The results were:

Number	1	2	3	4
Frequency	95	60	55	40

(a) Estimate the probability that the spinner lands on 1. [2]
(b) Estimate the probability that the spinner lands on an odd number. [2]
(c) The spinner is spun a further 60 times. Find the expected number of times it lands on 4. [2]

14. Events A and B satisfy $P(A) = 0.45$, $P(B) = 0.4$, and $P((A \cup B)') = 0.25$.

(a) Find $P(A \cap B)$. [2]
(b) Find the probability that A occurs but B does not. [1]
(c) Find $P(A \mid B)$. [2]
(d) Determine whether A and B are independent. [2]

15. A fair four-sided die (numbered 1–4) and a fair six-sided die (numbered 1–6) are rolled together, and the product of the two scores is recorded.

(a) Represent the outcomes in a sample space diagram. [2]
(b) Find the probability that the product is 12. [2]
(c) Find the probability that the product is odd. [2]
(d) Given that the product is even, find the probability that the four-sided die shows a 4. [3]

16. A football team's key striker is fit for a match with probability 0.85. The team wins with probability 0.7 if she is fit, and with probability 0.4 if she is not.

(a) Draw a tree diagram to represent this information. [2]
(b) Find the probability that the team wins the match. [2]
(c) Given that the team won, find the probability that the striker was fit. [3]

17. A box contains 4 red and 2 green counters. Three counters are drawn at random without replacement.

(a) Find the probability that all three counters are red. [2]
(b) Given that at least one counter is green, find the probability that exactly one is green. [3]
(c) Find the probability that the third counter drawn is the second green one. [2]

18. Events A and B satisfy $P(A) = 0.6$, $P(A \mid B) = 0.6$, and $P(A \cup B) = 0.84$.

(a) Explain why A and B are independent. [1]
(b) Find $P(B)$. [3]
(c) Write down $P(B \mid A')$, justifying your answer. [2]

19. The 120 members of a sports club are classified by age group and by whether they wear glasses: 18 of the 50 adults and 12 of the 70 juniors wear glasses.

(a) Display this information in a two-way table, including totals. [2]

(b) Find the probability that a randomly chosen member is a junior who does not wear glasses. [1]
(c) Given that a member wears glasses, find the probability that the member is an adult. [2]
(d) Determine, using conditional probability, whether the events "adult" and "wears glasses" are independent. [3]

20. An airline routes 50% of its flights through airport A, 30% through airport B, and 20% through airport C. The probabilities that a flight is delayed are 0.06, 0.03, and 0.10 respectively.

(a) Find the probability that a randomly chosen flight is delayed. [3]
(b) Given that a flight was delayed, find the probability that it was routed through airport A. [3]

21. Events A, B, and C are mutually exclusive and exhaustive, with $P(A) = 2k$, $P(B) = 3k$, and $P(C) = k + 0.1$.

(a) Find the value of k. [2]
(b) Find $P(A \cup B)$. [2]
(c) Find $P(A \mid B')$. [3]

22. GDC An archer hits the target with probability 0.9 on each shot, independently of all other shots.

(a) Find the probability that her first miss is on her third shot. [2]
(b) Find the probability that she hits at least four of her first five shots. [2]
(c) Find the least number of shots for which the probability of at least one miss exceeds 0.95. [3]

23. A family has two children. Assume each child is equally likely to be a boy or a girl, independently.

(a) List the sample space. [1]
(b) Given that at least one of the children is a boy, find the probability that both are boys. [3]
(c) Given instead that the *elder* child is a boy, find the probability that both are boys. [2]
(d) Explain why the answers to (b) and (c) differ. [1]

24. In a survey of 100 people, 50 like tea, 60 like coffee, and 35 like juice. Also, 30 like tea and coffee, 20 like tea and juice, 25 like coffee and juice, and 10 like all three.

(a) Find the number who like at least one of the three drinks. [3]
(b) Find the number who like none. [1]
(c) Find the number who like exactly one drink. [3]

25. A disease affects 0.5% of a population. A test is positive for 98% of those with the disease and for 3% of those without. A person tests positive *twice*, the two tests being independent given disease status.

(a) Find the probability the person has the disease after one positive test. [3]
(b) Find the probability the person has the disease after two positive tests. [4]
(c) Comment on the effect of the second test. [1]

26. Events A and B are independent. Prove that A and B' are also independent, that is, show $P(A \cap B') = P(A)\,P(B')$. [5]

27. Of the 200 students at a school, 130 cycle to school (C), 90 travel by train (T), and 40 do both. The remaining students walk.

(a) Draw a Venn diagram showing the number of students in each of the four regions. [3]
(b) Find the probability that a student chosen at random walks to school. [2]
(c) Find $P(C \mid T)$. [3]
(d) Determine whether C and T are independent, justifying your answer. [3]
(e) Two students are chosen at random, without replacement. Find the probability that both of them walk. [3]
(f) State whether your answer to part (e) would be larger or smaller if the two students were chosen with replacement, and explain why. [2]

28. A machine produces components, of which 6% are defective. Each component is inspected. The inspection rejects a defective component with probability 0.95, and wrongly rejects a component that is not defective with probability 0.02.

(a) Draw a tree diagram to represent this information. [3]
(b) Find the probability that a randomly chosen component is rejected. [3]
(c) Find the probability that a rejected component was in fact not defective. [3]
(d) Find the probability that a component which passes inspection is defective. [3]
(e) In a batch of 5000 components, find the expected number that are defective but pass inspection. [2]
(f) Using your answers to parts (d) and (e), comment on how well the inspection protects a customer. [2]

29. A box contains 10 batteries, of which 3 are flat. Batteries are taken from the box one at a time, without replacement, and tested until a working one is found. Let N be the number of batteries tested.

(a) Find $P(N = 1)$. [2]
(b) Find $P(N = 2)$. [3]
(c) Find $P(N \leq 3)$. [4]
(d) Explain why N can never be greater than 4. [2]

(e) Find $P(N = 3)$ and $P(N = 4)$, and verify that the four probabilities sum to 1. [4]

30. A lottery draws 5 numbers at random from 40. A ticket also carries 5 numbers.

(a) Find the number of possible draws. [2]
(b) Find the probability that a ticket matches all 5 numbers. [2]
(c) Find the probability that a ticket matches exactly 3 of the 5. [4]
(d) The organisers propose drawing 5 from 45 instead. Without computing the new probability in full, explain what happens to a ticket's chance of matching all five, and why lotteries are designed this way. [3]

31. A committee of 4 is chosen at random from 5 teachers and 4 students.

(a) Find the number of possible committees. [2]
(b) Find the probability that the committee has exactly 2 teachers and 2 students. [3]
(c) Find the probability that the committee contains at least one teacher and at least one student. [4]
(d) Explain why part (c) is quicker through its complement than by adding the cases directly. [2]

32. EXACT The letters of the word PARALLEL are arranged in a row, all arrangements being equally likely.

(a) Explain why the number of distinct arrangements is $\dfrac{8!}{2!\,3!}$, and evaluate it. [3]
(b) Find the probability that an arrangement begins with the letter P. [4]
(c) Find the probability that the three Ls are together. [4]

Challenge Questions

33. GDC **The birthday problem.** In a group of n people, assume that every birthday is equally likely among 365 days, that different people's birthdays are independent, and ignore leap years.

(a) Explain why, for $n = 3$, the probability that all three birthdays differ is $\dfrac{364}{365} \times \dfrac{363}{365}$, and write down an expression for the probability that all n birthdays differ. [4]
(b) Hence find, to three decimal places, the probability that at least two people share a birthday when $n = 10$, $n = 23$ and $n = 50$. [3]
(c) Find the smallest n for which the probability of a shared birthday exceeds $\frac{1}{2}$. [2]
(d) Two tempting guesses give much larger answers. The first is 183 people, about half the number of days. The second is the number of people needed for the probability that

someone shares *your* birthday to exceed $\frac{1}{2}$. Show that the second guess gives 254 people (you and 253 others). Explain why each guess answers a different question from part (c), and why the count that matters is the number of *pairs*, ${}^{n}C_{2}$. Evaluate ${}^{n}C_{2}$ for $n = 23$. [5]

(e) Each pair matches with probability $\frac{1}{365}$. Find the expected number of matching pairs in a group of 23, and comment on your answer. [2]

(f) A school issues every student a random four-digit code, so there are $10\,000$ equally likely codes. Find the smallest number of students for which two students sharing a code is more likely than not. [3]

(g) Now let there be N equally likely values. Using $1 - x \approx e^{-x}$ for small x, show that the probability that all n values differ is approximately $e^{-n(n-1)/(2N)}$. Hence show that, for large N, the smallest n for which a match is more likely than not is about $\sqrt{2N \ln 2}$. Evaluate this for $N = 365$ and $N = 10\,000$, and compare with parts (c) and (f). [4]

(h) Real birthdays are not spread evenly across the year. Investigate, using published birth data or a model of your own, whether unequal frequencies make a shared birthday more or less likely. Alternatively, investigate how large a group must be for *three* people to share a birthday with probability above $\frac{1}{2}$. [2]

34. Three doors. A prize is behind one of three closed doors, each door equally likely. You choose a door. The host knows where the prize is, and always opens one of the other two doors that does not hold the prize; if both of those doors are empty, because you chose the prize, he chooses between them at random. You may then keep your door or switch to the last closed door.

(a) Find the probability of winning if you always switch. Treat the case where your first choice is correct separately from the case where it is not. [5]

(b) State the probability of winning if you always stay, and say which strategy is better. [1]

(c) Now use Bayes' theorem. You chose door 1 and the host opened door 3. Write down $P(\text{host opens } 3 \mid \text{prize behind } 1)$, $P(\text{host opens } 3 \mid \text{prize behind } 2)$ and $P(\text{host opens } 3 \mid \text{prize behind } 3)$, and hence find $P(\text{prize behind } 2 \mid \text{host opens } 3)$. [5]

(d) A second host knows nothing. He opens one of the two doors you did not choose at random, and it happens to be empty. Repeat the calculation of part (c) for this host, and show that switching now wins with probability $\frac{1}{2}$. [4]

(e) The two hosts performed the same visible action: they opened door 3 and it was empty. Explain what makes the answers differ, and what it is that the first host's action tells you. [3]

(f) The game is now played with $n \geq 3$ doors, one prize and each door equally likely. After your choice, the knowing host opens $n - 2$ of the other doors, all of them empty, leaving your door and one other closed; if your door holds the prize, he chooses the other door to leave closed at random. Show that switching wins with probability $\dfrac{n-1}{n}$. [3]

(g) Return to three doors and a knowing host. When you have chosen the prize, this host opens door 3 with probability q and door 2 with probability $1 - q$, where $0 \leq q \leq 1$. You chose door 1 and he opened door 3. Show that

$P(\text{prize behind } 2 \mid \text{host opens } 3) = \dfrac{1}{1+q}$. [2]

(h) Deduce that switching is never worse than staying, whatever the value of q, and describe what the host's choice of door tells you when $q = 0$ and when $q = 1$. Suggest a further change to the host's rules that would be worth investigating. [2]

35. Independent in pairs, but not all together. Two events A and B are independent when $P(A \cap B) = P(A)P(B)$. Three events A, B and C are called *mutually independent* when each of the three pairs is independent and, in addition, $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Two fair coins are tossed. Let A be the event "the first coin shows heads", B the event "the second coin shows heads", and C the event "the two coins show different faces".

(a) Find $P(A)$, $P(B)$, $P(C)$, $P(A \cap B)$, $P(A \cap C)$ and $P(B \cap C)$, and show that each pair of the events A, B, C is independent. [4]

(b) Find $P(A \cap B \cap C)$, and show that A, B and C are not mutually independent. Explain in words why knowing whether A and B happened tells you whether C happened. [3]

(c) Let A, B and C be any three events that are independent in pairs, with $P(A \cap B) > 0$. Prove that $P(C \mid A \cap B) = P(C)$ if and only if $P(A \cap B \cap C) = P(A)P(B)P(C)$. [3]

(d) A fair eight-sided die, with faces numbered 1 to 8, is rolled once. Let $A = \{1, 2, 3, 4\}$, $B = \{1, 2, 3, 5\}$ and $C = \{1, 6, 7, 8\}$. Show that $P(A \cap B \cap C) = P(A)P(B)P(C)$, but that no two of these events are independent. [3]

(e) Explain what parts (b) and (d) together show about the definition of mutual independence. Mutual independence of four events requires the product rule for every collection of two or more of them. Show that this gives 11 equations, and find a formula for the number of equations needed for n events. [3]

(f) Now toss n fair coins, and for each pair of coins i and j let E_{ij} be the event that coins i and j show the same face. Investigate which collections of the events E_{ij} are mutually independent, starting with E_{12}, E_{13} and E_{23} for three coins, and state a conjecture. [2]

Chapter 12

Probability Distributions

12

SYLLABUS 4.7 4.8 4.9 4.12 4.14

If you measure the same kind of thing many times, you rarely get the same answer twice. Masses differ from packet to packet, and a repeated experiment gives a different count each time. A single number cannot describe a quantity that behaves this way. What describes it is the list of values the quantity can take, together with how likely each value is. That list is called a **distribution**, and this chapter covers the few that fit a great many situations.

A factory fills bags of rice, each meant to hold 500 grams, and weighs 5000 of them. No two masses are exactly equal. Group the readings, draw them as a graph, and a shape appears: most bags near 500 g, with fewer and fewer on either side, falling away evenly into one smooth hump. The shape is a bell.

The same curve appears where there is no rice at all. Astronomers found it in the *errors* of their telescope readings, and biologists in the heights of people. The reason is usually the same: a bag's mass is decided not by one thing but by many small effects, such as the size of the grains, a slight vibration in the machine and the exact moment the valve closes, and the end of the chapter shows why such a sum gives a bell.

A distribution turns uncertainty into something you can calculate with. From one you can predict an average result, measure how far individual results fall from that average, and put a number on the risk of an extreme value. That is why these curves are used in medicine, manufacturing and finance, and also why care is required: every answer depends on the distribution being the right description, and a wrong choice gives precise answers that are wrong.

By the end of this chapter you will be able to find the mean and variance of a discrete random variable, use the binomial distribution, find normal probabilities and work backwards with the inverse normal, standardise a value to compare across distributions, and work with a continuous random variable given by a density function.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

12.1 Discrete Random Variables

A **random variable** is a number whose value depends on chance. The number of heads in three coin tosses, the score on a die, the number of defective items in a box: each is a rule that assigns a number to a random outcome. We write random variables with capital letters, X, and particular values with lower case, x.

A random variable is **discrete** when it can only take separate, countable values. To describe it fully, you list every value it can take alongside the probability of each. This list is its **probability distribution**.

Here are five discrete random variables, each with the values it can take.

- $X =$ the number of heads in three coin tosses. Values $0, 1, 2, 3$.
- $X =$ the score on one roll of a fair die. Values $1, 2, 3, 4, 5, 6$.
- $X =$ the number of defective items in a box of twenty. Values $0, 1, \ldots, 20$.
- $X =$ the number of students absent from a class of thirty. Values $0, 1, \ldots, 30$.
- $X =$ the number of tosses of a coin until the first head appears. Values $1, 2, 3, \ldots$, with no largest value.

The last one deserves attention. A discrete variable does not have to take a finite number of values. It only has to take separate values that can be listed in order, and that list may continue for ever. What rules a variable out is taking *every* value across a range, as a height or a waiting time does. Those are continuous variables, and §12.4 treats them.

DEF A **discrete random variable** X takes countable values $x_1, x_2, \ldots$ Its **probability distribution** gives $P(X = x_i)$ for each value, subject to

$$P(X = x_i) \geq 0 \qquad \text{and} \qquad \sum_i P(X = x_i) = 1.$$

The second condition is the one the examples below use. Since some value must occur, all the probabilities add to exactly 1. This single fact lets you find a missing probability.

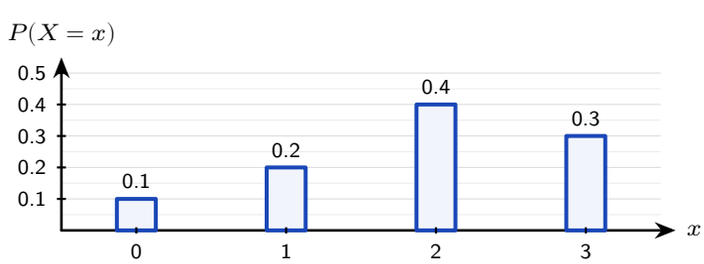


Figure 12.1 The probability distribution of a discrete random variable. Each bar is the probability of one value, and the four bars must total exactly 1.

Worked through. A discrete random variable X has the distribution below.

x	1	2	3	4
$P(X = x)$	0.1	0.3	k	0.2

To find k, use the fact that the probabilities must sum to 1:

$$0.1 + 0.3 + k + 0.2 = 1 \implies k = 0.4.$$

Then $P(X \geq 3) = P(X = 3) + P(X = 4) = 0.4 + 0.2 = 0.6$.

EXAMPLE 12.1

A random variable X takes values $1, 2, 3, 4$ with $P(X = x) = kx$. Find k.
Sum the probabilities and set equal to 1:

$$k(1) + k(2) + k(3) + k(4) = 10k = 1 \implies k = \frac{1}{10}.$$

When the distribution is given by a formula, the total-probability condition still determines the constant.

YOUR TURN 12.1 Discrete Random Variables

Discrete or continuous

1. State whether each random variable is discrete or continuous.

(a) the number of emails received in a day

(b) the time a runner takes to finish a race

(c) the number of rolls of a die until the first six

(d) the mass of an apple

Finding a missing probability

2.

(a) For $x = 1, 2, 3$ the probabilities are 0.2, 0.5 and k. Find k.

(b) A variable X has $P(X = x) = 0.15,\ 0.25,\ 0.35,\ k$ for $x = 0, 1, 2, 3$. Find k and $P(X \leq 1)$.

(c) A variable takes values $x = 0, 1, 2, 3$ with $P(X = x) = k(x + 1)$. Find k.

(d) A variable takes values $x = 1, 2, 3, 4$ with $P(X = x) = \dfrac{x^2}{30}$. Find $P(X = 3)$.

(e) For the distribution $P(X = x) = 0.1,\ 0.4,\ 0.3,\ 0.2$ at $x = 0, 1, 2, 3$, find $P(X > 1)$ and $P(1 \leq X \leq 2)$.

3.

(a) Show that $P(X = x) = \frac{1}{18}(4 + x)$ for $x = 1, 2, 3$ defines a probability distribution, and find $P(X \geq 2)$.

(b) A variable takes values $x = 1, 2, 3, 4$ with probabilities 0.1, a, 0.3, b. Given that $P(X \leq 2) = 0.4$, find a and b.

(c) Hence find $P(X > 2)$ for the distribution in part (b).

(d) A variable takes values $x = 1, 2, 3$ with $P(X = x) = k \cdot 2^x$. Find k.

Checking that a distribution is valid

4.

(a) Show that the values of $P(X = x) = \dfrac{2x - 1}{8}$ for $x = 0, 1, 2, 3$ add up to 1, and explain why this is still not a probability distribution.

(b) A variable has $P(X = 1) = c$, $P(X = 2) = 0.5 - c$ and $P(X = 3) = 0.5$. Find the set of values of c for which this is a probability distribution.

Concept check

5. A variable X takes the values $1, 2, 3, 4$ with probabilities 0.1, 0.3, 0.4 and 0.2. A student writes: "$P(X < 3) = P(X \leq 3) = 0.1 + 0.3 + 0.4 = 0.8$." Identify the error, and find the correct value of $P(X < 3)$.

12.2 Expected Value and Variance

If you played a game many times, what would you win on average per play? That long-run average is the **expected value** $E(X)$. It is not the most likely outcome, and it need not be a value that X can take at all. It is the balance point of the distribution, found by weighting each value by its probability.

KEY · IB

$$E(X) = \mu = \sum_i x_i \, P(X = x_i)$$

This is exactly the mean from Ch. 10, with probabilities playing the role of relative frequencies. Over many trials, the relative frequency of each value approaches its probability, so the average approaches $E(X)$.

Worked through. Suppose a game costs \$2 to play: you roll a fair die and win, in dollars, the number shown. To decide whether the game is worth playing, let X be the amount won. Each face is equally likely:

$$E(X) = \tfrac{1}{6}(1+2+3+4+5+6) = \tfrac{21}{6} = 3.5.$$

You expect to win \$3.50 per play but pay \$2, so the expected net gain is \$1.50. Over many games you gain money on average.

That last calculation names an idea the syllabus uses directly. A game is **fair** when the expected net gain is zero, so that neither side profits in the long run. Since the net gain is the winnings minus the stake, a fair stake is exactly the expected winnings.

KEY A game is **fair** when $E(\text{gain}) = 0$, that is, when the stake equals $E(\text{winnings})$.

In the worked-through game above, the stake of \$2 is less than the expected winnings of \$3.50, so the game is not fair: it favours the player.

The calculation so far ran forwards, from a complete distribution to its mean. Questions also run the other way. If a table hides two probabilities and the mean is given, you hold two facts about those unknowns: the probabilities total 1, and the weighted sum of the values is the stated mean. Two facts, two unknowns, so write both as equations and solve them together.

EXAMPLE 12.2

A discrete random variable X has the distribution below, and $E(X) = 1.8$. Find a and b.

x	**0**	**1**	**2**	**3**
$P(X = x)$	0.15	a	b	0.25

The total-probability condition gives one equation:

$$0.15 + a + b + 0.25 = 1 \implies a + b = 0.6.$$

The stated mean gives a second, by weighting each value by its probability:

$$0(0.15) + 1(a) + 2(b) + 3(0.25) = 1.8 \implies a + 2b = 1.05.$$

Subtracting the first equation from the second removes a and leaves $b = 0.45$, so $a = 0.6 - 0.45 = 0.15$.
Check both conditions on the finished table: $0.15 + 0.15 + 0.45 + 0.25 = 1$, and $0(0.15) + 1(0.15) + 2(0.45) + 3(0.25) = 1.8$, so the pair satisfies the two facts it was built from.

12.2.1 Variance and standard deviation HL

Expected value locates the centre, and **variance** measures the spread, exactly as in descriptive statistics. The variance is the expected squared distance from the mean. A short calculation turns that definition into a form that is easier to use.

KEY · IB HL

$$\text{Var}(X) = E((X-\mu)^2) = E(X^2) - \mu^2, \qquad \text{where } E(X^2) = \sum_i x_i^2 \, P(X = x_i)$$

The second form, $E(X^2) - \mu^2$, is usually the quicker one. Find the mean of the squares, then subtract the square of the mean. The standard deviation is $\sigma = \sqrt{\text{Var}(X)}$.

EXAMPLE 12.3

Find the variance and standard deviation of X:

x	0	1	2	3
$P(X = x)$	0.1	0.2	0.4	0.3

First the mean:

$$\mu = 0(0.1) + 1(0.2) + 2(0.4) + 3(0.3) = 1.9.$$

Then $E(X^2)$:

$$E(X^2) = 0^2(0.1) + 1^2(0.2) + 2^2(0.4) + 3^2(0.3) = 0.2 + 1.6 + 2.7 = 4.5.$$

So $\text{Var}(X) = 4.5 - 1.9^2 = 4.5 - 3.61 = 0.89$, and $\sigma = \sqrt{0.89} \approx 0.943$.

12.2.2 Linear transformations HL

If a random variable is scaled and shifted, its mean and variance change in a predictable way, exactly as those of a data set do. For constants a and b:

KEY · IB HL

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

The shift b moves the mean but leaves the spread unchanged. The scale factor a stretches the mean by a and the variance by a^2, because variance is built from squared distances.

EXAMPLE 12.4

For the variable above, $E(X) = 1.9$ and $\text{Var}(X) = 0.89$. Let $Y = 2X + 3$. Find $E(Y)$ and $\text{Var}(Y)$.

$$E(Y) = 2(1.9) + 3 = 6.8, \qquad \text{Var}(Y) = 2^2(0.89) = 3.56.$$

The $+3$ leaves the variance unchanged. Only the factor of 2 affects it, and it does so squared.

YOUR TURN 12.2 Expected Value and Variance

Expected value and fair games

6.

(a) Find $E(X)$ for $x = 1, 2, 3$ with $P(X = x) = 0.2,\ 0.3,\ 0.5$.

(b) Find $E(X)$ for $x = 0, 1, 2, 3$ with $P(X = x) = 0.1,\ 0.4,\ 0.3,\ 0.2$.

(c) A fair coin is tossed. The player wins \$3 for heads and loses \$2 for tails. Find the player's expected gain per toss.

(d) In a game you gain \$5 with probability $\frac{1}{8}$, and otherwise lose \$1. Find the expected gain, and state whether the game is fair.

7. A stall at a fair runs a prize wheel. The wheel pays \$8 with probability 0.1, \$2 with probability 0.3, and nothing otherwise.

(a) Find the expected amount won per spin, and the stake that would make the game fair.

(b) The stall charges \$2 to spin. Find the expected gain per spin for a player, and the expected loss over 50 spins.

8. A fair die is rolled. You win \$10 for a six, \$4 for a four or a five, and nothing otherwise. Find the stake that makes the game fair, and state whether a stake of \$4 favours the player or the organiser.

Unknown probabilities from the mean

9.

(a) A variable takes the values $1, 2, 3, 4$ with probabilities 0.2, a, b, 0.1, and $E(X) = 2.3$. Find a and b.

(b) A game pays \$10 with probability p, \$2 with probability q, and nothing with probability 0.5. The fair stake is \$3. Find p and q.

Variance and standard deviation

10. Find $E(X^2)$, $\text{Var}(X)$ and the standard deviation for each distribution.

(a) $x = 0, 1, 2, 3$ with $P(X = x) = 0.1,\ 0.4,\ 0.3,\ 0.2$

(b) $x = -1, 0, 2$ with $P(X = x) = 0.3,\ 0.2,\ 0.5$

Linear transformations

11.

(a) Given $E(X) = 4$ and $\text{Var}(X) = 2$, find $E(3X - 1)$ and $\text{Var}(3X - 1)$.

(b) A variable has $E(X) = 3$ and $\text{Var}(X) = 5$. Find $E(X^2)$, $\text{Var}(5 - 2X)$ and the standard deviation of $5 - 2X$.

Concept check

12. Given that $\text{Var}(X) = 4$, a student writes: "$\textit{Var}(2X + 3) = 2 \times 4 + 3 = 11$." Identify the two errors, and find the correct value.

13. A fair die is rolled and X is the score, so $E(X) = 3.5$. A student says: "*this must be wrong, because the die can never show* 3.5." Explain why $E(X) = 3.5$ is correct, and state what it means.

12.3 The Binomial Distribution

Many situations consist of the same experiment repeated: ten coin flips, twenty items checked for defects, twelve multiple-choice questions answered by guessing. When each repetition is independent and has only two possible outcomes, the number of "successes" follows one particular distribution, the **binomial**.

Four conditions define it. Check each one before reaching for the distribution.

- **Fixed number of trials.** The value of n is decided in advance, not by the results.
- **Two outcomes only.** Every trial ends as a success or a failure, with nothing in between.
- **Constant probability.** Every trial has the same probability of success p.
- **Independence.** The result of one trial does not change the probabilities for any other.

When all four hold, write $X \sim B(n, p)$.

DEF If X counts the successes in n independent trials each with success probability p, then $X \sim B(n, p)$, and

$$P(X = r) = {}^{n}C_{r}\, p^{r}(1-p)^{n-r}, \qquad r = 0, 1, 2, \ldots, n.$$

Every factor in that formula has a meaning. The term p^r accounts for the r successes, $(1-p)^{n-r}$ for the remaining trials failing, and ${}^{n}C_{r}$ for the number of different orders in which those successes can occur. In practice you evaluate the whole expression on a GDC: use the binomial pdf for $P(X = r)$ and the binomial cdf for $P(X \le r)$.

Worked through. A student guesses on a 10-question multiple-choice test, each question having 4 options. The number correct is $X \sim B(10, 0.25)$, since each guess is correct with probability $\frac{1}{4}$. The binomial pdf gives the probability of exactly 3 correct:

$$P(X = 3) = {}^{10}C_{3}(0.25)^{3}(0.75)^{7} \approx 0.250.$$

The binomial cdf gives the probability of at most 2 correct:

$$P(X \le 2) = P(0) + P(1) + P(2) \approx 0.526.$$

Even with 10 questions, scoring 2 or fewer is more likely than not. Guessing rarely produces a good score.

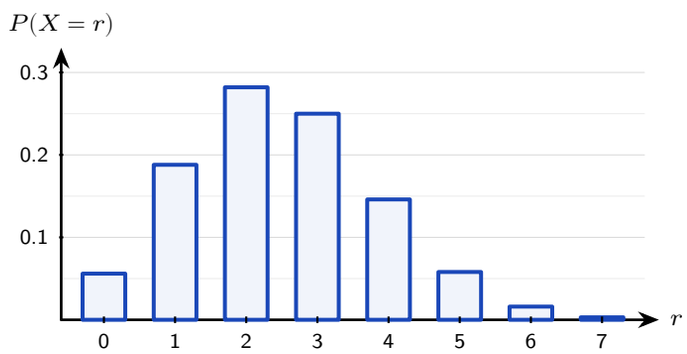


Figure 12.2 The distribution $B(10,\, 0.25)$, the guessing student above. The tallest bar is at $r = 2$ (0.282), with $r = 3$ close behind (0.250), and the long tail to the right shows why a high score by guessing is so unlikely.

12.3.1 Mean and variance of the binomial

The mean could be found by summing $r\,P(X = r)$, but a formula gives it directly. For n trials each succeeding with probability p, the results are clean.

KEY · IB For $X \sim B(n, p)$:

$$E(X) = np, \qquad \text{Var}(X) = np(1-p)$$

The mean is the result you would expect: 10 trials at probability 0.25 give 2.5 successes on average. The variance is largest when $p = 0.5$, which is the value at which a single trial is least predictable.

EXAMPLE 12.5

A factory's items are defective with probability 0.05, independently. A sample of 20 is taken. Find the expected number of defectives and the probability that at least one is defective.

Let $X \sim B(20, 0.05)$. The expected number is

$$E(X) = 20 \times 0.05 = 1.$$

For "at least one," use the complement:

$$P(X \geq 1) = 1 - P(X = 0) = 1 - (0.95)^{20} \approx 0.642.$$

Even with a low defect rate, a sample of 20 is more likely than not to contain a defective item. The probability of "at least one" rises quickly as the sample size increases.

CAUTION Check independence and a fixed n before using the binomial. Drawing items *without replacement* breaks independence, since each draw changes the next probability, so that count is not binomial.

12.3.2 Finding an unknown n or p

Every binomial calculation so far began with n and p and produced a probability. A question can supply the probability instead and ask for one of the parameters. Both reversals start in the same place, at

$$P(X = 0) = {}^{n}C_0\, p^0 (1-p)^n = (1-p)^n,$$

a term with no combinatorial factor, since ${}^{n}C_0 = 1$. It is a single power, so it can be solved by logarithms when n is the unknown and by a root when p is.

EXAMPLE 12.6

Each ticket in a charity draw wins a prize with probability 0.1, independently of every other ticket. Find the least number of tickets that must be bought for the probability of winning at least one prize to exceed 0.95.

Let n be the number of tickets, so the number of prizes is $X \sim B(n, 0.1)$ and the complement gives

$$P(X \geq 1) = 1 - P(X = 0) = 1 - (0.9)^n.$$

The requirement $1 - (0.9)^n > 0.95$ rearranges to $(0.9)^n < 0.05$. Take logarithms of both sides:

$$n \ln 0.9 < \ln 0.05 \implies n > \frac{\ln 0.05}{\ln 0.9} \approx 28.4,$$

where the inequality reverses because $\ln 0.9$ is negative. Tickets are counted in whole numbers, so $n = 29$.

Confirm at both ends: $(0.9)^{29} \approx 0.047$ and $(0.9)^{28} \approx 0.052$, so 29 tickets bring the failure probability below 0.05 and 28 do not.

EXAMPLE 12.7

The number of faulty components in a batch of six is $X \sim B(6, p)$. Given that $P(X = 0) = 0.262$, find p.

With $r = 0$ the binomial formula reduces to a single power:

$$(1 - p)^6 = 0.262.$$

Take the sixth root of both sides:

$$1 - p = 0.262^{1/6} \approx 0.800 \implies p \approx 0.200.$$

The root returns $1 - p$ rather than p, so do not stop at 0.800; substituting back gives $(0.8)^6 \approx 0.262$, which confirms both the value and the final subtraction.

YOUR TURN 12.3 The Binomial Distribution

Binomial probabilities

14. GDC

(a) $X \sim B(8, 0.5)$. Find $P(X = 4)$.
(b) $X \sim B(10, 0.3)$. Find $P(X = 2)$.
(c) $X \sim B(6, 0.4)$. Find $P(X \leq 2)$.
(d) $X \sim B(12, 0.25)$. Find $P(X \geq 1)$.

15. GDC In a factory, 8% of items are defective, independently of one another. A sample of 25 items is taken.

(a) Find the probability that exactly 2 items are defective.
(b) Find the probability that more than 3 items are defective.
(c) Find the expected number of defective items in the sample.

16. GDC A basketball player scores each free throw with probability 0.7, independently of the others. She takes 15 free throws.

(a) Find the probability that she scores at least 12.
(b) Find the probability that she scores between 8 and 11 inclusive.

Mean and variance

17.

(a) $X \sim B(20, 0.1)$. Find the mean and the variance.
(b) $X \sim B(n, p)$ with $E(X) = 6$ and $\text{Var}(X) = 4.2$. Find n and p.

When the binomial applies

18.

(a) A box holds 10 items, 3 of them defective, and 4 items are drawn without replacement. Explain why the number of defective items drawn is not binomial.
(b) A student counts how many of the 30 days in June are rainy. Give one reason why $B(30, p)$ may not be a good model for this count.

Finding an unknown n or p

19.

(a) Each item made by a machine is faulty with probability 0.03, independently. Find the least number of items that must be checked for the probability of finding at least one faulty item to exceed 0.9.

(b) $X \sim B(5, p)$ and $P(X = 0) = 0.07776$. Find p.

20.

(a) $X \sim B(n, 0.2)$. Find the least value of n for which $P(X = 0) < 0.01$.

(b) $X \sim B(4, p)$ and $P(X = 4) = 0.1296$. Find p.

Concept check

21. GDC $X \sim B(10, 0.4)$. A student writes:
"$P(X \geq 3) = 1 - P(X \leq 3) \approx 1 - 0.382 = 0.618$." Identify the error, and find the correct value of $P(X \geq 3)$.

12.4 Continuous Random Variables HL

A discrete variable jumps between separate values. A continuous one, such as a height, a mass or a waiting time, can take any value in a range. You cannot list a probability for each value, because there are infinitely many of them. Something has to replace the list.

The bags of rice show what replaces it. Group the 5000 masses into bars one gram wide. The probability of a bag falling in a given group is then the *area* of that group's bar. Now halve the bar width. There are twice as many bars, each half as wide and about as tall as before, and the total area is unchanged at 1. Halve it again, and again. The tops of the bars approach a smooth curve, and the area under that curve is still 1.

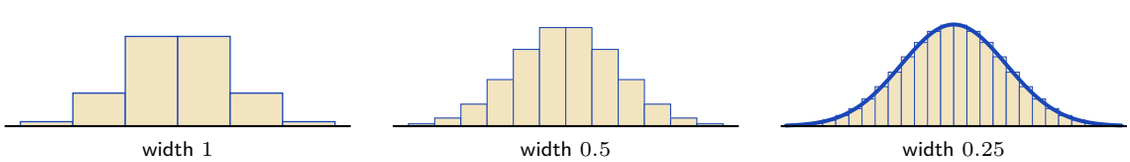


Figure 12.3 The same 5000 masses grouped into bars of decreasing width. Every version encloses total area 1; as the bars narrow, their tops approach a smooth curve. That curve is the density.

So probability for a continuous variable is **area**, and the curve whose area gives it is called a **probability density function**, written $f(x)$. Two conditions make f a valid density: it is never negative, and the total area under it is 1.

DEF HL A **probability density function** $f(x)$ satisfies

$$f(x) \geq 0 \quad \text{for all } x, \qquad \int_{-\infty}^{\infty} f(x)\,dx = 1,$$

and probabilities are areas: $P(a \leq X \leq b) = \int_a^b f(x)\,dx$.

The symbol $\int$ is new. It is the notation for **the area under a curve**, and $\int_a^b f(x)\,dx$ means the area under $y = f(x)$ between $x = a$ and $x = b$. Ch. 15 defines it and shows how to evaluate one, and the examples below use those methods; if you meet this section before Ch. 15, read each integral as an area and return to the calculations afterwards.

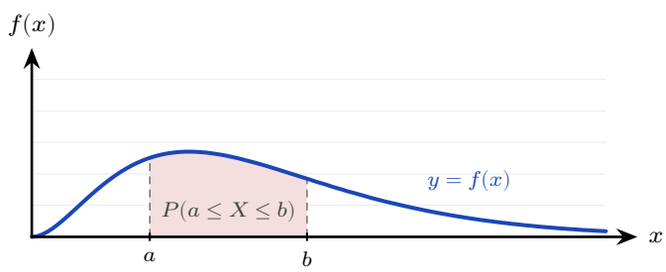


Figure 12.4 For a continuous variable, a probability is an area. The shaded region between a and b is $P(a \leq X \leq b)$, and the whole area under the curve is 1.

One consequence of this looks strange at first.

- A bar of zero width has zero area, so $P(X = c) = 0$ for every single value c.
- This does not mean the value is impossible. It means that no exact value carries a share of the total. A person's height is never exactly 165 cm to unlimited precision, only somewhere in a range near it.
- The endpoints therefore contribute nothing, so $P(a \leq X \leq b)$ and $P(a < X < b)$ are the same number. Here you may be careless about $<$ and $\leq$ in a way you cannot be with a discrete variable.

The expected value and variance carry over from the discrete case, with integrals replacing sums.

KEY · IB HL

$$E(X) = \int_{-\infty}^{\infty} x\,f(x)\,dx, \qquad \text{Var}(X) = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$$

Compare these with the discrete versions in §12.2. The structure is identical, and only the machinery has changed: a sum over a list of values has become an area over a range of them. As before, the standard deviation is $\sigma = \sqrt{\text{Var}(X)}$, and it is the number to quote

when you want a spread in the same units as X.

Three numbers describe where a density sits, and each is read off the curve in its own way.

- The **mode** is the value at which f is highest, that is, the peak of the curve.
- The **median** is the value that cuts the area in half, with 0.5 on each side.
- The **mean** is the balance point, $\int x f(x)\,dx$.

For a symmetric curve with a single peak all three sit together. For a lopsided one they do not, and the density worked through below shows all three.

Worked through. Take the density $f(x) = kx$ for $0 \le x \le 2$, and $f(x) = 0$ elsewhere. To find k, use the fact that the total area must be 1:

$$\int_0^2 kx\,dx = k\left[\tfrac{x^2}{2}\right]_0^2 = 2k = 1 \implies k = \tfrac{1}{2}.$$

Then

$$P(X < 1) = \int_0^1 \tfrac{x}{2}\,dx = \left[\tfrac{x^2}{4}\right]_0^1 = \tfrac{1}{4}.$$

The mean:

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

For the variance, first find $E(X^2)$, then subtract μ^2:

$$E(X^2) = \int_0^2 x^2 \cdot \tfrac{x}{2}\,dx = \int_0^2 \tfrac{x^3}{2}\,dx = \left[\tfrac{x^4}{8}\right]_0^2 = 2,$$

$$\text{Var}(X) = 2 - \left(\tfrac{4}{3}\right)^2 = 2 - \tfrac{16}{9} = \tfrac{2}{9}, \qquad \sigma = \sqrt{\tfrac{2}{9}} = \tfrac{\sqrt{2}}{3} \approx 0.471.$$

The **median** m splits the area in half:

$$\int_0^m \tfrac{x}{2}\,dx = \tfrac{1}{2} \implies \tfrac{m^2}{4} = \tfrac{1}{2} \implies m = \sqrt{2}.$$

The **mode** is where f is greatest. Since $f(x) = \frac{x}{2}$ increases across $[0, 2]$, the mode is at the right end, $x = 2$.

Collect the three centres: mode 2, median $\sqrt{2} \approx 1.414$, mean $\frac{4}{3} \approx 1.333$. Mean $<$ median $<$ mode: the density rises towards $x = 2$, so its long tail lies to the *left*. That is negative skew, in the language of Ch. 10.

CAUTION The mean, median, and mode of a continuous variable usually differ. The mode maximises $f(x)$; the median halves the area; the mean is the balance point $\int x f(x)\,dx$. For a symmetric density with a single peak all three coincide; for a skewed density they generally differ.

12.4.1 Piecewise densities

A density need not be a single formula. Any function that is non-negative and encloses total area 1 qualifies, including one built from pieces. The only new requirement is care with the arithmetic: integrals must be split at the joins, and the median may lie in either piece.

EXAMPLE 12.8

A continuous random variable has density

$$f(x) = \begin{cases} \dfrac{2x}{3} & 0 \le x \le 1, \\ \dfrac{3-x}{3} & 1 < x \le 3, \\ 0 & \text{otherwise.} \end{cases}$$

Verify that f is a valid density, find $P(X > 2)$, and find the median.
Both pieces are non-negative, and the total area splits at $x = 1$:

$$\int_0^1 \tfrac{2x}{3}\,dx + \int_1^3 \tfrac{3-x}{3}\,dx = \left[\tfrac{x^2}{3}\right]_0^1 + \tfrac{1}{3}\left[3x - \tfrac{x^2}{2}\right]_1^3 = \tfrac{1}{3} + \tfrac{2}{3} = 1. \checkmark$$

For $P(X > 2)$ only the second piece matters:

$$P(X > 2) = \int_2^3 \tfrac{3-x}{3}\,dx = \tfrac{1}{3}\left[3x - \tfrac{x^2}{2}\right]_2^3 = \tfrac{1}{6}.$$

For the median, first locate it: the area up to $x = 1$ is $\frac{1}{3} < \frac{1}{2}$, so the median lies in the second piece. Collect the remaining area there:

$$\tfrac{1}{3} + \int_1^m \tfrac{3-x}{3}\,dx = \tfrac{1}{2} \implies 3m - \tfrac{m^2}{2} - \tfrac{5}{2} = \tfrac{1}{2} \implies m^2 - 6m + 6 = 0,$$

giving $m = 3 - \sqrt{3} \approx 1.27$ (taking the root inside the interval). The mode is $x = 1$, where the two pieces meet at the peak of the triangle.

12.4.2 Modes by calculus

When a density rises and then falls, its mode lies inside the interval, and you find it by the standard method: differentiate and solve $f'(x) = 0$. A mode can also sit at an endpoint or at a corner, as in the two densities above, so always confirm where the maximum actually is before answering.

EXAMPLE 12.9

A continuous random variable has density $f(x) = kx^2(3 - x)$ for $0 \le x \le 3$, and zero elsewhere. Find k, the mode, and $E(X)$.

Total area 1:

$$\int_0^3 k(3x^2 - x^3)\,dx = k\Big[x^3 - \tfrac{x^4}{4}\Big]_0^3 = k\left(27 - \tfrac{81}{4}\right) = \tfrac{27k}{4} = 1 \implies k = \tfrac{4}{27}.$$

For the mode, maximise f:

$$f'(x) = k(6x - 3x^2) = 3kx(2 - x) = 0 \implies x = 0 \text{ or } x = 2.$$

Since f rises on $(0, 2)$ and falls on $(2, 3)$, the mode is $x = 2$. Finally,

$$E(X) = \int_0^3 k(3x^3 - x^4)\,dx = \tfrac{4}{27}\left[\tfrac{3x^4}{4} - \tfrac{x^5}{5}\right]_0^3 = \tfrac{4}{27} \cdot \tfrac{243}{20} = \tfrac{9}{5}.$$

Mean 1.8, mode 2: the longer tail of the density lies to the left of the peak, so the balance point sits below it.

YOUR TURN 12.4 Continuous Random Variables HL

Finding k and probabilities

22.

(a) A density is $f(x) = kx^2$ for $0 \le x \le 3$, and zero elsewhere. Find k.

(b) A density is $f(x) = k$ for $0 \le x \le 4$, and zero elsewhere. Find k and $P(X < 1)$.

(c) A density is $f(x) = \dfrac{x}{8}$ for $0 \le x \le 4$, and zero elsewhere. Find $P(X > 2)$.

(d) For the density in part (c), write down $P(X = 2)$, and explain why $P(X \le 2) = P(X < 2)$.

Mean, variance, median and mode

23.

(a) For the density $f(x) = \dfrac{x}{8}$, $0 \le x \le 4$, find $E(X)$.

(b) For the same density, find $\text{Var}(X)$.

(c) A density is $f(x) = 3x^2$ for $0 \le x \le 1$, and zero elsewhere. Find the median.

24. A continuous random variable has density $f(x) = kx(4 - x)$ for $0 \le x \le 4$, and zero elsewhere.

(a) Find k.

(b) Write down the mode of this density, and justify your answer.

(c) Find $E(X)$.

(d) Find $\text{Var}(X)$.

25. A continuous random variable has density $f(x) = k(4x - x^3)$ for $0 \le x \le 2$, and zero elsewhere.

(a) Find k.

(b) Use calculus to find the mode, giving your answer in exact form.

(c) Find $E(X)$.

Piecewise densities

26. A variable has density $f(x) = x$ for $0 \le x \le 1$, $f(x) = \frac{1}{2}$ for $1 < x \le 2$, and zero elsewhere.

(a) Verify that the total area is 1, and find $P(X < 1.5)$.

(b) Find the median.

27. A variable has density

$$f(x) = \begin{cases} \dfrac{x}{3} & 0 \le x \le 2, \\ \dfrac{2(3-x)}{3} & 2 < x \le 3, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Verify that f is a valid density.

(b) Find the median, giving your answer in exact form.

(c) Find $P(X > 2.5)$.

Concept check

28. A density is $f(x) = 2x$ for $0 \le x \le 1$, and zero elsewhere. A student says: "$f(0.8) = 1.6$, *which is greater than* 1, *so this cannot be a probability density.*" Explain why the student is wrong.

12.5 The Normal Distribution

If the histogram of bag masses is refined far enough, the tops of the bars approach a smooth curve, which is close to the **normal distribution**. Everything from the last section still applies to it: probability is area, and the total area is 1.

What earns this one curve a section of its own is how often it occurs. Heights, masses, measurement errors and examination marks often follow it approximately. The reason was given at the start of the chapter: when a quantity is the sum of many small independent effects, none of which dominates, its distribution is close to normal. That situation is common, so the curve is used widely in statistics.

The name can mislead. *Normal* here does not mean usual, correct or healthy. It is a nineteenth-century label that survived, and a great many real quantities are not normal at all. Incomes, city populations and waiting times are all strongly skewed. Before using the curve, check that the data actually resemble it.

Every normal curve, whatever its mean and spread, has the same four properties.

- It is symmetric about the mean, so the mean, median and mode are equal and all sit at the peak.
- It has a single maximum, at the centre, and falls away on both sides.
- The tails approach the horizontal axis but never meet it. Every interval of values, however far from the mean, therefore has a positive probability, however small.
- The total area beneath it is 1, as it is for any density.

Two numbers fix a normal distribution completely, and each does a different job.

- The **mean** μ locates the centre. Changing μ slides the whole curve left or right without altering its shape at all.
- The **standard deviation** σ sets the width. Increasing σ stretches the curve sideways, and because the area underneath must stay equal to 1, stretching it sideways also flattens it. A large σ therefore gives a wide, low bell and a small σ a narrow, tall one.

We write $X \sim N(\mu, \sigma^2)$. Note that the second entry is the *variance*, not the standard deviation, so take a square root before using σ.

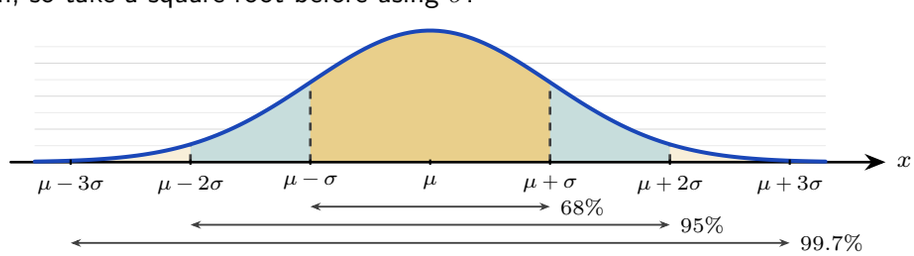


Figure 12.5 The normal curve, symmetric about μ. The three shaded bands reach one, two and three standard deviations either side of the mean, and hold about 68%, 95% and 99.7% of the values. Almost

nothing lies beyond three standard deviations.

The shape of the normal curve gives a quick approximation, the **68–95–99.7 rule**. It lets you judge, from μ and σ alone, where most of the data lie, without a calculator.

KEY **The 68–95–99.7 rule.** For $X \sim N(\mu, \sigma^2)$, approximately

$$68\% \text{ of values lie in } \mu \pm \sigma, \qquad 95\% \text{ in } \mu \pm 2\sigma, \qquad 99.7\% \text{ in } \mu \pm 3\sigma.$$

Read in reverse, the same rule measures how surprising a value is. Falling outside two standard deviations happens to about 5% of values, roughly one in twenty. Falling outside three happens to about 0.3%, roughly one in 370. That is why a reading three standard deviations from the mean is often treated, in quality control for example, as a signal that something has changed, rather than as ordinary variation.

The curve has an equation, and although you will not have to use it, seeing it explains where the shape comes from.

KEY **The normal density** (not assessed):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Read the exponent first, because it carries the whole shape. The quantity $(x-\mu)^2$ is a parabola: it is zero at $x = \mu$ and grows as x moves away from the mean, and it depends only on the *distance* from μ, not on the direction. The minus sign turns that growth into decay, since e^{-u} is largest when $u = 0$ and shrinks towards zero as u grows. So the curve is tallest at the mean, falls away at the same rate on both sides, and never quite reaches the axis. The $2\sigma^2$ underneath sets the speed of the fall: a large σ makes the exponent grow slowly, giving a wide, low bell. The factor $\dfrac{1}{\sigma\sqrt{2\pi}}$ in front changes no part of the shape; it is the number that makes the total area exactly 1.

This formula is not in the formula booklet, and you are not expected to use it. Normal probabilities in this course are found with a GDC, or from symmetry. It is here so that the bell is a consequence of something rather than a shape you are asked to accept.

Worked through. The heights of adult women are modelled by $X \sim N(165, 6^2)$ (in cm). Both $P(X < 172)$ and $P(158 < X < 172)$ are areas under the normal curve, read from a GDC with $\mu = 165$, $\sigma = 6$: $P(X < 172) \approx 0.878$ and $P(158 < X < 172) \approx 0.757$. The second answer is larger than the 68% of the rule of thumb, because 158 and 172 are about 1.17 standard deviations either side of the mean, a little more than one.

Questions ask for three shapes of area, and a calculator handles all three from the same two entries. An area to the left needs a lower limit far below the mean, an area to the right

needs an upper limit far above it, and an area between two values needs both limits as given.

EXAMPLE 12.10

The masses of eggs from a farm are modelled by $X \sim N(58, 4^2)$, in grams. Find (a) $P(X > 65)$, (b) $P(54 < X < 62)$.

Enter $\mu = 58$ and $\sigma = 4$ on a GDC.

(a) A right-hand tail has no natural upper limit, so supply one far above the mean, such as 10^{99}, and take the area from 65 upwards:

$$P(X > 65) \approx 0.0401.$$

(b) Here both limits are given:

$$P(54 < X < 62) \approx 0.683.$$

The limits in (b) sit exactly one standard deviation either side of 58, so the second answer is the 68% of the rule of thumb, computed rather than estimated.

A probability also answers a question about a whole batch. If each of N items follows the same normal distribution independently, the count that falls in some range is binomial with that probability, and its expected value is N times the probability. So the expected number in a batch is a normal calculation followed by one multiplication.

EXAMPLE 12.11

The lifetime of a light bulb is modelled by $X \sim N(1200, 100^2)$, in hours. A hospital installs 400 of these bulbs. Find (a) the probability that a bulb fails before 1000 hours, (b) the expected number of the 400 that fail before 1000 hours.

(a) With $\mu = 1200$ and $\sigma = 100$, a lower limit far below the mean gives

$$P(X < 1000) \approx 0.0228.$$

(b) Multiply the batch size by that probability, carried at more digits than the rounded 0.0228 of part (a):

$$400 \times 0.022750 = 9.100,$$

so about 9 bulbs are expected to fail before 1000 hours.

Round only at the end, and read the answer as a long-run average: a particular batch of 400 may contain any number of early failures.

YOUR TURN 12.5 The Normal Distribution

Shape and the 68–95–99.7 rule

29. The curves of $A \sim N(40, 3^2)$, $B \sim N(40, 6^2)$ and $C \sim N(55, 3^2)$ are drawn on the same axes.

(a) State which two curves have the same shape, and describe how they are related.
(b) State which curve has the lowest peak, and explain why.

30. Heights are modelled by $X \sim N(170, 8^2)$, in cm. Without a GDC, use the 68–95–99.7 rule to estimate the percentage of heights

(a) between 162 cm and 178 cm;
(b) between 146 cm and 194 cm;
(c) above 186 cm.

Normal probabilities

31. GDC

(a) $X \sim N(50, 10^2)$. Find $P(X < 60)$.
(b) $X \sim N(100, 15^2)$. Find $P(X > 120)$.
(c) $X \sim N(20, 4^2)$. Find $P(18 < X < 26)$.
(d) $Z \sim N(0, 1)$. Find $P(Z < 1.5)$.
(e) $X \sim N(500, 50^2)$. Find $P(X < 400)$.
(f) $X \sim N(30, 16)$. Find $P(X > 34)$.

32. GDC The masses of packets of rice are modelled by $X \sim N(500, 25^2)$, in grams.

(a) Find the probability that a packet weighs more than 540 g.
(b) Find $P(450 < X < 575)$, and compare your answer with the value the 68–95–99.7 rule gives.
(c) In a batch of 800 packets, find the expected number weighing more than 540 g.
(d) Write down the range of masses within which about 95% of packets lie.

Concept check

33. $X \sim N(50, 10^2)$. A student says: "*about 68% of values lie within one standard deviation of the mean, so $P(X < 60) \approx 0.68$.*" Identify the error, and estimate $P(X < 60)$ correctly without a GDC.

12.6 Standardisation and the Inverse Normal

Every normal distribution is the same bell, just shifted and scaled. **Standardising** expresses any value as the number of standard deviations it sits from its mean. That number is the **z-value**, and it converts any normal variable to the **standard normal** $Z \sim N(0, 1)$.

KEY · IB

$$z = \frac{x - \mu}{\sigma}$$

Standardising does two things at once. Subtracting μ moves the centre of the distribution to 0. Dividing by σ then rescales the spread, so that one unit of z means one standard deviation. The units cancel in that division, and that is what makes comparison possible: z is a pure number with no units attached, so a mass in grams and a time in seconds become directly comparable once standardised.

Read a z-value as follows.

- $z > 0$ means the value lies above the mean, and $z < 0$ means it lies below.
- $z = 0$ means the value is exactly the mean.
- The size of z is the distance from the mean, measured in standard deviations. A z-value of 1.5 means "one and a half standard deviations above the mean", whatever the original units were.

The 68–95–99.7 rule of §12.5 now reads very simply. About 68% of values have $|z| < 1$, about 95% have $|z| < 2$, and about 99.7% have $|z| < 3$. Standardising therefore lets you compare values drawn from different normal distributions, and it is also how you work backwards from a probability to a value.

Worked through. Exam scores are $N(60, 12^2)$. To see how many standard deviations above the mean a score of 78 lies, standardise it:

$$z = \frac{78 - 60}{12} = \frac{18}{12} = 1.5.$$

The score sits 1.5 standard deviations above the mean. A student scoring 78 on a different exam with $N(70, 4^2)$ would have $z = 2$, and so be higher *relative to their cohort*, despite the equal raw mark.

12.6.1 The inverse normal

Often the question runs in the other direction. Instead of being given a value and asked for a probability, you are given a probability and asked for the value that produces it. "What score is the 90th percentile?" is a question of this kind. The **inverse normal** on a GDC takes a probability and returns the corresponding value x.

One detail needs care. Unless you choose another tail setting, a calculator's inverse normal uses the area to the *left* of the unknown value, but examination questions are often phrased in terms of the area to the right. Converting between the two is a single subtraction, and forgetting it produces an answer on the wrong side of the mean.

KEY **Using the inverse normal.**

1. Write down the probability you are given, and decide whether it is the area to the left or to the right of the unknown value.
2. If it is a right-hand area, subtract it from 1. The calculator needs the left-hand area.
3. Enter that left-hand area together with μ and σ.
4. Read off x, then check that it falls on the expected side of the mean.

Step 4 is worth the few seconds it takes. An area greater than 0.5 must return a value above the mean, and an area less than 0.5 must return one below it. This single check catches the subtraction error before it costs marks.

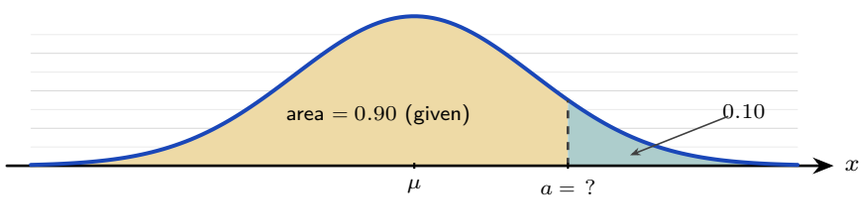


Figure 12.6 The inverse normal reverses the usual question. Instead of being given a and finding the area, you are given the area and must find a. A calculator expects the area to the *left* of a (gold), so a question phrased as a right-hand tail (teal), here 0.10, must be turned into its left-hand partner 0.90 first.

EXAMPLE 12.12

For heights $X \sim N(165, 6^2)$, find the height a below which 90% of women fall.
You need a with $P(X < a) = 0.90$. The inverse normal gives

$$a \approx 172.7 \text{ cm}.$$

Equivalently, the 90th percentile sits at $z \approx 1.282$, so $a = 165 + 1.282(6) \approx 172.7$, matching.

12.6.2 Finding an unknown μ or σ

The inverse normal also lets you recover a missing parameter. Suppose you know one probability and the value attached to it, but either μ or σ is missing. The method has three steps.

1. Use the inverse normal on the *standard* curve, $N(0, 1)$, to turn the probability into a z-value.
2. Substitute that z, the known value x, and whichever parameter you have into $z = \dfrac{x - \mu}{\sigma}$.

3. Solve the resulting equation for the one unknown that remains.

Multiplied through by σ, the equation becomes $x = \mu + z\sigma$, which is linear in μ and σ, so this last step is ordinary algebra.

EXAMPLE 12.13

A machine fills bottles so that volumes are normal with $\sigma = 5$ mL. Given that 80% of bottles contain less than 20 mL, find the mean μ.
Here $P(X < 20) = 0.80$. The inverse normal for the standard curve gives the matching z:

$$z = 0.8416, \qquad \text{so} \qquad 0.8416 = \frac{20 - \mu}{5}.$$

Solving, $20 - \mu = 4.21$, hence $\mu \approx 15.8$ mL.

12.6.3 Both parameters unknown

When neither μ nor σ is known, a single probability statement is not enough, because one equation cannot determine two unknowns. Two statements are required.

Rearranging $z = \dfrac{x - \mu}{\sigma}$ into $x = \mu + z\sigma$ shows why this works. Each statement produces one equation of that form, linear in μ and σ, and a pair of linear equations in two unknowns can be solved simultaneously in the usual way.

EXAMPLE 12.14

The masses of a species of fish are normally distributed. It is known that 5% of the fish have mass below 50 g and 10% have mass above 80 g. Find μ and σ.
Translate each statement into a z-value with the inverse normal:

$$P(X < 50) = 0.05 \implies z_1 = -1.645, \qquad P(X < 80) = 0.90 \implies z_2 = 1.282.$$

Each gives an equation $x = \mu + z\sigma$:

$$50 = \mu - 1.645\sigma, \qquad 80 = \mu + 1.282\sigma.$$

Subtracting, $30 = 2.926\sigma$, so $\sigma \approx 10.3$ g. Substituting back, $\mu = 50 + 1.645(10.25) \approx 66.9$ g.
Note the signs carefully: a value *below* the mean has a negative z. One wrong sign here changes both answers, so check each z against the picture before solving.

YOUR TURN 12.6 Standardisation and the Inverse Normal

Standardising and the inverse normal

34. GDC

(a) $X \sim N(60, 12^2)$. Find the z-value of $x = 84$.
(b) $X \sim N(70, 5^2)$. Find the value x for which $P(X < x) = 0.95$.
(c) $X \sim N(80, 10^2)$. Find the value x for which $P(X > x) = 0.15$.
(d) For $X \sim N(50, 8^2)$, find the lower quartile, the upper quartile, and the interquartile range.

Finding μ or σ

35. GDC

(a) $X \sim N(\mu, 8^2)$ with $P(X < 50) = 0.30$. Find μ.
(b) $X \sim N(\mu, 6^2)$ with $P(X > 40) = 0.20$. Find μ.
(c) $X \sim N(100, \sigma^2)$ with $P(X < 90) = 0.25$. Find σ.

36. GDC Scores are normally distributed with mean 50. Given that 10% of scores exceed 60, find the standard deviation σ.

Comparing with z-values

37. GDC Two examinations have marks $A \sim N(65, 10^2)$ and $B \sim N(72, 6^2)$. A student scores 75 in A and 80 in B. Find both z-values, and state which is the better performance relative to the cohort.

Both parameters unknown

38. GDC

(a) $X \sim N(\mu, \sigma^2)$ with $P(X < 20) = 0.10$ and $P(X < 35) = 0.90$. Find μ and σ.
(b) $X \sim N(\mu, \sigma^2)$ with $P(X < 30) = 0.20$ and $P(X > 52) = 0.05$. Find μ and σ.

Concept check

39. GDC $X \sim N(120, 15^2)$. To find x with $P(X > x) = 0.2$, a student enters an area of 0.2 into the inverse normal and obtains $x \approx 107.4$. Identify the error, find the correct value of x, and explain how a quick check would have caught the mistake.

Chapter 12 · Probability Distributions

Reflections

TOK The same situation can often be described by more than one model: a binomial count can be approximated by a normal curve (Challenge Question 26), and a real data set is never exactly either.

What criteria should decide between two models that both fit? Simplicity, accuracy, and agreement with existing theory can point different ways, and the choice is made by people.

IM In 1846 the Belgian astronomer Adolphe Quetelet found this curve in the chest measurements of 5738 Scottish soldiers, and took it as support for an idea he had introduced in 1835, *l'homme moyen*, the "average man." He treated this statistical fiction as an ideal that real people deviate from, although no actual person has all the average values at once. The normal curve is a fact about variation; the "average man" was an interpretation added to it. When we summarise a population by its mean, are we describing a real thing, or creating one?

RW The frequent appearance of the bell curve has an explanation: the central limit theorem. It says that the sum of many independent random effects, none of them dominant, tends toward a normal distribution, whatever the shape of the individual effects. This is why measurement errors, heights, and the means of large samples are often approximately normal. De Moivre found the curve as a limit of the binomial in 1733, and Laplace gave a first version of the theorem in 1810; general proofs came from Lyapunov (1901) and Lindeberg (1922).

EXT The binomial becoming normal is a theorem, not a loose comparison. Plot $B(n, 0.5)$ for increasing n and the discrete bars take the shape of the continuous bell. Galton built a machine, the quincunx, to show it physically: balls dropped through rows of pegs, each bounce an independent left-or-right trial, pile up at the bottom in a binomial distribution whose shape is close to normal. The binomial and normal distributions of this chapter are connected: as n increases, the first approaches the second.

End-of-Chapter Problems
Chapter 12: Probability Distributions

1. A discrete random variable X has the distribution below.

x	1	2	3	4
$P(X = x)$	0.2	k	0.35	0.15

(a) Find k. [2]
(b) Find $E(X)$. [2]
(c) Find Var(X). [3]

2. A charity raffle sells 500 tickets at \$2 each. There is one prize of \$300 and five prizes of \$50.

(a) Find the expected winnings (before cost) for a single ticket. [3]
(b) Hence find the expected gain or loss per ticket, and comment. [2]

3. GDC In a production line, 4% of items are faulty, independently. Items are packed in boxes of 40. Let X be the number of faulty items in a box.

(a) State the distribution of X. [1]
(b) Find $P(2 \leq X \leq 4)$. [2]
(c) Find the probability that X is greater than its mean. [2]
(d) Write down the most likely number of faulty items in a box. [1]

4. GDC A fair six-sided die is rolled 30 times. Let X be the number of sixes.

(a) State the distribution of X and find its mean. [2]
(b) Find $P(X = 5)$. [2]
(c) Find $P(X \geq 7)$. [2]

5. A continuous random variable has density $f(x) = k(x + 1)$ for $0 \leq x \leq 2$, and zero elsewhere.

(a) Find k. [2]
(b) Find $P(X < 1)$. [2]
(c) Find $E(X)$. [3]
(d) Find the median and the mode. [3]

6. GDC The heights of a species of plant are $X \sim N(170, 8^2)$ in cm.

(a) Find $P(X < 160)$. [2]
(b) Find $P(160 < X < 185)$. [2]
(c) Find the proportion of plants taller than 180 cm. [2]

7. GDC Test scores are normally distributed with mean 500 and standard deviation 80.

(a) Find the score exceeded by 30% of candidates. [2]
(b) The middle 90% of scores lie between a and b, symmetric about the mean. Find a and b. [3]

8. GDC A variable X is normally distributed. It is known that $P(X < 40) = 0.10$ and $P(X < 70) = 0.75$.

(a) Write down two equations involving μ and σ using standardised values. [3]
(b) Find μ and σ. [4]

9. The random variable X has $E(X) = 5$ and $\text{Var}(X) = 4$.

(a) Find $E(X^2)$. [1]
(b) The variable $Y = aX + b$, where $a > 0$, has $E(Y) = 0$ and $\text{Var}(Y) = 1$. Find a and b. [3]

10. GDC Reaction times are $X \sim N(60, 12^2)$ in milliseconds. A reaction is "slow" if it exceeds 78 ms.

(a) Find the probability that a single reaction is slow. [2]
(b) Thirty reactions are recorded independently. Find the expected number of slow reactions. [2]
(c) Find the probability that at least one of the thirty is slow. [3]

11. A continuous random variable has density $f(x) = \dfrac{1}{8}(4 - x)$ for $0 \le x \le 4$, and zero elsewhere.

(a) Verify that f is a valid probability density function. [3]
(b) Find $P(X < 2)$. [2]
(c) Find the mode. [1]

12. A discrete random variable X takes the values $1, 2, 3, 4$ with $P(X = 1) = 0.2$, $P(X = 2) = a$, $P(X = 3) = b$, and $P(X = 4) = 0.1$. It is known that $E(X) = 2.4$.

(a) Find the values of a and b. [4]
(b) Find $P(X > 2 \mid X \ge 2)$. [3]

13. A stall game costs $\$c$ to play. A bag holds 3 red and 7 blue counters, and the player draws two counters without replacement. The player wins \$10 if both counters are red, \$2 if

exactly one is red, and nothing otherwise.

(a) Find the expected winnings from a single play. [2]
(b) Find the value of c for which the game is fair. [1]
(c) The stall sets $c = 2$. Find the player's expected loss over 100 plays. [2]

14. GDC A gardener plants 12 seeds. Each seed germinates with probability 0.85, independently of the others. Let X be the number that germinate.

(a) State two conditions that make a binomial model appropriate here. [2]
(b) Find $P(X = 10)$. [2]
(c) Find $P(X \geq 10)$. [2]
(d) Find the mean and standard deviation of X. [2]

15. GDC A darts player hits the bullseye with probability 0.3 on each throw, independently.

(a) In 8 throws, find the probability of exactly 3 bullseyes. [2]
(b) Find the least number of throws needed for the probability of at least one bullseye to exceed 0.98. [3]

16. GDC The random variable $X \sim B(10, p)$ satisfies $P(X = 0) = 0.0282$.

(a) Find the value of p. [3]
(b) Find $P(X = 3)$. [2]
(c) Write down $E(X)$ and $\text{Var}(X)$. [2]

17. GDC The lifetime of a battery is $X \sim N(120, 15^2)$, in hours.

(a) Find the probability that a battery lasts less than 95 hours. [2]
(b) The manufacturer wants no more than 2% of batteries to fail before a guaranteed time of t hours. Find the largest possible value of t. [2]
(c) Given that a battery has lasted 130 hours, find the probability that it lasts more than 140 hours. [2]

18. GDC The masses of apples are $X \sim N(160, 20^2)$, in grams. The heaviest 8% are sold as premium, and the lightest 5% are rejected.

(a) Find the minimum mass of a premium apple. [2]
(b) Find the maximum mass of a rejected apple. [2]
(c) Write down the proportion of apples that are neither premium nor rejected. [1]

19. GDC Journey times are normally distributed with mean 25 minutes and standard deviation σ. It is known that $P(X < 30) = 0.85$.

(a) Find σ. [3]
(b) Hence write down $P(X < 20)$, justifying your answer. [2]

20. GDC The masses of loaves of bread are normally distributed with unknown mean μ and standard deviation σ. It is known that 2% of loaves have mass below 300 g and 5% have mass above 340 g.

(a) Write down two equations for μ and σ using standardised values. [3]
(b) Find μ and σ. [3]

21. A continuous random variable has density

$$f(x) = \begin{cases} k & 0 \leq x \leq 1, \\ k(2-x) & 1 < x \leq 2, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that $k = \frac{2}{3}$. [3]
(b) Sketch the graph of f. [2]
(c) Find $P(X > 1.5)$. [2]
(d) Find the median of X. [2]
(e) Find $E(X)$. [3]

22. A continuous random variable has density $f(x) = kx(2-x)^2$ for $0 \leq x \leq 2$, and zero elsewhere.

(a) Show that $k = \frac{3}{4}$. [2]
(b) Use calculus to find the mode of X. [2]
(c) Find $E(X)$. [2]
(d) Find Var(X). [3]

23. GDC Scores on an examination are $X \sim N(55, 10^2)$. The pass mark is 45.

(a) Find the probability that a randomly chosen candidate passes. [2]
(b) Six candidates are chosen at random. Find the probability that exactly five of them pass. [3]
(c) Find the probability that at least five of them pass. [2]

24. GDC In 1846 a Belgian astronomer published the chest measurements, in inches, of 5738 Scottish militiamen, and used them as evidence that human measurements follow a normal distribution. His table is reproduced below.

Chest (in)	33	34	35	36	37	38	39	40
Frequency	3	18	81	185	420	749	1073	1079
Chest (in)	41	42	43	44	45	46	47	48
Frequency	934	658	370	92	50	21	4	1

(a) Show that the mean chest measurement is 39.8 inches, to 3 significant figures. [3]

(b) The standard deviation of the measurements is 2.05 inches. Assuming $X \sim N(39.8,\, 2.05^2)$, find $P(X < 38.5)$. [2]

(c) Find the proportion of the 5738 soldiers who actually measured 38 inches or less, and compare it with your answer to part (b). [3]

(d) Find the proportion of soldiers whose measurement lies within one standard deviation of the mean, and compare it with the value predicted by the normal model. [3]

(e) Comment on how well the normal distribution describes this real data set. [2]

Challenge Questions

25. GDC **Being ahead after n games.** At a fair, a stall offers this game. You pay \$2, roll two fair dice, and are paid \$20 if the total is 12, \$5 if the total is 10 or 11, and nothing otherwise. Let G be your net gain from one game, the payment received minus the \$2 stake. The games are independent, and after n games you are *ahead* if your total net gain is greater than 0.

(a) Write down the probability distribution of G, and show that $E(G) = -\frac{3}{4}$ and $\text{Var}(G) = \frac{625}{48}$. [4]

(b) Find the probability that you are ahead after 1 game, and the probability that you are ahead after 5 games. [5]

(c) The table gives the probability of being ahead after n games for larger n, calculated exactly by computer.

n	10	50	200
$P(\text{ahead})$	0.198	0.0711	0.00342

Using these values and part (b), describe how the probability of being ahead changes as n increases, and make a conjecture about its behaviour for large n. [2]

(d) Let T_n be your total net gain after n games. You may use the facts that $E(T_n) = nE(G)$ and that the standard deviation of T_n is $\sigma\sqrt{n}$, where $\sigma^2 = \text{Var}(G)$. Assuming that T_n is approximately normally distributed, show that

$$P(\text{ahead after } n \text{ games}) \approx P\left(Z > \frac{0.75\sqrt{n}}{\sigma}\right),$$

where $Z \sim N(0, 1)$. Evaluate this for $n = 10$, 50 and 200, and compare with the table. [5]

(e) The standard deviation of T_n increases with n. Explain why the probability of being ahead nevertheless decreases towards 0. [2]

(f) Using the normal approximation, find the least number of games after which the stall is at least 99% sure to have made a profit. [3]

(g) Exact calculation shows that the stall needs about 156 games, not the number found in part (f), and at $n = 200$ the normal approximation gives about half the exact probability in the table. Suggest a reason for these differences. Then investigate how changing the payments, while keeping $E(G) = -\frac{3}{4}$, changes how often players finish ahead, and discuss which design a stall owner might prefer. [3]

26. GDC **From the binomial to the bell.** The Reflections claim that a binomial distribution comes to look normal as n grows. This question tests that claim, and finds the adjustment that makes it work.

Throughout, let $X \sim B(50, 0.5)$ and let $Y \sim N(\mu, \sigma^2)$ have the same mean and variance as X.

(a) Write down $E(X)$ and $\mathrm{Var}(X)$, and hence state μ and σ for Y. [2]

(b) Find $P(20 \le X \le 30)$ using the binomial. [2]

(c) Find $P(20 < Y < 30)$ using the normal. Comment on how well it matches part (b). [3]

(d) X takes whole-number values, while Y takes every value in between. The bar of the histogram for $X = 20$ covers the interval from 19.5 to 20.5. Using this idea, find $P(19.5 < Y < 30.5)$ and compare it with part (b). [4]

(e) Repeat parts (b), (c) and (d) for $X \sim B(10, 0.5)$ and the range $4 \le X \le 6$. [4]

(f) State which of the two normal calculations should be used, and describe how the quality of the approximation changes as n increases. [3]

(g) Many textbooks state that the normal approximation may be used when $np > 5$ and $n(1-p) > 5$. Test this rule: for several binomial distributions with np close to 5, such as $B(50, 0.1)$ and $B(10, 0.5)$, compare an exact binomial probability with its adjusted normal approximation. Discuss whether np alone decides how good the approximation is, and suggest how the rule might be improved. [3]

27. The most likely value. Let $X \sim B(n, p)$ with $0 < p < 1$. This question finds the mode of X, the value it takes most often.

(a) Show that for $r = 1, 2, \ldots, n$,

$$\frac{P(X = r)}{P(X = r-1)} = \frac{(n - r + 1)\,p}{r\,(1-p)}.$$

[3]

(b) Hence show that $P(X = r) \ge P(X = r-1)$ if and only if $r \le (n+1)p$. [3]

(c) Deduce that the mode of X is $\lfloor (n+1)p \rfloor$ when $(n+1)p$ is not an integer, and use this to find the most likely number of successes when $X \sim B(20, 0.3)$. [3]

(d) Explain what happens when $(n+1)p$ is an integer, and illustrate your answer with

$X \sim B(9, 0.5)$. [3]

(e) Show that the mode found in part (c) differs from the mean np by less than 1. Then investigate, for several values of n and p, how far the median of X can be from np, and state a conjecture. [3]

28. GDC **Mode, median and mean of one family.** For a constant $a > 0$, a continuous random variable X has probability density function

$$f(x) = \begin{cases} kx^a(1-x), & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Show that $k = (a+1)(a+2)$. [2]

(b) Show that the mode of X is $\dfrac{a}{a+1}$ and that $E(X) = \dfrac{a+1}{a+3}$. [4]

(c) Show that, for $0 \le x \le 1$, $P(X \le x) = (a+2)x^{a+1} - (a+1)x^{a+2}$. For $a = 2$ and for $a = \frac{1}{2}$, find the median of X and write the mode, median and mean in increasing order. Make a conjecture about the order when $a > 1$ and when $0 < a < 1$. [5]

(d) Prove that the mean is less than the mode if and only if $a > 1$, and state what happens when $a = 1$. [3]

(e) Show that $P\left(X \le \dfrac{a}{a+1}\right) = 2\left(\dfrac{a}{a+1}\right)^{a+1}$. You may assume that $g(t) = \left(1 - \frac{1}{t}\right)^t$ is increasing for $t > 1$, and that $g(t) \to e^{-1}$ as $t \to \infty$. Deduce that the median is less than the mode if and only if $a > 1$. Find the limit of $P\left(X \le \frac{a}{a+1}\right)$ as $a \to \infty$, and interpret it. [5]

(f) A proof that the mean is less than the median when $a > 1$ is harder. Check this numerically for several values of a. Then investigate either the wider family $f(x) = kx^a(1-x)^b$, or how well the rule of thumb mode $-$ mean ≈ 3(median $-$ mean) describes this family. [3]

Chapter 13

Limits & Derivatives

13

SYLLABUS 5.1–5.4 5.6 5.7 5.12 5.13 5.15

A car's speedometer reads 80 km/h. If you ask what that number means, you meet a problem that took mathematicians many centuries to settle.

Speed is distance divided by time. But the speedometer shows your speed *right now*, at a single instant. In one instant no time passes, and no distance is covered. Speed at an instant therefore appears to be $0 \div 0$, which has no value. How can the needle point to 80?

The solution is to stop asking about the instant, and to ask instead about short intervals around it.

Measure your average speed over the next hour. Then over the next minute. Then over the next second, and then over the next hundredth of a second. Each interval is shorter than the one before, and each average lies closer to one particular number. That number is the speed at the instant. The averages approach it, even though no calculation made at the instant itself can reach it.

This idea is called the **limit**. When a value cannot be reached directly, you find it by asking what nearby values approach. The rest of this chapter, and most of calculus, is built on it.

The **derivative** follows from the limit. It measures the exact rate at which a quantity changes: how steep a curve is at one point, or how fast an apple is falling at the moment it lands. This chapter builds the derivative from the limit, then develops methods for calculating it quickly.

By the end of this chapter you will evaluate limits, decide when a function is continuous and differentiable, and differentiate from first principles. You will apply the product, quotient and chain rules, differentiate the standard functions of the course, and find the equations of a tangent and a normal. You will evaluate limits of the form $\frac{0}{0}$ and $\frac{\infty}{\infty}$ with l'Hôpital's rule. You will also differentiate more than once, and meet the second derivative and the notation for every higher derivative. Calculus begins here because a derivative answers two questions at once. Geometrically it is the gradient of a curve at a point. Physically it is a rate of

© 2026 Hasan Khan ibxhasan.github.io/bettermath

change. Most later applications use one of those two readings.

13.1 Limits

A **limit** is the value a function approaches as its input approaches some point. It does not matter what the function does at that point itself. We write

$$\lim_{x \to a} f(x) = L$$

to mean: as x gets **arbitrarily close** (as close as you like) to a, $f(x)$ gets arbitrarily close to L. The value of f at a itself plays no part, so a limit can exist at a point where the function is undefined.

For most functions at most points, the limit is found by direct substitution. If f is defined at a and its graph passes through that point unbroken, then the value it approaches is simply the value it takes:

$$\lim_{x \to 2} (x^2 + 3x) = 2^2 + 3(2) = 10.$$

The cases that need a method are the ones where substitution fails. When substitution gives $\frac{0}{0}$, factorise the numerator and the denominator and cancel the common factor. The cancelled function agrees with the original everywhere except at the hole, so its value there is the limit.

Worked through. To find $\lim_{x \to 3} \frac{x^2 - 9}{x - 3}$, first try substituting $x = 3$. It gives $\frac{0}{0}$, which is undefined: the function has a hole at $x = 3$. But for every $x \neq 3$ we can factor and cancel:

$$\frac{x^2 - 9}{x - 3} = \frac{(x - 3)(x + 3)}{x - 3} = x + 3.$$

Away from the hole the function is simply $x + 3$, so as $x \to 3$ it approaches 6:

$$\lim_{x \to 3} \frac{x^2 - 9}{x - 3} = 6.$$

At $x = 3$ the function has no value at all, yet 6 is the value it approaches. The difference between "approaches" and "equals" is what a limit measures.

The form $\frac{0}{0}$ is called **indeterminate**. It does not mean the limit fails to exist. It means more work is needed.

One limit of this kind must be memorised, because the derivatives of the trigonometric functions depend on it:

$$\lim_{x \to 0} \frac{\sin x}{x} = 1.$$

When there is no algebraic method, you can *estimate* a limit. Calculate the function at values closer and closer to the target, from both sides. Here is the result for this one, with x in radians:

x	-0.1	-0.01	0.01	0.1
$\dfrac{\sin x}{x}$	0.99833	0.99998	0.99998	0.99833

Both sides approach 1. A table or a zoomed-in graph is a valid way to **conjecture** (guess, before proving) the value of a limit. On a calculator paper it is often the fastest way to check an algebraic answer.

The phrase "from both sides" needs its own notation. Write $x \to a^-$ when x approaches a through values *less* than a (from the left), and $x \to a^+$ when it approaches through values *greater* than a (from the right). The value that $f(x)$ approaches from each side is a **one-sided limit**:

$$\lim_{x \to a^-} f(x) \quad \text{(left-hand limit)}, \qquad \lim_{x \to a^+} f(x) \quad \text{(right-hand limit)}.$$

DEF The limit $\lim_{x \to a} f(x)$ exists only when the left-hand and right-hand limits both exist and are **equal**. Their common value is the limit.

The two graphs below show both cases. On the left, the graph has a jump at $x = 2$. Moving along the curve from the left, the height approaches 1; moving from the right, it approaches 3. The one-sided limits differ, so $\lim_{x \to 2} f(x)$ does not exist. On the right, the graph has only a hole at $x = 2$. Both sides approach the same height, 3, so the limit exists and equals 3, even though the function has no value at $x = 2$.

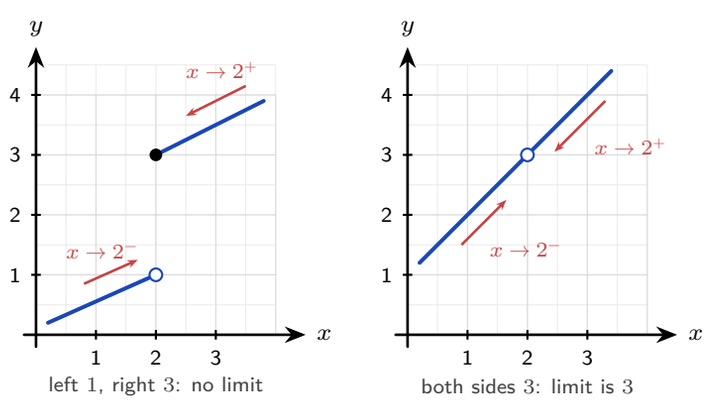


Figure 13.1 Left and right-hand limits. An open circle marks a point that is not on the graph. **Left:** at a jump the two sides approach different heights, so the limit does not exist. **Right:** at a hole both sides approach the same height, so the limit exists although $f(2)$ is undefined.

13.1.1 Reading a limit off the graph

A graph often shows why a limit takes the value it does. The two fractions $\dfrac{\sin x}{x}$ and $\dfrac{\cos x - 1}{x}$ both give $\frac{0}{0}$ at $x = 0$, yet they approach different values. The reason is how fast each numerator approaches 0 compared with the denominator x.

Sketch $y = \sin x$ and $y = x$ on the same axes. Near the origin the two graphs are almost the same: at $x = 0.1$, $\sin x = 0.0998$. The numerator and the denominator are almost equal, so their ratio approaches 1.

Now sketch $y = \cos x - 1$. Near the origin this graph is almost flat, because $\cos x$ has a maximum at $x = 0$. At $x = 0.1$, $\cos x - 1 = -0.0050$, while $x = 0.1$. The numerator is very small compared with the denominator, so their ratio approaches 0.

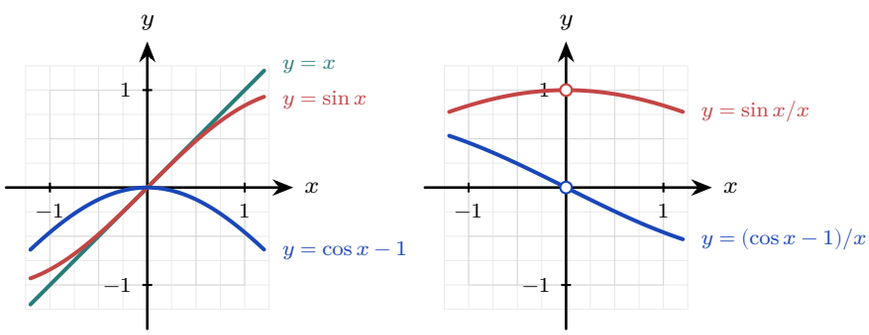


Figure 13.2 **Left:** near the origin $\sin x$ is almost the same as x, while $\cos x - 1$ is almost flat. **Right:** the two quotients. Each has a hole at $x = 0$, but from both sides $\frac{\sin x}{x}$ approaches 1 and $\frac{\cos x - 1}{x}$ approaches 0.

Both limits are needed later: they are the two facts from which the derivatives of $\sin x$ and $\cos x$ are built.

$$\lim_{x \to 0} \frac{\sin x}{x} = 1, \qquad \lim_{x \to 0} \frac{\cos x - 1}{x} = 0.$$

Not every limit exists. A limit **converges** when the values approach one finite number, and **diverges** when they do not. Three examples show the difference: $\frac{1}{x^2}$ increases without bound as $x \to 0$; $\frac{1}{x}$ increases without bound as $x \to 0^+$ but decreases without bound as $x \to 0^-$; while $\sin x$ oscillates between -1 and 1 as $x \to \infty$ and does not approach any single value. For $\frac{1}{x}$, writing that the limit "equals infinity" hides the disagreement between the two sides, so it is not an acceptable answer. You have met this distinction before: the infinite geometric series of Ch. 2 converges exactly when $|r| < 1$, because its partial sums approach a single value, and diverges otherwise.

Deciding which of the two happens as $x \to \infty$ needs a method, because there is no value of x to substitute. For a ratio of polynomials this method works: divide every term, above and below, by the highest power of x that appears. What is left is a set of constants together with terms like $\dfrac{1}{x}$ and $\dfrac{1}{x^2}$, and each of those approaches 0.

EXAMPLE 13.1

Evaluate (a) $\lim_{x\to\infty}\dfrac{5x-2}{x}$ and (b) $\lim_{x\to\infty}\dfrac{4x^2-x}{3x^2+5}$.

(a) The highest power is x, so divide both terms of the numerator by it:

$$\lim_{x\to\infty}\frac{5x-2}{x}=\lim_{x\to\infty}\left(5-\frac{2}{x}\right)=5-0=5.$$

(b) The highest power is x^2, so divide all four terms by x^2:

$$\lim_{x\to\infty}\frac{4x^2-x}{3x^2+5}=\lim_{x\to\infty}\frac{4-\dfrac{1}{x}}{3+\dfrac{5}{x^2}}=\frac{4-0}{3+0}=\frac{4}{3}.$$

Every term that approaches 0 came from a lower power of x than the one you divided by. When the top and the bottom have the same highest power, as here, what remains is the ratio of the leading coefficients: $\frac{5}{1}$ in (a) and $\frac{4}{3}$ in (b).

YOUR TURN 13.1 Limits

Direct substitution and factorising

1. Find each limit by direct substitution.

(a) $\lim_{x\to2}(3x-1)$

(b) $\lim_{x\to0}(x^2+2x+5)$

(c) $\lim_{x\to3}(x^2-9)$

(d) $\lim_{x\to1}\dfrac{x+1}{x+3}$

2. Each of these gives $\frac{0}{0}$. Factorise, cancel, then take the limit.

(a) $\lim_{x\to4}\dfrac{x^2-16}{x-4}$

(b) $\lim_{x\to1}\dfrac{x^2-1}{x-1}$

(c) $\lim_{x\to2}\dfrac{x^2-x-2}{x-2}$

(d) $\lim_{x\to-3}\dfrac{x^2+x-6}{x+3}$

One-sided limits

3. Explain why $\lim\limits_{x\to 0}\dfrac{|x|}{x}$ does not exist.

4. For each function, find the left-hand and right-hand limits at the value of x given, and state whether the limit exists there.

(a) $f(x)=\begin{cases}x^2, & x<1\\ 2x+1, & x\geq 1\end{cases}$ at $x=1$

(b) $g(x)=\begin{cases}x+2, & x<3\\ 8-x, & x>3\end{cases}$ with $g(3)=1$, at $x=3$. Compare the limit with $g(3)$.

5. Find $\lim\limits_{x\to 1^-}\dfrac{x^2-1}{|x-1|}$ and $\lim\limits_{x\to 1^+}\dfrac{x^2-1}{|x-1|}$, and hence state whether $\lim\limits_{x\to 1}\dfrac{x^2-1}{|x-1|}$ exists. *(Hint: $|x-1|=1-x$ when $x<1$)*

Limits at infinity

6. State whether each of the following converges or diverges as $x\to\infty$, and give the limit where one exists.

(a) $\dfrac{1}{x}$ (b) x^2

(c) e^{-x} (d) $2+\dfrac{5}{x}$

7. Evaluate each limit by dividing every term by the highest power of x.

(a) $\lim\limits_{x\to\infty}\dfrac{6x-1}{2x+3}$

(b) $\lim\limits_{x\to\infty}\dfrac{2x^2+3x}{5x^2-1}$

(c) $\lim\limits_{x\to\infty}\dfrac{x+4}{x^2+1}$

Estimating a limit

8. GDC Consider $y=\dfrac{\sin x}{x}$.

(a) Write down the value of $\lim\limits_{x\to 0}\dfrac{\sin x}{x}$.

(b) Graph the curve for $-10\leq x\leq 10$ on your GDC, and sketch what you see.

(c) Describe what the graph does at $x=0$, and what it does as $x\to\pm\infty$.

9. GDC Work in radians.

(a) Evaluate $\dfrac{\cos x-1}{x}$ for $x=\pm 0.1$ and $x=\pm 0.01$. State the value that its limit as $x\to 0$ appears to take.

(b) Estimate $\lim\limits_{x\to 0}\dfrac{1-\cos x}{x^2}$ by evaluating the expression at $x=0.1$ and $x=0.01$. State the exact value you think the limit takes.

Concept check

10. A student is asked for $\lim_{x\to 3} \dfrac{x^2-9}{x-3}$ and writes: "*substituting $x=3$ gives $\frac{0}{0}$, so the limit does not exist.*" Explain why this conclusion is wrong, and find the limit.

13.2 The Derivative

Calculus is the mathematics of change. You already know how to measure change: it is the gradient.

Rise over run gives the change in y for each unit of change in x. If the gradient doubles, the quantity changes twice as fast.

There is one problem. A gradient is defined for a *straight line* only. One number describes the whole line, because a line changes at the same rate everywhere.

Very few real quantities behave like that. A cooling cup, a thrown ball, a growing population: in each case the rate of change is itself changing. So the question "what is the gradient of this curve?" has no single answer. The answer is different at every point.

So ask a smaller question. Do not ask for the gradient of the curve. Ask for the gradient of the curve *at one point*. Once you can answer that for one point, you can answer it for every point, and the rate of change becomes a function of its own.

Here is the method. On the graph of $y=f(x)$, choose your point x and a second point $x+h$ close to it. Join the two points with a straight line, called the **secant**, and find its gradient. You can already calculate that:

$$\frac{f(x+h)-f(x)}{h},$$

the rise over the run. This is the average rate of change over the interval of width h.

The secant gradient is still not the answer. It is an average over an interval of width h, and the question was about a single point.

Now make h smaller and smaller. The second point moves towards the first, and the secant turns towards one limiting line, which touches the curve at the chosen point. That line is the **tangent**. Its gradient is the rate of change at a single point, and it is called the **derivative**.

That is the idea. A gradient is defined only for a line, so we find the line that matches the curve at that point, and use its gradient instead.

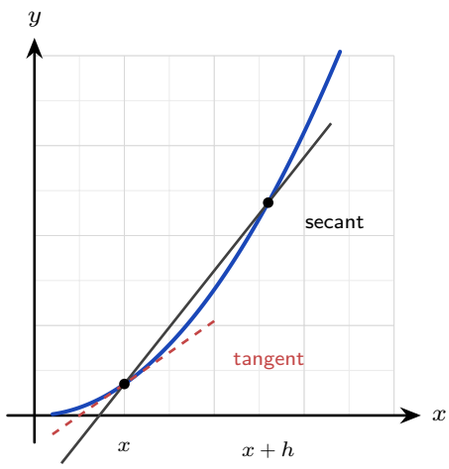


Figure 13.3 The secant joins two points on the curve and measures an average rate of change. As $h \to 0$ the second point moves to the first and the secant approaches the tangent.

DEF In words: the **derivative** of f at a point is the **gradient of the tangent** to the curve $y = f(x)$ at that point. It is the limit of the gradients of the secants through that point, as the second point moves towards it.

DEF · IB HL The **derivative** of f at x, written $f'(x)$, is

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

when this limit exists. Computing a derivative directly from this definition is called differentiation **from first principles**.

The derivative is itself a function, called the **gradient function**. For each value of x, the number $f'(x)$ is the gradient of the curve $y = f(x)$ at the point with that x-coordinate. For example, the derivative of $f(x) = x^2$ is $f'(x) = 2x$ (§13.3). So the gradient of $y = x^2$ is 2 at $x = 1$, 0 at $x = 0$ and -4 at $x = -2$. One formula, $2x$, gives the gradient at every point of the curve.

13.2.1 Notation, and where it came from

Calculus was developed independently by two people, and each was trying to solve a different problem.

- **Isaac Newton**, in England in the 1660s, was studying *motion*: falling bodies, and later the orbits of the planets under gravity. He needed the speed of a moving object at a single instant, which is the problem of the speedometer at the start of this chapter.
- **Gottfried Leibniz**, in Germany and France in the 1670s, was studying the *geometry of curves*: how to find the tangent to any curve, and the area under it. He wanted one set of symbols and rules that would work for every curve.

Both built on earlier methods for tangents by Pierre de Fermat and Isaac Barrow. Newton's own notation, a dot above the variable, is still seen in physics, but it is not used in this course. The notation used here comes from Leibniz.

The IB course uses these forms for the first derivative:

$$\frac{dy}{dx}, \qquad f'(x), \qquad \frac{dV}{dr}, \qquad \frac{ds}{dt}.$$

Leibniz wrote $\frac{dy}{dx}$, built from $\frac{\Delta y}{\Delta x}$ for a finite change, with the d marking the limit as the change approaches zero. The form $f'(x)$ names the function being differentiated. The form $\frac{d}{dx}f(x)$, used in the formula cards of this chapter, means "the derivative of $f(x)$ with respect to x".

The Leibniz form is the one that carries the most information, because it names both variables. When the quantities are not called x and y, the letters change with them. A volume V changing with radius r has derivative $\frac{dV}{dr}$, and a displacement s changing with time t has derivative $\frac{ds}{dt}$. Read each as a rate: how fast the top quantity changes per unit of the bottom one. Examination questions use these forms often, and choosing the right letters is an important part of setting up the problem.

One caution about $\frac{dy}{dx}$. It is written as a fraction, but it is not one. It is a single symbol standing for the limit of the fraction $\frac{\Delta y}{\Delta x}$, and dy and dx have no independent meaning at this level. The notation is still well designed, because the limit keeps some of the behaviour of a fraction. That is why the chain rule in §13.6 can be written as $\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$, with the du appearing to cancel. The cancelling is a memory aid, not a proof, but it works because the underlying limits behave that way.

Leibniz's notation became the standard one, largely for a practical reason: it shows the variable of differentiation, so it extends directly to the chain rule, to implicit differentiation, and to integration, where $\int f(x)\,dx$ carries the same dx.

13.2.2 Where derivatives are used

A derivative is a rate of change, and rates of change appear wherever one quantity depends on another.

- **Motion.** Differentiate displacement with respect to time to get velocity, and again to get acceleration. This is how a vehicle's braking distance is calculated, and it is the subject of Ch. 14.
- **Optimisation.** A curve is flat where its derivative is zero, so setting $f'(x) = 0$ locates maxima and minima. Manufacturers use this to find the packaging that uses the least material for a required volume.
- **Medicine.** The rate at which a drug leaves the bloodstream determines the dosing interval, and that rate is a derivative of the concentration with respect to time.

- **Economics.** *Marginal* cost is the derivative of total cost with respect to quantity produced, and it answers what one more unit costs to make.
- **Computer graphics.** A surface is shaded according to the direction it faces, which is found from derivatives of the surface, and this shading helps rendered objects look solid.

Every one of these is the same calculation, applied to a different pair of quantities.

To differentiate from first principles, substitute into $\dfrac{f(x+h)-f(x)}{h}$, expand, and simplify until the factor h in the denominator cancels. Only then let $h \to 0$, because the division by h is valid only while $h \neq 0$.

Worked through. To differentiate $f(x) = x^2 + 4x$ from first principles, form the difference quotient and expand:

$$\frac{f(x+h)-f(x)}{h} = \frac{(x+h)^2 + 4(x+h) - x^2 - 4x}{h} = \frac{2xh + h^2 + 4h}{h} = 2x + h + 4.$$

Now let $h \to 0$:

$$f'(x) = \lim_{h \to 0} (2x + h + 4) = 2x + 4.$$

You must cancel the h before taking the limit. While $h \neq 0$ the division is allowed. Only after cancelling is it safe to set h to zero.

EXAMPLE 13.2

Differentiate $f(x) = \dfrac{1}{x}$ from first principles.

$$\frac{f(x+h)-f(x)}{h} = \frac{1}{h}\left(\frac{1}{x+h} - \frac{1}{x}\right) = \frac{1}{h} \cdot \frac{x-(x+h)}{x(x+h)} = \frac{-h}{h\,x(x+h)} = \frac{-1}{x(x+h)}.$$

As $h \to 0$:

$$f'(x) = \lim_{h \to 0} \frac{-1}{x(x+h)} = -\frac{1}{x^2}.$$

TIP In examinations, first-principles differentiation questions usually involve polynomials. The $\frac{1}{x}$ example above shows how far the definition extends; it is not itself an examination task.

13.2.3 Continuity and differentiability HL

The derivative is defined as a limit, and a limit can fail to exist. So a function need not have a derivative at every point. Two ideas describe where it does.

A function f is **continuous** at $x = a$ when three things are true:

1. $f(a)$ is defined;
2. $\lim\limits_{x \to a} f(x)$ exists, so the left-hand and right-hand limits are equal (§13.1);
3. this limit equals $f(a)$.

On a graph, a continuous function has no hole and no jump at that point: you can draw

through it without lifting the pen.

A function f is **differentiable** at $x = a$ when the first-principles limit $f'(a)$ exists. On a graph, the curve is continuous there and also *smooth*: it has one non-vertical tangent at that point, with no corner.

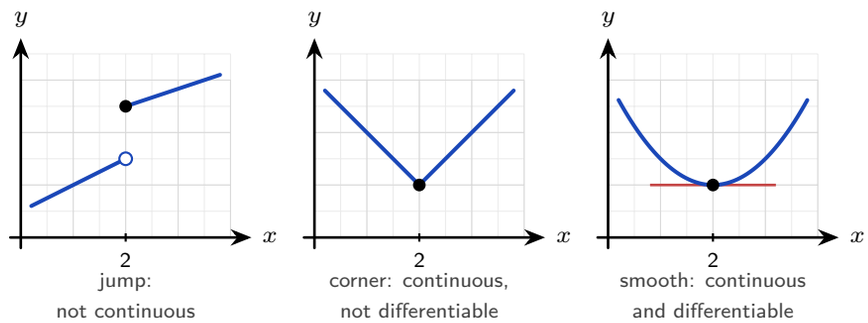


Figure 13.4 Three behaviours at $x = 2$. An open circle marks a point that is not on the graph. Only the smooth curve on the right has a single tangent (scarlet) at the point.

The two ideas are linked in one direction only.

- If f is differentiable at a point, then it is continuous there. A jump or a hole always prevents a derivative.
- If f is continuous at a point, it may still fail to be differentiable there. A corner is the usual reason.

The absolute value function shows the second case.

EXAMPLE 13.3

Show that $f(x) = |x - 1|$ is continuous at $x = 1$ but not differentiable there.

For continuity: $f(1) = 0$, and as $x \to 1$ from either side, $|x - 1| \to 0$. The limit exists and equals $f(1)$, so f is continuous at $x = 1$.

For differentiability, form the difference quotient at 1:

$$\frac{f(1+h) - f(1)}{h} = \frac{|h|}{h} = \begin{cases} 1 & h > 0, \\ -1 & h < 0. \end{cases}$$

The quotient approaches 1 from the right and -1 from the left. The one-sided limits are not equal, so the limit does not exist and f has no derivative at $x = 1$. On the graph, the V has a corner there: the gradient is -1 just to the left and 1 just to the right, so there is no single tangent.

So $|x - 1|$ is continuous at every real number, but it is not differentiable at $x = 1$.

TIP Examinations focus on the idea rather than on formal tests: you may differentiate only where the curve is smooth.

YOUR TURN 13.2 The Derivative

Notation

11.

(a) Write down three notations for the derivative of $y = f(x)$.

(b) A volume V depends on a radius r. Write down the notation for the rate of change of V with respect to r, and state in words what it measures.

First principles

12. Differentiate each from first principles, showing the limit.

(a) $f(x) = 3x^2$ (b) $f(x) = x^2 - 3x$

(c) $f(x) = 5 - 2x^2$

13. Differentiate $f(x) = x^3$ from first principles. *(Hint: $(x+h)^3 = x^3 + 3x^2h + 3xh^2 + h^3$)*

14. Differentiate $f(x) = \dfrac{2}{x}$ from first principles.

Estimating a gradient numerically

15. GDC The definition of the derivative can also be used numerically. For $f(x) = \sin x$, evaluate $\dfrac{f(0+h) - f(0)}{h}$ for $h = 0.1$, $h = 0.01$ and $h = 0.001$, working in radians. State the value the quotient is approaching, and what this suggests about the gradient of $y = \sin x$ at $x = 0$.

16. GDC Repeat the same calculation for $f(x) = e^x$ at $x = 0$, then at $x = 1$. Explain what the two answers suggest about the derivative of e^x.

Continuity and differentiability

17. For each function, state which of the three conditions for continuity at $x = 2$ fails, if any.

(a) $f(x) = \dfrac{x^2 - 4}{x - 2}$

(b) $f(x) = \begin{cases} x + 1, & x < 2 \\ x^2, & x \geq 2 \end{cases}$

(c) $f(x) = \begin{cases} x^2, & x \neq 2 \\ 1, & x = 2 \end{cases}$

18. Let $g(x) = \begin{cases} x + 1, & x < 1 \\ 3 - x, & x \geq 1. \end{cases}$

(a) Show that g is continuous at $x = 1$.

(b) Use the difference quotient at $x = 1$ to show that g is not differentiable at $x = 1$.

(c) A function h has a jump at $x = 4$. Explain why h cannot be differentiable at $x = 4$.

Concept check

19. A student differentiates $f(x) = x^2$ from first principles and writes $\dfrac{(x+h)^2 - x^2}{h} = \dfrac{x^2 + h^2 - x^2}{h} = h$, which tends to 0 as $h \to 0$, so $f'(x) = 0$. Identify the error and correct the working.

13.3 Differentiating Powers

First principles works, but it is slow. The two first-principles calculations above show a pattern: the x^2 term gave $2x$, and $\frac{1}{x} = x^{-1}$ gave $-x^{-2}$.

In both cases two things happened. The power decreased by one, and the original power became a multiplier in front. This is the **power rule**.

KEY · IB

$$\frac{d}{dx}\, x^n = nx^{n-1}$$

The Standard Level syllabus states this for integer powers in SL 5.3 and extends it to rational powers in SL 5.6. The rational form is what allows $\sqrt{x}$ to be differentiated below, and it is needed throughout §13.5 onwards.

The rule applies only to a term written as x^n, so rewrite every root and every reciprocal as a power first: $\sqrt{x} = x^{1/2}$ and $\frac{1}{x^2} = x^{-2}$.

Two more rules let you differentiate any polynomial. A constant multiplier is unchanged, and a sum is differentiated one term at a time:

$$\frac{d}{dx}\big(c\,f(x)\big) = c\,f'(x), \qquad \frac{d}{dx}\big(f(x)+g(x)\big) = f'(x)+g'(x).$$

A constant on its own has derivative zero, because its graph is a horizontal line with gradient zero.

Worked through. To differentiate $y = 3x^4 - 2x^2 + 7$, apply the power rule to each term:

$$\frac{dy}{dx} = 3(4x^3) - 2(2x) + 0 = 12x^3 - 4x.$$

The constant 7 has derivative zero, as every constant does.

EXAMPLE 13.4

Differentiate $y = \sqrt{x} + \dfrac{1}{x^2}$.

Rewrite both terms as powers before differentiating:

$$y = x^{1/2} + x^{-2} \implies \frac{dy}{dx} = \tfrac{1}{2}x^{-1/2} - 2x^{-3} = \frac{1}{2\sqrt{x}} - \frac{2}{x^3}.$$

The power rule requires the function in the form x^n. Convert every root and every reciprocal to a power before differentiating, and the rest of the question becomes routine.

EXAMPLE 13.5

Find the gradient of $y = x^3 - 5x$ at the point where $x = 2$.

Differentiate, then substitute:

$$\frac{dy}{dx} = 3x^2 - 5, \qquad \left.\frac{dy}{dx}\right|_{x=2} = 3(4) - 5 = 7.$$

The derivative is a function. A gradient *at a point* is that function evaluated at the point.

YOUR TURN 13.3 Differentiating Powers

Polynomials

20. Differentiate each of the following.

(a) $y = x^7$

(b) $y = 5x^3 - 2x$

(c) $y = x^4 - 3x^2 + 6$

(d) $y = 4x^5 + x$

(e) $y = 2x^3 - 7x^2 + x - 9$

(f) $y = 6$

Rewriting as a power

21. Rewrite each function as a power of x, then differentiate. In part (f), divide through first.

(a) $y = 2/x$
(b) $y = \sqrt{x}$
(c) $y = x^3 + 1/x$
(d) $y = 1/x^3$
(e) $y = \sqrt[3]{x}$
(f) $y = (x^2 + 1)/x$

22. Expand or divide to write each function as a sum of powers of x, then differentiate.

(a) $y = (2x + 1)^2$
(b) $y = \dfrac{x^3 - 4}{x^2}$
(c) $y = \dfrac{x + 1}{\sqrt{x}}$

Gradient at a point

23. Find the gradient of each curve at the point given.

(a) $y = x^2 - 5x$ at $x = 3$
(b) $y = x^3 - 4x$ at $x = -1$
(c) $y = \sqrt{x}$ at $x = 4$
(d) $y = 2x^3 + x$ at $x = 1$

24. Find the value or values of x at which each curve has the stated gradient.

(a) $y = x^2 - 6x$, gradient 0
(b) $y = x^3$, gradient 12
(c) $y = x^3 - 3x^2 + 1$, gradient 9

Concept check

25. A student differentiates $y = \dfrac{3}{x^2}$ and writes $y = 3x^{-2}$, so $\dfrac{dy}{dx} = -6x^{-1}$. Identify the error and give the correct derivative.

13.4 Tangents and Normals

The derivative gives the gradient of the tangent at any point, and a gradient plus a point determines a line. So you can now write the equation of the **tangent** to a curve at a given point. The **normal** is the line perpendicular to the tangent there, with gradient the negative reciprocal.

The tangent and the normal are both straight lines, so both are written with the same equation, the point-gradient form from Ch. 4:

$$y - y_1 = m(x - x_1),$$

where (x_1, y_1) is a known point on the line and m is its gradient. The point (x_1, y_1) is the same for both lines, because the tangent and the normal meet the curve at the same place. Only the gradient m is different.

KEY At the point (x_1, y_1) on the curve $y = f(x)$, both lines have the form $y - y_1 = m(x - x_1)$, with

$$m_{\text{tangent}} = f'(x_1), \qquad m_{\text{normal}} = -\frac{1}{f'(x_1)}$$

(for $f'(x_1) \neq 0$; if $f'(x_1) = 0$ the normal is the vertical line $x = x_1$).

So a question of this type follows three steps: find y_1 by substituting x_1 into f, find $f'(x_1)$ for the gradient, then substitute both into the point-gradient form.

On a calculator paper the same job can be done from the screen. A graphic display calculator will evaluate a numerical derivative at a point, and most will draw the tangent there and report its equation. Two cautions apply. Many calculators give only the *tangent*, so a normal still needs the $-1/f'(x_1)$ step by hand. And a numerical answer earns few marks on its own: show the gradient you used and the point you used it at, then quote the line. Use the screen to check the algebra, not to replace it.

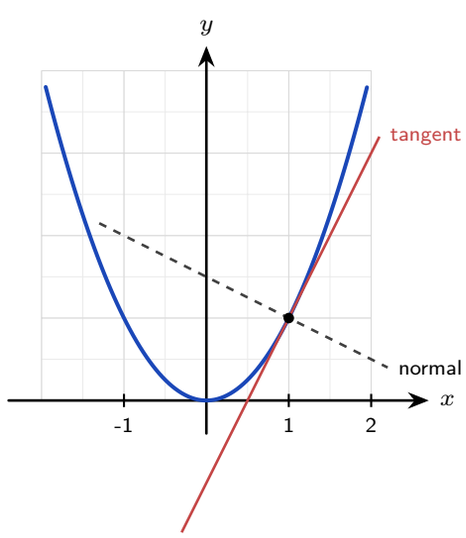


Figure 13.5 At a point on a curve the tangent has gradient $f'(x)$, and the normal is perpendicular to it, with gradient $-1/f'(x)$.

Worked through. To find the tangent and normal to $y = x^2 - 3x + 2$ at the point where $x = 2$, first find the point: $y(2) = 4 - 6 + 2 = 0$, so the point is $(2, 0)$. Then find the gradient: $\dfrac{dy}{dx} = 2x - 3$, so at $x = 2$ the tangent gradient is $2(2) - 3 = 1$. The tangent passes through $(2, 0)$ with gradient 1:

$$y - 0 = 1(x - 2) \implies y = x - 2.$$

The normal is perpendicular, so its gradient is -1:

$$y - 0 = -1(x - 2) \implies y = -x + 2.$$

Always find the y-coordinate first. A common error is to write the equation of the line using the x-value alone.

13.4.1 Tangents through a point off the curve

A harder question names a point the tangent must pass *through*, and that point is not on the curve. The point where the tangent touches the curve, the **point of contact**, is not given, so make it the unknown.

1. Let the point of contact be $(a, f(a))$.
2. Write the equation of the tangent there: $y - f(a) = f'(a)(x - a)$.
3. Substitute the coordinates of the given point, because the tangent passes through it. This gives an equation in a alone.
4. Solve for a, and substitute each value back into the tangent equation.

For a quadratic curve the equation in a is quadratic, so it usually has two roots, and each root gives a separate tangent.

EXAMPLE 13.6

Find the equations of both tangents to $y = x^2 - 2x + 3$ that pass through the point $(2, -1)$.
First check where the point sits: at $x = 2$ the curve is at $4 - 4 + 3 = 3$, not -1, so $(2, -1)$ is off the curve and is not the point of contact.
Let the contact point be $(a,\ a^2 - 2a + 3)$. Since $\dfrac{dy}{dx} = 2x - 2$, the tangent there has gradient $2a - 2$, so

$$y - (a^2 - 2a + 3) = (2a - 2)(x - a) \implies y = (2a - 2)x - a^2 + 3.$$

Now substitute the point $(2, -1)$, because the tangent passes through it:

$$-1 = 2(2a - 2) - a^2 + 3 \implies a^2 - 4a = 0 \implies a(a - 4) = 0,$$

so $a = 0$ or $a = 4$. Substituting each value back into $y = (2a - 2)x - a^2 + 3$:

$$a = 0: \quad y = -2x + 3, \qquad\qquad a = 4: \quad y = 6x - 13.$$

Two roots give two tangents, and both pass the check at the given point: $-2(2) + 3 = -1$ and $6(2) - 13 = -1$.

YOUR TURN 13.4 Tangents and Normals

Tangents

26. Find the equation of the tangent to each curve at the point given.

(a) $y = x^2$ at $(2, 4)$
(b) $y = x^2 - 4x + 3$ at $x = 1$
(c) $y = 1/x$ at $x = 1$
(d) $y = \sqrt{x}$ at $x = 9$

27. Find the coordinates of the point on $y = x^2 - 6x$ at which the tangent is horizontal.

Normals

28. Find the equation of the normal to each curve at the point given.

(a) $y = x^3$ at $(1, 1)$
(b) $y = \sqrt{x}$ at $x = 4$
(c) $y = \dfrac{2}{x}$ at $x = 1$

29.

(a) A tangent and a normal are drawn at the same point of a curve. Explain the relationship between their gradients.
(b) Find the equation of the normal to $y = x^2 - 4x + 7$ at the point where $x = 2$.

Tangents through a point off the curve

30. The tangent to $y = x^2$ at the point where $x = a$ passes through $(0, -4)$.

(a) Find the two possible values of a.
(b) Write down the two tangents, and check that each one passes through $(0, -4)$.

31. Find the equations of the two tangents to $y = x^2 + 3$ that pass through the point $(1, 0)$.

Concept check

32. A student finds the tangent to $y = x^2$ at the point $(3, 9)$ and writes $y - 9 = 2x(x - 3)$. Identify the error and find the correct equation of the tangent.

13.5 The Standard Derivatives

The power rule covers algebraic functions. The trigonometric, exponential and logarithmic functions have derivatives of their own. Each can be proved from first principles. You can also find each one by reading gradients from a graph. Do that once before memorising anything.

The idea is to move along the graph of f from left to right, and at each point plot its gradient as a height. The curve you plot is f'. In practice, sketch it in this order.

1. **Start with the peaks and troughs of f.** There the tangent is horizontal, so $f'(x) = 0$. Mark these x-values on the axis: the graph of f' meets the x-axis there.
2. **Where f is increasing**, its gradient is positive, so draw f' *above* the x-axis.
3. **Where f is decreasing**, its gradient is negative, so draw f' *below* the x-axis.
4. **Where f is steepest**, f' is furthest from the axis.

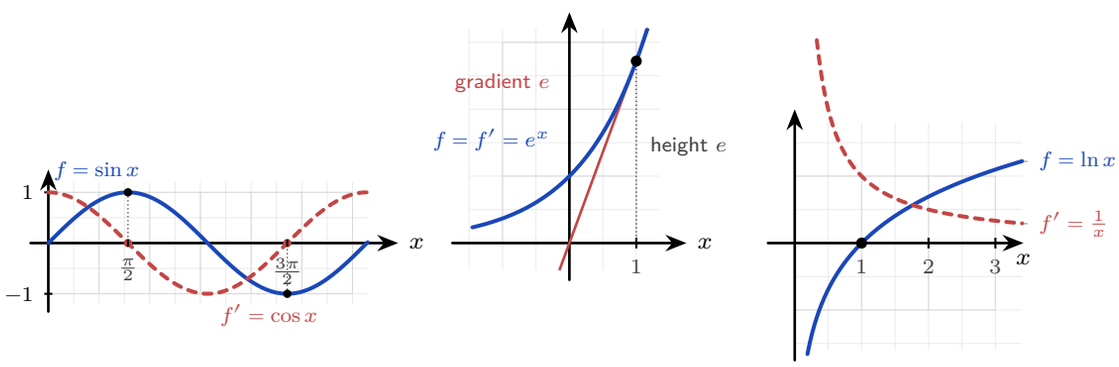


Figure 13.6 Each derivative read from the gradient of its own graph. **Left:** where $\sin x$ is flat, at $\frac{\pi}{2}$ and $\frac{3\pi}{2}$, the curve $\cos x$ crosses zero. **Centre:** the tangent to e^x at $x = 1$ has gradient e, which is also the height of the curve there. **Right:** $\ln x$ is steep near 0 and flattens as x grows, and $\frac{1}{x}$ does exactly that.

Read each panel separately.

Sine. At $x = 0$ the sine graph is increasing most steeply, so f' starts at its maximum. At the top of the curve, $x = \frac{\pi}{2}$, the tangent is horizontal, so f' is zero. After the top the graph decreases, so f' is negative. A curve that starts at 1, decreases through zero at $\frac{\pi}{2}$ and reaches -1 at π is exactly $\cos x$.

The exponential. Measure the gradient of $y = e^x$ at any point and you get the height of the curve at that point. That is the defining property of e, so here $f' = f$.

The logarithm. Just to the right of 0 the graph of $\ln x$ is almost vertical, so f' is very large. As x increases the curve becomes flatter, so f' decreases towards zero. It stays positive everywhere. That describes $\dfrac{1}{x}$.

Those are the four standard derivatives to learn.

KEY · IB

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

Note the exponential in particular: e^x is its own derivative. More precisely, the functions equal to their own derivative are exactly those of the form $y = Ce^x$, and e^x is the one

passing through $(0, 1)$. That property is one reason e matters so much in calculus. Note also that the results for $\sin$ and $\cos$ hold only when x is measured in *radians*, and there is a reason for that. The derivative of $\sin x$ is built on the limit $\lim_{x\to 0} \frac{\sin x}{x} = 1$ from §13.1, and that limit equals 1 only in radians. Measured in degrees the same limit is $\frac{\pi}{180}$, so the derivative would come out as $\frac{\pi}{180}\cos x$ instead. Radians are simply the unit that makes the constant equal to 1, which is why calculus uses them throughout.

At Higher Level a wider set of standard derivatives is provided. HL They are listed here for reference and used freely in the rules that follow. The reciprocal and inverse trigonometric functions in the table are defined in Ch. 8; only their derivatives are needed here.

KEY · IB HL

$$\frac{d}{dx}\tan x = \sec^2 x \qquad \frac{d}{dx}\sec x = \sec x \tan x$$

$$\frac{d}{dx}\operatorname{cosec} x = -\operatorname{cosec} x \cot x \qquad \frac{d}{dx}\cot x = -\operatorname{cosec}^2 x$$

$$\frac{d}{dx}a^x = a^x \ln a \qquad \frac{d}{dx}\log_a x = \frac{1}{x \ln a}$$

$$\frac{d}{dx}\arcsin x = \frac{1}{\sqrt{1-x^2}} \qquad \frac{d}{dx}\arccos x = -\frac{1}{\sqrt{1-x^2}}$$

$$\frac{d}{dx}\arctan x = \frac{1}{1+x^2}$$

Notice one pair. The derivative of $\arccos$ is the negative of the derivative of $\arcsin$, because the two graphs are reflections of each other.

Worked through. To differentiate $y = 3\sin x + 2e^x - 5\ln x$, apply the standard derivatives term by term:

$$\frac{dy}{dx} = 3\cos x + 2e^x - \frac{5}{x}.$$

Constant multipliers are unchanged, exactly as with powers.

The Higher Level table is used the same way. Read each derivative off it, keep the constant multipliers in place, and differentiate one term at a time.

EXAMPLE 13.7

Differentiate (a) $y = 3\tan x - 2\cos x$, (b) $y = 5^x + \ln x$, (c) $y = \arccos x + 3\arctan x$.

(a)

$$\frac{dy}{dx} = 3\sec^2 x + 2\sin x.$$

The derivative of $\cos x$ is $-\sin x$, so subtracting a cosine adds a sine.

(b)

$$\frac{dy}{dx} = 5^x \ln 5 + \frac{1}{x}.$$

The factor $\ln 5$ is a constant, and it appears because $5^x = e^{x\ln 5}$. For e^x the factor is $\ln e = 1$, which is why e^x has no extra factor.

(c)
$$\frac{dy}{dx} = -\frac{1}{\sqrt{1-x^2}} + \frac{3}{1+x^2}.$$

The leading minus sign is the only thing separating the $\arccos$ row of the table from the $\arcsin$ row. Neither inverse trigonometric derivative contains a trigonometric function; both are algebraic, which is what makes them the results you meet again as integrals in Ch. 15.

YOUR TURN 13.5 The Standard Derivatives

The four standard derivatives

33. Differentiate each of the following.

(a) $y = \sin x + \cos x$
(b) $y = e^x - 3\ln x$
(c) $y = 4\cos x$
(d) $y = 2\sin x - 5e^x$

34. The results $\dfrac{d}{dx}\sin x = \cos x$ and $\dfrac{d}{dx}\cos x = -\sin x$ hold only when x is measured in radians. Explain why the choice of unit matters.

The Higher Level table

35. Differentiate each of the following.

(a) $y = \tan x$
(b) $y = \arctan x$
(c) $y = 3^x$
(d) $y = \ln x + \tan x$
(e) $y = \sec x$
(f) $y = \arcsin x$

36. Differentiate each of the following.

(a) $y = \operatorname{cosec} x$
(b) $y = 2\cot x$
(c) $y = \log_2 x$
(d) $y = 3\arccos x - 4^x$

Gradients and tangents

37. Find the gradient of $y = \sin x$ at $x = \dfrac{\pi}{3}$, giving an exact answer.

38. Find the point on the curve $y = e^x$ at which the gradient is 1, and write down the equation of the tangent there.

39. Find the equation of the tangent to $y = \ln x$ at the point where $x = e$.

40. Find the values of x, for $0 \le x \le 2\pi$, at which the curve $y = x + 2\cos x$ has gradient zero.

Concept check

41. A student writes $\frac{d}{dx}\,2^x = x \cdot 2^{x-1}$. Explain why the power rule does not apply here, and give the correct derivative.

13.6 The Rules of Differentiation

Most functions are built from simpler ones, by composing, multiplying or dividing them. Three rules cover these three cases. With the standard derivatives, they let you differentiate almost any function in the course.

13.6.1 The chain rule

A **composite** function is a function inside a function, such as $(2x^2+3)^5$ or $\sin(2x)$. The **chain rule** differentiates it by working from the outside in: differentiate the outer function, then multiply by the derivative of the inner.

KEY · IB If $y = g(u)$ where $u = f(x)$, then

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$$

The meaning of the rule is that **rates multiply**. Suppose a car's fuel use y depends on its speed u, and its speed depends on time x. If fuel use rises 3 units for every unit of speed, and speed rises 2 units for every unit of time, then fuel use rises $3 \times 2 = 6$ units for every unit of time. That is exactly $\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$. A chain of dependencies multiplies its rates, in the same way that gears in a machine multiply their turning ratios.

The power rule is itself a case of the chain rule. In $y = x^7$ the inner function is just x, so the chain rule gives

$$\frac{dy}{dx} = 7x^6 \times \frac{d}{dx}(x) = 7x^6 \times 1 = 7x^6.$$

The factor 1 is the derivative of the inner function x, and it is not usually written. Replace the inner function x by $2x+1$ and the same rule gives

$$\frac{d}{dx}(2x+1)^7 = 7(2x+1)^6 \times 2 = 14(2x+1)^6.$$

Now the derivative of the inner function is 2, not 1, so it must be written.

The chain rule is needed whenever the function is a composite. Two habits make it reliable:

1. **Identify the inner function first**, before differentiating anything. Ask which function is

applied to x first: that one is the inner function.

2. **Multiply by the derivative of the inner function.** Differentiating the outer function alone and stopping there is a common error, so check for it at the end of every chain rule question.

The outer function need not be a power. Any of the standard functions of §13.5, such as e^u, $\ln u$ or $\sin u$, can be the outer function, and the rule does not change: differentiate the outer function, keep the inner function unchanged inside it, then multiply by the derivative of the inner function.

EXAMPLE 13.8

Differentiate (a) $y = \ln(2 + \cos x)$ and (b) $y = \sqrt{x^3 + 2}$.

(a) Outer is $\ln u$, inner is $u = 2 + \cos x$. The outer derivative is $\dfrac{1}{u}$, which must be written at $u = 2 + \cos x$:

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

(b) Rewrite the root as a power first, $y = (x^3 + 2)^{1/2}$, so the outer is $u^{1/2}$ and the inner is $u = x^3 + 2$:

$$\frac{dy}{dx} = \tfrac{1}{2}(x^3 + 2)^{-1/2} \cdot 3x^2 = \frac{3x^2}{2\sqrt{x^3 + 2}}.$$

The root itself needs $x^3 + 2 \geq 0$, and the derivative puts that root in a denominator, so this expression holds for $x^3 + 2 > 0$, that is $x > -\sqrt[3]{2}$.

Part (a) shows a common error to avoid: the outer derivative of a logarithm is $\dfrac{1}{\text{inner}}$, so it is $\dfrac{1}{2 + \cos x}$ here, never $\dfrac{1}{x}$.

13.6.2 The product rule

To differentiate a product of two functions, you cannot simply multiply their derivatives. The mistake is a natural one to make, and the reason it fails explains the correct rule.

Take $u = x$ and $v = x$, so that $y = uv = x^2$. The derivative of x^2 is $2x$. But $u' = 1$ and $v' = 1$, and $u'v' = 1$. The two answers do not agree, so multiplying the derivatives gives a wrong answer.

The reason is that a product changes for two separate reasons at the same time.

Suppose both u and v increase. Part of the increase in uv comes from the change in u, multiplied by the whole of v. The rest comes from the change in v, multiplied by the whole of u. The correct rule adds these two parts, which is why each term keeps one factor un-differentiated.

A rectangle makes this concrete. Let u and v be the two side lengths, so that uv is the area. Now stretch both sides slightly. The extra area has three parts. There is a thin strip along the top, of area (change in v) $\times\, u$, and a thin strip along the side, of area (change in u) $\times\, v$. There is also a tiny corner rectangle where the two strips meet. That corner is a product of

two small quantities, so its area approaches zero faster than the strips and adds nothing in the limit. The two strips are the two terms of the product rule.

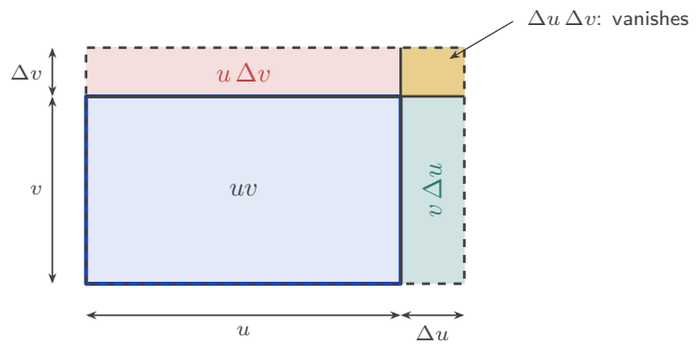


Figure 13.7 The product rule as an area. Growing both sides of a rectangle adds a strip along the top of area $u\,\Delta v$, a strip along the side of area $v\,\Delta u$, and a small corner of area $\Delta u\,\Delta v$. The two strips give the two terms of the rule; the corner is a product of two small quantities and vanishes in the limit.

KEY · IB If $y = uv$, then

$$\frac{dy}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}$$

Worked through. To differentiate $y = (x^2 + 1)e^x$, let $u = x^2 + 1$ and $v = e^x$, so $u' = 2x$ and $v' = e^x$:

$$\frac{dy}{dx} = (x^2 + 1)e^x + e^x(2x) = e^x(x^2 + 2x + 1) = (x + 1)^2 e^x.$$

Factorise the answer where you can. A factorised form makes later work much easier, for example finding where $\frac{dy}{dx} = 0$.

The rules combine, and many examination questions require it. When one factor of a product is itself a composite, the product rule gives the structure of the answer and the chain rule gives that factor's derivative. Find each of u' and v' completely before assembling anything.

EXAMPLE 13.9

Differentiate $y = x^3 e^{2x}$.

Let $u = x^3$ and $v = e^{2x}$, so $u' = 3x^2$. The second factor is a composite, with outer e^w and inner $w = 2x$, so the chain rule gives

$$v' = e^{2x} \cdot 2 = 2e^{2x}.$$

Now assemble the product rule:

$$\frac{dy}{dx} = x^3\,(2e^{2x}) + e^{2x}\,(3x^2) = e^{2x}\,(2x^3 + 3x^2) = x^2(2x + 3)e^{2x}.$$

The exponential factor is never zero, so the factorised form shows at once that the gradient is zero only at $x = 0$ and $x = -\frac{3}{2}$.

13.6.3 The quotient rule

For a quotient of two functions, the rule is similar, but the order of the terms now matters because a subtraction is involved.

The rule is not a new idea. A quotient $\dfrac{u}{v}$ is the product $u \times v^{-1}$, so the product and chain rules together produce the quotient rule.

The minus sign has a plain meaning. Increasing the top of a fraction increases its value, so the u term is positive. Increasing the bottom *decreases* the value, so the v term is negative. The v^2 in the denominator comes from differentiating v^{-1}: the chain rule gives $-v^{-2}\dfrac{dv}{dx}$.

KEY · IB If $y = \dfrac{u}{v}$, then

$$\frac{dy}{dx} = \frac{v\dfrac{du}{dx} - u\dfrac{dv}{dx}}{v^2}$$

EXAMPLE 13.10

Differentiate $y = \dfrac{x}{x^2+1}$.

Let $u = x$ and $v = x^2 + 1$, so $u' = 1$ and $v' = 2x$:

$$\frac{dy}{dx} = \frac{(x^2+1)(1) - x(2x)}{(x^2+1)^2} = \frac{1 - x^2}{(x^2+1)^2}.$$

The order in the numerator is: bottom times derivative of top, minus top times derivative of bottom. Reversing the two terms changes the sign of the answer, and it is a common error.

YOUR TURN 13.6 The Rules of Differentiation

The chain rule

42. Use the chain rule to differentiate each of the following.

(a) $y = (2x+1)^4$
(b) $y = \sin(x^2)$
(c) $y = e^{3x}$
(d) $y = \sqrt{3x+1}$
(e) $y = \ln(x^2+1)$
(f) $y = \ln(3x-2)$

43. Find the gradient of $y = (x^2-1)^3$ at the point where $x = 2$.

44. A student writes $\dfrac{d}{dx}\sin(3x) = \cos(3x)$. Identify the error and give the correct derivative.

45. Use the chain rule to differentiate each of the following.

(a) $y = (2x^2+3)^5$
(b) $y = e^{x^2}$

The product rule

46. Use the product rule to differentiate each of the following.

(a) $y = x\cos x$
(b) $y = x^2 e^x$
(c) $y = x^3 \ln x$
(d) $y = x\sin x$

47. Explain, with a counterexample, why the derivative of a product is not the product of the derivatives.

The quotient rule

48. Use the quotient rule to differentiate each of the following.

(a) $y = \dfrac{x}{x+1}$
(b) $y = \dfrac{x+1}{x-1}$
(c) $y = \dfrac{\sin x}{x}$
(d) $y = \dfrac{e^x}{x}$

Combining the rules

49. Each of these is a product in which one factor needs the chain rule. Differentiate, and factorise your answer where possible.

(a) $y = x^2 e^{3x}$
(b) $y = x\sin 2x$
(c) $y = x\sqrt{1+x^2}$

Concept check

50. A student differentiates $y = \dfrac{x^2}{x+1}$ with the quotient rule and writes $\dfrac{dy}{dx} = \dfrac{x^2\cdot 1 - (x+1)\cdot 2x}{(x+1)^2}$. Identify the error and give the correct derivative.

13.7 Higher Derivatives

A car that is speeding up has a velocity that is changing. To describe how fast that velocity changes, you need the rate of change of a rate of change. The derivative $f'(x)$ is itself a function, so it can be differentiated again.

Differentiating $f'(x)$ gives the **second derivative**. It is written $f''(x)$, or $\dfrac{d^2y}{dx^2}$ when $y = f(x)$.

DEF The **second derivative** of f is the derivative of f':

$$f''(x) = \frac{d}{dx}(f'(x)), \qquad \frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right).$$

The notation $\dfrac{d^2y}{dx^2}$ records the two steps. The operation $\dfrac{d}{dx}$ is applied twice, so the d on top carries a 2 and the dx below is written dx^2. Read it as "d two y by d x two". The 2 counts differentiations. It is not a power of y.

The second derivative measures how fast the gradient changes. For motion along a line this has a familiar name. If s is displacement at time t, then $\dfrac{ds}{dt}$ is velocity and $\dfrac{d^2s}{dt^2}$ is acceleration. See §14.4.

To find a second derivative, differentiate once, simplify, then differentiate the result. Every rule of §13.6 applies at each step, because each step is an ordinary differentiation.

Worked through. To find $f'(x)$, $f''(x)$ and $f''(2)$ for $f(x) = x^3 - 9x^2 + 5$, differentiate once, then differentiate the result:

$$f'(x) = 3x^2 - 18x, \qquad f''(x) = 6x - 18.$$

So $f''(2) = 12 - 18 = -6$. Each differentiation lowers the degree of a polynomial by one, so a cubic has a linear second derivative.

EXAMPLE 13.11

Find $\dfrac{d^2y}{dx^2}$ for $y = xe^{2x}$.

The product rule gives the first derivative:

$$\frac{dy}{dx} = e^{2x} + 2xe^{2x} = (1 + 2x)e^{2x}.$$

This is again a product, so apply the product rule a second time:

$$\frac{d^2y}{dx^2} = 2e^{2x} + (1 + 2x) \cdot 2e^{2x} = 4(1 + x)e^{2x}.$$

Taking out the common factor e^{2x} after each step keeps the expression short enough to differentiate again.

You can continue past the second derivative. Differentiating n times gives the **nth derivative**, written $\dfrac{d^n y}{dx^n}$ or $f^{(n)}(x)$. The brackets matter: $f^{(3)}(x)$ is the third derivative, while $f^3(x)$ usually means $\big(f(x)\big)^3$. For the first three derivatives, dashes are used instead: f', f'', f'''.

Some functions give derivatives that follow a pattern. Once the pattern is found, it gives every derivative at once.

$f(x)$	$f'(x)$	$f''(x)$	$f'''(x)$	$f^{(4)}(x)$	$f^{(n)}(x)$
e^{2x}	$2e^{2x}$	$4e^{2x}$	$8e^{2x}$	$16e^{2x}$	$2^n e^{2x}$
$\sin x$	$\cos x$	$-\sin x$	$-\cos x$	$\sin x$	repeats every 4
x^{-1}	$-x^{-2}$	$2x^{-3}$	$-6x^{-4}$	$24x^{-5}$	$(-1)^n\, n!\, x^{-(n+1)}$

Each pattern in the last column is a claim about every n, and proof by induction Ch. 3 is how such a claim is proved. Higher derivatives return in Ch. 16, where the derivatives of a function at one point are used to rebuild the function as a Maclaurin series. What the sign of f'' says about the shape of a graph is developed in Ch. 14.

YOUR TURN 13.7 Higher Derivatives

Second derivatives

51. Find the second derivative of each of the following.

(a) $f(x) = x^3 - 6x^2$

(b) $y = \sqrt{x}$

(c) $y = \cos 2x$

(d) $y = e^{-3x}$

52. Find $\dfrac{d^2y}{dx^2}$ for $y = \ln(2x)$.

53. A particle has displacement $s = t^3 - 6t^2$ metres at time t seconds. Find $\dfrac{d^2s}{dt^2}$ when $t = 3$.

54. In each case the first derivative is again a product, so the product rule is needed twice. Find $\dfrac{d^2y}{dx^2}$.

(a) $y = x^2 e^x$

(b) $y = x \sin x$

Third and higher derivatives

55. Find $f'''(x)$ for $f(x) = x^5 - 2x^3$.

56. For $f(x) = x^4$, find $f^{(3)}(x)$ and $\left(f(x)\right)^3$, and evaluate each at $x = 1$.

57. For $f(x) = e^{-x}$, find $f'(x)$, $f''(x)$ and $f'''(x)$, and hence write down $f^{(n)}(x)$.

58. Let $f(x) = \sin x$.

(a) Find $f'(x)$, $f''(x)$, $f'''(x)$ and $f^{(4)}(x)$.

(b) Hence write down $f^{(10)}(x)$ and $f^{(23)}(x)$.

Concept check

59. Let $f(x) = x^3 - 2x^2$. A student writes $f'(2) = 12 - 8 = 4$, then differentiates 4 to obtain $f''(2) = 0$. Identify the error and find the correct value of $f''(2)$.

13.8 L'Hôpital's Rule HL

This chapter began with the indeterminate form $\frac{0}{0}$, and resolved it by factorising. Factorising usually does not help when the functions are trigonometric, exponential or logarithmic.

There is a second method for those cases. If a limit gives $\frac{0}{0}$ or $\frac{\infty}{\infty}$, the derivatives of the top and the bottom give the answer. The rule is named after the French mathematician Guillaume de l'Hôpital, and the name is said *lo - pee - TAL*, with a silent h.

KEY **L'Hôpital's rule.** If $\lim\limits_{x \to a} \dfrac{f(x)}{g(x)}$ gives $\frac{0}{0}$ or $\frac{\infty}{\infty}$, then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)},$$

provided the limit on the right of this equation exists.

Differentiate the top and the bottom *separately*. This is not the quotient rule. Then try the limit again, and if it is still indeterminate, apply the rule a second time.

The reason the rule works is short, and it explains why it needs the $\frac{0}{0}$ form.

Suppose $f(a) = 0$ and $g(a) = 0$. Then near $x = a$ each function is close to its own tangent line at a, so

$$f(x) \approx f'(a)(x - a), \qquad g(x) \approx g'(a)(x - a).$$

Divide one by the other. The factor $(x-a)$ appears in both, so it cancels:

$$\frac{f(x)}{g(x)} \approx \frac{f'(a)(x-a)}{g'(a)(x-a)} = \frac{f'(a)}{g'(a)}.$$

So the quotient of the functions approaches the quotient of their gradients. This argument covers the $\frac{0}{0}$ case: only when both functions equal zero at a does each one reduce to a multiple of $(x-a)$, and only then does the cancelling work.

The limit $\lim_{x\to 0} \dfrac{\cos x}{x}$ shows what happens when they do not. Here $\cos 0 = 1$, not 0, so the form is $\frac{1}{0}$, which is not indeterminate and not one the rule covers. Differentiating top and bottom anyway would give

$$\lim_{x\to 0} \frac{-\sin x}{1} = 0,$$

and that is wrong: the limit does not exist, because the fraction increases without bound on one side of 0 and decreases without bound on the other. Check the form before you differentiate.

Worked through. In $\lim_{x\to 0} \frac{\sin x}{x}$, substituting gives $\frac{0}{0}$, so differentiate top and bottom:

$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\cos x}{1} = \cos 0 = 1.$$

This confirms the limit stated in §13.1. In the same way, substituting into $\lim_{x\to 0} \frac{\ln(1+x)-x}{x^2}$ gives $\frac{0}{0}$. One application gives $\dfrac{1/(1+x)-1}{2x}$, still $\frac{0}{0}$, so apply the rule again:

$$\lim_{x\to 0} \frac{\ln(1+x)-x}{x^2} = \lim_{x\to 0} \frac{\dfrac{1}{1+x}-1}{2x} = \lim_{x\to 0} \frac{-\dfrac{1}{(1+x)^2}}{2} = -\frac{1}{2}.$$

CAUTION L'Hôpital's rule applies *only* to the indeterminate forms $\frac{0}{0}$ and $\frac{\infty}{\infty}$. Always check that the limit is indeterminate before differentiating. Applying it to a form like $\frac{2}{0}$ or $\frac{3}{5}$ can give a wrong answer.

EXAMPLE 13.12

Evaluate $\lim_{x\to\infty} \dfrac{\ln x}{x}$.

As $x \to \infty$ this is $\frac{\infty}{\infty}$. Differentiate top and bottom:

$$\lim_{x\to\infty} \frac{\ln x}{x} = \lim_{x\to\infty} \frac{1/x}{1} = \lim_{x\to\infty} \frac{1}{x} = 0.$$

The denominator increases without bound while the numerator increases very slowly, so the ratio approaches zero. L'Hôpital's rule makes that comparison exact.

YOUR TURN 13.8 L'Hôpital's Rule HL

Applying the rule once

60. Evaluate each limit, checking first that it is indeterminate.

(a) $\displaystyle\lim_{x\to 0}\frac{e^x-1}{x}$

(b) $\displaystyle\lim_{x\to 0}\frac{\sin 2x}{x}$

(c) $\displaystyle\lim_{x\to 0}\frac{\tan x}{x}$

(d) $\displaystyle\lim_{x\to \infty}\frac{\ln x}{x^2}$

(e) $\displaystyle\lim_{x\to 1}\frac{\ln x}{x-1}$

61. Evaluate $\displaystyle\lim_{x\to\infty}\frac{3x^2+x}{2x^2-5}$ in two ways: by dividing by the highest power, and by L'Hôpital's rule. State which you would use in an examination, and why.

Applying the rule twice

62. Each of these needs the rule applied twice. Evaluate them.

(a) $\displaystyle\lim_{x\to 0}\frac{1-\cos 2x}{x^2}$

(b) $\displaystyle\lim_{x\to 0}\frac{e^x-1-x}{x^2}$

(c) $\displaystyle\lim_{x\to\infty}\frac{x^2}{e^x}$

Checking the form first

63. A student evaluates $\displaystyle\lim_{x\to 0}\frac{x+2}{x+1}$ by differentiating top and bottom, obtaining 1. Explain the error and give the correct value.

64. State the two indeterminate forms to which L'Hôpital's rule applies, and explain why the rule must not be used on a form such as $\frac{2}{0}$.

Concept check

65. Explain why L'Hôpital's rule requires the derivatives of the top and the bottom taken **separately**, and not the derivative of the quotient.

Chapter 13 · Limits & Derivatives

Reflections

TOK The early calculus was built on "infinitesimals," quantities treated as nonzero when dividing and as zero when convenient. The philosopher George Berkeley mocked them in 1734 as "ghosts of departed quantities," and his criticism was fair: the logic was not secure. The method gave correct answers for about 150 years before the limit gave it a secure logical foundation, and the limit brought a puzzle of its own: $\frac{0}{0}$ has no value, yet a limit of that form often has one exactly. Is a mathematical technique justified by its results, or only by its proof? And where does the meaning of a limit sit: in the symbols, or in the process that produced them?

IM Calculus was created twice, independently. Isaac Newton developed his "method of fluxions" in England in the 1660s; Gottfried Leibniz developed his version, with the $\frac{dy}{dx}$ notation we still use, in Paris in the 1670s. A long dispute over who was first divided European mathematics for about a century. Newton's notation was used in Britain and Leibniz's in the rest of Europe. British mathematics progressed more slowly in that period, partly because it kept the less convenient symbols. Notation is not neutral: a good symbol can make a whole subject easier to work in.

RW The derivative measures change, so it appears in any subject that studies how one quantity changes with another. Velocity and acceleration in physics, marginal cost in economics, reaction rates in chemistry, growth rates in biology: each is a derivative. When a report says prices are 'rising but more slowly', it is making a claim about a second derivative: prices are increasing, but their rate of increase is falling.

EXT L'Hôpital's rule is related to a deeper idea. Near a point, many functions are well approximated by a polynomial, the Maclaurin series: $\sin x \approx x - \frac{x^3}{6}$, $e^x \approx 1 + x + \frac{x^2}{2}$. Substituting these into an indeterminate limit resolves it without differentiating repeatedly. The same problem that L'Hôpital's rule solves leads to representing functions as infinite polynomials, an important idea in analysis Ch. 16.

End-of-Chapter Problems
Chapter 13: Limits & Derivatives

1. Differentiate $y = e^x \sin x$, and hence find the gradient of the curve at $x = 0$. [4]

2. Find the equation of the tangent to $y = x^2$ that is parallel to the line $y = 6x - 1$. [4]

3. Evaluate $\lim_{x \to 0} \dfrac{e^{2x} - 1}{x}$. [4]

4. Differentiate each function from first principles, showing every step of the limit.

(a) $f(x) = x^3$ [4]

(b) $f(x) = \sqrt{x}$, for $x > 0$. *Multiply the difference quotient by the conjugate.* [4]

(c) $f(x) = e^x$. You may use $\lim_{h \to 0} \dfrac{e^h - 1}{h} = 1$. [3]

(d) $f(x) = \sin x$. You may use the compound-angle formula for $\sin(x + h)$, together with $\lim_{h \to 0} \dfrac{\sin h}{h} = 1$ and $\lim_{h \to 0} \dfrac{\cos h - 1}{h} = 0$. [5]

5. Using the quotient rule on $\tan x = \dfrac{\sin x}{\cos x}$, prove that $\dfrac{d}{dx} \tan x = \sec^2 x$. [4]

6. Consider $f(x) = x^3 - 3x$.

(a) Find $f'(x)$. [1]

(b) Find the coordinates of the points where the tangent is horizontal. [4]

7. GDC The function $f(x) = \dfrac{e^x - 1}{x}$ is not defined at $x = 0$.

(a) Copy and complete the table of values of $f(x)$, giving 5 significant figures:
$x = -0.1,\ -0.01,\ 0.01,\ 0.1$. [2]

(b) Suggest the value of $\lim_{x \to 0} f(x)$. [1]

(c) Verify your conjecture using l'Hôpital's rule. [2]

8. Differentiate $f(x) = 3x^2 - 2x$ from first principles. [5]

9. Consider $f(x) = |x - 2|$.

(a) Sketch the graph of f. [1]

(b) Explain briefly why f is continuous at $x = 2$. [2]

(c) Explain why f is not differentiable at $x = 2$. [2]

10. Evaluate $\lim_{x\to 0} \dfrac{1-\cos 3x}{x^2}$. [5]

11. Find the coordinates of the points on the curve $y = x^3 - 3x$ where the tangent is parallel to the line $y = 9x + 1$. [5]

12. A curve has equation $y = \dfrac{e^{2x}}{1+x^2}$.

(a) Find $\dfrac{dy}{dx}$. [4]
(b) Find the gradient of the curve at $x = 0$. [1]

13. Evaluate $\lim_{x\to 0} \dfrac{x-\sin x}{x^3}$, justifying each application of L'Hôpital's rule. [5]

14. Consider $f(x) = x^2 - 2x$.

(a) Find $f'(x)$ from first principles. [4]
(b) Hence find the gradient of the curve at $x = 3$. [2]

15. The curve $y = x^2 - 2x$ is considered at the point where $x = 3$.

(a) Find the coordinates of the point. [1]
(b) Find the equation of the tangent at this point. [3]
(c) Find the equation of the normal at this point. [2]

16. Evaluate the following limits.

(a) $\lim_{x\to 0} \dfrac{1-\cos 2x}{x^2}$ [3]
(b) $\lim_{x\to 0} \dfrac{e^{2x}-1-2x}{x^2}$ [3]

17. A curve has equation $y = \dfrac{\ln x}{x}$ for $x > 0$.

(a) Find $\dfrac{dy}{dx}$. [3]
(b) Show that the curve has a horizontal tangent at $x = e$. [3]

18. Differentiate each of the following, simplifying your answers.

(a) $y = (2x^3 - 1)^4$ [2] (b) $y = x^2 \ln x$ [2]
(c) $y = \dfrac{\sin x}{x}$ [2] (d) $y = e^{3x}\cos x$ [3]
(e) $y = \sqrt{4x^2+1}$ [2] (f) $y = \dfrac{e^x}{x^2+1}$ [3]

19. Consider the curve $y = e^x$ at the point where $x = 1$.

(a) Show that the tangent there passes through the origin. [4]
(b) Find the equation of the normal at the same point. [2]

20. Differentiate each of the following.

(a) $y = \arctan(3x)$ [2] (b) $y = \sec(2x)$ [2]
(c) $y = \log_2 x$ [2] (d) $y = \arcsin(x^2)$ [2]
(e) $y = \operatorname{cosec} x$ [2] (f) $y = \cot(3x)$ [2]

21. Find the equations of *both* tangents to the curve $y = x^2$ that pass through the point $(0, -1)$. [6]

22. Differentiate each of the following, simplifying where possible.

(a) $y = (3x^2 + 1)^5$ [2] (b) $y = x^2 e^x$ [2]
(c) $y = \dfrac{x}{x+1}$ [3] (d) $y = \ln(3x^2 + 2)$ [2]
(e) $y = \dfrac{\cos x}{x^2}$ [3] (f) $y = (x+1)^3(2x-5)^4$ [3]

23. Differentiate each of the following.

(a) $y = \sec x$ [1] (b) $y = \ln(\cos x)$ [3]
(c) $y = x \arctan x$ [3] (d) $y = \arccos(2x)$ [2]
(e) $y = e^{\sin x}$ [2] (f) $y = \dfrac{\arctan x}{x}$ [3]

24. The displacement of a particle is $s(t) = t^3 - 6t^2 + 9t$ metres at time t seconds, for $t \geq 0$.

(a) Find the velocity $v(t) = s'(t)$. [2]
(b) Find the times at which the particle is momentarily at rest. [3]
(c) Find the acceleration at $t = 3$. [2]

25. Evaluate $\lim\limits_{x \to 0} \dfrac{\tan x - x}{x^3}$. You may use $\lim\limits_{x \to 0} \dfrac{\tan x}{x} = 1$ after establishing it. [7]

26. Let $g(x) = \dfrac{2x-1}{x+2}$, $x \neq -2$.

(a) Write down the equations of the vertical and horizontal asymptotes. [2]
(b) Use the quotient rule to show that $g'(x) = \dfrac{5}{(x+2)^2}$. [3]
(c) Hence explain why g is increasing on every interval of its domain. [2]
(d) Find the equation of the tangent to the curve at the point where $x = 0$. [2]
(e) Find the equation of the normal at the same point. [2]

Challenge Questions

27. The power rule, proved. Section 13.3 stated the power rule after two examples suggested it. Here you prove it for every positive integer n, using the binomial theorem, and then look at how far the argument reaches beyond the positive integers.

(a) Let $n \geq 2$ be an integer. Write down the first three terms of $(x+h)^n$ in ascending powers of h. [3]

(b) Hence show that, for every integer $n \geq 2$,

$$\frac{(x+h)^n - x^n}{h} = nx^{n-1} + {}^nC_2 x^{n-2}h + \cdots + h^{n-1},$$

and explain why every term after the first still contains h. [5]

(c) Deduce that $\dfrac{d}{dx}x^n = nx^{n-1}$ for every integer $n \geq 2$. Check separately that the result also holds for $n = 1$. [3]

(d) Check your working by expanding in full for $n = 4$. [3]

(e) For $n = -1$ and $x \neq 0$, the binomial series of Chapter 2 gives $(x+h)^{-1} = x^{-1}\left(1 + \dfrac{h}{x}\right)^{-1}$ as an infinite series, valid for $|h| < |x|$. Write down its first three terms, and hence the first two terms of $\dfrac{(x+h)^{-1} - x^{-1}}{h}$. Explain why the step used in part (c) is no longer enough on its own. Then confirm that $\dfrac{d}{dx}x^{-1} = -x^{-2}$ by differentiating $\dfrac{1}{x}$ from first principles without a series. [4]

(f) The rule also holds for fractional powers. For $x > 0$, prove it for $n = \frac{1}{2}$ by multiplying the numerator and denominator of $\dfrac{\sqrt{x+h} - \sqrt{x}}{h}$ by $\sqrt{x+h} + \sqrt{x}$. Then investigate how the factorisation

$$a^q - b^q = (a-b)\left(a^{q-1} + a^{q-2}b + \cdots + b^{q-1}\right)$$

extends the method to $n = \frac{1}{q}$ for any positive integer q, and how you might then reach every positive rational power. [3]

28. Where the derivative of sine comes from. The derivatives of $\sin x$ and $\cos x$ in §13.5 rest on two limits. Here you prove the first limit geometrically, deduce the second, and use both to prove that the derivative of $\sin x$ is $\cos x$. Angles are in radians throughout. You may use the *squeeze principle*: if $g(x) \leq f(x) \leq k(x)$ for all x close to a (with $x \neq a$), and $g(x) \to L$ and $k(x) \to L$ as $x \to a$, then $f(x) \to L$ as well.

(a) Let $0 < x < \frac{\pi}{2}$. On the unit circle with centre O, let $A = (1, 0)$ and $P = (\cos x, \sin x)$, and let the line OP meet the tangent to the circle at A at the point T. Show that $T = (1, \tan x)$. By comparing the areas of triangle OAP, sector OAP and triangle

OAT, show that

$$\sin x < x < \tan x.$$

[4]

(b) Hence show that $\cos x < \dfrac{\sin x}{x} < 1$ for $0 < x < \frac{\pi}{2}$, and explain why the same inequality holds for $-\frac{\pi}{2} < x < 0$. Deduce that $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$. [4]

(c) Show that $\dfrac{\cos x - 1}{x} = \dfrac{-\sin^2 x}{x(\cos x + 1)}$. Hence show that $\lim\limits_{x\to 0} \dfrac{\cos x - 1}{x} = 0$. *(Hint: multiply above and below by $\cos x + 1$)* [4]

(d) The compound-angle identity gives $\sin(x+h) = \sin x \cos h + \cos x \sin h$. Use it to show that

$$\frac{\sin(x+h) - \sin x}{h} = \sin x \left(\frac{\cos h - 1}{h}\right) + \cos x \left(\frac{\sin h}{h}\right).$$

[3]

(e) Hence prove from first principles that $\dfrac{d}{dx}\sin x = \cos x$. [3]

(f) Section 13.8 evaluates $\lim\limits_{x\to 0} \dfrac{\sin x}{x}$ in one line by L'Hôpital's rule. Explain why that calculation cannot serve as a *proof* of the limit. [2]

(g) Put $x = \frac{\pi}{n}$ in part (a), where $n \geq 3$ is an integer, to show that $n \sin \frac{\pi}{n} < \pi < n \tan \frac{\pi}{n}$, and interpret the two bounds as half the perimeters of regular n-sided polygons drawn inside and outside a circle of radius 1. Archimedes used these bounds with $n = 96$. Investigate how quickly the bounds close in on π as n is doubled, and how the bounds for $2n$ sides can be found from those for n sides without knowing π. [3]

29. When L'Hôpital's rule says nothing. The rule in §13.8 carries a condition: $\lim \dfrac{f}{g} = \lim \dfrac{f'}{g'}$ *provided the limit on the right exists*. This investigation finds limits where that condition fails, and then uses the rule, correctly, to compare how fast functions grow.

(a) Consider $\lim\limits_{x\to\infty} \dfrac{x + \sin x}{x}$. Show that it is of the form $\frac{\infty}{\infty}$, and that applying the rule once gives a limit that does not exist. Find the limit by another method, and explain why this does not contradict the rule. [5]

(b) Consider $\lim\limits_{x\to\infty} \dfrac{\sqrt{1+x^2}}{x}$. Show that applying the rule once gives $\dfrac{x}{\sqrt{1+x^2}}$, and that applying it again returns the original fraction. Find the limit by another method. [4]

(c) Use the rule to evaluate $\lim\limits_{x\to\infty} \dfrac{x^n}{e^x}$ for $n = 1$, 2 and 3. State a conjecture for every positive integer n, and prove it by induction on n. [5]

(d) Let $a > 0$. Show that $\lim\limits_{x\to\infty} \dfrac{\ln x}{x^a} = 0$. [2]

(e) Parts (c) and (d) place $\ln x$, then every positive power x^a, then e^x, in increasing order of growth as $x \to \infty$. Investigate where the functions x^{100}, 1.01^x, $x \ln x$, $e^{\sqrt{x}}$ and x^x fit into this order, and give a precise meaning to the statement "f grows faster than g". [3]

30. How smooth is a join? A curve built from two pieces can meet itself without a jump and still have a corner. This investigation measures how smooth a join is by counting how

many derivatives exist there. For each positive integer n define

$$f_n(x) = \begin{cases} x^n & x \geq 0, \\ 0 & x < 0. \end{cases}$$

(a) By finding the one-sided limits of $\dfrac{f_2(h) - f_2(0)}{h}$ as $h \to 0$, show that f_2 is differentiable at $x = 0$, and state $f_2'(0)$. [3]

(b) Write down $f_2'(x)$ for every real x. Show that f_2' is continuous at 0 but not differentiable there. [3]

(c) Investigate f_3 in the same way, and find how many times f_3 can be differentiated at $x = 0$. [3]

(d) Conjecture how many times f_n can be differentiated at $x = 0$, and prove your conjecture. [4]

(e) A straight railway track along the negative x-axis is joined at the origin to a track that follows $y = f_n(x)$ for $x \geq 0$. Where the track is nearly straight, y'' measures how sharply it bends, and a passenger feels that bending as a sideways push. Explain why the designer would choose $n = 3$ rather than $n = 2$. [2]

(f) Every f_n has only finitely many derivatives at 0. Investigate the function $g(x) = e^{-1/x}$ for $x > 0$, $g(x) = 0$ for $x \leq 0$. Use the substitution $t = \frac{1}{h}$ and Question 29(c) to show that $g'(0)$ exists, and then investigate whether g can be differentiated as many times as you like at $x = 0$. [3]

Chapter 14

Derivative Applications

14

SYLLABUS 5.2 5.7 5.8 5.9 5.14

Throw a ball straight up. It rises, it slows, and at the very top it stops for an instant before it falls back. At that instant its velocity is exactly zero. A moment earlier the ball was moving up; a moment later it is moving down. The top of the flight is the one place where it is doing neither.

That instant points to a method for a large class of problems. The top of the flight is where the height is greatest, and it is also where the velocity is zero. The same pattern holds widely: for a smooth quantity at an interior point, when it reaches its largest or smallest value its rate of change is zero at that moment, and the curve is flat.

In this chapter the derivative stops being something you only calculate and becomes something you use. By the end of this chapter you will be able to:

- find the turning points of a curve, and decide which is a maximum and which a minimum;
- read the shape of a graph from the signs of f' and f'';
- find the largest or smallest value of a quantity under a constraint;
- describe motion along a straight line;
- find the gradient of a curve that is not written as $y = f(x)$ (implicit differentiation), and link the rates of quantities that change together (related rates).

A derivative is not only a gradient: it is a rate of change, and many real questions ask how fast something changes or where a quantity is greatest or least.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

14.1 Stationary Points and Their Nature

When a quantity reaches its largest or smallest value, it has stopped increasing and has not yet started decreasing, or the reverse. The same condition, $f'(x) = 0$, marks that moment in many subjects. A projectile is at its greatest height when its vertical velocity is zero (§14.4). A profit is largest where its rate of change with respect to the number of items made is zero (§14.3). In physics a system is in equilibrium where its potential energy is stationary. In statistics the mode of a continuous distribution is at a maximum of its density function Ch. 12. To use the condition, first read what the sign of f' says about a curve.

14.1.1 Increasing and decreasing

The derivative gives the gradient, so its sign shows which way the curve is going. Where $f'(x) > 0$ the tangent rises from left to right, and a larger x gives a larger $f(x)$: the function is **increasing**. Where $f'(x) < 0$ the function is **decreasing**. Where $f'(x) = 0$ the tangent is horizontal.

KEY If $f'(x) > 0$ for every x in an interval, then f is **increasing** on that interval.
If $f'(x) < 0$ for every x in an interval, then f is **decreasing** on that interval.

To find where f is increasing, differentiate, factorise $f'(x)$, and solve the inequality $f'(x) > 0$. Solve $f'(x) < 0$ for where it is decreasing.

EXAMPLE 14.1

Find the intervals on which $f(x) = x^3 - 9x^2$ is increasing, and those on which it is decreasing.
Differentiate and factorise:

$$f'(x) = 3x^2 - 18x = 3x(x - 6).$$

The product is zero at $x = 0$ and $x = 6$. It is positive when both factors have the same sign, so $f'(x) > 0$ for $x < 0$ and for $x > 6$. Between the roots the factors have opposite signs, so $f'(x) < 0$ for $0 < x < 6$.
So f is increasing for $x < 0$ and for $x > 6$, and decreasing for $0 < x < 6$. The behaviour can change only where $f'(x) = 0$, and those points are the subject of the rest of this section.

14.1.2 Stationary points

A **stationary point** is a point where the gradient is zero: the tangent is horizontal. Since the derivative gives the gradient, stationary points are exactly the solutions of $f'(x) = 0$. They are the points where a curve stops increasing or stops decreasing.

For a smooth curve, every local maximum and minimum is a stationary point, so the condition $f'(x) = 0$ finds them all. It does not find an extreme value at a corner, where the

derivative does not exist: $f(x) = |x|$ has a minimum at $x = 0$ and no stationary point there. See Ch. 13.

DEF A **stationary point** of f occurs where $f'(x) = 0$. At a **local maximum** the curve changes from increasing to decreasing; at a **local minimum**, from decreasing to increasing.

The word local means highest or lowest among nearby points only. A local maximum need not be the largest value of the function: the curve below has a local maximum at $(-1, 2)$ and still increases without limit to the right.

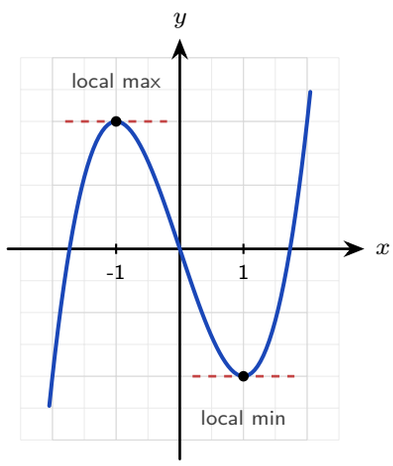


Figure 14.1 At each stationary point the tangent is horizontal. The curve changes from increasing to decreasing at the maximum, and from decreasing to increasing at the minimum.

A stationary point is a candidate, not an answer. A zero gradient says only that the curve is flat there. Flat can mean a local maximum, a local minimum, or neither: $y = x^3$ is flat at the origin and keeps increasing through it. Two tests decide which case you have.

14.1.3 The first derivative test

The sign of f' just to the left and just to the right of a stationary point shows what the curve does there. Just before a maximum the curve rises, so $f' > 0$; just after it, the curve falls, so $f' < 0$. A minimum is the reverse. If f' has the same sign on both sides, the curve keeps rising, or keeps falling, through the point. It is then a **stationary point of inflexion** (§14.2).

Sign of f'	Shape of the curve	Nature of the point
$+$ **then** $-$	rising, then falling	local maximum
$-$ **then** $+$	falling, then rising	local minimum
same sign	rising throughout, or falling throughout	stationary point of inflexion

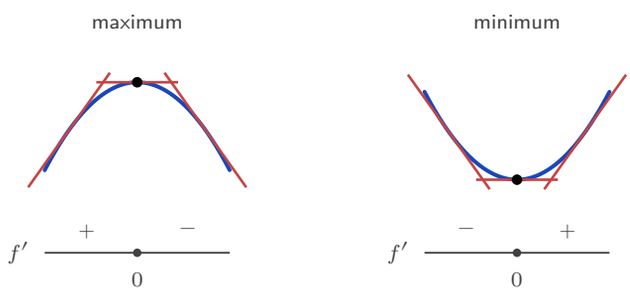


Figure 14.2 The first derivative test. The scarlet strokes are tangents: rising, flat, then falling at a maximum, and the reverse at a minimum. The strip below records the sign of f' on each side.

The method has three steps.

1. Solve $f'(x) = 0$ to find the stationary points.
2. Find the sign of f' just to the left and just to the right of each one. If $f'(x)$ is factorised, the sign of each factor gives the sign of f' without substituting any numbers.
3. Read the nature of each point from the table.

EXAMPLE 14.2

Find and classify the stationary points of $f(x) = 3x^4 - 4x^3$.

Differentiate and factorise:

$$f'(x) = 12x^3 - 12x^2 = 12x^2(x - 1).$$

So $f'(x) = 0$ at $x = 0$ and $x = 1$.

The factor $12x^2$ is positive on both sides of 0 and on both sides of 1. So the sign of f' is the sign of $x - 1$: negative for $x < 1$, positive for $x > 1$.

At $x = 0$ the sign of f' is negative, then negative again. The curve falls through the point, so $(0, 0)$ is a stationary point of inflexion.

At $x = 1$ the sign changes from negative to positive, so the point is a local minimum. Since $f(1) = 3 - 4 = -1$, it is $(1, -1)$.

A repeated factor such as x^2 does not change sign, so it cannot produce a maximum or a minimum on its own.

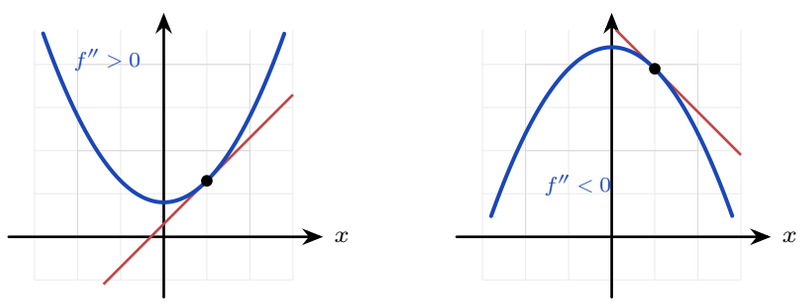
14.1.4 Concavity

The first derivative says whether a curve is going up or down. It does not say how the curve bends. Two curves can both rise from the same point, one more and more steeply, the other less and less steeply. The second derivative separates them, because f'' is the rate of change of the gradient §13.7.

Where $f''(x) > 0$ the gradient is increasing, and the curve is **concave up**: it bends upward, like a cup that holds water. Where $f''(x) < 0$ the gradient is decreasing, and the curve is **concave down**: it bends downward, like a cup turned upside down.

DEF The curve $y = f(x)$ is **concave up** on an interval where $f''(x) > 0$, and **concave down** on an interval where $f''(x) < 0$.

The tangent line gives a picture of the same fact. Where a curve is concave up it lies *above* its tangents, so a tangent touches the curve and the curve then rises away from it. Where the curve is concave down, it lies below its tangents.



concave up: curve *above* its tangent

concave down: curve *below* its tangent

Figure 14.3 The tangent picture of concavity. A tangent (scarlet) touches the curve at one point. Where the curve is concave up it stays above every tangent; where it is concave down it stays below.

Notice what f'' reports on: the gradient, not the height. Concave up means the gradient is increasing, and that can happen while the curve falls, because a gradient that goes from -5 to -1 has increased. So the signs of f' and f'' answer two separate questions.

Sign	f': going up or down?	f'': bending which way?
positive	**increasing**	**concave up** (bends upward)
negative	**decreasing**	**concave down** (bends downward)
zero	**stationary point**: the curve is flat there	**point of inflexion**, *if* the concavity also changes

The two signs are independent. A curve can be increasing and concave down at the same

time: it rises by less and less. It can also be decreasing and concave up: it falls by less and less.

14.1.5 The second derivative test

At a stationary point the tangent is horizontal. A curve that is concave up there lies above its horizontal tangent, so the point is a local minimum. A curve that is concave down lies below it, so the point is a local maximum. One substitution into f'' therefore decides the nature of the point. This is usually faster than the first derivative test.

KEY At a stationary point where $f'(x) = 0$:

$$f''(x) > 0 \Rightarrow \text{local minimum}, \qquad f''(x) < 0 \Rightarrow \text{local maximum}.$$

If $f''(x) = 0$ this test gives no answer. Use the first derivative test instead.

The case $f''(x) = 0$ is a real gap in the test, and it does not mean the point is an inflexion. Both $y = x^4$ and $y = x^3$ have $f'(0) = 0$ and $f''(0) = 0$. Yet x^4 has a minimum at the origin, and x^3 has a stationary point of inflexion there. Only the sign of f' on each side separates them: for x^4 it is negative then positive, while for x^3 it is positive on both sides.

Worked through. To find and classify the stationary points of $f(x) = x^3 + 3x^2 - 9x + 2$, differentiate and solve $f'(x) = 0$:

$$f'(x) = 3x^2 + 6x - 9 = 3(x + 3)(x - 1) = 0 \implies x = -3 \text{ or } x = 1.$$

The second derivative is $f''(x) = 6x + 6$. At $x = -3$, $f''(-3) = -12 < 0$, so the point is a local maximum; $f(-3) = 29$, so it is $(-3, 29)$. At $x = 1$, $f''(1) = 12 > 0$, so the point is a local minimum; $f(1) = -3$, so it is $(1, -3)$. This cubic increases without limit to the right, so 29 is not the largest value of f.

The method does not depend on the function being a polynomial. Any function you can differentiate twice can be tested in the same way. State the domain whenever the function restricts it, as $\ln x$ requires $x > 0$. If $f''(x) > 0$ across the whole domain, the curve is concave up everywhere, so a stationary point there is the lowest point of the entire curve and not only a local minimum.

EXAMPLE 14.3

Find and classify the stationary point of $f(x) = x \ln x$, for $x > 0$.

The product rule gives

$$f'(x) = 1 \cdot \ln x + x \cdot \frac{1}{x} = \ln x + 1.$$

Set $f'(x) = 0$: $\ln x = -1$, so $x = e^{-1} \approx 0.368$.

The second derivative is $f''(x) = \dfrac{1}{x}$, which is positive for every x in the domain. So the point is

a local minimum, and

$$f\left(\tfrac{1}{e}\right) = \tfrac{1}{e}\ln\tfrac{1}{e} = -\tfrac{1}{e} \approx -0.368.$$

The minimum is at $\left(\frac{1}{e}, -\frac{1}{e}\right)$. The domain $x > 0$ comes from the logarithm, and since $f'' > 0$ on all of it, this is the lowest point of the entire curve.

YOUR TURN 14.1 Stationary Points and Their Nature

Increasing and decreasing

1. Find the intervals on which each function is increasing, and those on which it is decreasing.

(a) $f(x) = x^2 - 4x$

(b) $f(x) = x^4 - 2x^2$

(c) $f(x) = \sin x$, for $0 \le x \le 2\pi$

(d) $f(x) = xe^{-x}$

Concavity

2. Find the intervals on which each curve is concave up, and those on which it is concave down.

(a) $y = x^3$

(b) $y = x^3 - 6x^2$

(c) $y = \cos x$, for $0 \le x \le 2\pi$

(d) $y = \ln x$, for $x > 0$

Finding and classifying stationary points

3. Find the stationary points of each curve.

(a) $y = x^2 - 6x + 5$

(b) $y = x^3 - 12x$

(c) $y = 2x^3 - 3x^2 - 12x + 1$

(d) $y = x^4 - 2x^2$

4.

(a) Use the second derivative to classify the stationary point of $y = x^2 - 6x + 5$.

(b) Use the second derivative to classify both stationary points of $y = x^3 - 12x$.

(c) Find the stationary point of $y = x + \dfrac{1}{x}$ for $x > 0$, and state its nature.

(d) A curve has $\dfrac{dy}{dx} = (x-1)(x-4)$. Write down the x-coordinates of its stationary points, and state the nature of each without differentiating again.

5. Find the stationary point of each curve and determine its nature.

(a) $y = e^x - 3x$

(b) $y = \dfrac{\ln x}{x}$, for $x > 0$

(c) $y = \sin x + \cos x$, for $0 \le x \le 2\pi$ (there are two)

(d) $y = xe^{-2x}$

The first derivative test

6. The curve $y = x^3 - 3x^2 + 3x$ has one stationary point.

(a) Show that $x = 1$ is a stationary point.

(b) Explain why the second derivative test cannot classify it.

(c) Use the sign of $\dfrac{dy}{dx}$ on either side to classify it.

7.

(a) A curve has $\dfrac{dy}{dx} = (x+1)^2(x-2)$. Find the x-coordinates of its stationary points, and use the first derivative test to classify each one.

(b) Find and classify the stationary points of $f(x) = x^4 - 4x^3$.

Local and global values

8. Let $f(x) = 2x^3 - 3x^2 - 12x$ for $-2 \le x \le 4$.

(a) Find the local maximum and the local minimum of f.

(b) Find the greatest and least values of f on the interval, and explain why the local maximum is not the greatest value.

Concept check

9. A student writes: "$f'(a) = 0$, *so* f *has a maximum or a minimum at* $x = a$." Give a counter-example, and state what must be checked before this conclusion can be drawn.

14.2 Points of Inflexion

The concavity of a curve can change, and the point where it changes often has a meaning. Suppose $f(t)$ is the total number of cases in an epidemic after t days. Then $f'(t)$ is the number of new cases per day. While $f'' > 0$ the daily count is growing. Once $f'' < 0$ the daily count is falling, even though the total still rises. The change from one to the other is the day on which the daily count is highest.

A **point of inflexion** is a point where the curve changes the direction of its bend: from concave up to concave down, or the reverse.

DEF A **point of inflexion** is a point on a curve where the concavity changes. For a function that can be differentiated twice, $f''(x) = 0$ at every point of inflexion. But $f''(x) = 0$ alone does not guarantee one: the concavity must actually change.

CAUTION $f''(x) = 0$ does not by itself give an inflexion. For $f(x) = x^4$, $f''(0) = 0$, yet the curve is

concave up on both sides: $x = 0$ is a minimum, not an inflexion. Always check that the concavity changes.

An inflexion can occur at any gradient, and the syllabus names two kinds. At a **stationary point of inflexion** the gradient is zero, as for $y = x^3$ at the origin. At a **non-stationary point of inflexion** the gradient is not zero, and the curve changes its bend while it is still rising or falling.

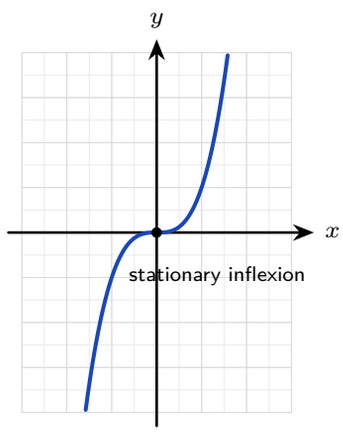


Figure 14.4 At the origin $y = x^3$ has both a horizontal tangent and a change of concavity, so this inflexion is also a stationary point.

To find the points of inflexion of a curve:

1. Find $f''(x)$ and solve $f''(x) = 0$.
2. Check that f'' changes sign at each solution, by testing a value on each side.
3. Substitute each confirmed value of x into f to find the point.
4. To name the kind, evaluate f' there: zero means stationary, non-zero means non-stationary.

Worked through. To find the point of inflexion of $f(x) = x^3 - 3x^2$, and whether it is stationary, differentiate twice:

$$f'(x) = 3x^2 - 6x, \qquad f''(x) = 6x - 6.$$

Set $f''(x) = 0$: $6x - 6 = 0 \implies x = 1$. For $x < 1$, $f'' < 0$ (concave down); for $x > 1$, $f'' > 0$ (concave up). The sign changes, so $x = 1$ gives a point of inflexion. Since $f(1) = -2$, the point is $(1, -2)$. The gradient there is $f'(1) = 3 - 6 = -3 \neq 0$, so this is a non-stationary inflexion: the curve changes its bend while it is falling.

14.2.1 Inflexions of other functions

Polynomials are not the only curves that bend both ways. Take $f(x) = \sin x$ and differentiate twice:

$$f'(x) = \cos x, \qquad f''(x) = -\sin x.$$

So $f''(x) = 0$ whenever $\sin x = 0$, that is at $x = 0, \pm\pi, \pm 2\pi, \ldots$. At each of these the concavity changes, because $-\sin x$ changes sign every time the sine crosses the axis. The sine curve therefore has a point of inflexion at every multiple of π. None is stationary, since $f'(x) = \cos x = \pm 1$ there. Each inflexion is also the steepest point on its part of the curve. This follows from the definition: an inflexion of f is where f'' changes sign, so it is a turning point of f', where the gradient is greatest or least.

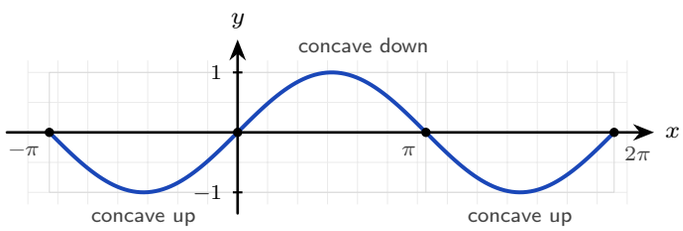


Figure 14.5 Every point where $y = \sin x$ crosses the axis is a point of inflexion: the bend reverses there, and the gradient is ± 1, the steepest the curve reaches. None of these inflexions is stationary.

Other familiar curves have no point of inflexion at all. For $y = e^x$, $f''(x) = e^x > 0$ for every x, so the curve is concave up everywhere. For $y = \ln x$, $f''(x) = -\dfrac{1}{x^2} < 0$ for every $x > 0$, so it is concave down across its whole domain. For $y = x^2$, $f''(x) = 2 > 0$ everywhere. In all three f'' never changes sign, so there is no inflexion to find.

EXAMPLE 14.4

Find the points of inflexion of $f(x) = e^{-x^2}$.

The chain rule gives $f'(x) = -2xe^{-x^2}$. Differentiating again needs the product rule:

$$f''(x) = -2e^{-x^2} + (-2x)(-2x)e^{-x^2} = (4x^2 - 2)\,e^{-x^2}.$$

Since $e^{-x^2} > 0$ for every x, the sign of f'' is decided by $4x^2 - 2$ alone. This is zero when $x^2 = \frac{1}{2}$, so $x = \pm\frac{1}{\sqrt{2}} \approx \pm 0.707$, and it changes sign at each: negative between the two values, positive outside them. Both are points of inflexion, and $f\left(\pm\frac{1}{\sqrt{2}}\right) = e^{-1/2} \approx 0.607$.

So the curve is concave down over the central hump and concave up in each tail. This is the bell shape of Ch. 12. The same calculation on $e^{-x^2/2}$, a constant multiple of the standard normal density, puts its inflexions at $x = \pm 1$: exactly one standard deviation from the mean.

14.2.2 Reading f, f' and f'' together

A common examination question shows one of these three graphs and asks about another. Below are $f(x) = x^3 - 3x$ and its two derivatives, drawn one above the other on the same x-axis. The vertical scales are not the same; only the zeros and the signs matter here.

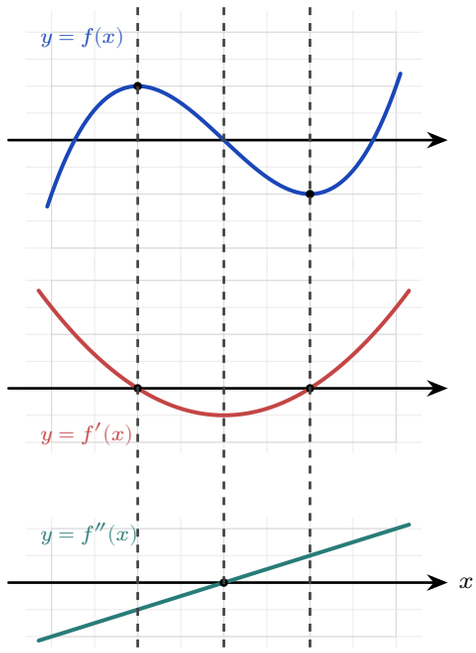


Figure 14.6 f, f' and f'' for $f(x) = x^3 - 3x$, stacked over a shared x-axis. Read down each dashed line: a turning point of f sits above a zero of f', and the inflexion of f sits above a turning point of f' and a zero of f''.

At every dashed line something happens in all three graphs at once. The three links are:

- A **turning point of** f sits above a **zero of** f'. At $x = -1$ the curve f has a maximum, and there f' crosses zero from positive to negative. At $x = 1$ the curve has a minimum, and there f' crosses zero from negative to positive.
- An **interval where** f **increases** sits above an **interval where** f' **is positive**, that is, where the graph of f' is above the axis.
- An **inflexion of** f sits above a **turning point of** f', and above a **zero of** f'' where f'' changes sign. In this example all three occur at $x = 0$.

Each link works in both directions, from f to f' and from f' back to f. So from any one of the three graphs you can sketch the other two.

YOUR TURN 14.2 Points of Inflexion

Finding points of inflexion

10. Find the point of inflexion of each curve, and state whether it is a stationary or a non-stationary point of inflexion.

(a) $y = x^3 - 6x^2 + 5$
(b) $y = x^3 + 6x^2 + 12x$
(c) $y = 2x^3 - 9x^2 + 12x$

11.

(a) Find the coordinates of the point of inflexion of $y = xe^x$.
(b) Find the x-coordinates of the points of inflexion of $y = 2x + \cos x$ for $0 \leq x \leq 2\pi$.
(c) Show that $y = \dfrac{x}{1 + x^2}$ has three points of inflexion, and find their coordinates.

When $f''(x) = 0$ is not enough

12.

(a) Show that $y = x^4 - 4x + 1$ has $\dfrac{d^2y}{dx^2} = 0$ at $x = 0$, and explain why there is no point of inflexion there.
(b) Determine, for $y = x^5$, whether the point where $f'' = 0$ is a point of inflexion, and if so whether it is stationary.
(c) Explain why $y = e^x + x^2$ has no point of inflexion.

Reading f, f' and f'' together

13. The graph of $y = f'(x)$ crosses the x-axis only at $x = -1$, from above to below, and at $x = 3$, from below to above. It has a single turning point, a minimum at $x = 1$. For the curve $y = f(x)$, state

(a) the x-coordinates of its stationary points, and the nature of each
(b) the interval on which f is decreasing
(c) the x-coordinate of its point of inflexion
(d) the interval on which it is concave up.

14. A function has $f'(x) = x^2(x - 3)$. Without finding $f(x)$,

(a) classify the stationary points of f
(b) find the x-coordinates of the points of inflexion of f, and state whether each is stationary or non-stationary.

Concept check

15. A student says: "*a point of inflexion is a point where the gradient is zero.*" Give a counter-example, and state the condition that defines a point of inflexion.

14.3 Optimisation

Optimisation means finding the largest or smallest value of a real quantity: the greatest volume, the least cost, the shortest time.

Without the derivative, you could only try inputs one at a time and compare the results. That can suggest an answer but cannot settle it, because between any two inputs you tried there are infinitely many you did not. The derivative replaces the search with an equation. At a maximum or a minimum inside the domain the curve is flat, so $f'(x) = 0$. Solving that one equation, together with checking any endpoints, gives the only candidates.

In practice, setting up the function is usually more work than the calculus.

Most optimisation questions follow the same five steps.

1. **Name the quantity to be optimised** and write an expression for it. At this stage it may contain more than one variable.
2. **Find the constraint**, the second equation linking those variables. It is given in the question, often as a fixed volume, a fixed length of material, or a fixed perimeter.
3. **Use the constraint to eliminate variables**, until the quantity is a function of one variable only. State the values that variable may take.
4. **Differentiate, set the derivative to zero, and solve.**
5. **Justify the nature of the point**, then answer the question that was actually asked.

Steps 3 and 5 are the ones to watch. Reducing to a single variable is where algebra often goes wrong, and skipping the justification can lose marks at the end.

TIP An optimisation question is not finished when you solve $f'(x) = 0$. Marks often depend on what follows: justify the nature of the point, using f'' or the sign of f', and then answer the question that was asked. If the question asks for the maximum volume, give the volume, not the value of x that produces it.

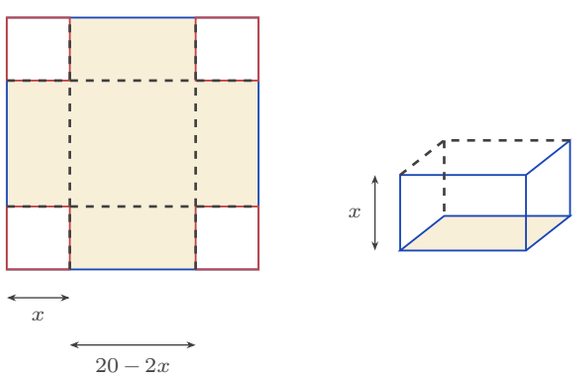


Figure 14.7 Cutting a square of side x from each corner of the card, then folding the flaps up. The base of the box measures $20 - 2x$ each way and its height is x, so $V = x(20 - 2x)^2$.

Worked through. To find the value of x that maximises the volume of the open box in the figure, made from a 20×20 cm square of card by cutting a square of side x from each corner and folding up the sides, note that after cutting the base is $(20 - 2x)$ by $(20 - 2x)$ and the height is x, so

$$V(x) = x(20 - 2x)^2, \qquad 0 < x < 10.$$

Differentiate with the product rule, then take out the common factor $(20 - 2x)$:

$$V'(x) = (20 - 2x)^2 + x \cdot 2(20 - 2x)(-2) = (20 - 2x)(20 - 6x).$$

Setting $V'(x) = 0$ gives $x = 10$ (rejected, no box) or $x = \frac{10}{3}$. Since $V' > 0$ before $\frac{10}{3}$ and $V' < 0$ after, this is a maximum by the first derivative test. The maximum volume is

$$V\left(\tfrac{10}{3}\right) = \frac{16000}{27} \approx 593 \text{ cm}^3.$$

The restriction $0 < x < 10$ comes from the geometry. If x is 0 the box has no height; if x is 10 it has no base.

EXAMPLE 14.5

A closed cylindrical can must hold $355\,\text{cm}^3$. Find the radius that minimises the surface area. Volume fixes the height: $\pi r^2 h = 355$, so $h = \dfrac{355}{\pi r^2}$. The surface area is two circles plus the curved side:

$$S = 2\pi r^2 + 2\pi r h = 2\pi r^2 + \frac{710}{r}.$$

Differentiate and set to zero:

$$\frac{dS}{dr} = 4\pi r - \frac{710}{r^2} = 0 \implies r^3 = \frac{710}{4\pi} \implies r \approx 3.84 \text{ cm}.$$

Since $\dfrac{d^2S}{dr^2} = 4\pi + \dfrac{1420}{r^3} > 0$, this is a minimum by the second derivative test. Either test earns the mark; choose the one whose working is shorter. Using the constraint to remove h is the step that reduces two variables to one.

CAUTION Setting $f' = 0$ finds only the best values *inside* an interval. If the variable is restricted, for example $0 < r \leq 3$, the best value may sit at an **endpoint** instead, where the derivative need not be zero. HL Always list the candidates: every stationary point in the interval, and both endpoints. Then compare their values and take the largest or the smallest.

EXAMPLE 14.6

For the can above, the manufacturer's machinery cannot roll a radius above 3 cm, so $0 < r \leq 3$. Find the radius that minimises the surface area now.

The unconstrained minimum at $r \approx 3.84$ is no longer allowed. Consider the derivative on $(0, 3]$:

$$\frac{dS}{dr} = 4\pi r - \frac{710}{r^2}.$$

Both terms increase as r increases, so $\dfrac{dS}{dr}$ is increasing on this interval. Testing the largest value is therefore enough: at $r = 3$, $12\pi - \frac{710}{9} \approx -41 < 0$. The derivative is negative at the right-hand end, and smaller still to the left, so it is negative throughout and S is decreasing across the whole interval. The minimum therefore sits at the endpoint $r = 3$, with $S = 18\pi + \frac{710}{3} \approx 293$ cm^2, even though $S'(3) \neq 0$.

Optimisation is not restricted to areas and volumes either. Any quantity that can be written as a differentiable function of one variable can be maximised or minimised the same way, and in applied work that function is often exponential.

EXAMPLE 14.7

After a dose is given, the concentration of a drug in the blood is modelled by $C(t) = 20t\,e^{-0.5t}$ mg L^{-1}, where $t \geq 0$ is the time in hours. Find when the concentration is greatest, and the maximum concentration.

The product rule gives

$$C'(t) = 20e^{-0.5t} + 20t(-0.5)e^{-0.5t} = 20e^{-0.5t}\,(1 - 0.5t)\,.$$

Since $e^{-0.5t}$ is never zero, $C'(t) = 0$ requires $1 - 0.5t = 0$, so $t = 2$ hours.

To justify that this is a maximum, look at the sign of C': the factor $e^{-0.5t}$ is always positive, so the sign is that of $1 - 0.5t$, which is positive for $t < 2$ and negative for $t > 2$. The concentration rises, then falls. The maximum is

$$C(2) = 40e^{-1} \approx 14.7 \text{ mg L}^{-1}.$$

The shape can be read from the two factors. For small t the factor t increases faster than $e^{-0.5t}$ decreases, so C rises. For large t the exponential factor decreases faster, so C falls. The maximum is where the two effects are equal, which is what $C'(t) = 0$ records.

YOUR TURN 14.3 Optimisation

Setting up and solving

16.

(a) Two numbers have a sum of 20. Find the maximum value of their product, and justify that it is a maximum.

(b) A rectangle has a perimeter of 40 cm. Find the dimensions giving the maximum area, and justify that it is a maximum.

(c) Find the minimum value of $S = x^2 + \dfrac{16}{x}$ for $x > 0$, justifying that it is a minimum.

17. A rectangular field is fenced on three sides, the fourth being a wall, using 100 m of fencing. Find the maximum area, and justify that it is a maximum.

Using a constraint to remove a variable

18. A closed box has a square base of side x cm and height h cm. Its total surface area is 150 cm^2.

(a) Show that the volume is $V = \dfrac{75x - x^3}{2}$, and state the possible values of x.

(b) Find the value of x that maximises the volume, and justify that it is a maximum.

(c) Find the maximum volume, and describe the shape of the box.

19. An open-topped cylindrical tank must hold 1000 cm^3. Show that its surface area is $S = \pi r^2 + \dfrac{2000}{r}$, and find the radius that minimises it.

20. Find the coordinates of the point on the curve $y = \sqrt{x}$ that is closest to the point $(4, 0)$, and the least distance. *(Hint: minimise the square of the distance)*

Exponential, logarithmic and trigonometric models

21.

(a) Find the minimum value of $y = x - \ln x$ for $x > 0$, justifying that it is a minimum.

(b) The concentration of a drug in the blood is $C(t) = 12t\,e^{-t/4}$ mg L^{-1}, where $t \geq 0$ is measured in hours. Find the time at which the concentration is greatest, and the maximum concentration.

22. GDC A rectangle has its base on the x-axis and its two upper vertices on the curve $y = \cos x$, symmetric about the y-axis, with $0 < x < \frac{\pi}{2}$. Show that its area is $A = 2x\cos x$, and find the value of x that maximises A.

Endpoints of a restricted domain

23. A rectangle has a perimeter of 40 cm, and one side, of length x cm, may be at most 8 cm long. Find the greatest possible area, and explain why it does not occur where $\dfrac{dA}{dx} = 0$.

24. For the drug in Question 21(b), blood samples are taken only during the first three hours, so $0 \le t \le 3$. Find the greatest concentration recorded in that time, and state when it occurs.

Concept check

25. A student is asked for the greatest value of $f(x) = x^3 - 3x$ for $0 \le x \le 3$, and writes: $f'(x) = 3x^2 - 3 = 0$, so $x = 1$ (as $x = -1$ is outside the interval), and the greatest value is $f(1) = -2$. Identify the two errors and find the greatest value.

14.4 Kinematics

Motion is change of position, so calculus describes it exactly. Let $s(t)$ be a particle's **displacement** from a fixed origin at time t. Its **velocity** is the rate of change of displacement, $v = \dfrac{ds}{dt}$, and its **acceleration** is the rate of change of velocity.

KEY · IB

$$a = \frac{dv}{dt} = \frac{d^2s}{dt^2}$$

It helps to read the chain $s \to v \to a$ in ordinary words. Displacement answers *where*. Velocity answers *how fast the where is changing*. Acceleration answers *how fast the how-fast is changing*. In a car on a straight road, s is the signed distance from the starting point (the odometer reads distance travelled, which is different), the speedometer reads $|v|$, and the push you feel in your seat is a. Each quantity is the derivative of the one before it, which is why acceleration is a rate of a rate, and so a second derivative.

Velocity carries a sign, so it records direction as well as how fast the particle moves. On a straight line there are only two directions, and the sign says which one. **Speed** is the size of the velocity, written $|v|$, and it is never negative. A particle with $v = -8$ is moving in the negative direction at a speed of 8.

The particle is **momentarily at rest** (at rest for an instant) when $v = 0$. This is the stationary-point condition of §14.1 applied to displacement: a turning point on the graph of $s(t)$ is the instant the particle stops and reverses. The ball at the top of its flight is exactly this point.

Speed follows a separate rule, and it is better understood than memorised. Think of acceleration as a push. If the push acts in the direction the particle is already moving, the particle gains speed; if it acts against the motion, the particle loses speed. In symbols, the particle **speeds up when v and a have the same sign** and slows down when the signs differ. Notice that this does not depend on a being positive: a particle moving in the negative direction with a negative acceleration is speeding up, because the push and the motion agree.

CAUTION $a = 0$ does not mean the particle has stopped. Acceleration is the derivative of velocity, so $a = 0$ marks a stationary point of $v(t)$: an instant where the velocity may be greatest or least (a candidate, as with $f'(x) = 0$ in §14.1), not where it is zero. A car cruising at a steady $100\,\text{km}\,\text{h}^{-1}$ has zero acceleration and is not stopped at all. The particle is at rest only when $v = 0$.

One quantity needs integration rather than differentiation: the **total distance travelled**, which keeps counting when the particle turns and comes back. Displacement over an interval is $\int v\,dt$ and distance travelled is $\int |v|\,dt$. Both belong to the same syllabus item as the derivatives above, and both are developed in Ch. 15.

Worked through. Take a particle whose displacement is $s(t) = t^3 - 12t^2 + 36t$ metres, for $t \geq 0$ seconds, and find when it is at rest and its acceleration at those times. Differentiate for velocity:

$$v(t) = 3t^2 - 24t + 36 = 3(t-2)(t-6).$$

The particle is at rest when $v = 0$: at $t = 2$ and $t = 6$ seconds. Acceleration is $a(t) = 6t - 24$. At $t = 2$, $a = -12\,\text{m}\,\text{s}^{-2}$; at $t = 6$, $a = 12\,\text{m}\,\text{s}^{-2}$. The sign change of velocity at each rest point means the particle reverses direction there.

Now ask whether the particle is speeding up at $t = 7$. There $v(7) = 3(5)(1) = 15 > 0$ and $a(7) = 6(7) - 24 = 18 > 0$. The two signs are the same, so the particle is speeding up: it moves in the positive direction and the acceleration acts the same way.

Not all motion is polynomial. Motion that repeats, such as a pendulum, a mass on a spring, a wave or a vibrating string, is often modelled by a sine or a cosine. When the acceleration is proportional to the displacement and acts in the opposite direction, so that $a = -k^2 s$ for a constant k, the motion is called **simple harmonic motion**. The particle is at rest at its greatest displacement, where the pull back towards the origin is strongest, and moves fastest as it passes through the origin, where the pull is zero.

EXAMPLE 14.8

A particle moves in a straight line with displacement $s = 4\sin 2t$ metres, for $t \geq 0$ seconds. Find its velocity and acceleration, the first time it is at rest, and show that $a = -4s$.

Differentiate twice, using the chain rule each time:

$$v = \frac{ds}{dt} = 8\cos 2t, \qquad a = \frac{dv}{dt} = -16\sin 2t.$$

The particle is at rest when $v = 0$, that is when $\cos 2t = 0$. The first solution is $2t = \frac{\pi}{2}$, so $t = \frac{\pi}{4} \approx 0.785$ s. At that instant $s = 4\sin\frac{\pi}{2} = 4$ m, the greatest displacement it reaches. Finally, since $s = 4\sin 2t$,

$$-4s = -4(4\sin 2t) = -16\sin 2t = a. \qquad \blacksquare$$

So $a = -k^2 s$ with $k = 2$, and the particle moves with simple harmonic motion: the further it is from the origin, the harder it is pulled back. The same relation describes a pendulum (for small swings), a mass on a spring and a sound wave.

TIP In kinematics questions, state the units with every answer: $\mathrm{m\,s^{-1}}$ for velocity, $\mathrm{m\,s^{-2}}$ for acceleration. A correct number without its unit may lose a mark.

TIP To find the greatest speed, compare $|v|$, not v. Because speed ignores direction, the greatest speed can occur where the velocity is most negative.

YOUR TURN 14.4 Kinematics

Velocity, acceleration and rest

26. A particle moves in a straight line. Its displacement is s metres and its velocity is v $\mathrm{m\,s^{-1}}$ at time t seconds, $t \geq 0$.

(a) Its displacement is $s = t^2 - 4t + 3$. Find when it is at rest, and its acceleration then.

(b) Its displacement is $s = 2t^3 - 9t^2 + 12t$. Find the times when it is at rest.

(c) Its displacement is $s = t^3 - 3t^2$. Find its velocity and acceleration at $t = 2$.

(d) Its velocity is $v = t^2 - 5t + 6$. Find when it is at rest, and its acceleration at those times.

27. A particle has velocity $v = 3t^2 - 12t + 9$ $\mathrm{m\,s^{-1}}$ at time t seconds, $t \geq 0$.

(a) Find the times when it is at rest.

(b) Find its acceleration at those times.

Speeding up or slowing down

28.

(a) For the particle in Question 27, determine whether it is speeding up or slowing down at $t = 1.5$, giving a reason.

(b) A second particle has velocity $v = t - 4$ m s^{-1} at time t seconds. State, with a reason, whether it is speeding up or slowing down at $t = 2$.

Speed

29. A particle has velocity $v = t^2 - 6t + 3$ m s^{-1} at time t seconds, for $0 \le t \le 5$.

(a) Find its speed at $t = 1$.

(b) Find the time at which its acceleration is zero. State, with a reason, whether the particle is at rest at that time.

(c) Find the greatest speed of the particle for $0 \le t \le 5$.

30. A particle has displacement $s = t^3 - 9t^2 + 24t$ metres at time t seconds, $t \ge 0$.

(a) Find the times when the particle is at rest.

(b) Find the time at which its acceleration is zero, and its speed at that time.

Other functions of time

31.

(a) A particle has displacement $s = 2\cos 3t$ metres, $t \ge 0$. Find v and a, and show that $a = -9s$.

(b) For the particle in part (a), find the first time after $t = 0$ at which it is at rest, and its displacement at that instant.

(c) A particle has velocity $v = 6e^{-0.5t}$ m s^{-1}. Find its acceleration at $t = 0$, and explain why the particle never comes to rest.

(d) A particle has displacement $s = t + 2\cos t$ metres, for $0 \le t \le 2\pi$. Find the times at which it is at rest.

Concept check

32. A student says: "*the acceleration is negative, so the particle is slowing down.*" State in one line when this is false.

33. Explain the difference between the *displacement* and the *distance travelled* over $0 \le t \le 5$, for a particle that reverses direction at $t = 3$.

14.5 Implicit Differentiation HL

So far every curve has been given as $y = f(x)$, with y written on its own. But an equation such as $x^2 + y^2 = 25$ also defines a curve, without y being isolated. **Implicit differentiation** finds the gradient of such a curve directly, without first solving for y. Every tangent and normal method of Ch. 13 then works on these curves too, including curves that are not the graph of a function.

The method is the chain rule. Treat y as an unknown function of x. A term in y then differentiates as usual, but it gains an extra factor $\frac{dy}{dx}$, because you are differentiating a function of y with respect to x. For example, $\frac{d}{dx}(y^2) = 2y\frac{dy}{dx}$.

The procedure has four steps.

1. Differentiate every term of the equation with respect to x, using the product rule where x and y are multiplied.
2. Collect the terms containing $\frac{dy}{dx}$ on one side.
3. Factorise $\frac{dy}{dx}$ out of those terms.
4. Divide to make $\frac{dy}{dx}$ the subject.

Worked through. For the circle $x^2 + y^2 = 25$, differentiate both sides with respect to x:

$$2x + 2y\frac{dy}{dx} = 0 \implies \frac{dy}{dx} = -\frac{x}{y}.$$

At $(3, 4)$ the gradient is $-\frac{3}{4}$. The answer depends on both coordinates, and it must: the gradient changes as you move round the circle.

EXAMPLE 14.9

Find $\frac{dy}{dx}$ for $x^2 + xy + y^2 = 7$.

The middle term needs the product rule. Differentiating term by term:

$$2x + \left(y + x\frac{dy}{dx}\right) + 2y\frac{dy}{dx} = 0.$$

Collect the $\frac{dy}{dx}$ terms:

$$\frac{dy}{dx}(x + 2y) = -(2x + y) \implies \frac{dy}{dx} = -\frac{2x + y}{x + 2y}.$$

Each of the four steps appears once: differentiate, collect, factorise, divide.

EXAMPLE 14.10

Find the equation of the tangent to the curve $x^2 + xy + y^2 = 7$ at the point $(1, 2)$.
First confirm the point lies on the curve: $1 + 2 + 4 = 7$. From the previous example,

$$\frac{dy}{dx} = -\frac{2x + y}{x + 2y} = -\frac{2 + 2}{1 + 4} = -\frac{4}{5} \quad \text{at } (1, 2).$$

The tangent is $y - 2 = -\frac{4}{5}(x - 1)$, or $4x + 5y = 14$. Implicit differentiation supplies the gradient. The rest is the usual equation of a line through a given point.

The same method handles exponential and trigonometric terms without any change. A term in y still gains a factor $\frac{dy}{dx}$, whatever function it sits inside.

EXAMPLE 14.11

The curve C has equation $e^y + xy = 1$. Show that $(0, 0)$ lies on C, and find the gradient there.
At $(0, 0)$: $e^0 + 0 = 1$, so the point lies on C.
Differentiate every term with respect to x. The first needs the chain rule and the second the product rule:

$$e^y \frac{dy}{dx} + \left(y + x\frac{dy}{dx}\right) = 0.$$

Collect the $\frac{dy}{dx}$ terms and factorise:

$$\frac{dy}{dx}(e^y + x) = -y \implies \frac{dy}{dx} = -\frac{y}{e^y + x}.$$

At $(0, 0)$ this gives $\frac{dy}{dx} = \frac{0}{1 + 0} = 0$, so the tangent there is horizontal.
Notice that this curve cannot be rearranged to give y in terms of the functions of this course: y appears both inside an exponential and multiplied by x. The curve still has a gradient at every point, and near $(0, 0)$ it still gives one y for each x. There is no expression in this course's functions to write down. Implicit differentiation is not a shortcut here; within the course, it is the most direct method available.

YOUR TURN 14.5 Implicit Differentiation HL

Finding the gradient function

34. Find $\dfrac{dy}{dx}$ for each curve.

(a) $x^2 + y^2 = 16$
(b) $xy = 12$
(c) $x^2 - y^2 = 9$
(d) $x^3 + y^3 = 6$

35. Each of these needs the product rule.

(a) Find $\dfrac{dy}{dx}$ for $x^2 + xy = 5$.
(b) Find $\dfrac{dy}{dx}$ for $x^2y + y^2 = 4$.

36. Find $\dfrac{dy}{dx}$ for each curve.

(a) $e^y = x^2 + y$
(b) $\sin y + x^2 = 4$
(c) $x \ln y = 4$

Gradients, tangents and normals

37.

(a) Find the gradient of $x^2 + y^2 = 169$ at the point $(5, -12)$.
(b) Find the equation of the tangent to $xy = 4$ at the point $(2, 2)$.
(c) On the curve $x^2 + 4y^2 = 16$, find the two points at which the tangent is horizontal.
(d) Explain why implicit differentiation gives a gradient in terms of both x and y, while differentiating $y = f(x)$ gives one in terms of x alone.

38.

(a) Find the gradient of $e^y + x = 3$ at the point $(2, 0)$.
(b) Show that $(1, 2)$ lies on the curve $x^3 + y^3 = 9$, and find the equation of the normal there.

39. Show that the point $(2, 1)$ lies on the curve $x^2 - 3xy + y^2 = -1$, and find the equations of the tangent and the normal at that point.

Concept check

40. A student differentiates $x^2 + y^2 = 25$ with respect to x and writes $2x + 2y = 0$. Identify the error, and find $\dfrac{dy}{dx}$.

14.6 Related Rates HL

When two quantities are linked by an equation and both change over time, their rates of change are linked too. **Related rates** problems use the chain rule to connect those rates. You differentiate the equation that relates the quantities with respect to time.

The method has three steps.

1. Write an equation relating the quantities.
2. Differentiate every term with respect to t, so that each variable contributes its own rate.
3. Substitute the values known at the instant in question, and solve for the rate you need.

If the equation contains a variable whose rate is neither given nor asked for, use a fixed relation from the shape, such as similar triangles, to remove that variable before differentiating.

TIP Substitute the given values only *after* differentiating. Putting $r = 5$ into the volume formula first turns a variable into a constant, and a constant has rate of change zero.

Worked through. Suppose air is pumped into a spherical balloon at $100\,\text{cm}^3\text{s}^{-1}$, and we need the rate at which the radius is increasing when $r = 5$ cm. Volume and radius are related by $V = \frac{4}{3}\pi r^3$. Differentiate with respect to t:

$$\frac{dV}{dt} = 4\pi r^2 \frac{dr}{dt}.$$

Substitute $\dfrac{dV}{dt} = 100$ and $r = 5$:

$$100 = 4\pi(25)\frac{dr}{dt} \implies \frac{dr}{dt} = \frac{1}{\pi} \approx 0.318 \text{ cm s}^{-1}.$$

The rate of change of volume is constant, but the radius increases more slowly as the balloon grows, because the same volume is spread over a larger surface.

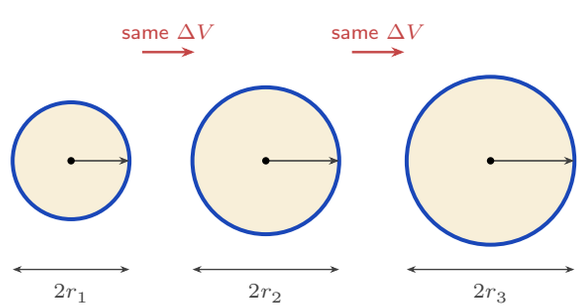


Figure 14.8 Equal volumes of air are added at each step, but the radius grows by less each time. The volume added is spread over the surface, and the surface area $4\pi r^2$ is itself growing, which is why $\frac{dr}{dt} = \frac{1}{4\pi r^2}\frac{dV}{dt}$ falls as r increases.

EXAMPLE 14.12

A 10 m ladder leans against a wall. The foot slides away from the wall at $0.6\,\text{m}\,\text{s}^{-1}$. How fast is the top sliding down when the foot is 6 m from the wall?

Let x be the distance of the foot from the wall and y the height of the top. Then $x^2 + y^2 = 100$. Differentiate with respect to t:

$$2x\frac{dx}{dt} + 2y\frac{dy}{dt} = 0 \implies \frac{dy}{dt} = -\frac{x}{y}\frac{dx}{dt}.$$

When $x = 6$, $y = \sqrt{100 - 36} = 8$. With $\dfrac{dx}{dt} = 0.6$:

$$\frac{dy}{dt} = -\frac{6}{8}(0.6) = -0.45 \text{ m}\,\text{s}^{-1}.$$

The negative sign means the top moves downward, as expected.

EXAMPLE 14.13

A water tank is an inverted cone with its vertex at the bottom, as tall as it is wide: at every level, the water surface has radius $r = \frac{h}{2}$, where h is the water depth. Water flows in at $2\,\text{m}^3\,\text{min}^{-1}$. How fast is the depth rising when $h = 3$ m?

Write the volume in terms of h alone, using the shape constraint $r = \frac{h}{2}$:

$$V = \frac{1}{3}\pi r^2 h = \frac{1}{3}\pi\left(\frac{h}{2}\right)^2 h = \frac{\pi h^3}{12}.$$

Differentiate with respect to time:

$$\frac{dV}{dt} = \frac{\pi h^2}{4}\frac{dh}{dt} \implies 2 = \frac{9\pi}{4}\frac{dh}{dt} \implies \frac{dh}{dt} = \frac{8}{9\pi} \approx 0.283 \text{ m}\,\text{min}^{-1}.$$

Remove r *before* differentiating. Differentiating an expression that still contains two linked variables is a common source of error.

The algebra also shows the physics: the same inflow raises the level more slowly as the cone widens.

The quantities linked need not be lengths and volumes. When the link is an angle, the relating equation is trigonometric, and the method is unchanged.

EXAMPLE 14.14

A balloon rises vertically at $3\,\text{m}\,\text{s}^{-1}$ from a point 40 m from an observer. Find the rate at which the angle of elevation of the balloon is increasing when the balloon is 40 m high.

Let h be the height of the balloon and θ the angle of elevation. The observer's distance is fixed at 40 m, so

$$\tan\theta = \frac{h}{40}.$$

Differentiate both sides with respect to t. The left side needs the chain rule:

$$\sec^2\theta\,\frac{d\theta}{dt} = \frac{1}{40}\,\frac{dh}{dt}.$$

When $h = 40$ the triangle is isosceles, so $\theta = \frac{\pi}{4}$ and $\sec^2\theta = 2$. Substituting this and $\dfrac{dh}{dt} = 3$:

$$2\,\frac{d\theta}{dt} = \frac{3}{40} \implies \frac{d\theta}{dt} = \frac{3}{80} = 0.0375 \text{ rad s}^{-1}.$$

The balloon rises at a constant rate, but the angle does not increase at a constant rate. As the balloon climbs, $\sec^2\theta$ grows, so $\dfrac{d\theta}{dt}$ falls: the observer's head turns more and more slowly, even though the balloon never slows down.

YOUR TURN 14.6 Related Rates HL

One formula, one rate

41.

(a) A square's side grows at $2\,\text{cm s}^{-1}$. Find the rate of change of its area when the side is 5 cm.

(b) A circle's radius grows at $3\,\text{cm s}^{-1}$. Find the rate of change of its area when $r = 4$ cm.

(c) A cube's side grows at $1\,\text{cm s}^{-1}$. Find the rate of change of its volume when the side is 10 cm.

(d) A circle's radius grows at 0.4 cm s^{-1}. Find the rate of increase of its circumference.

42. A spherical snowball melts so that its volume decreases at $20\,\text{cm}^3\,\text{min}^{-1}$. Find the rate of change of its radius when $r = 8$ cm, and explain what the sign of your answer means.

Pythagoras and similar triangles

43. A ladder 13 m long leans against a vertical wall. Its top slides down the wall at $0.5\,\text{m}\,\text{s}^{-1}$. Find the rate at which the foot moves away from the wall when the top is 5 m above the ground.

44. Water is poured at $20\,\text{cm}^3\text{s}^{-1}$ into a cone whose vertex points downwards. The water depth h is always twice the radius r of the water surface.

(a) Show that the volume of water is $V = \frac{2}{3}\pi r^3$.

(b) Find the rate at which r increases when $r = 5$ cm.

45. A conical tank, vertex down, is 6 m deep and has a top radius of 2 m. Water drains out at $0.5\,\text{m}^3\,\text{min}^{-1}$. Find the rate of change of the water depth when the depth is 3 m, and interpret its sign. *(Hint: use similar triangles to write r in terms of h)*

Angles

46. An aircraft flies horizontally at a height of 3 km, at $0.2\,\text{km}\,\text{s}^{-1}$, directly towards an observer on the ground. Find the rate of change of the angle of elevation of the aircraft when its horizontal distance from the observer is 4 km.

47. A searchlight rotates at $0.2\,\text{rad}\,\text{s}^{-1}$ and its beam sweeps along a straight wall 50 m away. Find the speed of the spot of light along the wall when the beam makes an angle of $\frac{\pi}{6}$ with the perpendicular to the wall.

Concept check

48. The radius of a circle grows at 0.5 cm s^{-1}. To find the rate of change of its area when $r = 6$ cm, a student writes $A = \pi(6)^2 = 36\pi$, so $\dfrac{dA}{dt} = 0$. Identify the error and find the correct rate.

Chapter 14 · Derivative Applications

Reflections

TOK This chapter is full of conditions that are necessary but not sufficient (they must hold, but they are not enough on their own): $f'(x) = 0$ holds at every smooth maximum, and, for a twice-differentiable function, $f''(x) = 0$ at every point of inflexion, yet neither condition guarantees one. What kind of knowledge is a necessary condition? Does knowing where something must be count as knowing where it is?

TOK Calculus makes exact claims about idealised objects. A ball is motionless at the top of its flight, but only for an instant, and an instant has no duration. The can that uses least metal has its height equal to its diameter, but real cans are taller and narrower, because the model ignores printing, shipping and the size of a hand. When a model is exact and its conclusion still differs from the world, which one is wrong: the model, or our reading of it?

IM The idea that nature minimises something appears across physics under many names. In seventeenth-century France, Pierre de Fermat proposed that light travels the path taking the least time, which explains refraction. Euler and Lagrange later generalised this into the *calculus of variations*, the study of which curve or path makes a quantity smallest, and the *principle of least action*, which restates much of classical mechanics as finding where a quantity is stationary. The belief that the world is in some sense optimal has produced a great deal of physics.

RW Many modern machine-learning models are trained by *gradient descent*, which is this chapter's idea at scale. A model has millions of adjustable numbers, and how wrong its predictions are is a function of those numbers; training means making that function as small as possible. With millions of variables the minimum cannot be found directly, so the computer measures the gradient at its current point, takes a small step in the direction that makes the function smaller, and repeats. As the gradient approaches zero the steps shrink and the values approach a minimum. That is $f' = 0$, applied billions of times.

End-of-Chapter Problems
Chapter 14: Derivative Applications

1. For $f(x) = x^4$, show that $f''(0) = 0$ but that $(0, 0)$ is *not* a point of inflexion. [4]

2. A stone dropped in a pond sends out a circular ripple whose radius grows at $0.5\,\text{m s}^{-1}$. Find the rate at which the enclosed area is increasing when the radius is 8 m. [4]

3. A particle's velocity is $v(t) = 2t^2 - 14t + 20\,\text{m s}^{-1}$, $t \geq 0$.

(a) Find the times when the particle is at rest. [2]
(b) Find the minimum velocity and the time at which it occurs. [3]

4. A company's profit from selling x units is $P(x) = -2x^2 + 120x - 1000$ dollars. Find the value of x that maximises the profit, justifying that it is a maximum, and state the maximum profit. [5]

5. The curve $\sin y + x^2 = 1$ passes through the point $(1, 0)$. Find the gradient of the curve at that point. [5]

6. A curve has equation $x^2 + xy + y^2 = 3$.

(a) Find $\dfrac{dy}{dx}$ in terms of x and y. [3]
(b) Find the equation of the tangent at the point $(1, 1)$. [3]

7. Water drains from an inverted cone whose radius always equals its height. The volume is $V = \frac{1}{3}\pi r^2 h$.

(a) Show that $V = \frac{1}{3}\pi h^3$. [2]
(b) Given the volume decreases at $2\,\text{cm}^3\text{s}^{-1}$, find the rate at which the depth h is falling when $h = 4$ cm. [4]

8. Consider $f(x) = x + 2\cos x$ on the closed interval $0 \leq x \leq 2\pi$.

(a) Find the stationary points in the interval and classify them. [3]
(b) Find the global maximum and minimum of f on the interval. [3]

9. Consider $g(x) = x + \dfrac{4}{x}$ on the interval $1 \leq x \leq 3$.

(a) Find the stationary point of g in the interval and classify it. [4]
(b) Find the maximum value of g on the interval, and state where it occurs. [2]

10. The curve C has equation $x^2 + y^2 - 4x + 2y = 15$.

(a) Show that the point $(0, 3)$ lies on C. [1]

(b) Use implicit differentiation to find $\dfrac{dy}{dx}$. [3]

(c) Find the equation of the tangent to C at $(0, 3)$. [2]

11. A street lamp is 5 m tall. A person of height 1.8 m walks away from it at $1.2\,\text{m}\,\text{s}^{-1}$. Find the rate at which the length of the person's shadow is increasing. [6]

12. A water tank is an inverted cone in which the water surface radius is always $\frac{3}{4}$ of the depth h. Water is pumped in at $1.5\,\text{m}^3\,\text{min}^{-1}$. Find the rate at which the depth is rising when $h = 2$ m. [6]

13. Find the point on the curve $y = \sqrt{x}$ closest to the point $(5, 0)$, and state the minimum distance. [6]

14. For the curve $x^2 + y^2 = 25$, use implicit differentiation twice to show that $\dfrac{d^2y}{dx^2} = -\dfrac{25}{y^3}$. [6]

15. A closed cylindrical can must hold $500\,\text{cm}^3$.

(a) Show that the surface area is $S = 2\pi r^2 + \dfrac{1000}{r}$. [3]

(b) Find the radius that minimises S, and justify it is a minimum. [4]

16. A rectangular pen of area $200\,\text{m}^2$ is built against a wall, needing fencing on three sides. Let the side perpendicular to the wall be x.

(a) Show that the length of fencing is $L = 2x + \dfrac{200}{x}$. [3]

(b) Find the dimensions that minimise the fencing used. [4]

17. The curve $y = x^3 + ax^2 + bx$ has stationary points at $x = 1$ and $x = -2$.

(a) Find the values of a and b. [4]

(b) Classify each stationary point. [3]

18. An open rectangular tank with a square base of side x must hold $32\,\text{m}^3$.

(a) Show that the surface area (base + four sides) is $A = x^2 + \dfrac{128}{x}$. [3]

(b) Find the value of x that minimises the material used. [4]

19. The population of a colony, in thousands, is modelled by $P(t) = 40t\,e^{-t/5}$ for $t \geq 0$, where t is measured in years.

(a) Show that $P'(t) = 8e^{-t/5}(5-t)$. [3]

(b) Find the time at which the population is largest, and the largest population, giving the population to three significant figures. [4]

20. A cone is inscribed in a sphere of radius R. Show that the cone of maximum volume has height $h = \frac{4R}{3}$. [7]

21. A 5 m ladder slides down a wall, its foot moving away at a constant $0.5\,\text{m s}^{-1}$.

(a) Find the rate at which the top descends when the foot is 3 m from the wall. [3]

(b) Find the rate at which the area of the triangle formed by the ladder, wall, and ground is changing at that instant. [4]

22. A particle moves along the curve $y = x^2$ so that $\dfrac{dx}{dt} = 2$ at all times. Find the rate at which its distance from the origin is increasing as it passes through the point $(1, 1)$. [7]

23. Consider $f(x) = x^3 - 3x^2 - 9x + 5$.

(a) Find and classify the stationary points. [4]

(b) Find the point of inflexion. [2]

(c) State the interval on which f is decreasing. [2]

24. A particle moves with displacement $s(t) = t^3 - 6t^2 + 9t$ metres, $t \geq 0$.

(a) Find the velocity and the times when the particle is at rest. [3]

(b) Find the acceleration when $t = 1$. [2]

(c) Determine whether the particle is speeding up or slowing down at $t = 1.5$. [3]

25. Consider $f(x) = e^{-x^2/2}$.

(a) Find $f'(x)$, and show that $f''(x) = (x^2 - 1)\,e^{-x^2/2}$. [5]

(b) Find the exact coordinates of the points of inflexion, and state the interval on which the curve is concave down. [3]

26. A particle moves with displacement $s(t) = t^3 - 9t^2 + 24t - 10$ metres, for $0 \leq t \leq 6$ seconds.

(a) Find the times at which the particle is at rest. [3]

(b) Find its acceleration at each of those times. [2]

(c) Find the intervals of time during which the particle's *speed* is increasing. [3]

27. Consider $f(x) = xe^{-x}$.

(a) Find $f'(x)$ and the coordinates of the stationary point. [3]

(b) Determine its nature using the second derivative. [3]
(c) Find the x-coordinate of the point of inflexion. [2]

28. Consider $f(x) = x^2 \ln x$ for $x > 0$.

(a) Find $f'(x)$, and factorise it. [2]
(b) Find the exact coordinates of the stationary point. [3]
(c) Find $f''(x)$, and use it to classify that point. [3]
(d) Find the exact coordinates of the point of inflexion, and show that the concavity changes there. [4]
(e) Describe the behaviour of $f(x)$ as $x \to 0^+$ and as $x \to \infty$. [3]

29. An open-topped box is made from a 16×10 cm rectangle of card by cutting a square of side x cm from each corner and folding up the sides.

(a) Show that the volume is $V = x(16 - 2x)(10 - 2x)$, and state the possible values of x. [3]
(b) Find $\dfrac{dV}{dx}$ and show that it is zero at $x = 2$ and $x = \frac{20}{3}$. [3]
(c) Explain why $x = \frac{20}{3}$ must be rejected. [2]
(d) Find the maximum volume, justifying that it is a maximum. [4]

30. A particle moves in a straight line with displacement $s(t) = 12\left(1 - e^{-t/3}\right)$ metres, for $t \geq 0$ seconds.

(a) Find expressions for the velocity and the acceleration. [3]
(b) Write down the initial velocity, and show that the velocity is never zero. [2]
(c) Find the displacement and the velocity at $t = 3$, each to three significant figures. [2]
(d) State what happens to s and to v as $t \to \infty$, and interpret the limiting value of s. [3]
(e) Find the exact time at which the speed has fallen to half its initial value. [3]

31. The curve C has equation $x^3 + y^3 = 9xy$.

(a) Verify that the point $(2, 4)$ lies on C. [2]
(b) Use implicit differentiation to show that $\dfrac{dy}{dx} = \dfrac{x^2 - 3y}{3x - y^2}$. [4]
(c) Find the gradient of C at $(2, 4)$. [2]
(d) Find the equation of the tangent to C at $(2, 4)$. [3]

32. Consider $f(x) = e^x - 2x$.

(a) Find $f'(x)$ and the coordinates of the stationary point. [3]
(b) Determine the nature of the stationary point. [2]
(c) Explain why f has no point of inflexion. [2]

33. A particle moves in a straight line with displacement $s(t) = 2\sin 3t$ metres, $t \geq 0$

seconds.

(a) Find expressions for the velocity and the acceleration. [2]
(b) Show that $a = -9s$. [2]
(c) Find the first two times at which the particle is at rest. [3]
(d) State the maximum speed of the particle, and the displacement at which it occurs. [3]

34. Consider $f(x) = \dfrac{\ln x}{x}$ for $x > 0$.

(a) Show that $f'(x) = \dfrac{1 - \ln x}{x^2}$. [3]
(b) Find the coordinates of the stationary point and determine its nature. [4]
(c) Show that $f''(x) = \dfrac{2\ln x - 3}{x^3}$, and hence find the x-coordinate of the point of inflexion. [4]

35. A gutter is made by bending a long strip of metal of width 30 cm into three equal sections of 10 cm. The two outer sections are turned up at an angle θ to the horizontal, where $0 < \theta < \frac{\pi}{2}$. The cross-sectional area is $A = 100\sin\theta\,(1 + \cos\theta)$.

(a) Find $\dfrac{dA}{d\theta}$. [3]
(b) Show that A is stationary when $2\cos^2\theta + \cos\theta - 1 = 0$, and find the value of θ that maximises A. [4]
(c) Find the maximum cross-sectional area. [2]

36. The curve C has equation $xe^y + y = 1$.

(a) Verify that the point $(1, 0)$ lies on C. [2]
(b) Show that $\dfrac{dy}{dx} = -\dfrac{e^y}{xe^y + 1}$. [4]
(c) Find the gradient of C at $(1, 0)$. [2]
(d) Find the equation of the tangent to C at $(1, 0)$. [3]

37. A lighthouse stands 2 km from a straight shoreline, and its beam rotates at one revolution per minute. Let θ be the angle between the beam and the perpendicular from the lighthouse to the shore, and let x be the distance from the foot of that perpendicular to the spot of light.

(a) Write x in terms of θ, and state the value of $\dfrac{d\theta}{dt}$ in radians per minute. [2]
(b) Show that $\dfrac{dx}{dt} = 4\pi\sec^2\theta$. [3]
(c) Find the speed of the spot of light along the shore when $\theta = 30°$. [3]
(d) Describe what happens to this speed as $\theta \to 90°$, and explain why. [2]

Challenge Questions

38. GDC **Modelling the best can.** The notes show that a closed cylinder of fixed volume uses least metal when its height equals its diameter. Real drink cans are taller than that. This investigation rebuilds the model around cost instead of area, and asks how far the improved model explains the cans on a shop shelf. Throughout, a closed can of radius r and height h holds a fixed volume V.

(a) Suppose the metal for the two ends costs k times as much per unit area as the metal for the curved side, where $k > 0$. Show that the cost of the can is proportional to $C = 2k\pi r^2 + \dfrac{2V}{r}$. Show that C has a minimum when $h = 2kr$, and hence that the cheapest can has $\dfrac{h}{d} = k$, where d is the diameter. [5]

(b) Now let the metal cost the same everywhere, but suppose each end is cut from its own square of side $2r$, and the corners are scrap that must still be paid for. Show that the cheapest can has $\dfrac{h}{d} = \dfrac{4}{\pi}$. [4]

(c) If instead the ends are cut from one sheet with their centres on a triangular grid, each disc uses a regular hexagon whose inscribed circle is the disc. Show that such a hexagon has area $2\sqrt{3}\,r^2$, and that the cheapest can then has $\dfrac{h}{d} = \dfrac{2\sqrt{3}}{\pi}$. [3]

(d) A typical 330 ml drink can is about 11.5 cm tall and 6.6 cm in diameter. Find the value of k in the model of part (a) that would make this shape the cheapest, and compare it with the values implied by parts (b) and (c). [2]

(e) Take $k = 1$ and write $h = xd$ with $x > 0$. Show that the surface area is

$$S = (2\pi)^{1/3}\, V^{2/3}\, \frac{1 + 2x}{x^{2/3}}.$$

Find the percentage of extra metal used by the drink-can shape of part (d) compared with the shape $x = 1$. Use the derivative $\dfrac{dS}{dx}$ at $x = 1$ to explain why a shape some way from the optimum costs only a little more. [4]

(f) List other features of a real can that a better model would include, and describe how you would investigate one of them. [3]

39. **Closest approach.** Finding the point of a curve nearest to a fixed point is an optimisation problem, and its answer has a geometric description. This investigation finds that description, states exactly when it holds, and uses it to count the normals that can be drawn to a parabola from a given point.

(a) Let $(x, \sqrt{x})$ be a point on $y = \sqrt{x}$ and let D be its distance from $(4, 0)$. Show that $D^2 = x^2 - 7x + 16$. [3]

(b) Explain why minimising D^2 gives the same answer as minimising D, and why using D^2 is easier. [2]

(c) Find the point on the curve closest to $(4,0)$, and the least distance. [4]

(d) Show that at this closest point the line joining it to $(4,0)$ is perpendicular to the tangent to the curve. [3]

(e) Let f be differentiable on an open interval, and let (a,b) be a point not on the curve $y=f(x)$. Suppose the distance from (a,b) to $(x,f(x))$ has a local minimum at a point x_0 of the interval. Show that

$$(x_0-a)+(f(x_0)-b)f'(x_0)=0,$$

and deduce that the line from (a,b) to $(x_0,f(x_0))$ is perpendicular to the tangent there. (This line is the *normal* to the curve at that point.) [4]

(f) Show that the conclusion of part (e) can fail when its conditions are dropped, using: (i) the curve $y=|x|$ and the point $(0,-1)$; (ii) the ray $y=x$, $x\geq 0$, and the point $(-2,0)$. In each case find the closest point and state which condition fails. [3]

(g) Use the condition in part (e) to show that the number of normals to $y=x^2$ that pass through the point $(0,b)$ is three if $b>\frac{1}{2}$ and one if $b\leq\frac{1}{2}$. [3]

(h) The value $b=\frac{1}{2}$ in part (g) is the radius of curvature of $y=x^2$ at its vertex (Question 40(e)). Investigate this connection, and the curve that separates the points from which three normals to $y=x^2$ can be drawn from those from which only one can. [2]

40. Quantifying curvature. A straight line does not bend; a small circle bends sharply. This investigation builds a number that measures how sharply a curve bends at a point. It is defined by

$$\kappa=\frac{|y''|}{\left(1+(y')^2\right)^{3/2}},$$

and κ is called the *curvature*.

(a) Show that every straight line has $\kappa=0$ at every point, and explain why the definition must give this. [3]

(b) For $y=x^2$, find κ as a function of x. Show that the curvature is greatest at the vertex, and find its value there. [4]

(c) Show that $\kappa\to 0$ as $x\to\pm\infty$ for $y=x^2$, and explain what this says about the shape of a parabola far from its vertex. [3]

(d) Use implicit differentiation on $x^2+y^2=R^2$ to show that a circle of radius R has $\kappa=\dfrac{1}{R}$ at every point. Explain why a constant curvature is exactly what you should expect. [5]

(e) The *radius of curvature* is defined as $\rho=\dfrac{1}{\kappa}$. Find ρ for $y=x^2$ at the origin, and state the circle that best matches the parabola there. [3]

(f) The second derivative alone will not do as a measure of bending. For $y=x^2$, $y''=2$ at every point, yet the curve clearly bends more sharply at the vertex than far out along an arm. State what the denominator $\left(1+(y')^2\right)^{3/2}$ corrects for, and use your answer to part (b) to support it. [3]

(g) Show that $\kappa=0$ at every point of inflexion of a twice-differentiable curve. Then explain

why a curve with $\kappa = 0$ at one point need not be a straight line, and give an example. [3]

(h) The curve $y = f(x)$ is stretched vertically by a factor $k > 0$, giving $y = k\,f(x)$. At a point where $f'(x) = 0$, show that the curvature of the stretched curve is k times the curvature of the original. Show also that a translation changes κ nowhere. What does this say about judging how sharply a curve bends from a graph drawn with unequal scales on the two axes? [4]

(i) Engineers join a straight railway track to a circular arc of radius R by a *transition curve*, along which the curvature increases steadily from 0 to $\frac{1}{R}$. For the cubic $y = cx^3$ with $c > 0$, show that $\kappa \approx 6cx$ when x is small and positive, and find the value of c for which this approximation reaches $\frac{1}{R}$ at $x = L$. Then investigate the curve whose curvature is exactly proportional to the distance travelled along it (the *clothoid*, or Euler spiral), and how closely the cubic approximates it. [3]

41. Filling a vessel. Water is poured at a constant rate Q into an empty vessel whose horizontal cross-section at height h above the lowest point has area $A(h)$. The shape of the vessel decides how the water level rises. This investigation finds how, and then reverses the question: given a required pattern of rising, what shape must the vessel have?

(a) The water between heights h and $h + \delta h$ forms a thin slice. Explain why its volume is approximately $A(h)\,\delta h$, and hence why $\dfrac{dV}{dh} = A(h)$. Deduce that $\dfrac{dh}{dt} = \dfrac{Q}{A(h)}$. [3]

(b) For a cylinder of radius R, show that the level rises at a constant rate. For a sphere of radius R filled from its lowest point, show that $A(h) = \pi(2Rh - h^2)$ for $0 \le h \le 2R$, and find the height at which the level rises most slowly. [4]

(c) Show that $\dfrac{d^2h}{dt^2} = -\dfrac{Q^2 A'(h)}{A(h)^3}$. [3]

(d) Hence state and prove a rule linking the concavity of the graph of h against t to whether the vessel is widening or narrowing at the water surface. Apply it to the sphere in part (b). [3]

(e) A bowl is formed by rotating the curve $y = cx^2$ $(c > 0)$ about the y-axis. Show that $A(h) = \dfrac{\pi h}{c}$, and verify that $h = \sqrt{\dfrac{2Qct}{\pi}}$ satisfies $\dfrac{dh}{dt} = \dfrac{Q}{A(h)}$ with $h = 0$ at $t = 0$. [4]

(f) An ancient water clock empties through a small hole in its base. When the water depth is h, the water leaves at a rate proportional to $\sqrt{h}$, so that $\dfrac{dh}{dt} = -\dfrac{\lambda\sqrt{h}}{A(h)}$ for a constant $\lambda > 0$. Find how $A(h)$ must depend on h for the level to fall at a constant rate, so that equal marks on the side show equal times, and investigate the shape of such a vessel. [3]

Chapter 15

Integral Calculus

15

SYLLABUS 5.5 5.9 5.10 5.11 5.15 5.16 5.17

In Ch. 13 a speedometer reading came from distance: the speed was the rate at which distance changed. Now reverse the question. Suppose you record the speedometer for an entire journey. Can you work out how far you travelled, using only the speed?

Over a tiny slice of time, your speed barely changes, so the distance covered in that slice is roughly speed times the slice of time. Add up all the slices and you have the total distance. As the slices become narrower, the rough sum approaches the exact distance. What you are computing is the area under the speed-time graph.

This is the second main idea of calculus, and it has two parts. The first is **reversing a derivative**: given a rate of change, recover the quantity it came from. The second is **accumulating**: adding up a great many thin slices to get a total. The area under a graph is exactly this kind of total.

They have no obvious connection. Reversing a derivative is a problem in algebra. Adding slices is a problem in geometry. Yet the car journey links them, because the distance travelled is both the reverse of the velocity and the area under the velocity graph.

The chapter is built in that order. Section 15.1 does the reversal on its own, with no mention of area. Section 15.2 defines the area on its own, with no mention of derivatives, and then proves that the two are the same calculation. That result is the Fundamental Theorem of Calculus. By the end of this chapter you will be able to find antiderivatives of the standard functions, evaluate definite integrals using the Fundamental Theorem, integrate by substitution and by parts, find areas between curves and volumes of revolution, and find displacement and distance travelled from velocity.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

15.1 Antiderivatives

This chapter reverses differentiation. Given a function f, the task is to find a function whose derivative is f. Such a function is called an **antiderivative** of f.

An antiderivative is never unique. The derivative of a constant is zero, so any constant may be added to an antiderivative without altering its derivative. Antiderivatives therefore occur in families whose members differ by a constant.

DEF F is an **antiderivative** of f when $F'(x) = f(x)$. If F is one antiderivative, then $F(x) + C$ is an antiderivative for every constant C, and on an interval these are all of them.

Naming the whole family in words each time is long, so it is given a symbol of its own.

KEY

$$\int f(x)\,dx = F(x) + C \qquad \text{means} \qquad F'(x) = f(x).$$

This is the **indefinite integral** of f, and C is the **constant of integration**. It is easy to forget, and an answer written without it names a single antiderivative rather than the whole family.

Reversing the power rule means raising the power by one and dividing by the new power.

KEY · IB

$$\int x^n\,dx = \frac{x^{n+1}}{n+1} + C, \qquad n \neq -1$$

The exception $n = -1$ is handled in §15.3. Constant multiples and sums are reversed term by term, just as they are differentiated term by term.

Worked through. To find $\int (3x^2 - 4x + 1)\,dx$, reverse each term, raising each power by one:

$$\int (3x^2 - 4x + 1)\,dx = x^3 - 2x^2 + x + C.$$

Check by differentiating: $\frac{d}{dx}(x^3 - 2x^2 + x + C) = 3x^2 - 4x + 1$. ✓ Differentiating the answer verifies it completely.

15.1.1 Boundary conditions

The $+C$ means that an antiderivative is not determined until further information is supplied. A single condition, such as a point through which the curve passes, fixes C and selects one member of the family.

EXAMPLE 15.1

A curve has gradient $\dfrac{dy}{dx} = 6x^2 - 2$ and passes through $(1, 4)$. Find its equation.

Antidifferentiate to recover y:

$$y = 2x^3 - 2x + C.$$

Use the point $(1, 4)$: $4 = 2 - 2 + C \implies C = 4$. So $y = 2x^3 - 2x + 4$.

YOUR TURN 15.1 Antiderivatives

Integrating term by term

1. Find each indefinite integral.

(a) $\displaystyle\int (4x^3 - 6x)\,dx$

(b) $\displaystyle\int (x^2 + 3x - 5)\,dx$

(c) $\displaystyle\int \frac{6}{x^3}\,dx$

(d) $\displaystyle\int \sqrt{x}\,dx$

(e) $\displaystyle\int \left(x^2 - \frac{1}{x^2}\right) dx$

(f) $\displaystyle\int \frac{x^3 - 2x}{x}\,dx$

(g) $\displaystyle\int x^{-1/2}\,dx$

(h) $\displaystyle\int (2x)^3\,dx$

Using a condition to find the constant

2.

(a) A curve has $\dfrac{dy}{dx} = 4x - 3$ and passes through $(2, 5)$. Find y.

(b) A curve has $\dfrac{dy}{dx} = 3\sqrt{x}$ and passes through $(4, 20)$. Find y.

(c) A curve has $\dfrac{d^2y}{dx^2} = 6x$, with $\dfrac{dy}{dx} = 1$ and $y = 2$ when $x = 0$. Find y. *(Hint: integrate twice, and use one condition after each integration)*

(d) A curve has $\dfrac{d^2y}{dx^2} = 12x^2 - 4$. Its gradient is 0 at the point $(1, 3)$. Find y.

(e) Explain why a gradient function alone does not determine a curve, and what extra information does.

(f) Explain why parts (c) and (d) needed two pieces of information rather than one.

Concept check

3. A student writes $\displaystyle\int \frac{1}{x^2}\,dx = \frac{1}{x^3/3} + C = \frac{3}{x^3} + C$. Identify the error, and find the correct integral.

15.2 Definite Integrals and Area

Set antiderivatives aside for the moment. This section takes up a second and apparently unrelated problem; the two are connected only later.

The problem is area. Regions with straight edges, and a few curved shapes such as the circle, have area formulas. A region under a general curve has none. The available approach is approximation.

Cut the interval from a to b into n strips, each of width Δx. Replace each strip by a rectangle whose height is the value of the function somewhere in that strip. Every rectangle has area height $\times$ width, so the total is

$$\sum_{i=1}^{n} f(x_i)\,\Delta x.$$

This is not the area. It is a sum of rectangles, and it differs from the area by whatever those rectangles fail to cover or cover in excess. Narrower strips leave less room for that error, and as n increases the sum approaches a single value.

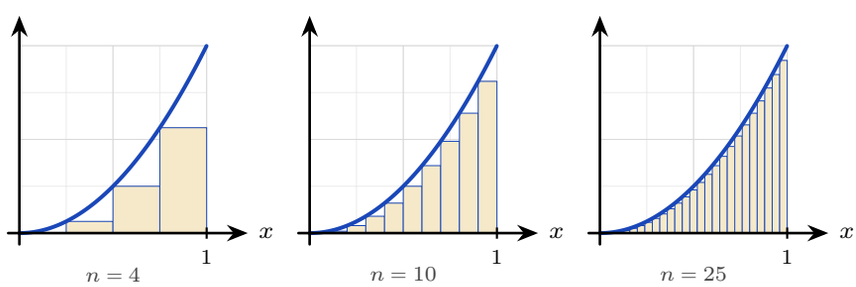


Figure 15.1 Rectangles under $y = x^2$ on $[0, 1]$, taking the height at the left of each strip. Four strips give 0.219, ten give 0.285 and twenty-five give 0.314; the true area is $\frac{1}{3}$. The missing pieces above each rectangle shrink as the strips narrow.

The definition allows the height to be read anywhere in the strip, but two choices are natural: the value at the left end, or the value at the right. Taking both is what makes the argument work, because one sum is too small and the other too large.

For a curve that rises across the interval, as $y = x^2$ does on $[0, 1]$, every left-hand rectangle falls short of the curve and every right-hand one overshoots it. The first sum therefore underestimates the area and the second overestimates it, so the true value lies between them. For a curve that falls across the interval the two exchange roles, and the bracketing still holds.

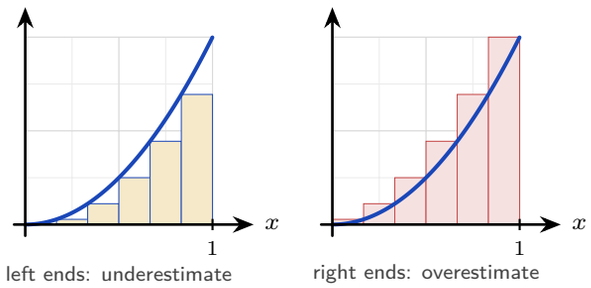


Figure 15.2 The same six strips, with the height of each rectangle read at the left of its strip and then at the right. On a rising curve the left-hand rectangles all sit below it and the right-hand ones all rise above it, so the area lies between the two sums. Narrowing the strips closes the gap from both sides at once.

The numbers behave as the picture predicts.

n	4	10	100	1000
low estimate	0.219	0.285	0.328	0.3328
high estimate	0.469	0.385	0.338	0.3338

The two rows converge, and the value they enclose is $\frac{1}{3}$. That limit is what is meant by the area, and it is called the **definite integral**.

DEF The **definite integral** of f from a to b is the limit of the rectangle sums as the strips become arbitrarily narrow:

$$\int_a^b f(x)\,dx = \lim_{\Delta x \to 0} \sum_i f(x_i)\,\Delta x.$$

For $f(x) \geq 0$ it is the area between the curve and the x-axis from a to b.

The notation can now be read directly. The elongated S denotes a sum, $f(x)\,dx$ represents a single rectangle of height $f(x)$ and very small width dx, and the integral accumulates them.

As it stands, however, the definition gives no practical way to calculate an area. Nothing in it indicates how such a limit is to be evaluated, and summing a thousand rectangles by hand is not a practical method.

The **Fundamental Theorem of Calculus** supplies the missing method. It states that this limit of sums can be evaluated by antidifferentiation, so the two ideas of the chapter are in fact one calculation.

KEY **The Fundamental Theorem of Calculus.** If f is continuous on $[a, b]$ and F is any antiderivative of f, then

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

Reversing a derivative and adding up rectangles give the same answer. Test it on the area the table was closing in on. An antiderivative of x^2 is $\frac{x^3}{3}$, so

$$\int_0^1 x^2\,dx = \Big[\tfrac{x^3}{3}\Big]_0^1 = \tfrac{1}{3} - 0 = \tfrac{1}{3},$$

the value the thousand rectangles were approaching, found in one line. The constant C cancels in the subtraction, which is why a definite integral is a plain number while an indefinite one carries a $+C$.

The theorem can also be read the other way round. Write the integrand as a derivative, $f = F'$:

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

Adding up a rate of change across an interval gives the total change over that interval: the area under the graph of F' equals the rise of the graph of F. This is the reading used in §15.9, where the area under a velocity graph is a displacement.

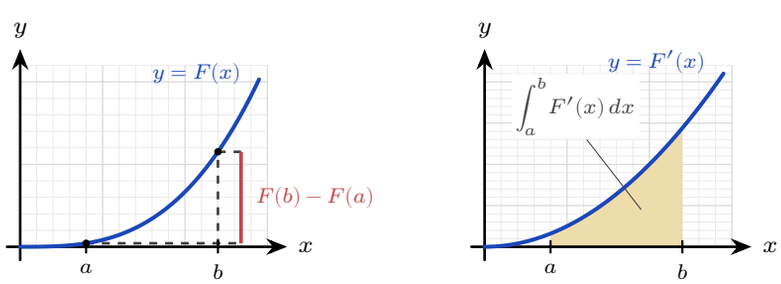


Figure 15.3 Left, $F(x) = \frac{x^3}{3}$ rises by $F(b) - F(a)$ between $x = a$ and $x = b$. Right, its derivative $F'(x) = x^2$, with the area over the same interval shaded. The rise on the left and the area on the right are the same number: for $a = 0.4$ and $b = 1.2$ both are 0.555 to 3 s.f.

The Fundamental Theorem should not be accepted on trust. Its proof uses only the definition of the derivative and one picture of a thin strip, and it shows exactly how the area under a curve and the rate of change are connected. It is given at the end of this chapter, after the Reflections. Read it once you can use the theorem: the proof is short, and it explains why the two halves of this chapter are one idea.

KEY · IB

$$A = \int_a^b y\,dx \qquad (y \geq 0)$$

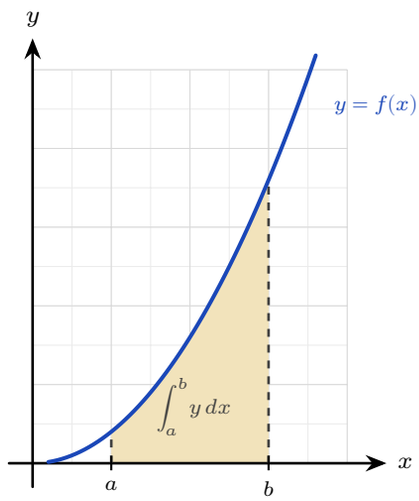


Figure 15.4 For $y \geq 0$ the definite integral measures the shaded area between the curve and the x-axis, from $x = a$ to $x = b$.

EXAMPLE 15.2

Evaluate $\int_0^2 3x^2\,dx$.

$$\int_0^2 3x^2\,dx = \Big[x^3\Big]_0^2 = 8 - 0 = 8.$$

TIP Your GDC evaluates definite integrals directly, and on a calculator paper that is often all a question requires. Even when analytic working is demanded, the GDC value is a free check of your antiderivative.

15.2.1 Areas below the axis

Where the curve goes below the x-axis, the integral is negative: it is a signed area. For a true geometric area you integrate the parts separately and take the absolute value of any negative piece.

KEY · IB

$$A = \int_a^b |y|\,dx$$

Without a calculator, find where $y = 0$ between a and b, integrate over each piece separately,

and add the absolute values of the results. With a GDC, evaluate $\int_a^b |y|\,dx$ directly.

The distinction is visible as soon as a curve crosses the axis more than once.

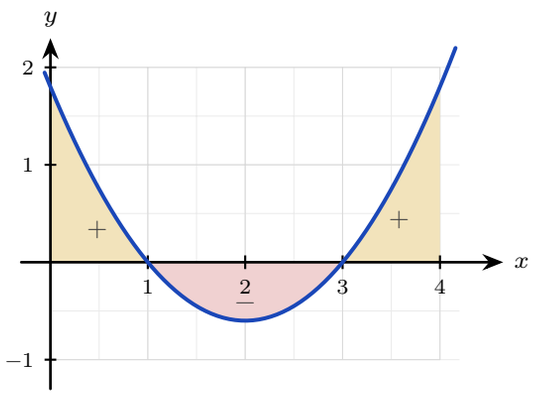


Figure 15.5 A definite integral accumulates *signed* area. The two gold regions lie above the axis and count positively; the scarlet region lies below and counts negatively. A single integral across the whole interval returns the sum of the three signed pieces, which is smaller than the total area enclosed.

EXAMPLE 15.3

Find the area enclosed between $y = x^2 - 4$ and the x-axis from $x = 0$ to $x = 2$.

On $[0, 2]$ the curve lies below the axis, so the integral is negative:

$$\int_0^2 (x^2 - 4)\,dx = \left[\tfrac{x^3}{3} - 4x\right]_0^2 = \tfrac{8}{3} - 8 = -\tfrac{16}{3}.$$

The area is the magnitude, $\dfrac{16}{3}$. An area is never negative, so $-\frac{16}{3}$ answers a different question from the one asked.

EXAMPLE 15.4

Find the area enclosed between $y = x^2 - 1$ and the x-axis from $x = 0$ to $x = 2$.

The curve crosses the axis at $x = 1$: it is below the axis on $(0, 1)$ and above it on $(1, 2)$. Split there:

$$\int_0^1 (x^2 - 1)\,dx = \left[\tfrac{x^3}{3} - x\right]_0^1 = -\tfrac{2}{3}, \qquad \int_1^2 (x^2 - 1)\,dx = \left[\tfrac{x^3}{3} - x\right]_1^2 = \tfrac{4}{3}.$$

The area is $\left|-\frac{2}{3}\right| + \frac{4}{3} = 2$. A single integral from 0 to 2 gives $-\frac{2}{3} + \frac{4}{3} = \frac{2}{3}$, because the piece below the axis cancels part of the piece above.

CAUTION Before finding an area, check where the curve crosses the x-axis. A curve that goes above and below will have positive and negative parts that partly cancel in a single integral, giving the wrong area. Split at the roots.

YOUR TURN 15.2 Definite Integrals and Area

Evaluating definite integrals

4. Evaluate each definite integral.

(a) $\displaystyle\int_1^2 4x\,dx$

(b) $\displaystyle\int_0^3 (x^2+1)\,dx$

(c) $\displaystyle\int_1^2 \frac{6}{x^3}\,dx$

(d) $\displaystyle\int_1^4 \frac{1}{\sqrt{x}}\,dx$

(e) $\displaystyle\int_0^2 (x+1)(x-3)\,dx$

(f) $\displaystyle\int_{-1}^2 (3x^2-2x)\,dx$

Areas and the x-axis

5. Each region lies on one side of the x-axis. Find its area.

(a) The region under $y=x^2$ between $x=1$ and $x=3$.

(b) The region enclosed between $y=x^2-1$ and the x-axis.

6. Each curve crosses the x-axis inside the interval. Find the total area between the curve and the x-axis over the interval given.

(a) $y=x^3$, from $x=-2$ to $x=2$.

(b) $y=x^2-x-2$, from $x=0$ to $x=3$. State also the value of $\displaystyle\int_0^3 (x^2-x-2)\,dx$, and explain why it is not the area.

(c) $y=x^3-x$, from $x=0$ to $x=2$.

Estimating an area with rectangles

7. The area under $y=x^2$ from $x=0$ to $x=2$ is to be estimated by four rectangles of equal width.

(a) Taking the height of each rectangle at the left of its strip, show that the estimate is 1.75.

(b) Taking the height at the right instead, show that the estimate is 3.75.

(c) Find the exact area, and confirm that it lies between the two estimates.

(d) Repeat part (a) with eight rectangles, and show that the left-hand estimate rises to 2.1875.

(e) State what happens to the two estimates as the number of rectangles increases.

Concept check

8. A student writes

$$\int_1^3 \frac{2}{x^3}\,dx = \left[-\frac{1}{x^2}\right]_1^3 = -\frac{1}{9} - 1 = -\frac{10}{9}.$$

Explain why a negative answer cannot be right here, identify the error, and find the correct value.

15.3 Standard Integrals

Every derivative established in Ch. 13 can be read in reverse. Doing so converts the table of derivatives into a table of integrals, and turns the question "what is the integral of this function?" into the easier question "what was differentiated to produce it?"

KEY · IB

$$\int \frac{1}{x}\,dx = \ln|x| + C, \qquad \int \sin x\,dx = -\cos x + C$$

$$\int \cos x\,dx = \sin x + C, \qquad \int e^x\,dx = e^x + C$$

None of these is a new fact. Each is a derivative already known, read from right to left, and each should be understood rather than memorised.

- $\int \frac{1}{x}\,dx = \ln|x| + C$ fills the one gap the power rule leaves. Raising $n = -1$ by one gives 0, and the rule then asks you to divide by that new power, which is impossible. So the antiderivative of x^{-1} cannot be a power at all. The logarithm supplies it, because $\frac{d}{dx}\ln x = \frac{1}{x}$. The modulus is there because $\frac{1}{x}$ is defined for negative x as well, and on that side $\frac{d}{dx}\ln(-x) = \frac{-1}{-x} = \frac{1}{x}$ by the chain rule: the same derivative on both sides of the origin.
- $\int e^x\,dx = e^x + C$ because e^x is equal to its own derivative. A function that is unchanged by differentiating is unchanged by antidifferentiating.
- $\int \cos x\,dx = \sin x + C$ is the direct reversal of $\frac{d}{dx}\sin x = \cos x$, with no sign to watch.
- $\int \sin x\,dx = -\cos x + C$ carries a minus sign that is easily dropped. Differentiating cosine introduces one: $\frac{d}{dx}\cos x = -\sin x$. To finish with $+\sin x$ you must therefore start from $-\cos x$.

The two trigonometric cases are easiest to remember as a cycle. Differentiating moves one step clockwise; integrating moves one step back.

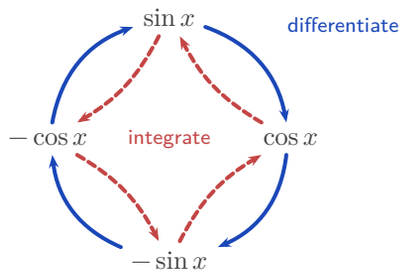


Figure 15.6 The four functions form a closed cycle. Differentiating advances one step clockwise (solid), integrating retreats one step anticlockwise (dashed). Reading the cycle backwards from $\sin x$ lands on $-\cos x$, which is where the minus sign in $\int \sin x\,dx$ comes from.

At HL, three more standard integrals are provided, the reverses of the exponential and inverse-trigonometric derivatives.

KEY · IB HL

$$\int a^x\,dx = \frac{a^x}{\ln a} + C, \qquad \int \frac{1}{a^2+x^2}\,dx = \frac{1}{a}\arctan\frac{x}{a} + C$$

$$\int \frac{1}{\sqrt{a^2-x^2}}\,dx = \arcsin\frac{x}{a} + C, \qquad |x| < a$$

These three reward the same reading. Differentiating a^x produces an extra factor: writing $a^x = e^{x\ln a}$ and applying the chain rule gives $\frac{d}{dx}a^x = a^x \ln a$. Dividing by $\ln a$ removes it. When $a = e$ the factor $\ln e$ is 1 and the formula becomes $\int e^x dx = e^x + C$, as it must.

The two inverse-trigonometric forms come from $\frac{d}{dx}\arctan x = \frac{1}{1+x^2}$ and $\frac{d}{dx}\arcsin x = \frac{1}{\sqrt{1-x^2}}$, with a acting as a scale. Their asymmetry is a common source of error: the arctangent form carries a factor $\frac{1}{a}$ and the arcsine form does not. In each case differentiating x/a contributes a chain-rule factor $\frac{1}{a}$. For the arcsine, $\frac{d}{dx}\arcsin\frac{x}{a} = \frac{1}{\sqrt{a^2-x^2}}$: the $\sqrt{a^2} = a$ that comes out of the square root cancels the $\frac{1}{a}$ exactly. For the arctangent, $\frac{d}{dx}\arctan\frac{x}{a} = \frac{a}{a^2+x^2}$: the denominator contributes a^2, which more than cancels the $\frac{1}{a}$ and leaves a factor a, so the formula needs $\frac{1}{a}$ in front to remove it.

Reading the rest of the HL derivative table backwards gives four more. They must be learned, because the formula booklet gives only the derivatives:

KEY HL

$$\int \sec^2 x\,dx = \tan x + C, \qquad \int \sec x\tan x\,dx = \sec x + C,$$

$$\int \operatorname{cosec}^2 x\,dx = -\cot x + C, \qquad \int \operatorname{cosec} x\cot x\,dx = -\operatorname{cosec} x + C$$

EXAMPLE 15.5

Find (a) $\displaystyle\int \frac{1}{x^2+25}\,dx$, (b) $\displaystyle\int \frac{1}{\sqrt{9-x^2}}\,dx$.

Match the standard forms with $a=5$ and $a=3$:

$$\text{(a) } \tfrac{1}{5}\arctan\frac{x}{5}+C, \qquad \text{(b) } \arcsin\frac{x}{3}+C.$$

The only work is recognising a^2 as 25 and as 9. After that the standard form gives the answer.

Definite integrals of these standard forms are evaluated exactly as in §15.2: find the antiderivative, then substitute the limits.

Worked through. The antiderivative of $\sin x$ is $-\cos x$, so

$$\int_0^{\pi/2} \sin x\,dx = \Big[-\cos x\Big]_0^{\pi/2} = -\cos\frac{\pi}{2}+\cos 0 = 0+1 = 1.$$

The area under one quarter-wave of the sine curve is exactly 1, with no π in the answer. In the same way, the antiderivative of $\dfrac{1}{x}$ is $\ln|x|$, and when both limits are positive

$$\int_1^{e^2} \frac{1}{x}\,dx = \Big[\ln x\Big]_1^{e^2} = \ln e^2 - \ln 1 = 2-0 = 2.$$

Stopping at $x=e$ instead would give an area of exactly one. This gives one way to define e: it is the upper limit that makes the area under $y=\frac{1}{x}$, starting at $x=1$, equal to one.

15.3.1 Linear composites

Sometimes the function inside is linear, as in e^{2x} or $\cos 3x$. Integrate as usual, then divide by the coefficient of x inside. This reverses the chain rule, whose extra factor was that same coefficient.

KEY

$$\int f(ax+b)\,dx = \frac{1}{a}F(ax+b)+C, \qquad \text{where } F'=f$$

EXAMPLE 15.6

Find (a) $\int e^{2x}\,dx$, (b) $\int (2x+1)^5\,dx$, (c) $\int \cos 3x\,dx$.

Each has a linear function inside, so integrate and then divide by the coefficient of x:

$$\text{(a) } \tfrac{1}{2}e^{2x}+C, \qquad \text{(b) } \frac{(2x+1)^6}{12}+C, \qquad \text{(c) } \tfrac{1}{3}\sin 3x+C.$$

In (b), reversing the power rule gives $\frac{(2x+1)^6}{6}$, then dividing by the coefficient 2 gives the 12.

The linear-composite rule extends to the whole table. For instance, $\int \sec^2(2x+5)\,dx = \frac{1}{2}\tan(2x+5)+C$: integrate as usual, then divide by the inner coefficient.

YOUR TURN 15.3 Standard Integrals

Reversing the standard derivatives

9. Find each indefinite integral.

(a) $\int (2\cos x - \sin x)\,dx$

(b) $\int \left(3e^x + \dfrac{1}{x}\right)dx$

Linear composites

10. Find each indefinite integral.

(a) $\int e^{3x}\,dx$

(b) $\int (3x-2)^4\,dx$

(c) $\int \sin 2x\,dx$

(d) $\int \sec^2(4x)\,dx$

(e) $\int \cos(3x+1)\,dx$

(f) $\int \dfrac{1}{2x+5}\,dx$

The HL standard forms

11. Find each indefinite integral, matching it to a standard form.

(a) $\int \dfrac{1}{x^2+9}\,dx$

(b) $\int \dfrac{1}{\sqrt{16-x^2}}\,dx$

(c) $\int 5^x\,dx$

(d) $\int \dfrac{2}{4+x^2}\,dx$

(e) $\int \dfrac{1}{\sqrt{9-4x^2}}\,dx$

12. Find each indefinite integral by reading a derivative in reverse.

(a) $\int \operatorname{cosec}^2 x\,dx$

(b) $\int \sec x\tan x\,dx$

(c) $\int 3\operatorname{cosec} 2x\cot 2x\,dx$

Definite integrals of standard forms

13.

(a) Evaluate $\displaystyle\int_1^{e^3} \frac{2}{x}\,dx$.

(b) Evaluate $\displaystyle\int_0^{\pi/6} \cos x\,dx$.

(c) Find the area under $y = \sin x$ from $x = 0$ to $x = \pi$.

(d) Evaluate $\displaystyle\int_0^{2\pi} \sin x\,dx$, and explain why the answer is not the area between the curve and the axis over that interval.

Concept check

14. A student writes $\displaystyle\int e^{x^2}\,dx = \frac{e^{x^2}}{2x} + C$, dividing by the derivative of the inner function as for $\displaystyle\int e^{3x}\,dx = \tfrac{1}{3}e^{3x} + C$. Show, by differentiating, that the answer is wrong, and explain why the method works for e^{3x} but not for e^{x^2}.

15.4 Rewriting Before Integrating HL

Many integrands do not match a standard form as they stand, but can be rewritten until they do. Four rewritings cover the cases met in this course: divide, complete the square, split into partial fractions, or replace a squared trigonometric function using a double angle. For the first three, deciding between them takes one look at the denominator. The table in §15.6.2 summarises how to choose between all the methods of this chapter.

15.4.1 Divide first

When the numerator has degree at least that of the denominator, the fraction is **improper** and no standard form applies. Divide before anything else. For instance $\dfrac{x^2}{x+1} = x - 1 + \dfrac{1}{x+1}$, so the integral is $\frac{x^2}{2} - x + \ln|x+1| + C$. Only when the numerator has the lower degree are the next two methods available.

15.4.2 Completing the square HL

A denominator does not always arrive in the form $a^2 + x^2$. A quadratic such as $x^2 + 2x + 5$ has that shape, but only after you complete the square.

EXAMPLE 15.7

Find $\displaystyle\int \frac{1}{x^2+2x+5}\,dx$.
Complete the square: $x^2+2x+5=(x+1)^2+4$. This is the arctangent shape with $a=2$ and the shifted variable $x+1$:

$$\int \frac{dx}{(x+1)^2+4} = \frac{1}{2}\arctan\left(\frac{x+1}{2}\right)+C.$$

Any quadratic x^2+bx+c with no real roots can be written in the form a^2+u^2 this way. The shift from x to $x+1$ does not change the integral, because the derivative of $x+1$ is 1.

15.4.3 Partial fractions HL

When the quadratic *does* factorise, use a different method. Split the fraction into partial fractions, as in Ch. 6. Each piece then has the form $\frac{1}{x+k}$, and integrates to a logarithm.

EXAMPLE 15.8
Find $\displaystyle\int \frac{1}{x^2+3x+2}\,dx$.
Factor and decompose: $\dfrac{1}{(x+1)(x+2)} = \dfrac{1}{x+1}-\dfrac{1}{x+2}$. Then

$$\int\left(\frac{1}{x+1}-\frac{1}{x+2}\right)dx = \ln|x+1|-\ln|x+2|+C = \ln\left|\frac{x+1}{x+2}\right|+C.$$

So the denominator decides the method. If it factorises, use partial fractions, which lead to logarithms. If it has no real roots, complete the square, which leads to an arctangent.

15.4.4 Using a double angle

There is no standard integral for $\sin^2 x$ or $\cos^2 x$, and neither is the derivative of anything simple. The method is to remove the square before integrating, using the double-angle identity from Ch. 8:

$$\cos 2x = 1-2\sin^2 x = 2\cos^2 x-1.$$

Rearranged, these give the two rewritings that matter here:

KEY

$$\sin^2 x = \frac{1-\cos 2x}{2}, \qquad \cos^2 x = \frac{1+\cos 2x}{2}$$

Each right-hand side is a constant plus a linear composite, and both of those integrate immediately.

EXAMPLE 15.9

Find (a) $\int \sin^2 x\,dx$, (b) $\int_0^{\pi} \sin^2 x\,dx$.

(a) Replace the square first:

$$\int \sin^2 x\,dx = \int \frac{1-\cos 2x}{2}\,dx = \frac{1}{2}\int (1-\cos 2x)\,dx = \frac{x}{2} - \frac{\sin 2x}{4} + C.$$

The $\frac{1}{4}$ comes from the linear composite: integrating $\cos 2x$ gives $\frac{1}{2}\sin 2x$, and the outer $\frac{1}{2}$ halves it again.

(b) Now use the limits:

$$\int_0^{\pi} \sin^2 x\,dx = \left[\frac{x}{2} - \frac{\sin 2x}{4}\right]_0^{\pi} = \left(\frac{\pi}{2} - 0\right) - (0-0) = \frac{\pi}{2}.$$

The answer comes out cleanly, and there is a reason for that. Over a full arch, $\sin^2 x$ averages exactly $\frac{1}{2}$, so the area is half the width π. The oscillating part, $\sin 2x$, contributes nothing over a whole number of half-periods.

The same rewriting handles $\cos^2 x$, and it is the standard way to integrate the square of a sine or cosine. Expect it when a volume of revolution involves a sine or cosine curve, since the volume formula squares the function.

YOUR TURN 15.4 Rewriting Before Integrating HL

Dividing first

15. Rewrite each integrand before integrating.

(a) $\int \dfrac{x^2+3x}{x}\,dx$

(b) $\int \dfrac{2x+1}{x}\,dx$

(c) $\int \dfrac{x^3-1}{x^2}\,dx$

(d) $\int \dfrac{2x+3}{x+1}\,dx$

(e) $\int \dfrac{x^2+1}{x+1}\,dx$

Partial fractions

16.

(a) Find $\int \dfrac{1}{x^2-9}\,dx$ using partial fractions.

(b) Find $\int \dfrac{4x+5}{(x+1)(x+2)}\,dx$.

Completing the square

17.

(a) Find $\displaystyle\int \frac{1}{x^2+4x+13}\,dx$.

(b) Find $\displaystyle\int \frac{1}{2x^2+4x+10}\,dx$.

(c) Explain how you decide, from the denominator alone, whether to use partial fractions or to complete the square.

Using a double angle

18.

(a) Find $\displaystyle\int \cos^2 x\,dx$.

(b) Find $\displaystyle\int \sin^2 3x\,dx$.

(c) Evaluate $\displaystyle\int_0^{\pi/6} \sin^2 x\,dx$, giving an exact answer.

Concept check

19. A student writes $\displaystyle\int \sin^2 x\,dx = \frac{\sin^3 x}{3} + C$. Identify the error, and find the correct integral.

20. A student writes $\displaystyle\int \frac{1}{x^2+4}\,dx = \ln(x^2+4) + C$. Identify the error, and find the correct integral.

15.5 Integration by Substitution

Substitution reverses the chain rule. Look for a composite function whose inner function also appears, up to a constant multiple, as a factor outside. Put u equal to that inner function, and a hard integral in x becomes an easy one in u.

The method has four steps.

1. Let u be the inner function.
2. Replace dx using $du = \dfrac{du}{dx}\,dx$, so that no x remains.
3. Integrate with respect to u.
4. Substitute back to write the answer in terms of x.

Worked through. In $\int 2x(x^2+1)^3\,dx$, let $u = x^2+1$, so $\dfrac{du}{dx} = 2x$, giving $du = 2x\,dx$. The integral becomes

$$\int u^3\,du = \frac{u^4}{4} + C = \frac{(x^2+1)^4}{4} + C.$$

The factor $2x$ outside is exactly the $\frac{du}{dx}$ required. Without it the substitution would leave an x behind and fail.

EXAMPLE 15.10

Find $\displaystyle\int \frac{2x}{x^2+1}\,dx$.

With $u = x^2+1$, $du = 2x\,dx$:

$$\int \frac{1}{u}\,du = \ln|u| + C = \ln(x^2+1) + C.$$

The numerator is the derivative of the denominator, and that is why the substitution works.

The same substitution, $u = f(x)$, works for every fraction whose numerator is the derivative of its denominator:

KEY

$$\int \frac{f'(x)}{f(x)}\,dx = \ln|f(x)| + C$$

CAUTION For a definite integral by substitution, either convert the limits to u-values, or substitute back to x before applying the original limits. Applying the x-limits to a u-expression gives a wrong answer.

EXAMPLE 15.11

Evaluate $\displaystyle\int_0^{\sqrt{3}} x\sqrt{x^2+1}\,dx$.

Let $u = x^2+1$, so $du = 2x\,dx$. Convert the limits too: $x = 0 \Rightarrow u = 1$ and $x = \sqrt{3} \Rightarrow u = 4$.

Then

$$\int_0^{\sqrt{3}} x\sqrt{x^2+1}\,dx = \frac{1}{2}\int_1^4 \sqrt{u}\,du = \frac{1}{2}\cdot\frac{2}{3}\Big[u^{3/2}\Big]_1^4 = \frac{1}{3}(8-1) = \frac{7}{3}.$$

Once the limits are converted, you never return to x at all; the whole calculation finishes in u.

Substitution is also the method for many exponential and trigonometric integrands that match no standard form.

EXAMPLE 15.12

Find (a) $\displaystyle\int xe^{-x^2}\,dx$, (b) $\displaystyle\int \sin^4 x\cos x\,dx$.

(a) The inner function is $-x^2$, and its derivative $-2x$ appears outside up to the constant -2.

Let $u = -x^2$, so $du = -2x\,dx$ and $x\,dx = -\frac{1}{2}\,du$:

$$\int xe^{-x^2}\,dx = -\frac{1}{2}\int e^u\,du = -\frac{1}{2}e^u + C = -\frac{1}{2}e^{-x^2} + C.$$

On its own, e^{-x^2} cannot be integrated in terms of the functions of this course. The factor x is what makes the integral possible, and without it these methods cannot find the integral.

(b) Here the inner function is $\sin x$, and its derivative $\cos x$ is sitting outside. Let $u = \sin x$, so $du = \cos x\,dx$:

$$\int \sin^4 x \cos x\,dx = \int u^4\,du = \frac{u^5}{5} + C = \frac{\sin^5 x}{5} + C.$$

In both parts the test is the same: is the derivative of the inner function present as a factor? If it is, substitution usually works. If it is not, another method is usually quicker.

YOUR TURN 15.5 Integration by Substitution

Indefinite integrals by substitution

21. Find each integral by substitution.

(a) $\displaystyle\int 3x^2(x^3+1)^4\,dx$

(b) $\displaystyle\int xe^{x^2}\,dx$

(c) $\displaystyle\int 2x\sqrt{x^2+1}\,dx$

(d) $\displaystyle\int \frac{x}{(x^2+4)^2}\,dx$

(e) $\displaystyle\int \frac{\ln x}{x}\,dx$

(f) $\displaystyle\int \sin^3 x \cos x\,dx$

Integrals of the form $\dfrac{f'(x)}{f(x)}$

22.

(a) Find $\displaystyle\int \frac{3x^2+2}{x^3+2x}\,dx$.

(b) Find $\displaystyle\int \frac{e^x}{e^x+3}\,dx$.

(c) By writing $\tan x = \dfrac{\sin x}{\cos x}$, show that $\displaystyle\int \tan x\,dx = -\ln|\cos x| + C$.

(d) Evaluate $\displaystyle\int_0^1 \frac{x+1}{x^2+2x+3}\,dx$, giving an exact answer.

Definite integrals by substitution

23.

(a) Evaluate $\displaystyle\int_0^2 x^2\sqrt{x^3+1}\,dx$ by converting the limits.

(b) Evaluate the same integral by substituting back to x first, and confirm that the answers agree.

(c) Evaluate $\displaystyle\int_0^{\pi/2} \sin^2 x\cos x\,dx$ by converting the limits.

Concept check

24. A student evaluates $\displaystyle\int_0^1 2x\left(x^2+1\right)^3\,dx$ with $u=x^2+1$, and writes

$$\int_0^1 u^3\,du = \left[\frac{u^4}{4}\right]_0^1 = \frac{1}{4}.$$

Identify the error, and find the correct value.

15.6 Integration by Parts HL

Substitution reverses the chain rule; **integration by parts** reverses the product rule. It handles products that substitution cannot, such as xe^x or $\ln x$.

KEY · IB HL

$$\int u\,\frac{dv}{dx}\,dx = uv - \int v\,\frac{du}{dx}\,dx$$

This is the form printed in the formula booklet. It names the factor to be integrated dv/dx, which is accurate but slow to use. It is quicker to call the two factors u and v directly, with v the one you integrate. Replace dv/dx by v in the booklet formula: its v then becomes $\int v\,dx$, and du/dx is u'. The formula becomes the working form used in this book.

KEY HL

$$\int u\,v\,dx = u\int v\,dx \;-\; \int\left(\int v\,dx\right)u'\,dx$$

The first function times the integral of the second, minus the integral of that integral times the derivative of the first.

The first factor, u, is copied once and differentiated once. The second, v, is only ever integrated, and the same $\int v\,dx$ appears twice. Work it out once and write it in both places.

Read from left to right, the formula says: multiply u by the integral of v, then subtract the

integral of that same $\int v\,dx$ times the derivative of u.

Everything now depends on which factor you call u. Choose the one that becomes simpler when differentiated, so that the second integral is easier than the one you started with. The order **ILATE** is a reliable guide: whichever type appears first in this list should be u.

I **Inverse trigonometric**: $\arcsin x$, $\arctan x$

L **Logarithmic**: $\ln x$

A **Algebraic**: x, x^2, $3x-1$

T **Trigonometric**: $\sin x$, $\cos x$

E **Exponential**: e^x, 2^x

In $\int xe^x\,dx$ the factors are algebraic and exponential, so A comes before E and $u = x$. In $\int x\ln x\,dx$ they are logarithmic and algebraic, so L comes before A and $u = \ln x$. The order works because differentiating reduces the first three types to something simpler, while trigonometric and exponential functions repeat themselves however often you differentiate them.

TIP ILATE is a guide, not a theorem. It gives the right choice for most integrals in this course, but the test that always applies is the one behind it: does the second integral end up easier than the first?

A single function that has no standard integral, such as $\ln x$ or $\arctan x$, can still be integrated by parts. Write it as the function times 1: take u as the function and $v = 1$, so $\int v\,dx = x$.

Worked through. In $\int xe^x\,dx$, let $u = x$ (so $u' = 1$) and $v = e^x$ (so $\int v\,dx = e^x$):

$$\int xe^x\,dx = xe^x - \int e^x \cdot 1\,dx = xe^x - e^x + C = e^x(x-1) + C.$$

Choosing $u = x$ is what makes the new integral, $\int e^x\,dx$, trivial. The other choice, $u = e^x$, would give $\int \frac{x^2}{2}e^x\,dx$, which has a higher power of x.

EXAMPLE 15.13

Find (a) $\displaystyle\int (x+1)\sin x\,dx$, (b) $\displaystyle\int_0^{\pi/2} (x+1)\cos x\,dx$.

(a) The factors are algebraic and trigonometric, so ILATE gives $u = x+1$ and $v = \sin x$, hence $\int v\,dx = -\cos x$ and $u' = 1$:

$$\int (x+1)\sin x\,dx = -(x+1)\cos x - \int (-\cos x)\,dx = -(x+1)\cos x + \sin x + C.$$

Watch the two minus signs. Integrating $\sin x$ gives $-\cos x$, and the formula then subtracts the integral of that, which turns the sign back.

(b) Same choice, now with $v = \cos x$ and $\int v\,dx = \sin x$:

$$\int (x+1)\cos x\,dx = (x+1)\sin x - \int \sin x\,dx = (x+1)\sin x + \cos x + C.$$

Now apply the limits:

$$\int_0^{\pi/2} (x+1)\cos x\,dx = \Big[(x+1)\sin x + \cos x\Big]_0^{\pi/2} = \left(\frac{\pi}{2} + 1 + 0\right) - (0+1) = \frac{\pi}{2}.$$

For a definite integral by parts, find the whole antiderivative first, then substitute the limits once at the end.

EXAMPLE 15.14

Find $\int \ln x\,dx$.

There is only one function, but write $u = \ln x$ and $v = 1$, so $u' = \dfrac{1}{x}$ and $\int v\,dx = x$:

$$\int \ln x\,dx = x\ln x - \int x \cdot \frac{1}{x}\,dx = x\ln x - x + C.$$

The same choice, $v = 1$, integrates $\arctan x$.

15.6.1 Repeated integration by parts

If one application of parts leaves an integral that still needs parts, apply it again. Each round of differentiation reduces the power of x by one.

EXAMPLE 15.15

Find $\int x^2 e^{-x}\,dx$.

First round, $u = x^2$ and $v = e^{-x}$, so $\int v\,dx = -e^{-x}$ and $u' = 2x$:

$$\int x^2 e^{-x}\,dx = -x^2 e^{-x} + 2\int x e^{-x}\,dx.$$

The remaining integral needs parts a second time, with $u = x$:

$$\int x e^{-x}\,dx = -x e^{-x} + \int e^{-x}\,dx = -e^{-x}(x+1) + C.$$

Substituting that back,

$$\int x^2 e^{-x}\,dx = -x^2 e^{-x} - 2e^{-x}(x+1) + C = -e^{-x}\left(x^2 + 2x + 2\right) + C.$$

A power x^n takes n rounds of parts to clear, since only the power changes from round to round; the exponential returns unchanged each time.

A different pattern appears when *neither* factor simplifies. For $e^x \sin x$, two rounds of parts bring the original integral back, and you solve for it like an equation.

EXAMPLE 15.16

Find $I = \int e^x \sin x \, dx$.

Parts with $u = \sin x$:

$$I = e^x \sin x - \int e^x \cos x \, dx.$$

Parts again on the new integral, $u = \cos x$:

$$\int e^x \cos x \, dx = e^x \cos x + \int e^x \sin x \, dx = e^x \cos x + I.$$

Substitute back:

$$I = e^x \sin x - e^x \cos x - I \implies 2I = e^x(\sin x - \cos x) \implies I = \tfrac{1}{2}e^x(\sin x - \cos x) + C.$$

The original integral reappeared on the right-hand side, which turned the problem into an equation for I. Watch for this whenever neither factor gets simpler: repeated rounds of parts often bring I back, and algebra gives the answer.

15.6.2 Choosing a method

Five methods are now available, and starting the wrong one costs time and marks. The choice is settled by looking at the integrand before writing anything.

What you see	Method	Example
A standard form exactly, or with a linear inside	read it off, dividing by the inner coefficient	$\cos 3x$, e^{2x}, $(2x+1)^5$
A function together with a multiple of its own derivative	substitution	$2x(x^2+1)^3$, $\dfrac{2x}{x^2+1}$
A product of two unlike functions, one of which simplifies when differentiated	parts	xe^x, $x\sin x$, $\ln x$
A fraction whose denominator factorises	partial fractions, then logarithms	$\dfrac{1}{x^2+3x+2}$
A fraction whose denominator is a quadratic with no real roots	complete the square, then arctangent	$\dfrac{1}{x^2+2x+5}$

Work down the list in that order, because the earlier tests are quicker to apply and the earlier methods are shorter. Two habits help further. Check first whether the fraction is improper, since dividing may remove the problem entirely. And if a substitution leaves an integral no easier than the one you started with, the substitution was the wrong one; go back rather than pressing on.

YOUR TURN 15.6 Integration by Parts HL

One application of parts

25. Find each integral by parts.

(a) $\int x\cos x\,dx$

(b) $\int xe^{2x}\,dx$

(c) $\int x\ln x\,dx$

(d) $\int \arctan x\,dx$

(e) $\int x^2\ln x\,dx$

(f) $\int x\sec^2 x\,dx$ *(Hint: $\int \tan x\,dx = -\ln|\cos x| + C$, from Question 22)*

Definite integrals and the choice of u

26.

(a) Evaluate $\displaystyle\int_0^{\pi} x\sin x\,dx$.

(b) Evaluate $\displaystyle\int_0^{1} xe^{x}\,dx$.

(c) State which factor you chose as u in part (b), and why the other choice would not help.

(d) Use ILATE to say which factor should be u in each of $\int x^2e^x\,dx$, $\int x\arctan x\,dx$ and $\int e^x\sin x\,dx$.

(e) Explain why the third integral in part (d) behaves differently from the first two.

Applying parts twice

27. Find each integral. Each needs integration by parts twice.

(a) $\int x^2e^x\,dx$

(b) $\int x^2\sin x\,dx$

When the original integral returns

28. Find each integral. After two applications of parts the original integral appears again; solve for it.

(a) $\displaystyle\int e^{x}\cos x\,dx$

(b) $\displaystyle\int e^{2x}\sin x\,dx$

Concept check

29. A student writes $\displaystyle\int xe^{x}\,dx=\frac{x^{2}}{2}\,e^{x}+C$, integrating each factor separately. Show, by differentiating, that this is wrong, and find the correct integral.

15.7 Areas Between Curves

To find the area between two curves, integrate the difference: the upper curve minus the lower. The signed-area problem disappears, because the difference is positive wherever the first curve is on top, regardless of the axis.

KEY

$$A=\int_{a}^{b}\left(y_{\text{top}}-y_{\text{bottom}}\right)dx$$

The limits a and b are usually the x-coordinates where the curves intersect, found by setting them equal. To decide which curve is on top, test one convenient point between the intersections. If the curves cross inside the interval, split the integral at each crossing.

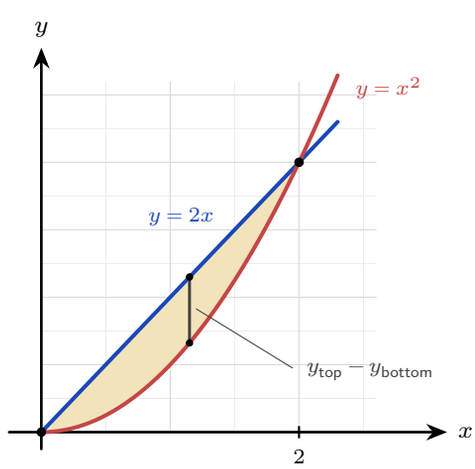


Figure 15.7 The region enclosed between $y=2x$ and $y=x^{2}$. The curves meet where $2x=x^{2}$, at $x=0$ and $x=2$, and the line is the upper boundary in between, so the area is $\int_{0}^{2}(2x-x^{2})\,dx$.

Worked through. For the region in the figure, the curves meet where $2x = x^2$, that is at $x = 0$ and $x = 2$. On $(0, 2)$ the line $y = 2x$ is above, so

$$A = \int_0^2 (2x - x^2)\,dx = \left[x^2 - \tfrac{x^3}{3}\right]_0^2 = 4 - \tfrac{8}{3} = \tfrac{4}{3}.$$

Nothing in the method requires the curves to be polynomials. The steps are the same for trigonometric and exponential boundaries: find where they cross, decide which is on top, and integrate the difference.

EXAMPLE 15.17

Find the area enclosed between $y = \sin x$ and $y = \cos x$ between their first two intersections for $x \geq 0$.

They meet where $\sin x = \cos x$, that is $\tan x = 1$, so the first two intersections are at $x = \frac{\pi}{4}$ and $x = \frac{5\pi}{4}$.

On $\left(\frac{\pi}{4}, \frac{5\pi}{4}\right)$ the sine is above: at $x = \frac{\pi}{2}$, $\sin \frac{\pi}{2} = 1$ while $\cos \frac{\pi}{2} = 0$. So

$$A = \int_{\pi/4}^{5\pi/4} (\sin x - \cos x)\,dx = \Big[-\cos x - \sin x\Big]_{\pi/4}^{5\pi/4}.$$

At the upper limit both $\cos x$ and $\sin x$ equal $-\frac{\sqrt{2}}{2}$, so the bracket is $\sqrt{2}$; at the lower limit both equal $\frac{\sqrt{2}}{2}$ and the bracket is $-\sqrt{2}$:

$$A = \sqrt{2} - \left(-\sqrt{2}\right) = 2\sqrt{2} \approx 2.83.$$

Testing the point $x = \frac{\pi}{2}$ decided which curve is on top, and so the sign of the integrand.

15.7.1 Area between a curve and the y-axis HL

Sometimes it is easier to integrate with respect to y: express x as a function of y and accumulate horizontal slices. This gives the area between the curve and the y-axis.

KEY · IB HL

$$A = \int_a^b |x|\,dy$$

EXAMPLE 15.18

Find the area between the curve $x = y^2$ and the y-axis from $y = 0$ to $y = 2$.

$$A = \int_0^2 y^2\,dy = \left[\tfrac{y^3}{3}\right]_0^2 = \tfrac{8}{3}.$$

In x the same area needs $\int_0^4 (2-\sqrt{x})\,dx$: the line $y=2$ minus the curve. Integrating in y is shorter.

YOUR TURN 15.7 Areas Between Curves

Finding an enclosed area

30. Consider the region enclosed by $y = x^2$ and $y = x + 2$.

(a) Find the x-coordinates where the two curves meet.
(b) State which curve is the upper boundary between those values.
(c) Find the area of the enclosed region.
(d) State what answer you would get by integrating $x^2 - (x+2)$ instead, and explain what has gone wrong.

31. Find the area of the region enclosed by each pair of curves.

(a) $y = x$ and $y = x^2$.
(b) $y = 6 - x^2$ and $y = 2$.
(c) $y = x^2$ and $y = 2 - x^2$.
(d) $y = x^2$ and $y = 3x + 4$.
(e) $y = e^x$ and $y = e^{-x}$, between $x = 0$ and $x = 1$.

Curves that cross inside the interval

32.

(a) Find the total area between $y = x^2$ and $y = x$ from $x = 0$ to $x = 2$.
(b) The curves $y = x^3$ and $y = 4x$ enclose two regions. Find their total area.

Trigonometric boundaries

33.

(a) Find the area between $y = \sin x$ and $y = \cos x$ from $x = 0$ to $x = \frac{\pi}{4}$.
(b) Find the total area between $y = \sin x$ and $y = \cos x$ from $x = 0$ to $x = \frac{\pi}{2}$.
(c) Find the area of the region enclosed by $y = \sin 2x$ and $y = \sin x$ for $0 \le x \le \frac{\pi}{3}$.

Integrating with respect to y

34.

(a) Find the area between the curve $x = y^2 + 1$ and the y-axis, from $y = 0$ to $y = 3$.
(b) Find the area between the curve $y = x^3$ and the y-axis, from $y = 1$ to $y = 8$.

Concept check

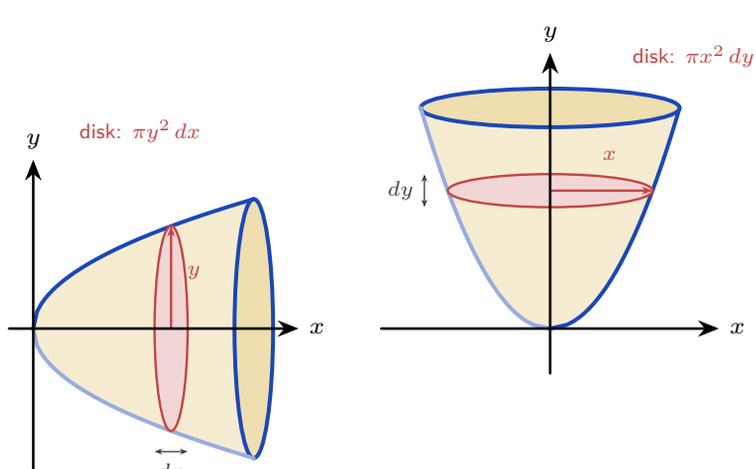
35. A student is finding the area enclosed by $y = x$ and $y = x^2 - 2$, and says: "*part of the region is below the x-axis, so I must split the integral where the curves cross the x-axis.*" Explain why no split is needed, and find the area.

15.8 Volumes of Revolution HL

Rotate the region under a curve through 360° about an axis and it sweeps out a solid. To find the volume, slice the solid into thin disks perpendicular to the axis. Each disk is a cylinder of radius y (the curve's height) and thickness dx, with volume $\pi y^2\,dx$. Summing the disks and letting their thickness shrink turns the sum into an integral.

KEY · IB HL

$$V = \int_a^b \pi y^2\,dx \quad \text{(about the } x\text{-axis)}, \qquad V = \int_a^b \pi x^2\,dy \quad \text{(about the } y\text{-axis)}$$



about the x-axis about the y-axis

Figure 15.8 The same curve rotated about each axis. Turning it about the x-axis gives disks stacked along x: each has radius y and thickness dx, so its volume is $\pi y^2\,dx$. Turning it about the y-axis gives disks stacked along y: each has radius x and thickness dy, so its volume is $\pi x^2\,dy$. In both cases the integral adds the disks from one end of the solid to the other.

About the y-axis, each disk is swept by a horizontal strip running from the y-axis out to the curve, so the region rotated is the one between the curve and the y-axis, not the region under the curve. The two regions give different volumes in general. Write x^2 in terms of y and use y-values as the limits.

Worked through. Rotate the region under $y = x^3$, for $0 \le x \le 2$, through 360° about the x-axis. The volume is

$$V = \int_0^2 \pi (x^3)^2 \, dx = \pi \int_0^2 x^6 \, dx = \pi \left[\tfrac{x^7}{7} \right]_0^2 = \frac{128\pi}{7}.$$

Square the function *before* integrating; squaring after is wrong.

EXAMPLE 15.19

The region between $y = x^2$ and the y-axis, from $y = 0$ to $y = 1$, is rotated about the y-axis. Find the volume.

About the y-axis, slice into disks of radius x and thickness dy, with $x^2 = y$:

$$V = \int_0^1 \pi x^2 \, dy = \pi \int_0^1 y \, dy = \pi \left[\tfrac{y^2}{2} \right]_0^1 = \frac{\pi}{2}.$$

The limits are now y-values, and x^2 is rewritten in terms of y.

The region rotated is the one between the curve and the y-axis, not the one under the curve.

Because the formula squares the function, a trigonometric curve produces $\sin^2$ or $\cos^2$, and those must be rewritten with a double angle before they can be integrated. See §15.4.

EXAMPLE 15.20

The region under $y = \sin 2x$ from $x = 0$ to $x = \frac{\pi}{4}$ is rotated 360° about the x-axis. Find the volume.

Apply the formula, squaring the function first:

$$V = \int_0^{\pi/4} \pi y^2 \, dx = \pi \int_0^{\pi/4} \sin^2 2x \, dx.$$

There is no standard integral for $\sin^2 2x$, so replace it using $\sin^2 2x = \dfrac{1 - \cos 4x}{2}$:

$$\pi \int_0^{\pi/4} \frac{1 - \cos 4x}{2} \, dx = \pi \left[\frac{x}{2} - \frac{\sin 4x}{8} \right]_0^{\pi/4} = \pi \left(\frac{\pi}{8} - 0 \right) = \frac{\pi^2}{8} \approx 1.23.$$

Two πs appear for different reasons, and they must be kept apart: one comes from the volume formula, the other from the upper limit.

EXAMPLE 15.21

The region under $y = e^x$ from $x = 0$ to $x = 2$ is rotated about the x-axis. Find the volume.

$$V = \pi \int_0^2 (e^x)^2 \, dx = \pi \int_0^2 e^{2x} \, dx = \pi \left[\tfrac{1}{2} e^{2x} \right]_0^2 = \frac{\pi}{2} \left(e^4 - 1 \right) \approx 84.2.$$

Squaring first is what turns e^x into e^{2x}, a linear composite. Squaring the integral instead of the function is a common error, and it gives a different answer entirely.

When the region lies between two curves, with y_{outer} further from the x-axis than y_{inner} across $a \le x \le b$, each slice is a disk with a disk-shaped hole in it, called a **washer**. Its volume is the volume of the outer disk minus that of the hole, $\pi y_{\text{outer}}^2\,dx - \pi y_{\text{inner}}^2\,dx$, so

$$V = \pi \int_a^b (y_{\text{outer}}^2 - y_{\text{inner}}^2)\,dx.$$

EXAMPLE 15.22

The region enclosed by $y = x$ and $y = x^2$ is rotated 360° about the x-axis. Find the volume.

The curves meet at $x = 0$ and $x = 1$, and on $(0, 1)$ the line is further from the axis, since $x > x^2$ there. So $y_{\text{outer}} = x$ and $y_{\text{inner}} = x^2$:

$$V = \pi \int_0^1 (x^2 - x^4)\,dx = \pi\left[\tfrac{x^3}{3} - \tfrac{x^5}{5}\right]_0^1 = \frac{2\pi}{15}.$$

CAUTION Square each function first, then subtract. Squaring the difference, $\pi \int_0^1 (x - x^2)^2\,dx = \frac{\pi}{30}$, finds a different quantity and gives the wrong volume.

YOUR TURN 15.8 Volumes of Revolution HL

Rotating about the x-axis

36. Each region under the curve is rotated 360° about the x-axis. Find the volume.

(a) $y = x^2$, $0 \le x \le 2$.

(b) $y = \sqrt{x}$, $0 \le x \le 4$.

(c) $y = 2x$, $0 \le x \le 3$.

(d) $y = \dfrac{1}{x}$, $1 \le x \le 3$.

(e) $y = e^{-x}$, $0 \le x \le \ln 2$.

(f) $y = e^{3x}$, $0 \le x \le \frac{1}{3}$.

Squares of sine and cosine

37. Each region under the curve is rotated $360°$ about the x-axis. Find the exact volume.

(a) $y = \cos x$, $0 \le x \le \frac{\pi}{2}$.
(b) $y = \sin x$, $0 \le x \le \frac{\pi}{4}$.
(c) $y = 1 + \cos x$, $0 \le x \le \pi$.
(d) Explain why the volumes in this question need a double-angle rewriting, while those in Question 36(e) and (f) did not.

Rotating about the y-axis

38. Each region is rotated $360°$ about the y-axis. Find the volume.

(a) The region between $y = x^3$ and the y-axis, from $y = 0$ to $y = 8$.
(b) The region between $y = x^2 + 1$ and the y-axis, from $y = 1$ to $y = 3$.
(c) The region between $x = y^2$ and the y-axis, from $y = 0$ to $y = 2$.
(d) Your answer to part (c) equals your answer to Question 36(a). Explain why.

Washers

39. The region enclosed by each pair of curves is rotated $360°$ about the x-axis. Find the volume.

(a) $y = \sqrt{x}$ and $y = x^2$.
(b) $y = x + 2$ and $y = x^2$.
(c) $y = 4 - x^2$ and $y = 3$.

Concept check

40. For the region enclosed by $y = \sqrt{x}$ and $y = x^2$ in Question 39(a), a student writes $V = \pi \displaystyle\int_0^1 \left(\sqrt{x} - x^2\right)^2 dx$. Explain the error, and give the correct integral and volume.

15.9 Kinematics by Integration

In Ch. 14 you differentiated displacement to get velocity, and velocity to get acceleration. Integration reverses those two steps. Integrate acceleration to get velocity, and velocity to get displacement, using an initial condition each time to fix the constant.

EXAMPLE 15.23

A particle starts from rest at the origin, and its acceleration is $a = 6t - 4\,\text{m s}^{-2}$. Find its velocity and displacement at $t = 2$.

Integrate the acceleration, and use $v = 0$ at $t = 0$ to find the constant:

$$v = \int (6t - 4)\,dt = 3t^2 - 4t + C, \qquad 0 = 0 - 0 + C \implies C = 0.$$

Integrate the velocity, and use $s = 0$ at $t = 0$:

$$s = \int (3t^2 - 4t)\,dt = t^3 - 2t^2 + D, \qquad D = 0.$$

At $t = 2$: $v = 12 - 8 = 4\,\text{m s}^{-1}$ and $s = 8 - 8 = 0$ m. Each integration introduces one constant, so each needs one initial condition.

KEY · IB

$$\text{displacement} = \int_{t_1}^{t_2} v\,dt, \qquad \text{distance travelled} = \int_{t_1}^{t_2} |v|\,dt$$

Displacement is the signed change in position; distance is the total path length. They differ whenever the particle reverses direction, because backward motion subtracts from displacement but adds to distance. On a velocity–time graph the difference is visible: area below the axis counts as negative for displacement, and as positive for distance.

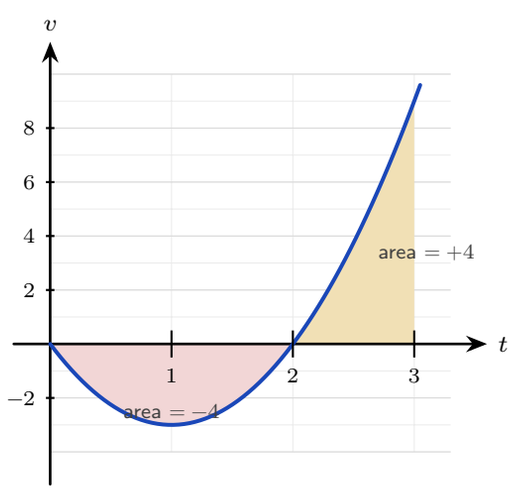


Figure 15.9 The velocity $v = 3t^2 - 6t$ over $0 \le t \le 3$. The particle moves backwards until $t = 2$ and forwards after it. Displacement adds the two signed areas, $-4 + 4 = 0$; distance adds their sizes, $4 + 4 = 8$ m.

Worked through. Take the velocity in the figure, $v(t) = 3t^2 - 6t\,\text{m s}^{-1}$ over $0 \le t \le 3$. Displacement integrates v directly:

$$\int_0^3 (3t^2 - 6t)\,dt = \Big[t^3 - 3t^2\Big]_0^3 = 27 - 27 = 0.$$

The particle ends where it began. For distance, note $v = 3t(t - 2)$ is negative on $(0, 2)$ and

positive on $(2,3)$, so split:

$$\left|\int_0^2 v\,dt\right| + \int_2^3 v\,dt = |-4| + 4 = 8 \text{ m}.$$

Zero displacement but 8 m travelled: the particle went out 4 m and came back.

Oscillating motion is handled the same way, and it makes the difference between displacement and distance especially clear.

EXAMPLE 15.24

A particle has velocity $v(t) = 4\cos 2t$ m s^{-1}. Find its displacement and the distance travelled over $0 \le t \le \frac{\pi}{2}$ seconds.

Displacement integrates v directly. The antiderivative of $4\cos 2t$ is $2\sin 2t$:

$$\int_0^{\pi/2} 4\cos 2t\,dt = \Big[2\sin 2t\Big]_0^{\pi/2} = 2\sin\pi - 2\sin 0 = 0.$$

The particle finishes where it started. For the distance, first find where it turns: $v = 0$ when $\cos 2t = 0$, so $2t = \frac{\pi}{2}$ and $t = \frac{\pi}{4}$. Split the interval there:

$$\int_0^{\pi/4} 4\cos 2t\,dt = \Big[2\sin 2t\Big]_0^{\pi/4} = 2, \qquad \int_{\pi/4}^{\pi/2} 4\cos 2t\,dt = 0 - 2 = -2.$$

Distance adds the sizes: $|2| + |-2| = 4$ m. So the particle travels 2 m out and 2 m back, ending at its starting point.

Without a GDC the split is necessary: integrating v straight through, instead of $|v|$, returns 0, which is the displacement, not the distance. With a GDC, evaluate $\int_0^{\pi/2} |4\cos 2t|\,dt$ directly.

YOUR TURN 15.9 Kinematics by Integration

Displacement from velocity

41. A particle moves in a straight line, and t is the time in seconds. Find its displacement over the interval given.

(a) $v = 2t + 1$ m s^{-1}, $0 \le t \le 3$.
(b) $v = t^2 - 4$ m s^{-1}, $0 \le t \le 3$.
(c) $v = 3\cos t$ m s^{-1}, $0 \le t \le \frac{\pi}{2}$.
(d) $v = 6e^{-2t}$ m s^{-1}, $0 \le t \le 1$.

Integrating acceleration

42. A particle moves in a straight line, and t is the time in seconds.

(a) Its acceleration is $a = 6t$ m s^{-2} and $v(0) = 1$ m s^{-1}. Find $v(t)$.

(b) Its acceleration is $a = \sin t$ m s^{-2} and $v(0) = 0$. Find $v(t)$.

(c) Its acceleration is $a = 6t - 2$ m s^{-2}. At $t = 0$ its velocity is 3 m s^{-1} and its displacement from a fixed point O is 1 m. Find its displacement from O at $t = 2$.

(d) Its acceleration is $a = -4\sin 2t$ m s^{-2}. At $t = 0$ it is at O with velocity 2 m s^{-1}. Find its displacement s from O in terms of t.

Distance travelled

43. A particle moves in a straight line, and t is the time in seconds.

(a) Its velocity is $v = 4 - 2t$ m s^{-1}. Find the distance travelled over $0 \le t \le 4$.

(b) For the particle in part (a), find the displacement over the same interval, and explain why the two answers differ.

(c) Its velocity is $v = 3t^2 - 12$ m s^{-1}. Find the time at which it changes direction, and the distance travelled over $0 \le t \le 3$.

(d) Its velocity is $v = 1 - 2\sin t$ m s^{-1}. Find the time in $0 \le t \le \frac{\pi}{2}$ at which it changes direction, and the exact distance travelled over that interval.

Concept check

44. A particle moves with velocity $v = t^2 - 4t + 3$ m s^{-1}. A student finds the distance travelled over $0 \le t \le 3$ as $\left|\int_0^3 v\,dt\right|$. Explain why this is wrong, and find the distance travelled.

Chapter 15 · Integral Calculus

Reflections

TOK The proof at the end of this chapter establishes *that* the area under a curve and the reverse of a derivative are the same calculation. It does not explain *why* the world should be arranged that way. Two ideas invented for different purposes, one geometric and one algebraic, turned out to be one idea.

The integral now has two descriptions, the limit of a sum and the reverse of a derivative, and a theorem says they agree. Which one is the definition, and does the choice change what an integral is?

IM Long before calculus, Archimedes (3rd century BCE, Syracuse) found the area of a parabolic segment by his "method of exhaustion": filling the region with ever-smaller triangles and summing them. He effectively computed an integral two thousand years before Newton and Leibniz gave the operation a name and a reverse-differentiation shortcut. The idea of accumulation by slicing is older than the algebra we use to carry it out, and his triangles did the same job as the rectangles in §15.2.

RW Integration is summation made continuous, so it appears wherever a total is built from a varying rate. The total charge from a varying current, the work done by a changing force, the area of an irregular plot of land, the probability accumulated under a distribution curve (Ch. 12): each is an integral. Whenever a quantity is the running total of a rate, integration computes it.

A shape made by spinning a profile about an axis, such as a bottle, a lens or a turbine housing, is fixed completely by that one curve, so its volume is a single integral rather than a measurement. Engineers can size fuel tanks and machined parts this way, and the same integral gives the mass once the density is known.

EXT Not every function has an antiderivative in terms of familiar functions: $\int e^{-x^2}\,dx$, central to statistics, cannot be written with powers, exponentials, and trigonometric functions. Yet the definite integral still exists as a perfectly real area, and is computed numerically. The reach of "find the area" is wider than the reach of "reverse the derivative."

The Fundamental Theorem of Calculus

The Fundamental Theorem was stated in §15.2 and used throughout the chapter. Here it is proved. It is not examined, but it is the reason the chapter works. The result, reached by Barrow, Newton and Leibniz in the 17th century, states that if f is continuous on $[a, b]$ and F is any antiderivative of f, then

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

Assume first that $f(x) \geq 0$, so that every integral below is an area; the last step removes this assumption.

The **area function** $A(x) = \int_a^x f(t)\,dt$ is the area under the curve from a to a moving right end x. The letter t is only the variable of integration. The area of any band is then a difference:

$$\text{area from } x_1 \text{ to } x_2 \;=\; A(x_2) - A(x_1).$$

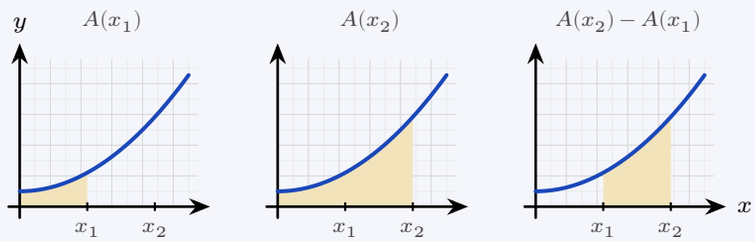


Figure 15.10 Shade up to x_2, remove what was shaded up to x_1, and the band between them remains.

So what kind of function is A? When x increases by a small amount h, the extra area $A(x+h) - A(x)$ is one thin strip of width h.

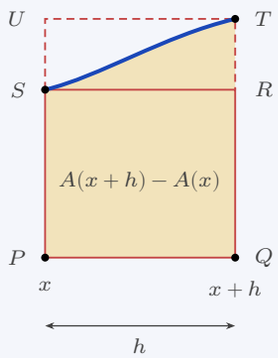


Figure 15.11 One strip, enlarged. Its area $A(x+h) - A(x)$ contains the solid rectangle $PQRS$ and fits inside the dashed rectangle $PQTU$, both of width h.

The short rectangle has height $f(x)$ and the tall one $f(x+h)$, so

$$f(x)\,h \;\leq\; A(x+h) - A(x) \;\leq\; f(x+h)\,h \quad\Longrightarrow\quad f(x) \;\leq\; \frac{A(x+h) - A(x)}{h} \;\leq\; f(x+h).$$

As $h \to 0$, $f(x+h) \to f(x)$ because f is continuous. The middle term, which is the

definition of $A'(x)$, is squeezed between them:

$$A'(x) = f(x).$$

The same squeeze works for $h < 0$. So the area function is **an antiderivative of** f. If f falls across the strip the rectangles swap roles, and if it rises and falls, use its least and greatest values on the strip.

Now let F be any antiderivative of f. Then $(A - F)' = 0$ on the interval, so $A = F + C$ for a constant C. The area from a to a is zero, so $C = -F(a)$, and at $x = b$

$$\int_a^b f(x)\,dx \;=\; A(b) \;=\; F(b) - F(a).$$

If f takes negative values, apply the argument to $f + k$, with k chosen so that $f + k \geq 0$ on $[a, b]$. Its antiderivative is $F + kx$, and the extra rectangle area $k(b - a)$ appears on both sides and cancels. ∎

The proof also shows that

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x).$$

Accumulating f and then differentiating returns f. This is often called the first part of the Fundamental Theorem. It is not examined.

End-of-Chapter Problems
Chapter 15: Integral Calculus

1. A curve has $\dfrac{dy}{dx} = 3x^2 - 4x + 1$ and passes through $(1, 2)$.

(a) Find the equation of the curve. [3]

(b) Find the value of y when $x = 2$. [2]

2. Consider $y = x^2 - 4$.

(a) Find the x-intercepts. [2]

(b) Find the area enclosed between the curve and the x-axis. [4]

3. The curves $y = x^2$ and $y = 8 - x^2$ intersect at two points.

(a) Find the coordinates of the intersection points. [3]

(b) Find the area enclosed between the curves. [4]

4.

(a) Using the substitution $u = x^2 + 4$, find $\displaystyle\int \frac{x}{x^2 + 4}\,dx$. [3]

(b) Hence evaluate $\displaystyle\int_0^2 \frac{x}{x^2 + 4}\,dx$, giving an exact answer. [3]

5.

(a) Find $\displaystyle\int xe^{2x}\,dx$ using integration by parts. [3]

(b) Evaluate $\displaystyle\int_0^1 xe^{2x}\,dx$, giving an exact answer. [2]

6. The region bounded by $y = \sqrt{x}$, the x-axis, and $x = 4$ is rotated 360° about the x-axis.

(a) Write down an integral for the volume. [2]

(b) Find the volume. [2]

7. A particle starts at the origin with velocity $5\,\text{m s}^{-1}$ and has acceleration $a = 6t - 4\,\text{m s}^{-2}$.

(a) Find the velocity $v(t)$. [2]

(b) Find the displacement $s(t)$. [2]

(c) Find the displacement when $t = 2$. [2]

8. Consider the region enclosed by $y = x^2$, the y-axis, and $y = 4$.

(a) Find the area of the region. [3]

(b) Find the volume when this region is rotated about the y-axis. [3]

9.

(a) Find $\displaystyle\int \frac{1}{x^2+4}\,dx$. [2]

(b) Evaluate $\displaystyle\int_0^2 \frac{1}{x^2+4}\,dx$, giving an exact answer. [3]

10. The velocity of a particle is $v(t) = 3t^2 - 12t + 9\,\text{m s}^{-1}$, $t \geq 0$.

(a) Find the times when the particle is at rest. [2]

(b) Find the total distance travelled in the first 3 seconds. [5]

11. A curve has gradient function $\dfrac{dy}{dx} = 8x^3 - 6x$ and passes through the point $(-1, 4)$.

(a) Find the equation of the curve. [3]

(b) Write down its y-intercept. [1]

12. Find the following integrals.

(a) $\displaystyle\int e^{3x}\,dx$ [2]

(b) $\displaystyle\int (4x-1)^6\,dx$ [2]

(c) $\displaystyle\int \sin 2x\,dx$ [2]

13. Evaluate $\displaystyle\int_1^4 \left(\frac{2}{\sqrt{x}} + 1\right) dx$. [4]

14. Consider the curve $y = x^2 - 5x + 4$ on the interval $0 \leq x \leq 2$.

(a) Find where the curve crosses the x-axis in this interval. [1]

(b) Find the total area enclosed between the curve and the x-axis on the interval. [5]

15. Find the area of the region enclosed by the curve $y = 4x - x^2$ and the line $y = x$. [5]

16. Using the substitution $u = x^2 + 1$, evaluate $\displaystyle\int_0^2 x(x^2+1)^3\,dx$. [5]

17. Evaluate $\displaystyle\int_0^2 \frac{x^2}{x^3+1}\,dx$, giving your answer in the form $a \ln b$. [5]

18. Find $\displaystyle\int \frac{x+7}{(x-1)(x+3)}\,dx$. [6]

19. Find $\displaystyle\int \frac{1}{x^2+8x+20}\,dx$. [5]

20. Find the following integrals.

(a) $\displaystyle\int x\cos x\,dx$ [3]

(b) $\displaystyle\int x^2\ln x\,dx$ [4]

21. The region under $y=\sqrt{x}$ from $x=0$ to $x=4$ is rotated about the x-axis.

(a) Find the volume generated. [3]

(b) The same curve, written $x=y^2$ for $0\le y\le 2$, is rotated about the y-axis. Find the volume generated. [4]

22. A particle moves with velocity $v(t)=t^2-4t+3\,\text{m s}^{-1}$ for $0\le t\le 4$.

(a) Find the displacement of the particle over the interval. [3]

(b) Find the total distance travelled. [5]

23. The area under $y=x^2+1$ from $x=0$ to $x=3$ is to be estimated by three rectangles of width 1.

(a) Find the estimate obtained by taking the height of each rectangle at the left of its strip. [3]

(b) Find the estimate obtained by taking the height at the right. [2]

(c) Evaluate $\displaystyle\int_0^3 (x^2+1)\,dx$ and verify that it lies between the two estimates. [3]

(d) State, with a reason, whether the left estimate would increase or decrease if six rectangles were used instead. [2]

24. The curve $y=x^2$ and the line $y=2-x$ enclose a region R.

(a) Find the x-coordinates of the two points of intersection. [3]

(b) Sketch the curve and the line, and shade R. [2]

(c) Show that the area of R is $\dfrac{9}{2}$. [5]

(d) Explain why integrating $y_{\text{curve}}-y_{\text{line}}$ instead would give the wrong sign. [2]

25. This question uses integration by parts throughout.

(a) Find $\displaystyle\int x\ln x\,dx$. [4]

(b) Hence evaluate $\int_1^e x \ln x \, dx$, giving your answer in terms of e. [3]

(c) Evaluate $\int_0^{\pi/2} x \cos x \, dx$. [5]

(d) In part (c), state which factor you took as u, and explain what would go wrong with the other choice. [2]

26. Consider the curve $y = x^3$ for $0 \leq x \leq 1$.

(a) The region between the curve and the x-axis is rotated 360° about the x-axis. Show that the volume is $\frac{\pi}{7}$. [4]

(b) The region between the curve and the y-axis is rotated 360° about the y-axis. Find this volume, giving an exact answer. [5]

(c) State which of the two solids has the greater volume, and explain the answer by describing the two regions. [3]

27. A particle moves in a straight line with acceleration $a = 6 - 2t \text{ m s}^{-2}$, and is at rest at $t = 0$.

(a) Show that $v = 6t - t^2$. [3]

(b) Find the time at which the particle is next at rest. [2]

(c) Find the maximum velocity, and the time at which it occurs. [3]

(d) Find the distance travelled from $t = 0$ until the particle is next at rest. [4]

(e) Explain why, for this motion, the distance travelled equals the displacement. [2]

28. Name the method for each integral before using it.

(a) $\int \frac{x^2 + 1}{x + 1} \, dx$ [5]

(b) $\int_1^2 \frac{2x - 1}{x^2 - x + 1} \, dx$ [4]

(c) $\int_0^{\pi/2} \sin x \cos^2 x \, dx$ [4]

(d) $\int \frac{1}{x^2 - 4x + 13} \, dx$ [4]

29. The region R is enclosed by $y = \sin x$, $y = \cos x$ and the y-axis.

(a) Show that the two curves first meet at $x = \frac{\pi}{4}$. [2]

(b) Find the area of R. [4]

(c) R is rotated 360° about the x-axis. Show that the volume is $\frac{\pi}{2}$. [6]

30. The region under $y = \cos x$ from $x = 0$ to $x = \frac{\pi}{2}$ is rotated 360° about the x-axis.

(a) Write down an integral for the volume. [2]

(b) Use a double angle to show that the volume is $\dfrac{\pi^2}{4}$. [5]

(c) The same region under $y = \sin x$ gives the same volume. Explain why, without integrating. [2]

31. A particle moves in a straight line with velocity $v(t) = 6\sin 2t$ m s^{-1}, for $0 \le t \le \pi$ seconds.

(a) Find the times at which the particle is at rest. [3]

(b) Find the displacement over $0 \le t \le \pi$. [3]

(c) Find the total distance travelled over the same interval. [4]

(d) Explain why the two answers differ. [2]

32. A particle starts from rest at the origin with acceleration $a(t) = 4e^{-2t}$ m s^{-2}.

(a) Show that $v(t) = 2 - 2e^{-2t}$. [3]

(b) Find the displacement over $0 \le t \le 1$. [4]

(c) State the velocity the particle approaches as t grows large, and explain what is happening physically. [2]

33. Consider the region enclosed by $y = e^{x}$, $y = e^{-x}$ and the line $x = 1$.

(a) Show that the two curves meet at $x = 0$. [1]

(b) Find the area of the region. [4]

(c) The region is rotated 360° about the x-axis. Find the volume, giving your answer in terms of e and π. [5]

34. Name the method for each integral before using it.

(a) $\displaystyle\int x e^{3x}\,dx$ [4]

(b) $\displaystyle\int \cos^2 3x\,dx$ [4]

(c) $\displaystyle\int \frac{\sin x}{\cos x}\,dx$ [3]

(d) $\displaystyle\int_0^{\ln 2} \frac{e^{x}}{1+e^{x}}\,dx$ [4]

Challenge Questions

35. Solids you already know. Several volume formulas from Ch. 7 can be recovered by rotating a curve about the x-axis. This question derives three of them, and then finds a

single formula that contains two of them as special cases.

(a) The line $y = r$ for $0 \le x \le h$ is rotated about the x-axis. Show that the volume is $\pi r^2 h$, and name the solid. [3]

(b) The line $y = \dfrac{r}{h}x$ for $0 \le x \le h$ is rotated about the x-axis. Show that the volume is $\frac{1}{3}\pi r^2 h$. [4]

(c) The semicircle $y = \sqrt{r^2 - x^2}$ for $-r \le x \le r$ is rotated about the x-axis. Show that the volume is $\frac{4}{3}\pi r^3$. [4]

(d) Let $m > 0$. The curve $y = r\left(\dfrac{x}{h}\right)^m$ for $0 \le x \le h$ is rotated about the x-axis. Show that the volume is

$$V = \frac{\pi r^2 h}{2m + 1}.$$

[4]

(e) Show that the formula in part (d) also gives the results of parts (a) and (b), taking $m = 0$ for the line $y = r$. Use it to explain where the factor $\frac{1}{3}$ in the volume of the cone comes from. [3]

(f) Every solid in part (d) fits inside the cylinder of part (a). Find the value of m for which the solid has exactly half the volume of that cylinder, and describe the curve that is rotated. [2]

(g) Many containers, such as a bowl, a vase or a glass, have a profile close to one of the curves in part (d). Choose one, and describe how you would find a value of m that models it and test the model against a measured volume. [2]

36. Damped oscillation. A pendulum swinging in air moves in a way that the graph of $y = e^{-x}\sin x$ describes: it oscillates, and each swing is smaller than the last. This question finds the area under such curves by two methods and then adds up the areas of all the swings. Throughout, a and b are non-zero real constants, and

$$I = \int e^{ax}\cos bx\,dx, \qquad J = \int e^{ax}\sin bx\,dx.$$

(a) Apply integration by parts once to each integral, differentiating e^{ax} each time, to show that, apart from constants of integration,

$$bI = e^{ax}\sin bx - aJ \qquad \text{and} \qquad bJ = -e^{ax}\cos bx + aI.$$

[4]

(b) Solve these two equations simultaneously to show that

$$I = \frac{e^{ax}\left(a\cos bx + b\sin bx\right)}{a^2 + b^2} + C,$$

and find J in the same form. [4]

(c) A second route uses complex numbers (Ch. 17). You are given that

$$e^{(a+ib)x} = e^{ax}(\cos bx + i\sin bx) \qquad \text{and} \qquad \int e^{cx}\,dx = \frac{e^{cx}}{c} + C,$$

where the second result holds when c is a complex constant. Deduce that $I + iJ = \dfrac{e^{(a+ib)x}}{a+ib} + C$. By multiplying the numerator and denominator by $a - ib$, recover both results of part (b) at once. [4]

(d) Let $f(x) = e^{-x}\sin x$ and, for $k = 0, 1, 2, \ldots$, let $A_k = \displaystyle\int_{k\pi}^{(k+1)\pi} f(x)\,dx$, the signed area of the kth swing. Show that $A_0 = \frac{1}{2}(1 + e^{-\pi})$, and find A_1 and A_2 exactly. [3]

(e) Conjecture a relationship between A_{k+1} and A_k. Prove it for every $k \geq 0$ by using the substitution $x = u + \pi$ in A_{k+1}. [4]

(f) Hence show that $\displaystyle\int_0^{\infty} e^{-x}\sin x\,dx = \frac{1}{2}$, and confirm this value using your expression for J from part (b). [3]

(g) For the curve $y = e^{-cx}\sin x$ with $c > 0$, investigate how the ratio of the areas of successive swings depends on c. Suggest how timing and measuring the successive swings of a real pendulum could be used to estimate its value of c. [2]

37. A reduction formula. Let $I_n = \displaystyle\int_0^{\pi/2} \sin^n x\,dx$ for $n \geq 0$.

(a) Evaluate I_0 and I_1 directly. [3]

(b) Writing $\sin^n x = \sin^{n-1} x \cdot \sin x$ and using integration by parts, show that

$$I_n = \frac{n-1}{n} I_{n-2}, \qquad n \geq 2.$$

[6]

(c) Use the formula to find I_2, I_3, I_4 and I_5 exactly. [4]

(d) Explain why I_n involves π when n is even but not when n is odd. [3]

(e) Explain why $I_{n+1} \leq I_n$ for every $n \geq 0$. Investigate what happens to the ratio $\dfrac{I_{2m}}{I_{2m+1}}$ as m increases, and what this suggests about writing $\frac{\pi}{2}$ as an infinite product of fractions. [3]

38. Exact areas from rectangles. The notes estimate the area under $y = x^2$ on $[0, 1]$ with rectangles, and then evaluate it exactly with the Fundamental Theorem of Calculus. This question finds the same exact value from the rectangles alone. Divide $[0, 1]$ into n strips of equal width $\frac{1}{n}$, and let U_n and L_n be the sums of the rectangles whose heights are read at the right and at the left of each strip.

(a) Show that $U_n = \dfrac{1}{n^3}\displaystyle\sum_{r=1}^{n} r^2$, and write L_n as a sum in the same way. [3]

(b) Given that $\displaystyle\sum_{r=1}^{n} r^2 = \frac{n(n+1)(2n+1)}{6}$, show that $U_n = \dfrac{(n+1)(2n+1)}{6n^2}$, and find L_n in

the same form. Evaluate L_{10} and U_{10}. [4]

(c) Show that $U_n - L_n = \dfrac{1}{n}$. Hence, without using the Fundamental Theorem, show that the area under $y = x^2$ on $[0, 1]$ is exactly $\frac{1}{3}$. [3]

(d) Given that $\displaystyle\sum_{r=1}^{n} r^3 = \frac{n^2(n+1)^2}{4}$, use the same method to show that the area under $y = x^3$ from $x = 0$ to $x = b$, where $b > 0$, is $\frac{1}{4}b^4$. [4]

(e) Conjecture the area under $y = x^p$ on $[0, 1]$ for a positive integer p. State what the leading term of a formula for $\displaystyle\sum_{r=1}^{n} r^p$ would have to be for the method to give your conjecture. [3]

(f) The method depends on having a formula for the sum. Investigate a curve for which a different kind of sum can be evaluated exactly, for example $y = 2^x$ on $[0, 1]$, whose rectangle sum is a geometric series (Ch. 2). [2]

Chapter 16

Differential Equations

16

SYLLABUS 5.18 5.19 HL ONLY

A cup of coffee cools fastest when it is hottest. Take it from the machine at $80°$C and its temperature falls quickly at first; near room temperature it hardly changes. The *rate* at which it cools depends on how far above the room its temperature currently is.

That sentence is a differential equation. Written in symbols, the rate of cooling $\frac{d\theta}{dt}$ is proportional to the gap $\theta - 20$ between the coffee and the 20-degree room:

$$\frac{d\theta}{dt} = -k(\theta - 20).$$

Notice what kind of equation this is. The unknown is not a number. It is a *function*, $\theta(t)$, and the equation links that function to its own derivative. Many laws of nature have this form: they say how fast something grows, cools, decays or moves.

A **differential equation** relates a function to its derivatives. To **solve** one is to find the function itself, recovering $\theta(t)$ from a statement about $\frac{d\theta}{dt}$.

That is the point of solving one: it turns a rule about change into a prediction. A law of nature tells you what is happening at this instant; solving its differential equation tells you how the quantity will change from then on.

By the end of this chapter you will be able to solve such an equation exactly by three methods: separating the variables, substituting to make them separate, and using an integrating factor. You will also write a function as an infinite polynomial, called a Maclaurin series. That gives a fourth route to a solution, and a way to approximate a function. When no exact solution exists, you will solve the equation approximately by Euler's method. The coffee equation is solved in the first section.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

A calculator is assumed throughout this chapter. The EXACT badge marks the questions whose answer must be left exact.

16.1 Separable Variables

Differential equations appear wherever a law describes a rate of change, and that covers much of science. A few examples show the range.

- **Cooling.** $\dfrac{d\theta}{dt} = -k(\theta - \theta_{\text{room}})$: a hot drink, an engine, or a body found at a crime scene.
- **Radioactive decay.** $\dfrac{dN}{dt} = -\lambda N$: the basis of carbon dating.
- **Population growth.** $\dfrac{dn}{dt} = kn(M - n)$: growth that slows as the population fills the space available.
- **Medicine.** $\dfrac{dC}{dt} = -kC$: a drug leaving the bloodstream, which sets how often a dose must be repeated.
- **Falling through air.** $\dfrac{dv}{dt} = g - kv^2$: why a skydiver reaches a steady terminal speed.
- **Epidemics.** Three linked differential equations predict how an infection spreads through a population.

Each one says the same thing in different words: the rate is known, and the quantity is wanted.

Two words describe the equations in this chapter, and both are settled by looking at the equation.

- The **order** is the highest derivative that appears. Every equation here contains $\dfrac{dy}{dx}$ and no higher derivative, so all of them are **first order**.
- An equation is **linear** when y and its derivatives appear only to the first power, and are never multiplied by one another. So $\dfrac{dy}{dx} + xy = 1$ is linear, and $\dfrac{dy}{dx} = 2xy^2$ is not.

The three exact methods in this chapter are chosen by which form an equation takes, so these two points are checked first.

16.1.1 Separating the variables

The simplest case can be rearranged so that each variable sits on its own side. A first-order equation is **separable** if it can be written as a function of y times a function of x:

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

Then you separate and integrate both sides. Treating $\frac{dy}{dx}$ as a ratio is informal but valid here: gather all the y terms with dy and all the x terms with dx.

KEY

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

One unknown constant appears, giving the **general solution**, a family of curves. An **initial condition** fixes the constant and selects the **particular solution**. A particular solution holds only on the interval that contains the initial point and on which its formula is defined: if the formula is undefined at some value of x, the solution stops there.

When integration gives $\ln|y-a| = f(x) + C$, exponentiate both sides and write $y - a = Ae^{f(x)}$. The constant $A = \pm e^{C}$ replaces the modulus sign and the old constant together.

Worked through. To solve $\frac{dy}{dx} = 2xy^2$ given $y(0) = 1$, separate the variables and integrate:

$$\int \frac{1}{y^2}\,dy = \int 2x\,dx \implies -\frac{1}{y} = x^2 + C.$$

Apply $y(0) = 1$: $-1 = 0 + C \implies C = -1$. So $-\frac{1}{y} = x^2 - 1$, giving

$$y = \frac{1}{1 - x^2}, \qquad -1 < x < 1.$$

The solution is valid on $-1 < x < 1$, the interval containing the starting point $x = 0$ on which $1 - x^2 \neq 0$. Apply the initial condition as early as possible; carrying C through more steps than necessary makes errors more likely.

EXAMPLE 16.1

A cup of coffee obeys $\frac{d\theta}{dt} = -k(\theta - 20)$ with $\theta(0) = 80$. Solve for $\theta(t)$.
Separate and integrate:

$$\int \frac{1}{\theta - 20}\,d\theta = \int -k\,dt \implies \ln|\theta - 20| = -kt + C.$$

Exponentiating, $\theta - 20 = Ae^{-kt}$. The condition $\theta(0) = 80$ gives $A = 60$, so

$$\theta(t) = 20 + 60e^{-kt}.$$

The temperature decays exponentially toward the room's 20 degrees, never quite reaching it. This is the behaviour described at the start of the chapter.

Separation also solves a standard population model in ecology. In the 1960s, once the cattle disease rinderpest was eliminated around the Serengeti, the wildebeest herd began to grow,

and its growth obeyed the **logistic equation**: growth proportional both to the herd itself and to the room left below the **carrying capacity**, the largest population the land can support.

EXAMPLE 16.2

The Serengeti wildebeest population n (in millions) satisfies $\dfrac{dn}{dt} = kn(1.4 - n)$, with $n(0) = 0.25$ in 1961. Solve for $n(t)$.

Separate, and split the left side with the partial fractions of Ch. 6:

$$\int \frac{dn}{n(1.4-n)} = \int k\,dt, \qquad \frac{1}{n(1.4-n)} = \frac{1}{1.4}\left(\frac{1}{n} + \frac{1}{1.4-n}\right).$$

Integrating,

$$\frac{1}{1.4}\ln\left|\frac{n}{1.4-n}\right| = kt + C \implies \frac{n}{1.4-n} = Be^{1.4kt}.$$

The initial condition gives $B = \dfrac{0.25}{1.15} \approx 0.217$, and solving for n:

$$n(t) = \frac{1.4}{1 + 4.6\,e^{-1.4kt}}.$$

This is the logistic S-curve: growth that is nearly exponential at first, then approaches the carrying capacity of 1.4 million. The bracket $(1.4 - n)$ is what slows it, since it decreases towards zero as n approaches 1.4. The real herd behaved this way, climbing to about 1.4 million by 1977 and fluctuating around that level since.

YOUR TURN 16.1 Separable Variables

Recognising the type of equation

1. Each equation is first order. State whether it is linear and whether it is separable, giving a reason for each.

(a) $\dfrac{dy}{dx} = x^2 y$

(b) $\dfrac{dy}{dx} = y^2 + x$

(c) $\dfrac{dy}{dx} + y\cos x = \sin x$

(d) $y\dfrac{dy}{dx} = x$

General solutions

2. Find the general solution of each equation.

(a) $\dfrac{dy}{dx} = \dfrac{x}{y}$

(b) $\dfrac{dy}{dx} = y$

(c) $\dfrac{dy}{dx} = \dfrac{1+y}{1+x}$

(d) $\dfrac{dy}{dx} = e^{x-y}$

Particular solutions and where they are valid

3. Find the particular solution of each equation.

(a) $\dfrac{dy}{dx} = xy$ with $y(0) = 2$.

(b) $\dfrac{dy}{dx} = 1 + y^2$ with $y(0) = 1$. State the interval containing $x = 0$ on which the solution is valid.

(c) $\dfrac{dy}{dx} = y\tan x$ with $y(0) = 4$, for $0 \le x < \frac{\pi}{2}$.

(d) $\dfrac{dy}{dx} = y^2\cos x$ with $y(0) = 1$. State the interval containing $x = 0$ on which the solution is valid.

(e) Explain why a general solution contains one arbitrary constant, and what the initial condition does to it.

Models

4. A metal bar is taken from a furnace at $300°$C into a room at $25°$C. Its temperature $\theta\,°$C after t minutes satisfies $\dfrac{d\theta}{dt} = -0.05(\theta - 25)$.

(a) Solve the equation to find θ in terms of t.

(b) Find the time the bar takes to cool to $100°$C.

5. The number N of bacteria in a culture grows at a rate proportional to N. Initially $N = 500$, and after 2 hours $N = 800$.

(a) Write down a differential equation for N, and solve it to find N in terms of t, the time in hours.

(b) Find N after 5 hours.

6. A population of n thousand animals satisfies $\dfrac{dn}{dt} = 0.3n(2-n)$, where t is the time in years, and $n(0) = 0.5$.

(a) Use partial fractions to show that $n = \dfrac{2}{1+3e^{-0.6t}}$.

(b) Find the time at which $n = 1.5$, and state the value that n approaches as t increases.

Concept check

7. A student solves $\dfrac{dy}{dx} = xy$, for $y > 0$, and writes

$$\ln y = \tfrac{1}{2}x^2 + C \quad \Rightarrow \quad y = e^{x^2/2} + C.$$

Identify the error, and give the correct general solution.

16.2 Homogeneous Differential Equations

Some equations cannot be separated as they stand, but a substitution makes them separable.

A first-order equation is **homogeneous** if the right-hand side depends only on the ratio $\dfrac{y}{x}$, and not on x and y separately.

Equations of this kind appear when a problem has no fixed size, only directions and proportions.

- **Pursuit curves.** A ferry always aims at the dock while a current pushes it sideways. At each moment only the *direction* from the ferry to the dock matters, and that direction is set by $\dfrac{y}{x}$ alone. You will solve the ferry in the challenge problems.
- **Reflector shapes.** The curve of a car headlight or a satellite dish is fixed by requiring every ray from the focus to leave parallel. That condition involves only angles, so the equation depends on the ratio.
- **Scale-free rules.** If multiplying both quantities by the same factor leaves a rule unchanged, the rule can only involve their ratio, and the equation is homogeneous.

The substitution $y = vx$ turns such an equation into a separable one. Since $y = vx$, the product rule gives $\dfrac{dy}{dx} = v + x\dfrac{dv}{dx}$.

KEY For $\dfrac{dy}{dx} = f\left(\dfrac{y}{x}\right)$, substitute $y = vx$, so $\dfrac{dy}{dx} = v + x\dfrac{dv}{dx}$.

After the substitution the equation becomes $v + x\dfrac{dv}{dx} = f(v)$, that is $x\dfrac{dv}{dx} = f(v) - v$. This has v on one side and x on the other, so it is always separable. The method has four steps.

1. Check that the right side depends only on $\dfrac{y}{x}$: divide its top and bottom by the highest power of x.
2. Substitute $y = vx$ and $\dfrac{dy}{dx} = v + x\dfrac{dv}{dx}$.
3. Separate the variables and integrate.

4. Replace v by $\dfrac{y}{x}$.

Worked through. To solve $\frac{dy}{dx} = \frac{x+y}{x}$, rewrite the right side in terms of $\frac{y}{x}$: $\dfrac{x+y}{x} = 1 + \dfrac{y}{x}$. Substitute $y = vx$, $\dfrac{dy}{dx} = v + x\dfrac{dv}{dx}$:

$$v + x\frac{dv}{dx} = 1 + v \implies x\frac{dv}{dx} = 1.$$

This is separable:

$$\int dv = \int \frac{1}{x}\,dx \implies v = \ln|x| + C.$$

Finally replace $v = \dfrac{y}{x}$:

$$y = x(\ln|x| + C).$$

Here the v on each side cancelled. In general it does not, but $x\dfrac{dv}{dx} = f(v) - v$ is still separable.

EXAMPLE 16.3

Solve $\dfrac{dy}{dx} = \dfrac{2x^2 + y^2}{xy}$ for $x > 0$.

Divide the top and bottom by x^2 to show that the right side depends only on the ratio:

$$\frac{2x^2 + y^2}{xy} = \frac{2 + (y/x)^2}{y/x}.$$

Substitute $y = vx$, so $\dfrac{dy}{dx} = v + x\dfrac{dv}{dx}$:

$$v + x\frac{dv}{dx} = \frac{2 + v^2}{v} \implies x\frac{dv}{dx} = \frac{2 + v^2}{v} - v = \frac{2}{v}.$$

Again the v terms cancel, leaving a separable equation:

$$\int v\,dv = \int \frac{2}{x}\,dx \implies \frac{v^2}{2} = 2\ln x + C.$$

Replace $v = \dfrac{y}{x}$ and clear the fraction:

$$y^2 = 2x^2(2\ln x + C).$$

YOUR TURN 16.2 Homogeneous Differential Equations

Recognising a homogeneous equation

8. State whether each equation is homogeneous, giving a reason.

(a) $\dfrac{dy}{dx} = \dfrac{x+y}{x-y}$

(b) $\dfrac{dy}{dx} = \dfrac{x^2+y}{x}$

(c) $\dfrac{dy}{dx} = \dfrac{y^2+xy}{x^2}$

General solutions

9. Use the substitution $y = vx$ to solve each equation, for $x > 0$.

(a) $\dfrac{dy}{dx} = \dfrac{y}{x}$

(b) $\dfrac{dy}{dx} = \dfrac{y-x}{x}$

(c) $\dfrac{dy}{dx} = \dfrac{x-y}{x}$

(d) $\dfrac{dy}{dx} = \dfrac{x+2y}{x}$

10.

(a) Show that $\dfrac{dy}{dx} = \dfrac{x^2+3y^2}{2xy}$ is homogeneous, by writing the right-hand side in terms of $\dfrac{y}{x}$.

(b) Hence find its general solution, for $x > 0$.

Particular solutions

11. Find the particular solution of each equation, for $x > 0$.

(a) $\dfrac{dy}{dx} = \dfrac{2y-x}{x}$ with $y(1) = 3$.

(b) $\dfrac{dy}{dx} = \dfrac{x^2+y^2}{xy}$ with $y(1) = 2$.

Concept check

12. A student solves $\dfrac{dy}{dx} = \dfrac{x+y}{x}$, for $x > 0$, with the substitution $y = vx$, and writes $\dfrac{dy}{dx} = \dfrac{dv}{dx}$, so that $\dfrac{dv}{dx} = 1 + v$. Identify the error, and find the correct general solution.

16.3 The Integrating Factor

A linear equation often appears when a quantity is added and removed at the same time, and each rate depends at most linearly on the amount present.

- **Mixing tanks.** Liquid flows into a tank and drains out of it, and the dissolved substance obeys a linear equation. When the volume itself changes, the equation is linear but not separable, so the integrating factor is the method that applies.
- **Drug infusion.** A drug enters the blood at a steady rate through a drip, while the body clears it at a rate proportional to the amount already present.
- **Electrical circuits.** In a circuit with a resistor and a capacitor, the charge satisfies a linear equation driven by the applied voltage.
- **Savings with deposits.** An account earns interest continuously while money is paid in at a steady rate.

A **linear** first-order differential equation has the form

$$\frac{dy}{dx} + P(x)\,y = Q(x).$$

It is usually not separable. The method is to multiply through by a chosen function, the **integrating factor**, which turns the left side into the derivative of a single product. The coefficient of $\dfrac{dy}{dx}$ must be 1 first: if the equation starts as $x\dfrac{dy}{dx} + 2y = x^3$, divide by x before reading off $P(x)$.

KEY · IB For $\dfrac{dy}{dx} + P(x)y = Q(x)$, the integrating factor is

$$I = e^{\int P(x)\,dx}.$$

After multiplying by I, the left side becomes exactly $\dfrac{d}{dx}(Iy)$. The reason is the product rule: since $I = e^{\int P\,dx}$, its derivative is $I' = PI$, so

$$\frac{d}{dx}(Iy) = I\frac{dy}{dx} + I'y = I\frac{dy}{dx} + PIy = I\left(\frac{dy}{dx} + Py\right).$$

You then integrate both sides directly.

Worked through. In $\frac{dy}{dx} + \frac{1}{x}y = x^2$, for $x > 0$, $P(x) = \frac{1}{x}$, so the integrating factor is

$$I = e^{\int \frac{1}{x}\,dx} = e^{\ln x} = x.$$

Multiply through by x:

$$x\frac{dy}{dx} + y = x^3 \implies \frac{d}{dx}(xy) = x^3.$$

The left side is now a single derivative. Integrate both sides:

$$xy = \frac{x^4}{4} + C \implies y = \frac{x^3}{4} + \frac{C}{x}.$$

The method exists to create that $\dfrac{d}{dx}(xy)$ on the left; everything after is ordinary integration.

EXAMPLE 16.4

Solve $\dfrac{dy}{dx} + y = e^{x}$.

$P(x) = 1$, so $I = e^{\int 1\,dx} = e^{x}$. Multiply through:

$$\frac{d}{dx}(e^{x}y) = e^{x}\cdot e^{x} = e^{2x}.$$

Integrate: $e^{x}y = \tfrac{1}{2}e^{2x} + C$, hence $y = \tfrac{1}{2}e^{x} + Ce^{-x}$.

The last worked example of this section is a mixing problem of the first kind. To model a well-stirred mixing tank, write $\dfrac{dS}{dt} = (\text{rate in}) - (\text{rate out})$, where S is the amount of dissolved substance. It enters at the inflow concentration times the inflow rate, and leaves at the current concentration $\dfrac{S}{V}$ times the outflow rate. When the two flow rates differ, the volume changes too: $V(t) = V_0 + (\text{inflow rate} - \text{outflow rate})\,t$.

EXAMPLE 16.5

A 100-litre vat in a brewery holds 5 kg of dissolved sugar. Syrup carrying 0.2 kg of sugar per litre flows in at 1 litre per minute, the vat is kept well stirred, and the mixture drains at 2 litres per minute. Find the sugar $S(t)$ in the vat at time t minutes.

The volume shrinks: $V(t) = 100 - t$. Sugar arrives at $0.2 \times 1 = 0.2$ kg min^{-1} and leaves at concentration times outflow, $\dfrac{S}{100-t}\times 2$. So

$$\frac{dS}{dt} = 0.2 - \frac{2S}{100-t} \;\Longrightarrow\; \frac{dS}{dt} + \frac{2}{100-t}\,S = 0.2.$$

The integrating factor is

$$I = e^{\int \frac{2}{100-t}\,dt} = e^{-2\ln(100-t)} = (100-t)^{-2}.$$

Multiplying through turns the left side into a single derivative:

$$\frac{d}{dt}\Big(S\,(100-t)^{-2}\Big) = 0.2\,(100-t)^{-2} \;\Longrightarrow\; S\,(100-t)^{-2} = \frac{0.2}{100-t} + C.$$

So $S = 0.2(100-t) + C(100-t)^{2}$, and $S(0) = 5$ gives $C = -0.0015$:

$$S(t) = 0.2(100-t) - 0.0015(100-t)^{2}.$$

Check the endpoints: $S(0) = 5$, and $S(100) = 0$, an empty vat holds no sugar. In between, the sugar first *rises*, to about 6.7 kg near $t = 33$, and then falls. That rise is not obvious from the description, and it is the equation rather than intuition that reveals it.

16.3.1 Choosing a method

Three exact methods are now available, and a question may not say which to use. The form of the equation decides it, and the test is quick.

What the equation looks like	Method	Example
The variables can be split, $\frac{dy}{dx} = g(x)h(y)$	separate and integrate	$\frac{dy}{dx} = 2xy^2$
The right side depends only on $\frac{y}{x}$	substitute $y = vx$	$\frac{dy}{dx} = \frac{x+y}{x}$
Linear in y: $\frac{dy}{dx} + P(x)y = Q(x)$	integrating factor $e^{\int P\,dx}$	$\frac{dy}{dx} + \frac{1}{x}y = x^2$
None of the above	Euler's method, numerically	$\frac{dy}{dx} = x^2 + y^2$

Try them in that order, because each test is quicker than the one after it. Note that the first two overlap: an equation can be both separable and homogeneous, in which case separating directly is the shorter route. A linear equation can overlap too. When Q/P is constant, $\frac{dy}{dx} + P(x)y = Q(x)$ rearranges to $\frac{dy}{dx} = P(x)\left(\frac{Q}{P} - y\right)$ and separates at once, which is how $\frac{dy}{dx} + y = 4$ and the cooling equation of §16.1 are solved. In general, though, Q/P is not constant, the equation does not separate, and the integrating factor is the method to use.

YOUR TURN 16.3 The Integrating Factor

Solving a linear equation

13. Solve each linear equation.

(a) $\dfrac{dy}{dx} + y = 4$

(b) $\dfrac{dy}{dx} + 2y = 6$

(c) $\dfrac{dy}{dx} + 2y = 4x$

(d) $\dfrac{dy}{dx} - y = e^{2x}$

14.

(a) Write down the integrating factor for $\dfrac{dy}{dx} + \dfrac{3}{x}y = x^2$, for $x > 0$.

(b) Solve $\dfrac{dy}{dx} + \dfrac{2}{x}y = x$ for $x > 0$.

(c) Solve $\dfrac{dy}{dx} - \dfrac{1}{x}y = x^2$ with $y(1) = 1$, for $x > 0$.

(d) Explain what the integrating factor is chosen to achieve on the left-hand side.

Rearranging into standard form first

15. Divide by the coefficient of $\dfrac{dy}{dx}$ first, then solve each equation.

(a) $x\dfrac{dy}{dx} + 3y = x^2$, for $x > 0$.

(b) $(x^2 + 1)\dfrac{dy}{dx} + 2xy = 4x$

(c) $\dfrac{dy}{dx} + y\tan x = \cos x$, for $0 \le x < \frac{\pi}{2}$.

(d) $\cos x\dfrac{dy}{dx} + y\sin x = 1$, for $0 \le x < \frac{\pi}{2}$.

A mixing tank

16. A tank holds 50 litres of water in which 2 kg of salt is dissolved. Brine containing 0.1 kg of salt per litre flows in at 3 litres per minute, the tank is kept well stirred, and the mixture drains at 2 litres per minute. Let S kg be the salt in the tank after t minutes.

(a) Show that $\dfrac{dS}{dt} + \dfrac{2S}{50 + t} = 0.3$.

(b) Solve the equation to find S in terms of t.

(c) Find the salt in the tank after 10 minutes.

Choosing a method

17. For each equation, state which method applies, and find the general solution.

(a) $\dfrac{dy}{dx} = xe^{-y}$

(b) $\dfrac{dy}{dx} - \dfrac{2y}{x} = x^3$, for $x > 0$.

(c) $\dfrac{dy}{dx} = \dfrac{y^2 + 2xy}{x^2}$, for $x > 0$.

Concept check

18. A student solves $\dfrac{dy}{dx} + \dfrac{y}{x} = 1$, for $x > 0$, with the integrating factor x, and writes

$$xy = \tfrac{1}{2}x^2 + C \quad \Rightarrow \quad y = \tfrac{1}{2}x + C.$$

Identify the error, and show that the student's answer does not satisfy the equation unless $C = 0$.

16.4 Maclaurin Series

Many functions can be written as an infinite polynomial. This idea stands apart from the three methods above, and it has two uses: it approximates a function, and it gives a fourth route to solving a differential equation.

There is a reason a polynomial is wanted at all. A polynomial is evaluated using only addition and multiplication, which are fast and exact to program. A function such as $\sin x$ has no such direct recipe, so computer software often evaluates it from a polynomial approximation built in this way. (Many calculators use a different method based on repeated addition and halving, but the aim is the same: reduce the function to arithmetic.)

The method is to match the function at a single point. Near $x = 0$, we look for a polynomial with the same value as the function, then the same gradient, then the same concavity, then the same higher derivatives. The more derivatives agree, the wider the interval around 0 on which the polynomial stays close to the curve. The result is the **Maclaurin series**.

The coefficients are completely determined by the function. Suppose the series exists and write it as

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots$$

Putting $x = 0$ makes every term except the first equal to zero, so $a_0 = f(0)$. Differentiating gives $f'(x) = a_1 + 2a_2x + 3a_3x^2 + \cdots$, and $x = 0$ now leaves $a_1 = f'(0)$. Differentiating again gives $f''(0) = 2a_2$, and once more $f'''(0) = 6a_3 = 3!\,a_3$. Each differentiation removes the constant term and multiplies each remaining coefficient by its power, so in general

$f^{(n)}(0) = n!\,a_n$.

KEY · IB

$$f(x) = f(0) + f'(0)\,x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots$$

Worked through. To find the Maclaurin series for $f(x) = e^x$, note that every derivative of e^x is e^x, and $e^0 = 1$, so $f^{(n)}(0) = 1$ for all n:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

This series can be taken as the definition of e^x, and it computes the function to any accuracy you want. It also explains why e^x is its own derivative: differentiate the series term by term and you get the same series back.

The series below are provided in the formula booklet. Each one equals its function exactly on an interval around 0: for every x in the case of e^x, $\sin x$ and $\cos x$, but only for $-1 < x \leq 1$ for $\ln(1+x)$ and for $|x| \leq 1$ for $\arctan x$. Truncating a series gives a polynomial approximation that is very accurate near $x = 0$.

KEY · IB

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \qquad \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots \qquad \ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$$

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots$$

One more expansion is named by the syllabus but printed elsewhere in the booklet, so it is easy to overlook. The binomial family $(1+x)^p$ is a Maclaurin series like any other, and it is the extended binomial theorem of Ch. 2 (booklet 1.10) seen from this side.

KEY For $p \in \mathbb{Q}$ and $|x| < 1$,

$$(1+x)^p = 1 + px + \frac{p(p-1)}{2!}x^2 + \frac{p(p-1)(p-2)}{3!}x^3 + \cdots.$$

Like $\ln(1+x)$ and $\arctan x$, it carries a validity condition, because, unless p is a non-negative integer, it does not converge for every x.

EXAMPLE 16.6

Use a Maclaurin series to approximate $\sqrt{1.1}$ to four decimal places.

Write $\sqrt{1.1} = (1+x)^{1/2}$ with $x = 0.1$, which satisfies $|x| < 1$. With $p = \frac{1}{2}$,

$$(1+x)^{1/2} = 1 + \tfrac{1}{2}x - \tfrac{1}{8}x^2 + \tfrac{1}{16}x^3 - \cdots.$$

Substituting $x = 0.1$ gives

$$1 + 0.05 - 0.00125 + 0.0000625 = 1.0488125.$$

The true value is $\sqrt{1.1} = 1.0488088\ldots$, so four terms agree to four decimal places, 1.0488. Each term is smaller than the last by roughly a factor of x, which is why these series are useful: the smaller x is, the fewer terms you need.

The figure below shows the first three odd-degree truncations of the sine series: the line $y = x$, the cubic, and the quintic. Each new term extends the interval over which the polynomial and the curve agree.

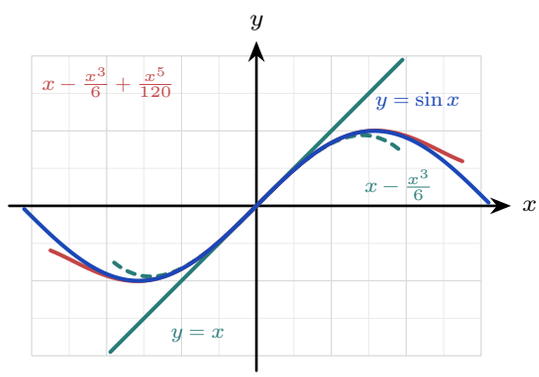


Figure 16.1 The first three odd-degree truncations of the sine series against $y = \sin x$. Each extra term widens the interval on which the polynomial and the curve agree, but every truncation eventually leaves the curve.

The same convergence can be watched numerically. Below are the successive truncations of the sine series at $x = 1$, against the true value $\sin 1 = 0.841471$.

terms up to	x	x^3	x^5	x^7
estimate	1	0.833333	0.841667	0.841468
error	0.159	0.0081	0.0002	0.0000027

Each new term divides the error by a factor of tens, and that factor grows as more terms are taken: the ratios along this row are about 20, 40 and 70. So a short polynomial is easy to compute, and each extra term gives more accuracy for a little more work.

16.4.1 Building new series

New series are usually not built from the definition. They are made from the standard five by four operations, all of which the syllabus names.

- **Substitute.** Replacing x with $-x^2$ in the series for e^x gives

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \cdots$$

- **Differentiate.** Differentiating the sine series term by term produces exactly the cosine series: the identity $\frac{d}{dx}\sin x = \cos x$ in polynomial form.
- **Integrate.** The geometric series of Ch. 2 gives $\dfrac{1}{1+x^2} = 1 - x^2 + x^4 - \cdots$ for $|x| < 1$, and integrating term by term yields $\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots$. The booklet's arctangent series is simply the geometric series integrated.
- **Multiply.** For a product, expand each factor and collect powers up to the degree you need.

EXAMPLE 16.7

Find the Maclaurin series of $e^x \sin x$ up to the term in x^3.

Multiply the two standard series, keeping only terms of degree ≤ 3:

$$\left(1 + x + \frac{x^2}{2} + \frac{x^3}{6}\right)\left(x - \frac{x^3}{6}\right) = x + x^2 + \left(\frac{1}{2} - \frac{1}{6}\right)x^3 + \cdots = x + x^2 + \frac{x^3}{3} + \cdots$$

The work is careful collection of terms. Decide the target degree first, and discard anything above it before multiplying rather than after.

16.4.2 Small angles, limits, and series from differential equations

Truncating at the first term or two gives the **small-angle approximations** used throughout physics. For θ near 0, in radians,

$$\sin\theta \approx \theta, \qquad \cos\theta \approx 1 - \frac{\theta^2}{2}, \qquad \tan\theta \approx \theta.$$

These are not conveniences invented by physicists; they are the first one or two terms of the series, with the rest discarded. The pendulum formula depends on the first of them. For small swings the restoring force involves $\sin\theta$, and replacing $\sin\theta$ by θ is exactly what turns the equation of motion into simple harmonic motion, which can then be solved.

The series also measures the error of the approximation. The first term discarded is $\frac{\theta^3}{6}$, so that is roughly the error, and the relative error stays under 1% for swings up to about $14°$. Beyond that the pendulum's real period moves away from the textbook formula, by an amount large enough to measure.

The same truncation resolves indeterminate limits: replace each function by the first terms

of its series, and the limit can be read off.

EXAMPLE 16.8

Use Maclaurin series to evaluate $\lim_{x\to 0} \dfrac{e^x - 1 - x}{x^2}$.

Expand the numerator with the series for e^x:

$$e^x - 1 - x = \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots\right) - 1 - x = \frac{x^2}{2} + \frac{x^3}{6} + \cdots$$

Divide by x^2:

$$\frac{e^x - 1 - x}{x^2} = \frac{1}{2} + \frac{x}{6} + \cdots \xrightarrow{x\to 0} \frac{1}{2}.$$

The series resolves an indeterminate limit as neatly as L'Hôpital's rule, and it shows *why* the answer is $\frac{1}{2}$: that number is the leading coefficient.

A Maclaurin series can also be built directly *from a differential equation*, by differentiating the equation itself to obtain the derivatives at 0. This joins the two halves of the chapter.

EXAMPLE 16.9

A function satisfies $\dfrac{dy}{dx} = 1 + y^2$ with $y(0) = 0$. Find the Maclaurin series of y up to the term in x^3.

Take the derivatives at 0 from the equation itself. First, $y'(0) = 1 + 0 = 1$. Differentiating the equation,

$$y'' = 2y\,y' \implies y''(0) = 0, \qquad y''' = 2(y')^2 + 2y\,y'' \implies y'''(0) = 2.$$

Substitute these into the Maclaurin formula:

$$y = 0 + x + \frac{0}{2!}x^2 + \frac{2}{3!}x^3 + \cdots = x + \frac{x^3}{3} + \cdots$$

You may recognise the answer: $y = \tan x$ solves this equation, since $\frac{d}{dx}\tan x = \sec^2 x = 1 + \tan^2 x$, and these are indeed the first terms of its series. The equation was never solved. Each derivative at 0 was read off from the equation itself, one at a time, and that was enough to build the series.

TIP Maclaurin series suggest several good IA topics. Two worth exploring: Madhava's series $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots$, which comes from setting $x = 1$ in the arctangent series, and why it converges so slowly that a better choice such as $\arctan\frac{1}{\sqrt{3}}$ reaches the same accuracy in a fraction of the terms; or how software evaluates $\sin x$ quickly, by first reducing the angle to a small interval and then using a short polynomial.

YOUR TURN 16.4 Maclaurin Series

Using the standard series

19. Find the first three non-zero terms of the Maclaurin series for each function.

(a) e^{2x}
(b) $\cos 3x$
(c) $\ln(1+2x)$
(d) e^{-x^2}

20. Find the first two non-zero terms of the Maclaurin series for each function.

(a) $\sin 2x$
(b) $\sin(x^2)$
(c) $\arctan 2x$
(d) $\sqrt{1+x}$

Series from the definition

21. Use the Maclaurin formula $f(x) = f(0) + f'(0)x + \dfrac{f''(0)}{2!}x^2 + \cdots$, differentiating each function yourself, to find its series up to the term in x^4.

(a) $f(x) = \cos 2x$
(b) $f(x) = \ln(\cos x)$

Products of series

22. Find the Maclaurin series of each product up to the term in x^3.

(a) $e^x \cos x$
(b) $e^x \ln(1+x)$

Differentiating and integrating a series

23.

(a) Differentiate the series $\arctan x = x - \dfrac{x^3}{3} + \dfrac{x^5}{5} - \cdots$ term by term, and verify that the result is the expansion of $\dfrac{1}{1+x^2}$.

(b) Differentiate the series for $\ln(1+x)$ term by term, and verify that the result is the expansion of $\dfrac{1}{1+x}$.

(c) Integrate the geometric series $\dfrac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots$ term by term to find the first four terms of the series for $-\ln(1-x)$, for $|x| < 1$.

(d) Use the first three non-zero terms of the series for e^{-x^2} to estimate $\displaystyle\int_0^{0.5} e^{-x^2}\,dx$, correct to four decimal places.

Approximating values

24.

(a) Use the first three terms of the series for e^x to estimate $e^{0.1}$.

(b) Use the first three terms of the cosine series to estimate $\cos 0.2$ to four decimal places.

(c) Compare each estimate with the true value, and state which is the more accurate. Explain why.

25. The small-angle approximations are $\sin\theta \approx \theta$, $\cos\theta \approx 1 - \dfrac{\theta^2}{2}$ and $\tan\theta \approx \theta$, for θ in radians close to 0.

(a) Find the value that $\dfrac{\sin 3\theta \tan\theta}{1-\cos 2\theta}$ is close to when θ is small.

(b) Show that $\cos\theta + \theta\sin\theta \approx 1 + \dfrac{\theta^2}{2}$ when θ is small.

Limits by series

26. Use Maclaurin series to evaluate each limit.

(a) $\displaystyle\lim_{x\to 0} \frac{1-\cos x}{x^2}$

(b) $\displaystyle\lim_{x\to 0} \frac{\ln(1+x) - x}{x^2}$

(c) $\displaystyle\lim_{x\to 0} \frac{x - \arctan x}{x^3}$

Series from a differential equation

27. A function satisfies $\dfrac{dy}{dx} = x + y$ with $y(0) = 1$.

(a) Write down $y'(0)$.

(b) By differentiating the equation, find $y''(0)$ and $y'''(0)$.

(c) Hence find the Maclaurin series for y up to the term in x^3.

(d) The exact solution is $y = 2e^x - x - 1$. Use the series to estimate $y(0.3)$, and compare it with the exact value.

28. A function satisfies $\dfrac{dy}{dx} = x - y^2$ with $y(0) = 1$.

(a) Find $y'(0)$, $y''(0)$ and $y'''(0)$.

(b) Hence find the Maclaurin series for y up to the term in x^3.

Concept check

29. A student replaces x with $3x$ in the cosine series and writes $\cos 3x = 1 - \dfrac{3x^2}{2} + \dfrac{3x^4}{24} - \cdots$. Identify the error, and give the correct first three terms.

30. A student estimates $\ln 3$ by putting $x = 2$ into $\ln(1+x) = x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \cdots$, and gets $2 - 2 + \frac{8}{3} \approx 2.67$. Explain why this estimate is not valid, and why taking more terms does not help.

16.5 Euler's Method

Not every differential equation can be solved exactly, and many of the equations that describe the real world cannot. When no exact solution exists, the solution can still be approximated numerically. **Euler's method** does this by moving along the curve in short straight steps, taking the gradient of each step from the differential equation itself.

The idea is the tangent line of Ch. 13. Near a known point the curve is close to its tangent, so moving a small distance h along the tangent lands close to the curve:

$$y \approx y_0 + \underbrace{f(x_0, y_0)}_{\text{gradient}} \times \underbrace{h}_{\text{step}}.$$

The differential equation supplies the gradient at whatever point you have reached, so the same move can be repeated from the new point, and again from the one after that.

KEY · IB

$$y_{n+1} = y_n + h\, f(x_n, y_n), \qquad x_{n+1} = x_n + h$$

Here h is the fixed **step length**. Each step has the correct gradient only at its starting point and moves away from the curve after that, so shorter steps stay closer, at the cost of needing more of them.

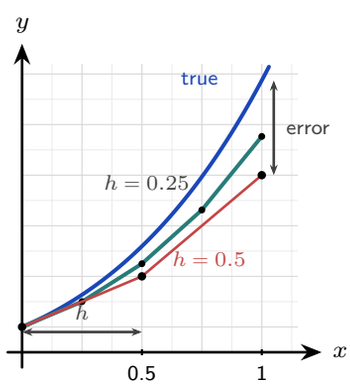


Figure 16.2 Euler's method on $\frac{dy}{dx} = x + y$ with $y(0) = 1$. Each straight step leaves the curve, which bends upward away from it, so the estimate falls short and the gap at $x = 1$ is the error. Halving the step length cuts that gap substantially, here from about 0.94 to 0.55.

In the figure, the gap at the far end is the *error*. The teal path takes steps half as long, and it ends closer to the curve.

The direction of the error follows from the concavity. Where the solution is concave up, each tangent step lies below the curve; where it is concave down, each step lies above it. For a small step length this is the usual pattern: Euler's method under-estimates a concave-up solution and over-estimates a concave-down one, so the sign of $\frac{d^2y}{dx^2}$ tells you which way your answer is wrong before you check it.

Worked through. Take the equation in the figure, $\frac{dy}{dx} = x + y$ with $y(0) = 1$, and estimate $y(0.3)$ with $h = 0.1$. Apply $y_{n+1} = y_n + h(x_n + y_n)$ three times. A table is the clearest way to organise the work, and it shows each step clearly to whoever marks the work.

n	x_n	y_n	$f = x_n + y_n$	$h\,f$
0	0	1	1	0.1
1	0.1	1.1	1.2	0.12
2	0.2	1.22	1.42	0.142
3	0.3	1.362		

Each y_n is the previous one plus the $h\,f$ on the line above. So $y(0.3) \approx 1.362$, against an exact value of about 1.3997.

The estimate is too small, and the reason is visible in the figure: this curve is concave up, so a tangent step lies below it and the estimate falls short, as the concavity pattern above predicts.

The method matters most when no exact solution is available. Make the resistance on a falling body depend on the shape of the object and on the air density, which changes with height, and the equation has no solution among the functions of this course, but Euler's method still produces a table of values. This is, in miniature, how weather forecasts are computed: equations too hard to solve exactly, stepped forward numerically, millions of times.

A quantity stays at a steady value only where its rate of change is zero. For an equation $\dfrac{dv}{dt} = f(v)$, solve $f(v) = 0$ to find the terminal or **equilibrium** value; this does not require solving the differential equation.

EXAMPLE 16.10

A skydiver falls from rest, and air resistance gives $\dfrac{dv}{dt} = 9.8 - 0.1v^2$ (in m s^{-1} and seconds). Use Euler's method with $h = 0.5$ to estimate the speed after 1.5 seconds, and compare with the terminal velocity.

Apply $v_{n+1} = v_n + 0.5(9.8 - 0.1v_n^2)$ three times from $v_0 = 0$:

$$v_1 = 0 + 0.5(9.8) = 4.9$$
$$v_2 = 4.9 + 0.5(9.8 - 0.1(4.9)^2) = 4.9 + 0.5(7.399) = 8.60$$
$$v_3 = 8.60 + 0.5(9.8 - 0.1(8.60)^2) = 8.60 + 0.5(2.404) = 9.80$$

So $v(1.5) \approx 9.80$ m s^{-1}. The terminal velocity is where the acceleration is zero: $9.8 - 0.1v^2 = 0 \implies v = \sqrt{98} \approx 9.90$ m s^{-1}. The exact solution gives $v(1.5) \approx 8.93$ m s^{-1}, so this estimate is about 10% too high. The reason is the concavity pattern stated earlier in this section: the graph of v against t is concave down, so each tangent step lies above the curve and Euler's method overestimates. A smaller step length reduces the error. Each bracket $9.8 - 0.1v_n^2$ is the net force per unit mass, and it decreases towards zero as the speed increases.

YOUR TURN 16.5 Euler's Method

Carrying out Euler's method

31. Use Euler's method for each of the following.

(a) $\dfrac{dy}{dx} = 2x$, $y(0) = 1$; use $h = 0.5$ to estimate $y(1)$.

(b) $\dfrac{dy}{dx} = y$, $y(0) = 1$; use $h = 0.1$ to estimate $y(0.2)$.

(c) $\dfrac{dy}{dx} = x - y$, $y(0) = 1$; use $h = 0.2$ to estimate $y(0.4)$.

(d) $\dfrac{dy}{dx} = xy$, $y(1) = 1$; use $h = 0.1$ to estimate $y(1.2)$.

32.

(a) For $\dfrac{dy}{dx} = xy$ with $y(1) = 2$, take one step with $h = 0.2$ to estimate $y(1.2)$.

(b) For $\dfrac{dy}{dx} = x^2 + y$ with $y(0) = 0$, take two steps with $h = 0.5$ to estimate $y(1)$.

Accuracy and step length

33. This question uses the equation $\dfrac{dy}{dx} = y$ with $y(0) = 1$ from Question 31(b).

(a) Solve the equation exactly, and compare the exact $y(0.2)$ with your answer to Question 31(b).

(b) State whether Euler's method over- or under-estimates the solution, and explain why in terms of the shape of the curve.

(c) Estimate $y(0.2)$ again, using $h = 0.05$.

(d) Compare the errors of the two estimates, and state the effect of halving the step length.

Equilibrium values

34.

(a) A falling object has $\dfrac{dv}{dt} = 10 - 0.4v$, with v in m s^{-1}. Find its terminal velocity.

(b) A skydiver falls from rest, with $\dfrac{dv}{dt} = 9.8 - 0.2v^2$ and v in m s^{-1}. Find the terminal velocity.

(c) For the skydiver in part (b), use Euler's method with $h = 0.5$ to estimate $v(1)$.

(d) Explain why the estimate in part (c) cannot be accurate.

Concept check

35. A student says: "*Euler's method underestimates, because each step follows a tangent.*" Use Euler's method on $\dfrac{dy}{dx} = -2x$, $y(0) = 0$, with $h = 0.5$, to estimate $y(1)$. Hence show that the claim is false, and state when Euler's method does underestimate.

Chapter 16 · Differential Equations

Reflections

TOK Many laws of physics are differential equations: Newton's second law, radioactive decay, the spread of an epidemic. For the equations of this chapter, a rule for change together with one starting value determines the whole future of the system. This is mathematical determinism. Yet the same kind of equation can be so sensitive to its starting value that long-term prediction becomes impossible in practice, the phenomenon of *chaos*. Determined but unpredictable: what does that tell us about the limits of knowledge?

TOK A truncated series is deliberately wrong. It is chosen because it is simple enough to compute, and every extra term bought accuracy at the cost of that simplicity: four terms of the sine series give $\sin 1$ correct to five decimal places, and one term gives it to none.

Is there always a trade-off between accuracy and simplicity? And when a model is knowingly approximate, as every model in this chapter is, what makes it good enough to act on?

IM The series of this chapter came first from India. Mathematicians of the Kerala school, notably Madhava in the 14th century, found the expansions of $\sin x$, $\cos x$ and $\arctan x$ roughly three centuries before Newton, Leibniz and Maclaurin; the arctangent series, usually named after Gregory and Leibniz, was known in Kerala generations earlier. Numerical methods have a similar story. In the 15th century Jamshīd al-Kāshī, working at the observatory in Samarkand, needed $\sin 1^\circ$ to great precision for his astronomical tables. No formula gave it, so he wrote down an equation the value satisfied and solved it by repeated approximation, each round feeding its answer into the next, reaching sixteen correct decimal places. Neither idea was the work of one place or one name.

EXT The five standard series behave differently at the edges. The series for e^x, $\sin x$ and $\cos x$ converge for every x, but $\arctan x$ only for $|x| \le 1$ and $\ln(1+x)$ only for $-1 < x \le 1$, while the geometric series they are built from needs $|x| < 1$. Each series has a *radius of convergence*, and outside it the polynomial does not merely lose accuracy: it diverges completely. Finding that radius, and explaining why the radius is finite for $\ln(1+x)$ but infinite for e^x, is the beginning of complex analysis, where the answer depends on where the function is undefined in the complex plane. See Ch. 17.

Maclaurin series turn calculus into algebra: integrals with no elementary antiderivative, like $\int e^{-x^2}\,dx$, can be integrated term by term once the function is a series. Series also solve differential equations that no other method in this course can reach, by assuming the answer is a power series and matching coefficients. The infinite polynomial is a far-reaching idea in analysis.

End-of-Chapter Problems

1. A population grows so that $\dfrac{dP}{dt} = 0.2P$, with $P = 500$ when $t = 0$.

(a) Solve the differential equation for $P(t)$. [4]

(b) Find the population when $t = 10$. [2]

2. A body cools according to $\dfrac{d\theta}{dt} = -0.1(\theta - 25)$, with $\theta(0) = 90$ (in °C).

(a) Solve for $\theta(t)$. [4]

(b) Find the temperature after 10 minutes. [2]

(c) State the temperature the body approaches in the long run. [1]

3. Consider $\dfrac{dy}{dx} = \dfrac{x^2 + 3y^2}{2xy}$ for $x > 0$.

(a) Use the substitution $y = vx$ to show that $x\dfrac{dv}{dx} = \dfrac{1 + v^2}{2v}$. [4]

(b) Hence find the general solution. [3]

4. Consider $\dfrac{dy}{dx} + \dfrac{1}{x}y = 3x$ for $x > 0$.

(a) Find the integrating factor. [2]

(b) Solve the equation given $y(1) = 2$. [4]

5. Consider $\dfrac{dy}{dx} = x + y$, with $y(0) = 1$.

(a) Use Euler's method with $h = 0.25$ to estimate $y(0.5)$, showing each step. [4]

(b) Solve the equation exactly (integrating factor) and find the true $y(0.5)$. [4]

(c) Comment on the size and direction of the error. [2]

6. Let $f(x) = \cos x$.

(a) Find $f(0), f'(0), f''(0), f'''(0), f^{(4)}(0)$. [3]

(b) Hence write the Maclaurin series of $\cos x$ up to the term in x^4. [2]

(c) Use it to estimate $\cos(0.2)$. [2]

7. EXACT

(a) Write down the Maclaurin series for $\sin x$ up to the term in x^5. [2]

(b) Hence evaluate $\lim\limits_{x \to 0} \dfrac{x - \sin x}{x^3}$. [3]

8.

(a) Write down the Maclaurin series for e^x up to the term in x^3. [1]

(b) By substituting $-x^2$, find the series for e^{-x^2} up to x^4. [2]

(c) Hence estimate $\displaystyle\int_0^1 e^{-x^2}\,dx$ using terms up to x^4. [4]

9. A bowl of miso soup at $90°$C is left in a $25°$C room and obeys $\dfrac{d\theta}{dt} = -k(\theta - 25)$. After 10 minutes it has cooled to $65°$C.

(a) Solve the differential equation, expressing θ in terms of k and t. [4]

(b) Find k. [2]

(c) Find how long, from the start, the soup takes to reach $40°$C. [2]

10. Consider $\dfrac{dy}{dx} = \dfrac{y^2 + xy}{x^2}$ for $x > 0$, with $y(1) = 1$.

(a) Use the substitution $y = vx$ to show that $x\dfrac{dv}{dx} = v^2$. [3]

(b) Hence solve for y in terms of x. [4]

11. Solve each linear equation.

(a) $\dfrac{dy}{dx} - \dfrac{y}{x} = x^2$ for $x > 0$, given $y(1) = 1$. [6]

(b) $\dfrac{dy}{dx} + 2y = e^{-x}$ with $y(0) = 0$. [6]

12. Solve $\dfrac{dy}{dx} + (\tan x)\,y = \sec x$ for $-\frac{\pi}{2} < x < \frac{\pi}{2}$, given $y(0) = 2$. [7]

13. A fish population n (in thousands) in a lake satisfies the logistic equation $\dfrac{dn}{dt} = 0.4n(2 - n)$, with $n(0) = 0.5$.

(a) Using partial fractions, show that the solution is $n = \dfrac{2}{1 + 3e^{-0.8t}}$. [6]

(b) Find the time at which the population reaches 1500 fish. [3]

14. Consider $\dfrac{dy}{dx} = y - x^2$ with $y(0) = 1$.

(a) Use Euler's method with $h = 0.25$ to estimate $y(0.5)$. [5]

(b) The exact solution is concave up on $0 \le x \le 0.5$. State, with a reason, whether your estimate is above or below the true value. [2]

15. A skydiver falls from rest with $\dfrac{dv}{dt} = 9.8 - 0.2v^2$.

(a) Use Euler's method with $h = 0.5$ to estimate $v(1.5)$. [4]

(b) Find the terminal velocity, and comment on the behaviour of the Euler estimates around it. [3]

16. Find the Maclaurin series of $\ln(1+2x)$ up to the term in x^3

(a) by substitution into a standard series; [3]
(b) by computing derivatives from the definition. [3]

17.

(a) Find the Maclaurin series of $e^{-x}\sin x$ up to the term in x^3. [6]
(b) Use your series to estimate $e^{-0.1}\sin 0.1$, and compare it with the value from your calculator. [3]

18. By setting $x=1$ in the Maclaurin series for $\arctan x$:

(a) write down a series for $\dfrac{\pi}{4}$; [2]
(b) estimate π using the first four terms; [2]
(c) comment on the rate of convergence. [1]

19. EXACT Use Maclaurin series to evaluate $\displaystyle\lim_{x\to 0}\frac{1-\cos x}{x\sin x}$. [5]

20. A function satisfies $\dfrac{dy}{dx}=x+y^2$ with $y(0)=1$. Find the Maclaurin series of y up to the term in x^3. [7]

21. A drug is infused into the bloodstream at a rate that raises its concentration by r mg L^{-1} per hour, and the body removes it at a rate proportional to the concentration C, so that

$$\frac{dC}{dt}+kC=r,$$

where t is measured in hours and $C(0)=0$.

(a) Explain why this equation is linear, and write down its integrating factor. [3]
(b) Solve the equation to show that $C=\dfrac{r}{k}\left(1-e^{-kt}\right)$. [4]
(c) A patient is given $r=6$ and has $k=0.3$. Find the concentration after 4 hours. [2]
(d) Find the concentration the patient approaches in the long run, and explain what this means for the treatment. [3]
(e) Find how long it takes to reach 90% of that level. [2]
(f) Once that level is reached the infusion is stopped, so that $\dfrac{dC}{dt}=-kC$. Solve this equation and find how long the concentration then takes to halve. [3]

22. In this question you will build a series for logarithms that converges quickly.

(a) Find the Maclaurin series of $\ln(1+x)$ up to the term in x^4. [4]

(b) Write down the Maclaurin series of $\ln(1-x)$ up to the term in x^4. [2]

(c) Hence show that $\ln\left(\dfrac{1+x}{1-x}\right) = 2\left(x + \dfrac{x^3}{3} + \dfrac{x^5}{5} + \cdots\right)$. [3]

(d) By choosing a suitable value of x, use the three terms shown to estimate $\ln 3$. [3]

(e) The series for $\ln(1+x)$ is valid only for $-1 < x \leq 1$. State the value of x that would be needed to reach $\ln 3$ from that series alone, and explain why the series in part (c) is the more useful one here. [2]

23. Water drains from a tank through a hole in its base. The depth h metres at time t minutes satisfies

$$\frac{dh}{dt} = -k\sqrt{h},$$

where k is a positive constant.

(a) Separate the variables to show that $2\sqrt{h} = c - kt$ for some constant c. [3]

(b) Given that the depth is h_0 when $t = 0$, show that $h = \left(\sqrt{h_0} - \frac{1}{2}kt\right)^2$. [2]

(c) A tank starts with $h_0 = 1.44$ and has $k = 0.2$. Find how long it takes to empty. [3]

(d) Find the depth after 6 minutes. [2]

(e) Use Euler's method with steps of 3 minutes to estimate the depth after 6 minutes, and find the error in this estimate. [4]

(f) Describe the shape of the graph of h against t, and use it to explain why the Euler estimate lies below the exact value. [2]

Challenge Questions

24. GDC **Coffee in a room that warms and cools.** The chapter opened with a cup of coffee cooling in a room held at $20°$. A real room is not held fixed. Suppose instead that the room's temperature is $20 + 5\sin\omega t$ degrees, where t is the time in minutes and $\omega > 0$ is a constant, so that the room moves between $15°$ and $25°$ once every $\frac{2\pi}{\omega}$ minutes. Newton's law of cooling then gives

$$\frac{d\theta}{dt} = -k\left(\theta - 20 - 5\sin\omega t\right),$$

where θ is the temperature of the coffee and $k > 0$ is a constant. The coffee is poured at $80°$, so $\theta(0) = 80$.

(a) Let $u = \theta - 20$. Show that $\dfrac{du}{dt} + ku = 5k\sin\omega t$, and explain why this equation cannot be solved by separating the variables. [2]

(b) Using an integrating factor, and given that

$$\int e^{kt}\sin\omega t\,dt = \frac{e^{kt}\left(k\sin\omega t - \omega\cos\omega t\right)}{k^2 + \omega^2} + C,$$

show that $u = \dfrac{5k\,(k\sin\omega t - \omega\cos\omega t)}{k^2+\omega^2} + Ce^{-kt}$, and find C. [5]

(c) Show that $k\sin\omega t - \omega\cos\omega t = \sqrt{k^2+\omega^2}\,\sin(\omega t - \alpha)$, where $\tan\alpha = \dfrac{\omega}{k}$ and $0 < \alpha < \frac{\pi}{2}$. Hence show that, once t is large, $\theta \approx 20 + R\sin(\omega t - \alpha)$, where $R = \dfrac{5k}{\sqrt{k^2+\omega^2}}$, and explain why the coffee's temperature reaches each maximum $\frac{\alpha}{\omega}$ minutes after the room's. [5]

(d) Take $k = 0.05$. Find R and the delay $\frac{\alpha}{\omega}$ for a room that completes one cycle a day, $\omega = \frac{2\pi}{1440}$, and for a room whose heater switches on and off every 10 minutes, $\omega = \frac{2\pi}{10}$. [4]

(e) Using your answers to part (d), conjecture what happens to R and to the delay $\frac{\alpha}{\omega}$ as $\omega \to 0$ and as $\omega \to \infty$. Prove your conjectures for R, and use the Maclaurin series $\arctan z = z - \frac{z^3}{3} + \cdots$ to show that the delay tends to $\frac{1}{k}$ as $\omega \to 0$. Explain both results in terms of the coffee and the room. [4]

(f) A real room's temperature is not a single sine wave. Explain why, because the equation in part (a) is linear, a room temperature of 20 plus a sum of sine waves produces a long-run coffee temperature that is 20 plus the sum of the separate responses. Describe how you could use this to predict, and then test, the temperature of a drink left in a real room. [2]

25. GDC **How wrong is Euler?** Take $\dfrac{dy}{dx} = y$ with $y(0) = 1$, whose exact solution is $y = e^x$, so that $y(1) = e$.

(a) Show that one Euler step of length h gives $y_{n+1} = (1+h)y_n$, and hence that after n steps $y_n = (1+h)^n$. [3]

(b) Taking $h = \dfrac{1}{n}$, deduce that Euler's estimate of $y(1)$ is $\left(1+\dfrac{1}{n}\right)^n$. [2]

(c) Evaluate this estimate for $n = 2, 4, 8, 16$, and tabulate the error in each case. [4]

(d) Describe what happens to the error each time h is halved, and state what this says about the accuracy of the method. [3]

(e) Use the Maclaurin series for $\ln(1+x)$ to show that, for $n \geq 2$,

$$n\ln\left(1+\frac{1}{n}\right) = 1 - \frac{1}{2n} + \frac{1}{3n^2} - \cdots.$$

[3]

(f) Hence show that, for large n, the error $e - \left(1+\dfrac{1}{n}\right)^n$ is approximately $\dfrac{e}{2n}$. Compare this with your table, and deduce the limit of $\left(1+\dfrac{1}{n}\right)^n$ as $n \to \infty$. [3]

(g) The error is close to a known multiple of h. Suggest how two Euler estimates, one with n steps and one with $2n$ steps, could be combined into a more accurate estimate, and test your suggestion on the values in your table. [3]

26. **The ferry and the current.** A river flows in the positive y-direction, and one bank is the line $x = 0$, with a dock at the origin. A ferry sets off from $(1, 0)$, moving through the water at a constant speed of 1 unit per minute and always aiming straight at the dock, while

the current carries it in the positive y-direction at k units per minute, where $k > 0$. Its velocity therefore has components

$$\frac{dx}{dt} = -\frac{x}{\sqrt{x^2+y^2}}, \qquad \frac{dy}{dt} = -\frac{y}{\sqrt{x^2+y^2}} + k,$$

and its path satisfies

$$\frac{dy}{dx} = \frac{y - k\sqrt{x^2+y^2}}{x}, \qquad x > 0.$$

(a) Show the equation is homogeneous, and that $y = vx$ leads to $x\dfrac{dv}{dx} = -k\sqrt{1+v^2}$. [3]

(b) Given that $\displaystyle\int \frac{dv}{\sqrt{1+v^2}} = \ln\left(v + \sqrt{1+v^2}\right) + C$, show that $v + \sqrt{1+v^2} = x^{-k}$. [3]

(c) Hence show that $y = \frac{1}{2}\left(x^{1-k} - x^{1+k}\right)$. [4]

(d) Taking $k = \frac{1}{2}$, find the point that the ferry approaches as $x \to 0$. [2]

(e) Show that, as $x \to 0$, the ferry approaches the dock if and only if $k < 1$. Describe what happens to the path when $k = 1$ and when $k > 1$. [4]

(f) Using part (b), show that $\sqrt{1+v^2} = \frac{1}{2}\left(x^{-k} + x^{k}\right)$. Hence, using the expression for $\dfrac{dx}{dt}$, show that for $0 < k < 1$ the ferry reaches the dock after

$$T = \frac{1}{2}\int_0^1 \left(x^{-k} + x^{k}\right) dx = \frac{1}{1-k^2} \text{ minutes},$$

and explain what the integral shows when $k \geq 1$. [5]

(g) Aiming at the dock is not the only way to steer. Suggest a different strategy for $0 < k < 1$, find the time it takes, and compare it with T. [2]

27. GDC **The series for $\tan x$.** In the notes, the equation $\dfrac{dy}{dx} = 1 + y^2$ with $y(0) = 0$ gave the first two terms of the Maclaurin series of its solution, $y = \tan x$, by repeated differentiation. This question finds more terms by a quicker method and asks how far the series can be trusted. You may assume that $y = \tan x$ is the only solution with $y(0) = 0$ on $-\frac{\pi}{2} < x < \frac{\pi}{2}$, and that it is equal to its Maclaurin series near $x = 0$.

(a) Show that if $y = g(x)$ satisfies $\dfrac{dy}{dx} = 1 + y^2$ with $g(0) = 0$, then so does $y = -g(-x)$. Deduce that $\tan x$ is an odd function, so that its series contains only odd powers of x. [3]

(b) Write $y = a_1x + a_3x^3 + a_5x^5 + a_7x^7 + \cdots$. Substitute into the equation and compare the coefficients of x^0, x^2 and x^4 to show that $a_1 = 1$, $a_3 = \frac{1}{3}$ and $a_5 = \frac{2}{15}$. Find a_7. [6]

(c) Use the four terms found in part (b) to estimate $\tan 0.5$, $\tan 1$ and $\tan 1.5$, and compare each estimate with the value from your calculator. [3]

(d) Using the graph of $y = \tan x$, explain why no polynomial can stay close to $\tan x$ as $x \to \frac{\pi}{2}$, and suggest the largest interval centred on 0 on which the series could converge. [3]

(e) Apply the method of part (b) to $\dfrac{dy}{dx} = 1 - y^2$ with $y(0) = 0$, as far as the term in x^7. Conjecture how its coefficients are related to those of $\tan x$, and explain why the relationship continues for every power. [3]

(f) The coefficients $1, \frac{1}{3}, \frac{2}{15}, \frac{17}{315}, \frac{62}{2835}, \ldots$ follow no obvious pattern. Investigate the ratio of successive coefficients, and suggest how its behaviour is connected to your answer to part (d). [2]

Chapter 17

Complex Numbers

17

SYLLABUS 1.12–1.14 HL ONLY

In the sixteenth century, Italian mathematicians learned to solve cubic equations. Their formula produced something unexpected. It sometimes asked for the square root of a negative number, even when the equation had three perfectly ordinary answers. No number they knew could give a negative square. Yet if they wrote the impossible quantity down and kept calculating, it disappeared before the end, and the answers were correct.

This chapter begins with that impossible quantity. We define a number i with $i^2 = -1$. No real number does this, so i is new. For a long time many mathematicians distrusted it, and the name it was given, "imaginary", records that distrust. The name was unfair. Once i is allowed, every non-constant polynomial equation has a solution, a result now called the **Fundamental Theorem of Algebra**.

These new numbers are also geometric. Each one is a point in a plane, with the real numbers along one axis. Adding two of them moves a point across the plane. Multiplying two of them rotates and stretches. Multiplying by i rotates by a quarter-turn. In ordinary coordinates a rotation takes several steps; here it is one multiplication. This is one reason complex numbers are used widely in physics and engineering.

By the end of this chapter you will be able to write a complex number in three forms: $a + bi$, $r\,\mathrm{cis}\,\theta$ and $re^{i\theta}$. You will multiply and divide in each form, find powers with De Moivre's theorem, and find the nth roots of a complex number, including the roots of unity. You will also use complex roots to factorise polynomials and to prove trigonometric identities.

Three forms may seem like more than you need. Each one makes a different calculation easy. Adding is simplest in $a + bi$ form. A power that would fill a page of working in Cartesian form is immediate in polar form. Choosing the form before you start is a skill this chapter practises throughout.

We start with that algebra: numbers of the form $a + bi$.

© 2026 Hasan Khan ibxhasan.github.io/bettermath

17.1 Cartesian Form

Everything starts from a single definition that no real number satisfies.

DEF The **imaginary unit** i satisfies $i^2 = -1$. A **complex number** is any $z = a + bi$ with $a, b \in \mathbb{R}$; $a = \mathrm{Re}(z)$ is the **real part** and $b = \mathrm{Im}(z)$ the **imaginary part**. Of these the booklet prints only the form $z = a + bi$.

Complex numbers are added, subtracted and multiplied exactly like binomials in i. There is one extra rule: replace i^2 by -1 wherever it appears.

That one rule is enough to find every power of i, because the powers repeat in a cycle of four:

$$i^1 = i, \qquad i^2 = -1, \qquad i^3 = i^2 \cdot i = -i, \qquad i^4 = (i^2)^2 = 1,$$

and $i^5 = i^4 \cdot i = i$ starts the cycle again. To simplify any power, divide the exponent by 4 and keep only the remainder: $i^{2026} = i^{4(506)+2} = i^2 = -1$.

Worked through. Given $z = 2 + i$ and $w = 5 - 3i$, adding is done part by part: $z + w = (2 + 5) + (1 - 3)i = 7 - 2i$. Multiplying, expand and then use $i^2 = -1$:

$$zw = (2 + i)(5 - 3i) = 10 - 6i + 5i - 3i^2 = 10 - i + 3 = 13 - i.$$

The whole of complex arithmetic is ordinary algebra plus the one substitution $i^2 = -1$.

17.1.1 The conjugate and division

Every complex number has a mirror image across the real axis, called its **conjugate**. The conjugate is what makes division possible.

DEF The **complex conjugate** of $z = a + bi$ is $z^* = a - bi$.

The conjugate matters because a number times its own conjugate is always real: $zz^* = (a + bi)(a - bi) = a^2 + b^2$. This is exactly the tool for division, in the same way that a surd is rationalised. To divide, multiply top and bottom by the conjugate of the denominator, turning the denominator real.

EXAMPLE 17.1
Find $\dfrac{3 + 2i}{1 - i}$.

Multiply numerator and denominator by the conjugate $1+i$:

$$\frac{3+2i}{1-i}\cdot\frac{1+i}{1+i}=\frac{3+3i+2i+2i^2}{1^2+1^2}=\frac{3+5i-2}{2}=\frac{1+5i}{2}=\tfrac{1}{2}+\tfrac{5}{2}i.$$

Multiplying by the conjugate removes i from the denominator every time; the answer is then read off as real and imaginary parts.

One more Cartesian tool: two complex numbers are equal exactly when their real parts agree *and* their imaginary parts agree. A single complex equation therefore carries two real equations at once, and **equating parts** solves for unknowns.

EXAMPLE 17.2

Find the complex number z satisfying $z+2z^*=9+2i$.
Write $z=a+bi$, so $z^*=a-bi$:

$$(a+bi)+2(a-bi)=3a-bi=9+2i.$$

Equate real parts: $3a=9\implies a=3$. Equate imaginary parts: $-b=2\implies b=-2$. So $z=3-2i$. One equation was enough for two unknowns, because the real and imaginary parts each give an equation.

17.1.2 The Argand diagram

A complex number carries two real numbers, so it is naturally a *point* in a plane: the horizontal axis for the real part, the vertical for the imaginary. This picture, the **Argand diagram**, is what makes the subject geometric.

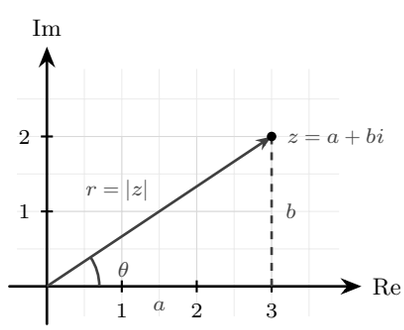


Figure 17.1 A complex number $z=a+bi$ drawn as the point (a,b) of the Argand diagram. Its distance from the origin is the modulus r, and the angle its arrow makes with the positive real axis is the argument θ.

The picture immediately explains the arithmetic. **Addition is vector addition**: to add w to z, place w's arrow tail-to-tip on z's, exactly as in Ch. 9, so $z+w$ is the far corner of the parallelogram the two arrows span. And **conjugation is reflection**: z^* is z flipped across the real axis. This is the picture behind $zz^*=a^2+b^2$ being real, a fact already checked algebraically above.

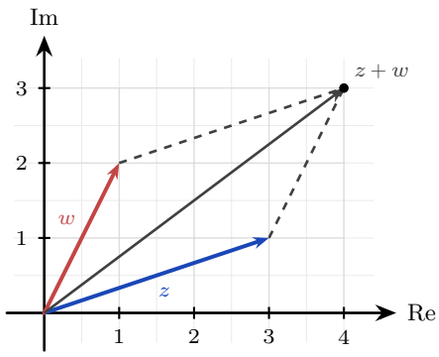


Figure 17.2 Addition is vector addition. Placing w's arrow tail-to-tip on z's puts $z + w$ at the far corner of the parallelogram the two arrows span.

Conjugation is the second picture. Later in the chapter it explains why the non-real roots of a polynomial with real coefficients come in pairs placed symmetrically about the real axis.

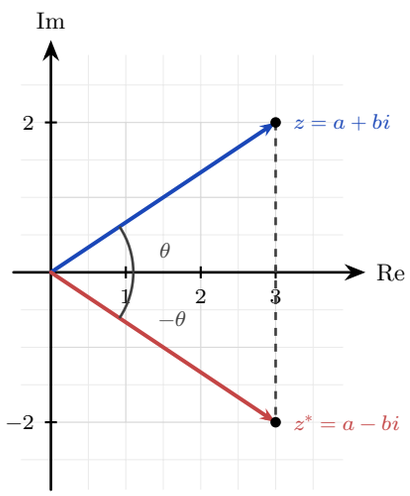


Figure 17.3 The conjugate is a reflection in the real axis. The two numbers share a modulus, and their arguments are equal and opposite. Because conjugating respects addition and multiplication and leaves every real number alone, the roots of a polynomial with real coefficients appear in mirror-image pairs. See §17.4.3.

There are two natural quantities to attach to a point in a plane: how far it is from the origin, and in what direction. These are the modulus and the argument, and they are the subject of the next section.

YOUR TURN 17.1 Cartesian Form

Arithmetic in Cartesian form

1. Evaluate each expression, giving your answer in the form $a + bi$.

(a) $(3 + 2i) + (1 - 4i)$
(b) $(5 - i) - (2 + 3i)$
(c) $(2 + 3i)(1 - i)$
(d) $(2 + i)(2 - i)$
(e) $(1 + 2i)^2(3 - i)$

2. Simplify each power of i.

(a) i^3
(b) i^4
(c) i^{15}
(d) i^{100}

3. Let $z = 7 - 4i$.

(a) Write down $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$.
(b) Write down z^*.
(c) Evaluate zz^*.
(d) Evaluate $z + z^*$.

Division

4. Write each quotient in the form $a + bi$.

(a) $\dfrac{4 + 3i}{2 - i}$
(b) $\dfrac{1}{1 + i}$
(c) $\dfrac{2 - 3i}{i}$
(d) $\dfrac{(1 + i)^2}{2 - i}$

Equations involving z^*

5. Find the complex number z in each case.

(a) $z + 3z^* = 8 - 4i$
(b) $2z - z^* = 3 + 6i$
(c) $zz^* + 2z = 12 + 4i$ (there are two answers)

The Argand diagram

6. Let $z = 2 + i$ and $w = -1 + 3i$.

(a) Plot z, w and $z + w$ on an Argand diagram, and state the shape formed by the points 0, z, $z + w$ and w.
(b) Plot z^* and w^*, and describe the single transformation that maps each number to its conjugate.
(c) Find $z - w$, and describe how it appears on the diagram in relation to the points z and w.

Concept check

7. A student writes $(3 + 2i)^2 = 9 + 4i^2 = 5$. Identify the error, and find $(3 + 2i)^2$ correctly.

17.2 Modulus and Argument

The distance of z from the origin is its **modulus** $|z|$, straight from Pythagoras; the angle its arrow makes with the positive real axis is its **argument** $\arg(z)$, measured anticlockwise.

KEY For $z = a + bi$: $\quad |z| = \sqrt{a^2 + b^2}, \qquad \tan(\arg z) = \dfrac{b}{a}.$

CAUTION Finding the argument needs care: $\tan\theta = \frac{b}{a}$ alone cannot tell which quadrant z is in, because tan repeats every π. Always sketch z on the Argand diagram first, then choose the angle in the correct quadrant, conventionally in the range $-\pi < \theta \leq \pi$ (the **principal argument**).

The danger is easiest to see with an example. The numbers $1 + i$ and $-1 - i$ both give $\frac{b}{a} = 1$, so a calculator returns $\arctan 1 = \frac{\pi}{4}$ for both. Only one of them is correct.

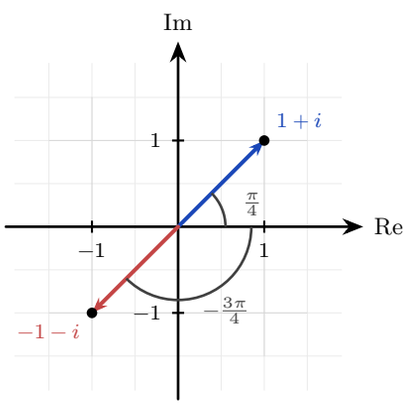


Figure 17.4 Both numbers give $\frac{b}{a} = 1$, so both return $\arctan 1 = \frac{\pi}{4}$ from a calculator. The sketch settles it: $1 + i$ lies in the first quadrant and has argument $\frac{\pi}{4}$, while $-1 - i$ lies in the third and has principal argument $\frac{\pi}{4} - \pi = -\frac{3\pi}{4}$.

The modulus also measures the distance between two points, not only the distance from the origin. Since $z - w$ is the vector running from w to z Ch. 9, its length $|z - w|$ is the distance between the points representing z and w.

Knowing the modulus r and argument θ is enough to rebuild the number, since $a = r\cos\theta$ and $b = r\sin\theta$. This gives the **modulus–argument** (or polar) form, and the equivalent exponential version, **Euler's form**.

KEY · IB **Polar and Euler form:** $\quad z = r(\cos\theta + i\sin\theta) = r\,\mathrm{cis}\,\theta = re^{i\theta},$
where $r = |z|$ and $\theta = \arg(z)$.

17.2.1 Euler's form

That $e^{i\theta} = \cos\theta + i\sin\theta$ is unexpected at first sight, and you already have the tools to see why it is true. Take the Maclaurin series for e^x from Ch. 16 and substitute $x = i\theta$, using the cycle of powers of i from §17.1:

$$e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \cdots = \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots\right) + i\left(\theta - \frac{\theta^3}{3!} + \cdots\right).$$

The even terms, where i has become ± 1, form the cosine series. The odd terms, which still carry an i, form the sine series. So once complex inputs are allowed, the exponential function contains cosine and sine: at $i\theta$ its real part is $\cos\theta$ and its imaginary part is $\sin\theta$. Setting $\theta = \pi$ gives $e^{i\pi} + 1 = 0$, which links e, i, π, 1 and 0 in a single line. It is a calculation, not a coincidence.

Worked through. To write $z = -1 - \sqrt{3}\,i$ in polar and Euler form, with $-\pi < \theta \le \pi$, start with the modulus. It is $r = \sqrt{(-1)^2 + (-\sqrt{3})^2} = \sqrt{4} = 2$. The point $(-1, -\sqrt{3})$ lies in the third quadrant, where $\tan\theta = \sqrt{3}$ gives a reference angle of $\frac{\pi}{3}$, so the principal argument is $\theta = \frac{\pi}{3} - \pi = -\frac{2\pi}{3}$. Hence

$$z = 2\left(\cos\left(-\tfrac{2\pi}{3}\right) + i\sin\left(-\tfrac{2\pi}{3}\right)\right) = 2\,\mathrm{cis}\left(-\tfrac{2\pi}{3}\right) = 2e^{-2i\pi/3}.$$

Sketch first, then read off the quadrant. Taking $\arctan$ without checking the quadrant gives the wrong angle in the second or third quadrant.

The reverse direction is easier, because no quadrant decision is needed: substitute the angle and evaluate.

EXAMPLE 17.3

Write (a) $4\,\mathrm{cis}\,\dfrac{2\pi}{3}$ and (b) $2e^{-i\pi/2}$ in Cartesian form.

(a) Expand the cis and evaluate the two exact values:

$$4\left(\cos\tfrac{2\pi}{3} + i\sin\tfrac{2\pi}{3}\right) = 4\left(-\tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i\right) = -2 + 2\sqrt{3}\,i.$$

(b) Euler form is the same thing written differently, so read $r = 2$ and $\theta = -\frac{\pi}{2}$:

$$2\left(\cos\left(-\tfrac{\pi}{2}\right) + i\sin\left(-\tfrac{\pi}{2}\right)\right) = 2(0 - i) = -2i.$$

Going from polar to Cartesian is always this direct. It is the other direction that needs the sketch.

YOUR TURN 17.2 Modulus and Argument

Modulus and argument

8. Find the modulus of each number.

(a) $3+4i$ (b) $5-12i$
(c) $1+i$ (d) $-2i$

9. Find the principal argument of each number, with $-\pi < \theta \le \pi$.

(a) $1+i$ (b) i
(c) -3 (d) $-1-i$
(e) $-\sqrt{3}+i$ (f) $-4-4\sqrt{3}\,i$

Converting between forms

10. Write each number in modulus–argument form and in Euler form.

(a) $2i$ (b) -3
(c) $1-i$ (d) $-1+\sqrt{3}\,i$

11. Write each number in Cartesian form.

(a) $4\,\mathrm{cis}\,\frac{\pi}{3}$ (b) $2e^{-i\pi/4}$
(c) $3\,\mathrm{cis}\,\pi$ (d) $e^{i\pi/2}$

Distance on the Argand diagram

12. Let $z = 3+4i$ and $w = -1+2i$.

(a) Find $|z-w|$, giving an exact answer.
(b) State what $|z-w|$ measures on the Argand diagram.
(c) Verify that $zz^{*} = |z|^{2}$.

Concept check

13. A student writes $\arg\left(-1-\sqrt{3}\,i\right) = \arctan\dfrac{-\sqrt{3}}{-1} = \arctan\sqrt{3} = \dfrac{\pi}{3}$. Identify the error, and find the principal argument.

17.3 Products, Quotients and De Moivre's Theorem

Polar form needs extra notation, and it repays that cost by making multiplication easy. When two complex numbers are multiplied, their moduli multiply and their arguments *add*, which is just the law of exponents applied to $re^{i\theta}$.

KEY $|zw| = |z|\,|w|$ and $\arg(zw) = \arg z + \arg w$; $\left|\dfrac{z}{w}\right| = \dfrac{|z|}{|w|}$ and $\arg\left(\dfrac{z}{w}\right) = \arg z - \arg w$.

With principal arguments these rules hold up to a multiple of 2π: if the sum or difference falls outside $-\pi < \theta \le \pi$, add or subtract 2π. Read geometrically, multiplying by a complex number *rotates* by its argument and *scales* by its modulus. So multiplying by $i = 1 \cdot e^{i\pi/2}$, whose modulus is 1, is a quarter-turn anticlockwise with no stretching at all.

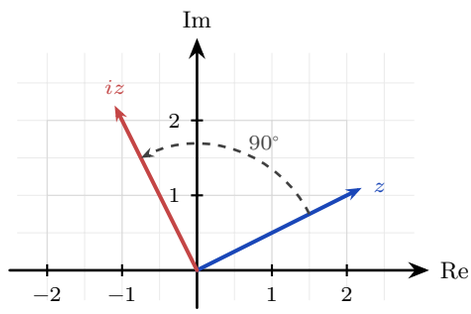


Figure 17.5 Multiplying by i turns a complex number through a quarter-turn anticlockwise about the origin, leaving its modulus unchanged.

The quotient rule with $z = 1$ gives the reciprocal: for $r > 0$,

$$\frac{1}{r\,\mathrm{cis}\,\theta} = \frac{1}{r}\,\mathrm{cis}(-\theta).$$

Taking the reciprocal replaces the modulus by its reciprocal and reverses the argument, so for $|z| = 1$ the reciprocal is the conjugate z^*.

Worked through. Take $z = 6\,\mathrm{cis}\,\frac{2\pi}{3}$ and $w = 3\,\mathrm{cis}\,\frac{\pi}{6}$. For the product, multiply the moduli and add the arguments:

$$zw = (6)(3)\,\mathrm{cis}\left(\tfrac{2\pi}{3} + \tfrac{\pi}{6}\right) = 18\,\mathrm{cis}\,\tfrac{5\pi}{6}.$$

For the quotient, divide the moduli and subtract the arguments:

$$\frac{z}{w} = \frac{6}{3}\,\mathrm{cis}\left(\tfrac{2\pi}{3} - \tfrac{\pi}{6}\right) = 2\,\mathrm{cis}\,\tfrac{\pi}{2} = 2i.$$

Compare this with doing the same multiplication in Cartesian form, which would need four products and a collection of terms. Once the numbers are in polar form the work is one multiplication and one addition.

If multiplying adds arguments, then raising to a power *multiplies* the argument. This is **De Moivre's theorem**, and the roots, the unit-circle results and the trigonometric identities in the rest of the chapter all follow from it.

KEY · IB **De Moivre's theorem:**
$(r(\cos\theta + i\sin\theta))^n = r^n(\cos n\theta + i\sin n\theta) = r^n e^{in\theta} = r^n \operatorname{cis} n\theta.$

The syllabus requires the theorem for $n \in \mathbb{Z}$ and extends it to *rational* exponents. Past the integers the statement has to be read more carefully: a power like $z^{1/n}$ is multi-valued, so $\sqrt[n]{r}\operatorname{cis}\frac{\theta}{n}$ is one of its n values rather than the whole answer. That is exactly the story of the nth roots in the next section. For positive integers the theorem is proved by the induction of Ch. 3. The inductive step is one multiplication: assuming $(r\operatorname{cis}\theta)^k = r^k \operatorname{cis} k\theta$, multiplying both sides once more by $r\operatorname{cis}\theta$ scales the modulus to r^{k+1} and, by the compound-angle identities of Ch. 8, adds θ to the argument to give $(k+1)\theta$. Every step is that same identity applied once, and you carry the induction out in full in End-of-Chapter Problem 22.

EXAMPLE 17.4

The theorem is claimed for every integer exponent. Taking the case $n \in \mathbb{Z}^+$ as proved, show that it also holds for every negative integer n, for $r > 0$.

Start with the reciprocal of a number of modulus 1. Multiplying adds the arguments, so

$$\operatorname{cis}\phi \cdot \operatorname{cis}(-\phi) = \operatorname{cis}(\phi - \phi) = \operatorname{cis} 0 = 1, \qquad \text{hence} \qquad \frac{1}{\operatorname{cis}\phi} = \operatorname{cis}(-\phi).$$

Dividing by a number of modulus 1 simply reverses its angle.

Now let n be a negative integer, and write $n = -m$ with $m \in \mathbb{Z}^+$. Applying the positive case to m,

$$(r\operatorname{cis}\theta)^n = \frac{1}{(r\operatorname{cis}\theta)^m} = \frac{1}{r^m \operatorname{cis}(m\theta)} = r^{-m}\operatorname{cis}(-m\theta) = r^n \operatorname{cis}(n\theta). \quad \blacksquare$$

The exponent $n = 0$ needs no work either, since $z^0 = 1 = r^0 \operatorname{cis} 0$. So one reciprocal carries the theorem from the positive integers to the whole of $\mathbb{Z}$, which is the range the syllabus asks for.

Raising a complex number to a high power means expanding a bracket again and again in Cartesian form. In polar form there is nothing to expand: raise the modulus to the power, multiply the angle by the power, and the work is finished.

EXAMPLE 17.5

Find $(1+i)^8$.

In polar form $1 + i = \sqrt{2}\operatorname{cis}\frac{\pi}{4}$. By De Moivre,

$$(1+i)^8 = (\sqrt{2})^8 \operatorname{cis}(8 \cdot \tfrac{\pi}{4}) = 16\operatorname{cis}(2\pi) = 16.$$

Seven repeated multiplications reduce to one power of the modulus and one multiple of the angle.

Draw the powers and you can watch the theorem work. Each multiplication by $1 + i$ turns the arrow 45° anticlockwise and stretches it by $\sqrt{2}$, so the powers lie on a spiral that winds

outward from the origin:

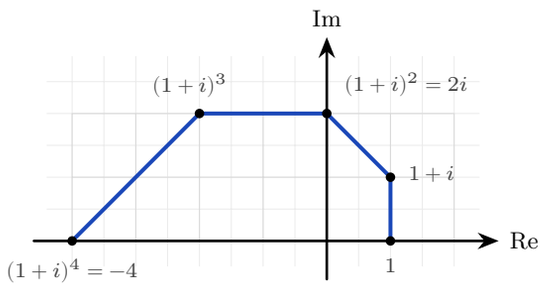


Figure 17.6 The powers $(1+i)^0$ through $(1+i)^4$. Each multiplication turns the arrow through 45° and stretches it by $\sqrt{2}$, so the powers lie on a spiral winding outward from the origin.

Four more steps continue the spiral through the lower half-plane and back to the positive real axis at $(1+i)^8 = 16$: a full turn of $8 \times 45^\circ = 360^\circ$, with the modulus grown to $(\sqrt{2})^8 = 16$. The worked example above is this picture written in algebra.

17.3.1 Numbers on the unit circle

One case of De Moivre is used often enough to state separately. If $|z| = 1$ then $z = \operatorname{cis}\theta$, and by the reciprocal result above, $\dfrac{1}{z} = \operatorname{cis}(-\theta) = z^*$. Add the two and the imaginary parts cancel while the real parts double; subtract them and the opposite happens.

KEY For $z = \operatorname{cis}\theta$ and any $n \in \mathbb{Z}$:

$$z^n + \frac{1}{z^n} = 2\cos n\theta, \qquad z^n - \frac{1}{z^n} = 2i\sin n\theta.$$

Read from left to right this converts powers into multiple angles, and that is what makes it useful: an expression like $\cos^6\theta$, awkward to integrate as it stands, becomes a sum of $\cos 2\theta$, $\cos 4\theta$ and $\cos 6\theta$ terms that can be integrated directly. That construction is §17.5.

EXAMPLE 17.6

Let $z = \operatorname{cis}\theta$. Expand $\left(z + \dfrac{1}{z}\right)\left(z - \dfrac{1}{z}\right)$ in two ways, and hence express $\sin\theta\cos\theta$ in terms of $\sin 2\theta$.

Taking $n = 1$ in the two formulas, $z + \dfrac{1}{z} = 2\cos\theta$ and $z - \dfrac{1}{z} = 2i\sin\theta$, so the product is

$$\left(z + \frac{1}{z}\right)\left(z - \frac{1}{z}\right) = (2\cos\theta)(2i\sin\theta) = 4i\sin\theta\cos\theta.$$

Expanding algebraically instead, the two cross terms cancel:

$$\left(z + \frac{1}{z}\right)\left(z - \frac{1}{z}\right) = z^2 - 1 + 1 - \frac{1}{z^2} = z^2 - \frac{1}{z^2} = 2i\sin 2\theta,$$

using the second formula again with $n = 2$. The two expressions are equal, so

$$4i \sin\theta\cos\theta = 2i\sin 2\theta \implies \sin\theta\cos\theta = \frac{\sin 2\theta}{2}.$$

This is the double-angle identity for the sine, from Ch. 8, now obtained without using the compound-angle formula directly. Evaluating one product in two ways is the technique of §17.5.1.

YOUR TURN 17.3 Products, Quotients and De Moivre's Theorem

Products and quotients

14. Evaluate each expression, giving your answer in modulus–argument form.

(a) $\left(2\operatorname{cis}\frac{\pi}{6}\right)\left(3\operatorname{cis}\frac{\pi}{3}\right)$ (b) $\dfrac{6\operatorname{cis}\frac{2\pi}{3}}{2\operatorname{cis}\frac{\pi}{3}}$

15. Given $|z| = 3$, $|w| = 4$, $\arg z = \frac{\pi}{4}$ and $\arg w = \frac{\pi}{6}$, write down

(a) $|zw|$ (b) $\arg(zw)$

(c) $\left|\dfrac{z}{w}\right|$ (d) $\arg\left(\dfrac{z}{w}\right)$

16. Give each answer in modulus–argument form, with the principal argument $-\pi < \theta \le \pi$.

(a) Find zw, where $z = 2\operatorname{cis}\frac{5\pi}{6}$ and $w = 3\operatorname{cis}\frac{\pi}{3}$.

(b) Find $\dfrac{z}{w}$, where $z = 4\operatorname{cis}\left(-\frac{2\pi}{3}\right)$ and $w = 2\operatorname{cis}\frac{\pi}{2}$.

De Moivre's theorem

17. Use De Moivre's theorem for each part.

(a) Write $(\cos\theta + i\sin\theta)^4$ in the form $\cos n\theta + i\sin n\theta$.

(b) Evaluate $\left(\operatorname{cis}\frac{\pi}{5}\right)^5$, in modulus–argument form.

(c) Evaluate $\left(2\operatorname{cis}\frac{\pi}{6}\right)^3$, in modulus–argument form.

(d) Evaluate $(1+i)^6$.

(e) Evaluate $\left(1-\sqrt{3}\,i\right)^5$.

Negative powers and reciprocals

18.

(a) Find $\left(2\operatorname{cis}\frac{\pi}{6}\right)^{-2}$ in the form $a + bi$.

(b) Find $(1-i)^{-4}$.

(c) Write $\dfrac{1}{\sqrt{3}+i}$ in modulus–argument form.

Multiplying as a rotation

19.

(a) Let $z = 3 + 2i$. Find iz, and describe the transformation of the Argand diagram that maps z to iz.

(b) The point $4 + 2i$ is rotated anticlockwise about the origin through $\frac{\pi}{3}$. Find its image in the form $a + bi$.

(c) $OABC$ is a square on the Argand diagram, with vertices labelled anticlockwise, O at the origin and A at $2 + i$. Find the numbers at C and at B.

Numbers on the unit circle

20. Let $z = \operatorname{cis}\theta$. Write each expression in terms of a single cosine or sine.

(a) $z + \dfrac{1}{z}$ (b) $z^3 + \dfrac{1}{z^3}$

(c) $z^2 - \dfrac{1}{z^2}$

Concept check

21. A student uses De Moivre's theorem and writes $(1+i)^4 = \operatorname{cis}(4 \times \frac{\pi}{4}) = \operatorname{cis}\pi = -1$. Identify the error, and find $(1+i)^4$.

17.4 Roots of Complex Numbers

Solving $z^3 = 125$ over the real numbers gives one answer. Over the complex numbers it gives three, and they are not scattered: they sit at equal spacing on a circle. The first part of this section is built on that picture.

De Moivre's theorem, used in reverse, finds these roots. A nonzero complex number has exactly n distinct nth roots. They all share one modulus, and their arguments are equally spaced. To find them, write the number in polar form, take the nth root of the modulus, divide the argument by n, then step around by $\frac{2\pi}{n}$ until all n are collected.

KEY The nth roots of $z = r\operatorname{cis}\theta$ are $\sqrt[n]{r}\ \operatorname{cis}\left(\dfrac{\theta + 2\pi k}{n}\right)$, for $k = 0, 1, \ldots, n-1$.

Because the roots share a modulus $\sqrt[n]{r}$ and their arguments differ by equal steps of $\frac{2\pi}{n}$, for $n \geq 3$ they sit at the vertices of a **regular n-gon** centred at the origin (for $n = 2$, at the two ends of a diameter). One root therefore locates all the others: find any single root, and the rest follow by turning it through $\frac{2\pi}{n}$ at a time.

KEY To solve $z^n = w$:

1. Write w in polar form, $w = r\,\mathrm{cis}\,\theta$.
2. The modulus of every root is $\sqrt[n]{r}$.
3. Take $\dfrac{\theta}{n}$ as the first argument, then add $\dfrac{2\pi}{n}$ repeatedly until you have n roots.
4. Convert back to Cartesian form only if the question asks for it.

Worked through. To find the three cube roots of 125, write $125 = 125\,\mathrm{cis}\,0$. The modulus of each root is $\sqrt[3]{125} = 5$, and the arguments are $\frac{0+2\pi k}{3}$ for $k = 0, 1, 2$:

$$5\,\mathrm{cis}\,0 = 5, \qquad 5\,\mathrm{cis}\,\tfrac{2\pi}{3} = -\frac{5}{2} + \frac{5\sqrt{3}}{2}i, \qquad 5\,\mathrm{cis}\,\tfrac{4\pi}{3} = -\frac{5}{2} - \frac{5\sqrt{3}}{2}i.$$

The three roots form an equilateral triangle on the circle of radius 5, one real and a conjugate pair. Only the first argument needed any thought; the other two came from adding $\frac{2\pi}{3}$.

That calculation hid one step, because 125 has argument 0 and dividing 0 by 3 changes nothing. When the number has a non-zero argument, the division does real work.

EXAMPLE 17.7

Find the three cube roots of $-27i$, giving each in polar form and in Cartesian form.

First write $-27i$ in polar form. It lies on the negative imaginary axis, so its modulus is 27 and its argument is $-\frac{\pi}{2}$:

$$-27i = 27\,\mathrm{cis}\left(-\tfrac{\pi}{2}\right).$$

Each root has modulus $\sqrt[3]{27} = 3$. The first argument is the given argument divided by 3:

$$\frac{-\pi/2}{3} = -\frac{\pi}{6}.$$

The others follow by adding $\frac{2\pi}{3}$ each time:

$$-\frac{\pi}{6}, \qquad -\frac{\pi}{6} + \frac{2\pi}{3} = \frac{\pi}{2}, \qquad \frac{\pi}{2} + \frac{2\pi}{3} = \frac{7\pi}{6} \equiv -\frac{5\pi}{6}.$$

The third has been reduced by 2π to bring it back into $-\pi < \theta \le \pi$. In polar form the roots are $3\,\mathrm{cis}\left(-\frac{\pi}{6}\right)$, $3\,\mathrm{cis}\,\frac{\pi}{2}$ and $3\,\mathrm{cis}\left(-\frac{5\pi}{6}\right)$, and converting each,

$$\frac{3\sqrt{3}}{2} - \frac{3}{2}i, \qquad 3i, \qquad -\frac{3\sqrt{3}}{2} - \frac{3}{2}i.$$

Check the easiest one: $(3i)^3 = 27i^3 = -27i$. ✓ Note that adding $\frac{2\pi}{3}$ repeatedly can carry an argument past π; subtract 2π to return it to the principal range if the question asks for principal arguments.

17.4.1 Roots of Unity

One case is used often enough to study separately: $w = 1$. Solving $z^n = 1$ gives the **roots of unity**. These are n points on the *unit* circle. One of them is 1, and together they sit at the vertices of a regular n-gon.

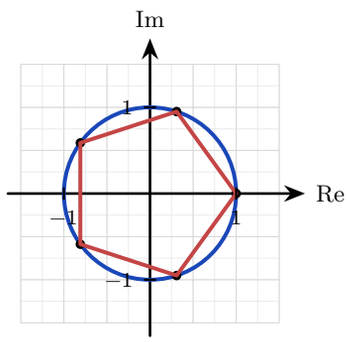


Figure 17.7 The five fifth roots of unity, at the vertices of a regular pentagon inscribed in the unit circle with one vertex at 1.

One root generates all the others. Write $\omega = \operatorname{cis}\dfrac{2\pi}{n}$ for the first root anticlockwise from 1. Then $\omega^k = \operatorname{cis}\dfrac{2\pi k}{n}$, which is the kth root round the circle. Every power of ω is again a root, because $\left(\omega^k\right)^n = \left(\omega^n\right)^k = 1^k = 1$. The n roots are therefore

$$1,\ \omega,\ \omega^2,\ \ldots,\ \omega^{n-1},$$

one number and its powers, stepping round the polygon and returning to the start.

The case $n = 3$ is worth memorising. Converting each root to Cartesian form, the cube roots of unity are 1, $\omega = \operatorname{cis}\frac{2\pi}{3} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $\omega^2 = \operatorname{cis}\frac{4\pi}{3} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$: a real root and a conjugate pair on the unit circle.

For $n \geq 2$ they also **sum to zero**. There are two ways to see this. Geometrically, the n arrows point in n equally spaced directions, so they cancel. Algebraically, the geometric series of Ch. 2 gives the result directly, since $\omega \neq 1$:

$$1 + \omega + \omega^2 + \cdots + \omega^{n-1} = \frac{\omega^n - 1}{\omega - 1} = \frac{1 - 1}{\omega - 1} = 0.$$

The same conclusion follows from Ch. 6. The polynomial $z^n - 1$ has no z^{n-1} term, so the sum of its roots is zero. For $n = 3$ this gives the identity $1 + \omega + \omega^2 = 0$, which the next example uses twice.

EXAMPLE 17.8

Let $\omega = \operatorname{cis}\dfrac{2\pi}{3}$, so that $\omega^3 = 1$ and $1 + \omega + \omega^2 = 0$. Simplify (a) $\omega^{10} + \omega^{20}$, and (b) $(1 + \omega)(1 + \omega^2)$.

(a) A power of ω is reduced by dividing the exponent by 3 and keeping the remainder, since

$\omega^3 = 1$. Here $10 = 3(3) + 1$ and $20 = 3(6) + 2$, so

$$\omega^{10} + \omega^{20} = \omega + \omega^2 = -1,$$

the last step because $1 + \omega + \omega^2 = 0$.
(b) Expand, and use $\omega^3 = 1$ again:

$$(1 + \omega)(1 + \omega^2) = 1 + \omega^2 + \omega + \omega^3 = (1 + \omega + \omega^2) + 1 = 0 + 1 = 1.$$

Both parts turned on two facts: $\omega^3 = 1$ reduces any power of ω, and $1 + \omega + \omega^2 = 0$ removes what is left. Apply them in that order.

These evenly spaced points have a large practical use. The *discrete Fourier transform* separates a signal, such as a sound, into the pure frequencies it is made of. It does this by multiplying the n samples by powers of these n points on the circle and adding the results. A fast method of computing it, the fast Fourier transform, uses the symmetry of the polygon to cut the work, and it is one of the methods that let a phone process sound as it arrives.

17.4.2 Square Roots Without Polar Form

Polar form finds roots of every order, but it has one drawback. It needs the argument, and unless the angle is one of the standard ones, $\arctan$ returns a decimal. The answer is then no longer exact. For *square* roots there is a second route that stays in Cartesian form and keeps the surds intact. Assume the root is $a + bi$, square it, and equate real and imaginary parts as in §17.1.

EXAMPLE 17.9

Find the square roots of $7 + 24i$ in the form $a + bi$.
Let $(a + bi)^2 = 7 + 24i$. Expanding the left side gives $a^2 - b^2 + 2abi$, so equating parts,

$$a^2 - b^2 = 7 \qquad \text{and} \qquad 2ab = 24.$$

The second gives $b = \dfrac{12}{a}$. Substituting into the first,

$$a^2 - \frac{144}{a^2} = 7 \implies a^4 - 7a^2 - 144 = 0 \implies (a^2 - 16)(a^2 + 9) = 0.$$

Since a is real, $a^2 = 16$, so $a = \pm 4$ and correspondingly $b = \pm 3$. The square roots are $4 + 3i$ and $-4 - 3i$, that is $\pm(4 + 3i)$.

Two features of that calculation repeat every time. When the imaginary part is non-zero, the substitution always produces a **quadratic in a^2** with one positive and one negative solution, and the negative one is rejected, because a is real. The two answers always differ by a sign, since square roots lie at opposite ends of a diameter, half a turn apart.

CAUTION Use Cartesian equating for square roots when the answer should be exact. For cube roots and higher, go back to polar form: equating parts would give a cubic or an equation of even higher degree, while De Moivre handles roots of any order directly.

17.4.3 Complex Roots of Polynomials

Every nth root above was a solution of $z^n = w$, which is a polynomial equation of degree n with exactly n solutions. That is not a coincidence of this one family. This is the **Fundamental Theorem of Algebra**: every polynomial of degree $n \geq 1$ has exactly n roots in the complex numbers, counted with multiplicity. The roots of a polynomial need not lie on the real line, but none of them can lie outside the complex plane.

The simplest case has already appeared, in a different form. In Ch. 4, a quadratic with negative discriminant had "no solutions." It always had two roots; they were simply not on the real line.

EXAMPLE 17.10

Solve $z^2 - 4z + 13 = 0$.

The quadratic formula works unchanged:

$$z = \frac{4 \pm \sqrt{16 - 52}}{2} = \frac{4 \pm \sqrt{-36}}{2} = \frac{4 \pm 6i}{2} = 2 \pm 3i.$$

The negative discriminant simply produces an i: the two roots are $2 + 3i$ and $2 - 3i$, a conjugate pair placed symmetrically above and below the real axis.

That symmetry is no accident. When a polynomial has *real* coefficients, its non-real roots always come in conjugate pairs.

KEY **Conjugate root theorem.** If a polynomial has real coefficients and $z = a + bi$ is a root, then $z^* = a - bi$ is also a root.

This is why a cubic with real coefficients always has at least one real root: its three roots are either all real, or one real root plus a conjugate pair. A non-real root never appears alone in a polynomial with real coefficients.

To use the theorem, turn the pair into a real factor. A conjugate pair $a \pm bi$ gives the real quadratic factor $(z - a)^2 + b^2 = z^2 - 2az + (a^2 + b^2)$. Divide the polynomial by this factor, or compare coefficients, to find the remaining roots.

EXAMPLE 17.11

Given that $2 + i$ is a root of a polynomial with real coefficients, state another root and find a real quadratic factor.

By the conjugate root theorem, $2 - i$ is also a root. The two together give a quadratic factor

with real coefficients:

$$(z-(2+i))(z-(2-i)) = (z-2)^2 - i^2 = z^2 - 4z + 5.$$

This is the real quadratic factor $(z-a)^2+b^2$ with $a=2$ and $b=1$; dividing it out of the polynomial gives the remaining roots.

YOUR TURN 17.4 Roots of Complex Numbers

How many roots, and where

22.

(a) Write down a square root of -9.

(b) State the modulus of each cube root of 27.

(c) State how many distinct fifth roots a non-zero complex number has.

(d) State the shape formed on the Argand diagram by the nth roots ($n \geq 3$) of a non-zero complex number.

Roots of unity

23.

(a) State the three cube roots of unity, in Cartesian form.

(b) State the four fourth roots of unity.

(c) State the six sixth roots of unity, in modulus–argument form.

(d) Let ω be a cube root of unity with $\omega \neq 1$. Show that $1+\omega+\omega^2 = 0$.

24. Let $\omega = \operatorname{cis}\frac{2\pi}{3}$, so that $\omega^3 = 1$ and $1+\omega+\omega^2 = 0$.

(a) Simplify $(1-\omega)(1-\omega^2)$.

(b) Show that $\left(1+\omega^2\right)^4 = \omega$.

Finding nth roots

25. Find, in modulus–argument form with principal arguments,

(a) the three cube roots of 8;

(b) the four fourth roots of 16;

(c) the three cube roots of i.

26. Find the square roots of each number in the form $a+bi$ by equating real and imaginary parts.

(a) $5+12i$ (b) $-8+6i$

(c) $-5+12i$

Polynomials with real coefficients

27. Solve each equation.

(a) $z^2 + 9 = 0$
(b) $z^2 - 2z + 5 = 0$
(c) $z^2 + 2z + 2 = 0$
(d) $2z^2 - 2z + 5 = 0$

28.

(a) Given that $3 + i$ is a root of a polynomial with real coefficients, state another root.
(b) Write down a quadratic with real coefficients whose roots are $2i$ and $-2i$.
(c) Write down a quadratic with real coefficients whose roots are $-2 + 5i$ and $-2 - 5i$.

29.

(a) The equation $z^2 + bz + c = 0$ has real coefficients and $3 - 2i$ as a root. Find b and c.
(b) Explain why a cubic with real coefficients must have at least one real root.

30.

(a) Given that $1 - i$ is a root of $z^3 - 4z^2 + 6z - 4 = 0$, find the other two roots.
(b) Given that $2i$ is a root of $z^4 - 2z^3 + 5z^2 - 8z + 4 = 0$, solve the equation.
(c) The equation $z^3 + az^2 + bz - 26 = 0$, where $a, b \in \mathbb{R}$, has $2 - 3i$ as a root. Find a and b, and the real root.

Concept check

31. A student solves $z^3 = 8i$ and writes $z = 2\,\mathrm{cis}\left(\frac{\pi}{6} + 2k\pi\right)$ for $k = 0, 1, 2$. Explain why this gives only one root, and find all three roots in modulus–argument form, with principal arguments.

32. A student says: "*i is a root of $z^2 - 3iz - 2 = 0$, so $-i$ is also a root, because complex roots come in conjugate pairs.*" Show that $-i$ is not a root, and state the condition the student has overlooked.

17.5 Trigonometric Identities from De Moivre

The compound-angle rules of Ch. 8 give $\cos 2\theta$ and $\sin 2\theta$ without much work. For $\cos 3\theta$ they take considerably more. Each multiple angle has to be built from the one before it, so the work grows with every step. One expansion with complex numbers gives any multiple angle at once, and it gives the sine identity and the cosine identity together.

The method comes from one observation. De Moivre's theorem says that $(\cos\theta + i\sin\theta)^n$ and $\cos n\theta + i\sin n\theta$ are the same number. So one number has been written in two ways. Expand the first way with the binomial theorem of Ch. 2. The two ways must then match in

both parts: the real part of one equals the real part of the other, and the imaginary parts match in the same way.

KEY To find an identity for $\cos n\theta$ or $\sin n\theta$:

1. Expand $(\cos\theta + i\sin\theta)^n$ by the binomial theorem.
2. Collect the terms into a real part and an imaginary part, using $i^2 = -1$.
3. Equate the real part to $\cos n\theta$ and the imaginary part to $\sin n\theta$.
4. Replace $\sin^2\theta$ by $1 - \cos^2\theta$, or the reverse, to leave one function only.

Each expansion is compared in two parts, so it always produces two identities at once. Divide one identity by the other and you have a third: the imaginary part over the real part is $\tan n\theta$, and the real part over the imaginary part is $\cot n\theta$.

Worked through. To show that $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$, and to find a matching identity for $\sin 3\theta$, write $c = \cos\theta$, $s = \sin\theta$, and expand by the binomial theorem:

$$(c + is)^3 = c^3 + 3c^2(is) + 3c(is)^2 + (is)^3 = (c^3 - 3cs^2) + i(3c^2 s - s^3).$$

By De Moivre this equals $\cos 3\theta + i\sin 3\theta$, so the parts must match separately. Real parts, using $s^2 = 1 - c^2$:

$$\cos 3\theta = c^3 - 3c(1 - c^2) = 4\cos^3\theta - 3\cos\theta.$$

Imaginary parts, using $c^2 = 1 - s^2$:

$$\sin 3\theta = 3(1 - s^2)s - s^3 = 3\sin\theta - 4\sin^3\theta.$$

One expansion gives two identities, and the same method works for any n. Results that took several steps with compound angles in Ch. 8 now follow from a single binomial expansion.

A third identity comes from the same expansion and needs no new work: divide one part by the other.

EXAMPLE 17.12

Use the expansion above to show that $\cot 3\theta = \dfrac{\cot^3\theta - 3\cot\theta}{3\cot^2\theta - 1}$.

Take the two parts of the expansion *before* either was rewritten in a single function:

$$\cos 3\theta = c^3 - 3cs^2, \qquad \sin 3\theta = 3c^2 s - s^3.$$

Since $\cot 3\theta$ is the first divided by the second,

$$\cot 3\theta = \frac{c^3 - 3cs^2}{3c^2 s - s^3}.$$

Now divide every term, top and bottom, by s^3. Each division turns a c over an s into a $\cot\theta$:

$$\cot 3\theta = \frac{\dfrac{c^3}{s^3} - \dfrac{3cs^2}{s^3}}{\dfrac{3c^2 s}{s^3} - \dfrac{s^3}{s^3}} = \frac{\cot^3\theta - 3\cot\theta}{3\cot^2\theta - 1}. \qquad \blacksquare$$

Two points make this reliable. Use the parts in their original mixed form, since replacing s^2 by $1 - c^2$ first would leave a mixture that no division tidies. And divide by the highest power present, here s^3, so that every term becomes a power of a single ratio.

17.5.1 Powers of Cosine and Sine

The method above turns a multiple angle into powers of $\cos\theta$. The same idea works in the other direction. A power such as $\cos^6\theta$ becomes a sum of multiple angles, and a sum of multiple angles can be integrated directly with the methods of Ch. 15.

The tool is the pair of results from §17.3. For $z = \operatorname{cis}\theta$,

$$z + \frac{1}{z} = 2\cos\theta, \qquad\qquad z^n + \frac{1}{z^n} = 2\cos n\theta,$$

and the same expressions with a minus sign give $2i\sin\theta$ and $2i\sin n\theta$. So raise $z + \frac{1}{z}$ to a power and expand it. Each pair of terms then combines into a cosine of a multiple angle.

EXAMPLE 17.13

Express $\cos^6\theta$ in terms of $\cos 2\theta$, $\cos 4\theta$ and $\cos 6\theta$.
Start from $2\cos\theta = z + \dfrac{1}{z}$ and raise both sides to the sixth power, with binomial coefficients $1, 6, 15, 20, 15, 6, 1$:

$$64\cos^6\theta = \left(z + \frac{1}{z}\right)^6 = z^6 + 6z^4 + 15z^2 + 20 + \frac{15}{z^2} + \frac{6}{z^4} + \frac{1}{z^6}.$$

Now pair the terms that are reciprocals of each other:

$$64\cos^6\theta = \left(z^6 + \frac{1}{z^6}\right) + 6\left(z^4 + \frac{1}{z^4}\right) + 15\left(z^2 + \frac{1}{z^2}\right) + 20 = 2\cos 6\theta + 12\cos 4\theta + 30\cos 2\theta + 20.$$

Divide by 64:

$$\cos^6\theta = \tfrac{1}{32}\cos 6\theta + \tfrac{3}{16}\cos 4\theta + \tfrac{15}{32}\cos 2\theta + \tfrac{5}{16}.$$

The pairing is what makes the method work. In the expansion, the first term z^6 is the reciprocal of the last term $\frac{1}{z^6}$ and has the same binomial coefficient, and the same holds for the second term and the second from last. Every pair therefore combines into a cosine.

TIP A power of $\sin\theta$ works the same way, starting from $2i\sin\theta = z - \frac{1}{z}$. Take one extra care. Raising i to the power leaves a factor that must be divided out at the end, and an odd power gives sines rather than cosines. Check your answer by putting $\theta = \frac{\pi}{2}$, where $\sin\theta = 1$.

TIP Complex numbers offer several good IA topics. Three directions to explore: iterate $z \mapsto z^2 + c$ starting from $z = 0$ and investigate which values of c make the sequence escape to infinity (this is how the Mandelbrot set is defined); investigate which regular polygons can be constructed with ruler and compass, a question the Greeks asked, which Gauss advanced at eighteen when he used exactly the roots of unity above to construct the regular 17-gon, and which Wantzel completed in 1837; or measure phase shift in a simple AC circuit and model it with rotating complex numbers.

YOUR TURN 17.5 Trigonometric Identities from De Moivre

Multiple angles as powers

33. Use De Moivre's theorem and the binomial theorem for each part.

(a) Show that $\cos 2\theta = 2\cos^2\theta - 1$.

(b) Find a matching identity for $\sin 2\theta$.

(c) Show that $\cos 4\theta = 8\cos^4\theta - 8\cos^2\theta + 1$.

(d) Hence, or otherwise, express $\tan 2\theta$ in terms of $\tan\theta$.

Powers as multiple angles

34. Let $z = \operatorname{cis}\theta$. Use $z \pm \dfrac{1}{z}$, where $z + \dfrac{1}{z} = 2\cos\theta$ and $z - \dfrac{1}{z} = 2i\sin\theta$.

(a) Express $\cos^2\theta$ in terms of $\cos 2\theta$.

(b) Express $\sin^2\theta$ in terms of $\cos 2\theta$.

(c) Express $\cos^3\theta$ in terms of $\cos 3\theta$ and $\cos\theta$.

35.

(a) Show that $\sin^3\theta = \frac{1}{4}(3\sin\theta - \sin 3\theta)$.

(b) Express $\cos^5\theta$ in terms of cosines of multiple angles.

Integrating a power of sine or cosine

36.

(a) Use your answer to Question 34(c) to evaluate $\displaystyle\int_0^{\pi/2} \cos^3\theta\, d\theta$.

(b) Use the result of Question 35(a) to find $\displaystyle\int \sin^3\theta\, d\theta$.

Concept check

37. A student uses $z = \operatorname{cis}\theta$ and writes

$$\left(z - \frac{1}{z}\right)^2 = (2i\sin\theta)^2 = 4\sin^2\theta \quad \text{and} \quad \left(z - \frac{1}{z}\right)^2 = z^2 - 2 + \frac{1}{z^2} = 2\cos 2\theta - 2,$$

and concludes that $\sin^2\theta = \frac{1}{2}(\cos 2\theta - 1)$. Show that this result is false by taking $\theta = \frac{\pi}{2}$, identify the error, and correct the result.

Chapter 17 · Complex Numbers

Reflections

TOK Euler's identity $e^{i\pi} + 1 = 0$ has often been called the most beautiful equation in mathematics. What is being judged when a mathematician calls a result beautiful, and is that judgement evidence of anything? Beauty in mathematics is often described as economy, surprise, or unexpected connection, all of which this identity has.

IM The name "imaginary" was meant as a dismissal. Descartes coined it in the seventeenth century for numbers he regarded as fictions, and Leibniz called them "an amphibian between being and non-being." Their rehabilitation was slow and international. The Norwegian-born Danish surveyor Caspar Wessel and the Swiss bookkeeper Jean-Robert Argand independently drew the plane picture that made them concrete. The German mathematician Carl Friedrich Gauss did much to make them respectable: his 1799 proof of the Fundamental Theorem of Algebra was the first to be widely accepted, although it contained a gap that was closed only in 1920, and he popularised the term "complex" itself. Gauss also remarked that these numbers would have met less resistance if they had been called "lateral" rather than imaginary. Does the name of an idea change how readily it is believed?

RW Complex numbers are a standard tool for describing oscillations. An alternating current, a radio wave or a vibration is modelled as a complex number rotating about the origin, and its real part is the signal you measure. Engineers add and scale these "phasors" directly, in place of pages of trigonometry. Quantum mechanics goes further: a particle's state is a complex-valued wave, and i appears explicitly in the Schrödinger equation, which describes how that state changes over time. Phone signal processing, power-grid impedance and quantum computing all use the numbers of this chapter.

EXT De Moivre's theorem connects two families of functions. The identity $e^{i\theta} = \cos\theta + i\sin\theta$ shows that the exponential and trigonometric functions, which look nothing alike on the real line, are closely related once complex inputs are allowed: $\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2}$ and $\sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}$. Exponential growth and oscillation both come from the one function e^z. Push this further and you reach complex analysis, the study of functions of a complex variable. Such functions turn out to be far more tightly constrained than real ones, and some problems about ordinary integers, such as how the prime numbers are distributed, have been partly answered by studying functions over the complex plane.

End-of-Chapter Problems

1. Given $z = 3 + 4i$ and $w = 1 - 2i$:

(a) find $z + w$; [1]
(b) find zw; [2]
(c) find $\dfrac{z}{w}$; [2]
(d) find $|zw|$. [1]

2. Express $\dfrac{2+i}{3-i}$ in the form $a + bi$. [4]

3. Find the modulus and the principal argument of

(a) $z = -1 + i$; [3]
(b) $z = \sqrt{3} - i$. [3]

4. Write $z = 1 + \sqrt{3}\,i$ in modulus–argument form and in Euler form. [4]

5. Use De Moivre's theorem to evaluate

(a) $(1 + \sqrt{3}\,i)^6$; [3]
(b) $(1 + i)^{10}$; [3]
(c) $\left(\dfrac{1+i}{\sqrt{2}}\right)^{12}$. [3]

6. Find the square roots of each number in the form $a + bi$.

(a) $3 + 4i$ [4]
(b) $-21 - 20i$ [4]

7. Given $z = 3\,\mathrm{cis}\,\frac{3\pi}{4}$, giving each answer in modulus–argument form with $-\pi < \theta \leq \pi$:

(a) find z^2; [2]
(b) find z^3. [2]

8. The equation $z^2 + bz + c = 0$ has real coefficients and $-2 + 5i$ as a root. Find b and c. [4]

9. Find the three cube roots of $8i$, giving each

(a) in modulus–argument form; [3]

(b) in exact Cartesian form. [3]

10. Simplify

(a) $i^{2027} + i^{2028}$; [2]
(b) i^{-7}. [2]

11. Find all solutions of $z^3 = -4\sqrt{3} + 4i$, giving each in modulus–argument form with $-\pi < \theta \leq \pi$. [4]

12. Find real numbers x and y such that $(x + yi)(2 - i) = 3 + i$. [4]

13.

(a) Express -8 in the form $re^{i\theta}$, with $-\pi < \theta \leq \pi$. [3]
(b) The complex number z has $|z| = 6$ and $\arg z = \frac{3\pi}{4}$. Write z in Cartesian form. [3]

14.

(a) Find the four fourth roots of -16, giving each in the form $a + bi$. [5]
(b) Find all solutions of $z^6 = 64$, giving each in the form $a + bi$, and show that they sum to zero. [4]

15. Solve $2z^2 + 2z + 5 = 0$. [4]

16. Given $z = \cos\theta + i\sin\theta$ with $\cos\theta \neq 0$, show that $\dfrac{z^2 - 1}{z^2 + 1} = i\tan\theta$. [4]

17. The points representing 1, i, -1 and $-i$ are plotted on an Argand diagram. State the shape they form and find its area. [3]

18. Given that $-1 + 4i$ is a root of a polynomial with real coefficients, find a real quadratic factor of that polynomial. [4]

19. Solve $z^2 = 2i$, giving the roots in the form $a + bi$. [4]

20. Given $z = 2 + 2i$, find $|z|$ and $\arg z$, and hence write z in Euler form. [4]

21. Find the complex number z satisfying $3z + iz^* = 5 + 7i$. [4]

22. Prove De Moivre's theorem, $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$, for all $n \in \mathbb{Z}^+$ by mathematical induction. [7]

23. Use De Moivre's theorem to show that $\sin 5\theta = 16\sin^5\theta - 20\sin^3\theta + 5\sin\theta$. [6]

24. Use De Moivre's theorem to show that $\tan 4\theta = \dfrac{4t - 4t^3}{1 - 6t^2 + t^4}$, where $t = \tan\theta$. [7]

25. One cube root of a complex number w is $1 + \sqrt{3}\,i$. Find w and the other two cube roots. [6]

26. GDC In this question you will investigate the fifth roots of unity, the solutions of $z^5 = 1$.

(a) Write down the five roots in the form $\operatorname{cis}\theta$. [3]
(b) Show that the five roots sum to zero, using the sum of a geometric series. [3]
(c) Let $\omega = \operatorname{cis}\frac{2\pi}{5}$. Show that every fifth root of unity is a power of ω. [2]
(d) Deduce that $\omega + \omega^2 + \omega^3 + \omega^4 = -1$. [2]
(e) The five roots are plotted on an Argand diagram and joined to form a regular pentagon. Show that its area is $\frac{5}{2}\sin\frac{2\pi}{5}$, and evaluate this to three significant figures. [5]

27. Let $z = \operatorname{cis}\theta$.

(a) Show that $z^n + \dfrac{1}{z^n} = 2\cos n\theta$ for any $n \in \mathbb{Z}^+$. [3]
(b) Expand $\left(z + \dfrac{1}{z}\right)^4$ using the binomial theorem, and hence show that $16\cos^4\theta = 2\cos 4\theta + 8\cos 2\theta + 6$. [5]
(c) Deduce that $\cos^4\theta = \frac{1}{8}(\cos 4\theta + 4\cos 2\theta + 3)$. [2]
(d) Hence evaluate $\displaystyle\int_0^{\pi/2} \cos^4\theta\,d\theta$ exactly. [4]
(e) Explain what this method achieved that integrating $\cos^4\theta$ directly would not. [2]

28. The polynomial $P(z) = z^4 - 2z^3 + 14z^2 - 18z + 45$ has real coefficients, and $1 + 2i$ is a root.

(a) Write down another root of $P(z)$, and give a reason. [2]
(b) Find a real quadratic factor of $P(z)$. [3]
(c) Find the other quadratic factor. [3]
(d) Hence write down all four roots of $P(z)$. [3]
(e) The four roots are plotted on an Argand diagram. Describe the symmetry of the picture, and state the feature of $P(z)$ that guarantees it. [3]

Challenge Questions

29. Chords of a regular polygon. The nth roots of unity, $1, \omega, \omega^2, \ldots, \omega^{n-1}$ with $\omega = \operatorname{cis}\frac{2\pi}{n}$, are the vertices of a regular n-sided polygon inscribed in the unit circle. From

the vertex 1, draw the $n-1$ chords to the other vertices. The length of the chord to ω^k is $|1-\omega^k|$. This question asks for the product of those $n-1$ lengths, for every $n \geq 2$.

(a) Take $n=3$. Find $|1-\omega|$ and $|1-\omega^2|$ in exact form, and show that their product is 3. [3]

(b) Take $n=4$, so the vertices are $1, i, -1, -i$. Find the three chord lengths from 1 and their product. Using (a) and (b), conjecture the product of the $n-1$ chord lengths for a general n. [3]

(c) Explain why $z^n - 1 = (z-1)(z-\omega)(z-\omega^2)\cdots(z-\omega^{n-1})$ for every complex z. By also factorising $z^n - 1 = (z-1)\left(1+z+z^2+\cdots+z^{n-1}\right)$, deduce that

$$(z-\omega)(z-\omega^2)\cdots(z-\omega^{n-1}) = 1+z+z^2+\cdots+z^{n-1}$$

for every complex z, including $z=1$. [4]

(d) Hence prove that $\displaystyle\prod_{k=1}^{n-1}(1-\omega^k) = n$, and deduce your conjecture from (b). [3]

(e) Show that $1-\operatorname{cis}\theta = -2i\sin\frac{\theta}{2}\operatorname{cis}\frac{\theta}{2}$, and deduce that $|1-\omega^k| = 2\sin\dfrac{k\pi}{n}$ for $k=1,2,\ldots,n-1$. Hence show that

$$\prod_{k=1}^{n-1}\sin\frac{k\pi}{n} = \frac{n}{2^{n-1}},$$

and check this directly for $n=6$. [5]

(f) Now let $P = \operatorname{cis}\theta$ be any point on the circle. Investigate the product of the distances from P to all n vertices: find a formula in terms of θ, its greatest value, and where on the circle P must be to attain it. [3]

30. **Adding waves.** Two loudspeakers play the same pure note. At a listening point the two arrivals are $A\cos\omega t$ and $B\cos(\omega t+\phi)$, where $A>0$, $B>0$, and ϕ is the phase difference caused by the extra distance one sound has travelled. Adding them with trigonometry is lengthy; with complex numbers each wave becomes a vector in the Argand diagram, and the waves add as vectors do.

(a) Show that $A\cos\omega t + B\cos(\omega t+\phi) = \operatorname{Re}\left[(A+B\operatorname{cis}\phi)\operatorname{cis}\omega t\right]$. [3]

(b) Deduce that the combined sound is a wave of the same frequency, with amplitude $|A+B\operatorname{cis}\phi|$ and phase $\arg(A+B\operatorname{cis}\phi)$. [2]

(c) Two speakers each have amplitude 3 and the phase difference is $\frac{\pi}{3}$. Find the amplitude and phase of the combined sound, in exact form. [4]

(d) Now let both amplitudes equal A, and take $-\pi \leq \phi \leq \pi$. Show that $A + A\operatorname{cis}\phi = 2A\cos\dfrac{\phi}{2}\operatorname{cis}\dfrac{\phi}{2}$. Deduce that the combined amplitude is $2A\cos\dfrac{\phi}{2}$, that the phase is $\dfrac{\phi}{2}$ when $-\pi < \phi < \pi$, and state what happens when $\phi = \pm\pi$. [4]

(e) Use (d) to explain what a listener hears when $\phi = 0$ and when $\phi = \pi$, and why noise-cancelling headphones work by generating a wave with $\phi = \pi$. [3]

(f) Now n speakers in a row each have amplitude A, and each arrival is a phase ϕ behind

the one before, so the arrivals are $A\cos(\omega t + k\phi)$ for $k = 0, 1, \dots, n-1$, with $0 < \phi < 2\pi$. Use the sum of a geometric series in $\operatorname{cis}\phi$ to show that the combined amplitude is

$$A\left|\frac{\sin(n\phi/2)}{\sin(\phi/2)}\right|.$$

(Hint: $\operatorname{cis}\alpha - 1 = 2i\sin\frac{\alpha}{2}\operatorname{cis}\frac{\alpha}{2}$, found by taking out the factor $\operatorname{cis}\frac{\alpha}{2}$ as in (d)) [4]

(g) For $n = 4$, find every ϕ in $0 < \phi < 2\pi$ that gives silence. Then investigate how the loud region around $\phi = 0$ changes as n increases, and suggest why an array of many speakers can aim sound in one direction. [3]

31. From a circle to an aeroplane wing. The map $w = z + \dfrac{1}{z}$ is called the **Joukowski transformation**. Early aerodynamicists used it because it turns circles, whose airflow is easy to calculate, into wing-shaped curves, whose airflow is not.

(a) Let $z = \operatorname{cis}\theta$, so $|z| = 1$. Show that $w = 2\cos\theta$, and hence that the unit circle maps onto the line segment from -2 to 2. [4]

(b) Now let $z = r\operatorname{cis}\theta$ with $r > 1$. Writing $w = x + iy$, show that

$$x = \left(r + \tfrac{1}{r}\right)\cos\theta, \qquad y = \left(r - \tfrac{1}{r}\right)\sin\theta,$$

and deduce that $\dfrac{x^2}{\left(r+\frac{1}{r}\right)^2} + \dfrac{y^2}{\left(r-\frac{1}{r}\right)^2} = 1$. [5]

(c) That equation describes an ellipse with semi-axes $r + \frac{1}{r}$ and $r - \frac{1}{r}$. Find both for $r = 2$. [2]

(d) Describe what happens to the ellipse as $r \to 1^+$, and reconcile your answer with part (a). [3]

(e) A wing needs a rounded front and a sharp trailing edge. Let $0 < a < 1$ and let Γ be the circle with centre $-a$ and radius $1 + a$. Show that Γ passes through $z = 1$, and that $z = -1$ lies inside Γ. [2]

(f) The derivative $\dfrac{dw}{dz} = 1 - \dfrac{1}{z^2}$ is zero only at $z = \pm 1$, and at a point where it is zero the map doubles angles instead of preserving them. Use this and (e) to explain why the image of Γ has a sharp point at $w = 2$ and is smooth everywhere else. [3]

(g) Real wings are curved (cambered) as well as thick. Investigate the image of the circle through $z = 1$ whose centre is $-a + bi$, with $a > 0$ and $b > 0$: state its radius, and describe how a and b separately change the shape of the image. [2]

32. How a phone hears a chord. A microphone records a sound as a list of numbers, one per sampling instant. To find which frequencies are present, the recording is compared against the roots of unity. For four samples x_0, x_1, x_2, x_3, the **discrete Fourier transform** is

$$X_k = \sum_{n=0}^{3} x_n\, i^{-nk}, \qquad k = 0, 1, 2, 3,$$

where $i = \operatorname{cis}\frac{2\pi}{4}$ is the fourth root of unity next to 1. Each X_k measures how strongly the frequency k is present, meaning k complete cycles across the four samples.

(a) Write down the four fourth roots of unity and verify that they sum to zero. [2]

(b) A steady tone with no variation gives samples $(1, 1, 1, 1)$. Compute X_0, X_1, X_2, X_3. [4]

(c) Explain, using part (a), why every X_k with $k \neq 0$ had to vanish in (b) (for $k = 2$, note that $i^{-2} = -1$), and say what X_0 measures. [3]

(d) A signal that rises and falls once across the four samples gives $(1, 0, -1, 0)$. Compute X_0, X_1, X_2, X_3. [4]

(e) Interpret your answer to (d): which frequency is present, and what does $X_0 = 0$ tell you about the signal? [3]

(f) A chord is two notes at once. The samples $(2, -1, 0, -1)$ are the sum of the signal in (d) and the signal $(1, -1, 1, -1)$. Compute the transform of $(2, -1, 0, -1)$, and explain how it shows both notes. [3]

(g) Prove that the samples can be recovered from the transform by

$$x_n = \frac{1}{4}\sum_{k=0}^{3} X_k\, i^{nk}, \qquad n = 0, 1, 2, 3,$$

and check this for x_0 in (f). [4]

(h) Suggest how the transform and the formula in (g) extend to N samples. Then investigate what the transform gives for a tone that does not complete a whole number of cycles across the samples. [2]

33. The escape radius of the Mandelbrot set. Fix a complex number c, start at $z_0 = 0$, and iterate

$$z_{n+1} = z_n^2 + c.$$

For some c the sequence stays bounded forever; for others it grows without bound. The set of c for which it stays bounded is the **Mandelbrot set**, and a computer draws it by testing one c per pixel. This question finds the circle outside which no point of the set can lie, and the test a computer uses to stop.

(a) Take $c = i$. Compute z_1, z_2, z_3 and z_4, and show that the sequence then repeats with period 2. Is i in the set? [4]

(b) Take $c = 1$. Compute z_1, z_2, z_3, z_4 and say what is happening. [2]

(c) Show that if $|z| > 2$ and $|z| \geq |c|$ then $|z^2 + c| \geq |z|(|z| - 1) > |z|$. *(Hint: $|z^2 + c| \geq |z|^2 - |c|$)* [4]

(d) Suppose $|c| > 2$. Show that z_1 satisfies both conditions of (c), and prove by induction that $|z_n| \geq |c|(|c| - 1)^{n-1}$ for $n \geq 1$. Deduce that every c in the Mandelbrot set satisfies $|c| \leq 2$. [4]

(e) Now suppose $|c| \leq 2$, and that $|z_m| > 2$ for some m. Use (c) to show that the sequence then grows without bound. Combining this with (d), explain why a computer can stop testing a pixel, whatever c is, as soon as $|z_n| > 2$. [3]

(f) Investigate which real numbers c lie in the Mandelbrot set. Use the fixed points of $z \mapsto z^2 + c$ and a few iterations to suggest the interval, and justify at least one of its endpoints. [3]

BetterMath

Intuitive. Clear. Uncompromising.

IB Diploma Programme

Mathematics

Analysis & Approaches

Higher Level

Hasan Khan

Teacher of IB Mathematics and Physics · IB Examiner · National curriculum contributor
M.Sc. Particle Physics, University of Jyväskylä · PGD Education
Training with CERN, ESA and Wolfram Research

> I have taught for about ten years, in Denmark, Tanzania and now Japan, and this book grew out of those classrooms and my students' questions. Graphic design is a hobby I take far too seriously, which is why the book looks the way it does. Corrections and suggestions are always welcome.
>
> Away from teaching, I am a *pathfinder*, more interested in the route than the arrival.

ibxhasan.github.io/bettermath

Updated September 2026

· © 2026 Hasan Khan

All rights reserved · permission requests: ibxhasan@gmail.com

Not affiliated with or endorsed by the International Baccalaureate Organization.