MATH · ANALYSIS & APPROACHES SL

Paper 1
everything you need

A complete revision outline for the non-calculator paper — the five syllabus topics, the exam craft that turns understanding into marks, and the traps that cost students marks they have already earned.

90 minutes · 80 marks · no calculator · formula booklet provided

AA SL syllabus, first assessment 2021 · current for May 2026

Generated 14 May 2026 · private revision document

Contents

Read §How Paper 1 Works first — most of the recoverable marks live in exam craft, not topic content. Then the five topics. Every formula is flagged as given-in-the-booklet or must-memorise. All worked examples are original.

How Paper 1 Works — and How to Not Lose Marks You've Earned

A revision guide is only half the job. The other half is exam craft: knowing what the paper is, how it is marked, and where students who *understand the maths* still bleed marks. Read this section first. Most of the recoverable marks on Paper 1 live here, not in the topic content.

The paper

Paper 1 is **90 minutes, 80 marks, no calculator**. A clean copy of the formula booklet is provided. Two sections:

- **Section A** — roughly six shorter questions, around 50 marks total. The paper escalates: Q1 is the gentlest, the questions lengthen as you go.
- **Section B** — three extended-response questions of roughly 13–17 marks each, around 30 marks total. Each is a *chain* of subparts (a), (b), (c)... where later parts often depend on earlier ones.

The average pace is **1.125 minutes per mark**. Section A is mark-rich (you can move fast). Section B closers — kinematics, optimisation, trig-identity chains — are the time-sinks; budget extra. The hardest question on the paper is almost always the last one, Q9, not Q8.

Paper 1 is 40% of your final SL grade (Paper 2 is the other 40%, the IA is 20%).

The six habits that move your score

Drill these until they are automatic. They are not knowledge — they are reflexes, and each one is worth marks on almost every paper.

1. **Announce every method in writing.** "By the chain rule." "By the product rule." "Since $|r| < 1$." "By the discriminant." "By the cosine rule." Each phrase banks a method mark that survives even when the arithmetic that follows is wrong.

1. **State $f''(x_0)$ as a number and its sign before declaring a maximum or minimum.** Writing "this is a maximum" without showing the second-derivative value forfeits the reasoning mark. Stationary-point classification appears on essentially every paper.

1. **When you see sin, cos, or tan of nx, write the working domain for nx before you solve.** If $x \in [0, 2\pi]$, then $2x \in [0, 4\pi]$ — so there are four solutions for 2x, not two. Find them all, then divide and filter back.

1. **In Section B, attempt every part (a) and (b) before any part (d) or (e).** Mark-mass is front-loaded: roughly 60% of a chain's marks live in its first half. Skipping the opening parts to "save time for the hard bits" trades easy marks for hard ones.

1. **Default to exact form on every numerical answer.** The paper is non-calculator; $\pi/6$, $\sqrt{3}/2$, ln 2 are *final answers*, not intermediate forms to be approximated. A correct decimal where exact form was

required is still a wrong answer.

1. **If a "show that" stumps you, write the target result down and use it for the next part.** "Show that" almost always feeds the next subpart. The follow-through marks on that next part are recoverable even when the show-that itself is not. Never abandon a "show that" — write the conclusion line and march on.

The grammar of a mark scheme

Every mark on a Paper 1 mark scheme is one of three letters. Once you internalise what each rewards, your working starts to write itself in the form the examiner needs.

M — method mark (`M1`, occasionally `M2`). Rewards a *correct approach*, whether or not the arithmetic finishes correctly. The M is earned by the visible setup — the chain rule written with its substitution, the quadratic formula with a, b, c plugged in, the conditional-probability formula with events named — *before* simplifying. A correct method that lives only in your head is, for marking purposes, no method at all. Phrases like "by the product rule" take ten seconds to write and explicitly nail down the M.

A — accuracy mark (`A1`, sometimes `A2`). Rewards a correct result: the right derivative, the right intersection point, the right probability — in the form the question asked for. A marks are usually *gated* by a preceding M: skip the method and write only the answer, and you can forfeit both.

R — reasoning mark (`R1`). Rewards an explicit *justification* in words or symbols. "Since the discriminant is negative, there are no real solutions." "$f''(2) = -6 < 0$, so this is a maximum." R marks appear in "show that" arguments, "explain why" parts, the final step of a proof, and anywhere a conclusion needs to be *shown* to follow. **The R mark is the single most-missed mark by competent students** — the conclusion feels obvious to them, so they don't write the sentence. The examiner does not award marks for the obvious; the examiner awards marks for the written.

FT — follow-through. If you make an error early and then correctly propagate that wrong value through the rest of the question, you can still earn the later marks — but *only if the examiner can see the working*. No visible intermediate values, no follow-through. Write down every intermediate value you use, even when the next step feels trivial. It costs ten seconds and buys back marks a single arithmetic slip would otherwise cost you.

AG — answer given. Some questions print the answer and ask you to *derive* it ("show that..."). You must show every step, you must work *forwards* (starting from the given answer and reverse-engineering earns nothing), and the final AG line itself is unmarked — the marks live in the argument above it.

OE / AEF — or equivalent / any equivalent form. If your answer looks different from the printed one but is algebraically the same, you still get the mark. When self-marking, check equivalence before docking yourself.

The implication of all this: a student who shows a full method and slips on the arithmetic typically out-scores a student who shows no method and gets the right number. The IB grades mathematical understanding, not arithmetic luck.

Command terms — what each one demands on the page

The command term tells you the *genre* of working the examiner expects. The two most-misread are **hence** (binding — you must use the previous part) and **sketch** (not a doodle — features must be labelled with their values).

- **Calculate / Find** — the requested quantity, with method visible. Formula, substitute, evaluate.
- **Determine** — slightly stronger than "find": method visible *and* the conclusion stated explicitly.
- **Solve** — all solutions in the requested domain, each one stated, with the route shown.
- **Show / Show that** — a forward derivation from premise to the given conclusion, every step visible. Cannot be done backwards. Generally needs no calculator.
- **Prove** — a rigorous argument with justification at every step; stronger than "show that".
- **Hence** — you *must* use the previous part's result. An independent method earns nothing.
- **Hence or otherwise** — you *may* use the previous part, or any other valid method; the previous-part route is usually the intended, easier one.
- **State / Write down** — a short answer, no working required. One line.
- **Sketch** — a hand-drawn graph with the qualitative features labelled: every intercept with its coordinates, every asymptote drawn dashed and labelled with its equation, every turning point indicated. An unlabelled sketch earns almost none of its marks.
- **Explain / Justify** — a sentence (or two) of reasoning. An R mark is in play; don't just compute, say *why*.

Cross-cutting traps

These lose marks in every topic:

- **Premature rounding.** Round only at the final step. An intermediate value truncated and re-used contaminates the later answer — the mark scheme will withhold the accuracy mark on the *later* answer too. On a non-calculator paper, carry exact forms; don't manufacture decimals by approximating π as 3.14 mid-working.
- **Decimal when exact was required.** "Give your answer in exact form" / "in terms of π" / "in the form $a + b\sqrt{c}$" are non-negotiable instructions. A correct decimal still loses the mark.
- **Missing units.** When the quantity is physical — length, area, volume, time, velocity — the units must appear in the final answer.
- **Rule used but not announced.** Method marks reward the *visible* application of a named rule. Mental algebra with only the result written down is routinely caught out by a downstream slip that retrospectively erases the method credit.
- **Working backwards in a "show that".** Forward derivation only. Substituting both sides into something and arriving at a tautology earns nothing.
- **Unlabelled sketches.** Intercepts, asymptotes, and turning points must carry their values. The shape alone is a fraction of the marks.
- **Answering a different question.** Re-read the question. If it asked for the *equation* of the tangent and you gave the *gradient*, that is zero marks, not partial credit. If it asked for the maximum *volume*

and you gave the value of x that produces it, you have not finished the question.

The self-marking discipline

When you practise past papers, the marking is half the value. After each question, before you tally:

1. For every M in the scheme, is there a corresponding visible attempt on your page?
2. Is your answer in the form demanded — exact, decimal to 3 s.f., radians, units?
3. Did you round only at the end?
4. Did you justify every "show", "prove", "explain" with a sentence?
5. Where the command was "hence", did you actually use the previous part?
6. Is a different-looking answer actually equivalent? Don't dock yourself for a valid alternative form.
7. Is enough intermediate working visible that follow-through could rescue an early slip?
8. On sketches — are the features labelled with values?
9. Did you answer the question that was actually asked?

Tag every lost mark into a category: *didn't know it* (least recoverable), *no method visible* (recoverable — it's a habit), *wrong form* (recoverable), *unjustified reasoning* (recoverable — write the sentence), *arithmetic slip* (medium). The recoverable categories are where overnight progress lives. You can fix a habit overnight; you cannot learn a topic overnight.

The 72-hour and exam-morning approach

What the evidence on learning actually supports for the final stretch:

- **Produce, don't consume.** Re-reading notes and highlighting feel productive and transfer poorly. Closing the book and reconstructing a method from a blank page is what builds exam-ready recall.
- **Mix topics rather than blocking them.** A session of one question from each of the five topics transfers better than six questions all on differentiation — the exam doesn't tell you which method to use, so practise *choosing*.
- **At least one full timed paper in exam conditions** — pen, paper, no calculator, no looking up, phone in another room. The skill of doing maths under those exact conditions is itself a thing to train.
- **Sleep is a study technique.** Memory consolidation happens during sleep. With a paper tomorrow, a normal night's sleep is not a luxury — cutting it to add study is a negative-return trade.
- **Exam morning is for recognition, not learning.** Re-read your six habits aloud. Skim the mark-scheme grammar above. Glance at the formula booklet to refresh *where* things sit, not to memorise. Re-read one worked example you've already done. Then stop, eat normally, and arrive early enough to sit down calm. No fresh timed papers, no new notes — there is no time left to consolidate anything new, and a paper that goes badly an hour before the exam only spikes anxiety.

The deepest finding of the science of learning is also the most boring: there is no trick. The only way to get better at producing maths from a blank page under time pressure is to practise producing maths from a blank page under time pressure. The rest is bookkeeping.

1. Number & Algebra

Number & Algebra is the most "rules-and-procedures" topic on the course: sequences, exponents, logs, the binomial theorem, and a first taste of formal proof. It carries roughly **17% of the SL syllabus** and shows up reliably on Paper 1 — usually as one full sequences-or-series question, log/exponent work folded into equation-solving, and a short binomial expansion. Because Paper 1 is **non-calculator**, the marks here go to clean algebraic manipulation and **exact-form answers** (leave things as `ln 3`, `2^(5/2)`, fractions — never a rounded decimal unless the question asks for accuracy).

1.1 Scientific notation, significant figures, accuracy

Key facts

- Scientific (standard) notation: `a × 10^k` where `1 ≤ a < 10` and `k` is an integer. *Not in the booklet — it's a convention you must know.*
- Significant figures count from the first non-zero digit. Trailing zeros after a decimal point count; leading zeros never do.
- Default reporting accuracy on IB papers is **3 significant figures**, unless an exact answer is available or the question states otherwise.

Method (writing a number in standard form)

1. Place the decimal point after the first non-zero digit to get `a`.
2. Count how many places it moved: that magnitude is `|k|`.
3. Moved left ⇒ `k` positive; moved right ⇒ `k` negative.

Worked micro-example Write `0.000642` in scientific notation to 2 s.f. The first non-zero digit is `6`; the point moves 4 places right, so `k = −4`. To 2 s.f., `6.42 → 6.4`. Answer: `6.4 × 10^−4`.

Traps

- `a` must satisfy `1 ≤ a < 10` — writing `64.2 × 10^−5` is not standard form.
- Don't round mid-calculation. Carry exact values or extra digits and round **once** at the end.
- `0.0500` has 3 s.f., not 1 — the trailing zeros are significant.

1.2 Arithmetic sequences and series

Key facts / formulas

- nth term: `u_n = u_1 + (n − 1)d` — *GIVEN in the formula booklet.*
- Sum of n terms: `S_n = (n/2)(2u_1 + (n − 1)d)` and equivalently `S_n = (n/2)(u_1 + u_n)` — *GIVEN.*
- Common difference: `d = u_{n+1} − u_n`, constant for all `n`.

- Recursive form: $u_{n+1} = u_n + d$ with u_1 stated. Explicit form is the u_n formula above. You must be able to move between them.

Method (typical "find a term / find a sum" question)

1. Identify u_1 and d from the information given (often two simultaneous equations from two known terms).
2. Solve for the unknowns.
3. Substitute into u_n or S_n as required.

Worked micro-example The 3rd term of an arithmetic sequence is 11 and the 7th term is 27. Find u_1, d, and S_{10}. $u_3 = u_1 + 2d = 11$ and $u_7 = u_1 + 6d = 27$. Subtract: $4d = 16$, so $d = 4$. Then $u_1 = 11 - 8 = 3$. $S_{10} = (10/2)(2\cdot 3 + 9\cdot 4) = 5(6 + 36) = 5 \cdot 42 = 210$.

Traps

- **Indexing.** The IB uses u_1 as the first term. If a problem hands you u_0, the nth term shifts to $u_n = u_0 + nd$. Read carefully.
- $(n - 1)d$, not nd, in the nth term — the first term has *zero* common differences added.
- "Sum of the first n terms" S_n is not the same as "the nth term" u_n. The question wording decides which.
- When solving simultaneously, subtract the equations to eliminate u_1 cleanly rather than substituting.

1.3 Geometric sequences and series

Key facts / formulas

- nth term: $u_n = u_1 \cdot r^{(n-1)}$ — *GIVEN.*
- Sum of n terms: $S_n = u_1(r^n - 1)/(r - 1) = u_1(1 - r^n)/(1 - r)$, for $r \neq 1$ — *GIVEN.*
- Sum to infinity: $S_\infty = u_1/(1 - r)$, valid **only if** $|r| < 1$ — *GIVEN.*
- Common ratio: $r = u_{n+1}/u_n$, constant for all n.
- Recursive form: $u_{n+1} = r \cdot u_n$.

Method (sum to infinity / convergence)

1. Find r by dividing consecutive terms.
2. **Check** $|r| < 1$ before you write the S_∞ formula — this is a reasoning (R) mark.
3. Substitute u_1 and r into $S_\infty = u_1/(1 - r)$.

Worked micro-example A geometric series has first term 12 and common ratio $1/3$. Find S_4 exactly and S_∞. $S_4 = 12(1 - (1/3)^4)/(1 - 1/3) = 12(1 - 1/81)/(2/3) = 12 \cdot (80/81) \cdot (3/2) = 18 \cdot 80/81 = 1440/81 = 160/9$. Since $|1/3| < 1$, $S_\infty = 12/(1 - 1/3) = 12/(2/3) = 18$.

Traps

- **Forgetting the $|r| < 1$ check.** Stating S_∞ without justifying convergence loses the reasoning mark every time.
- Negative r: $(-2)^3 = -8$ but $(-2)^4 = 16$ — keep the sign inside the power and watch the alternating signs.
- $r^{(n-1)}$ in the nth term vs r^n in the sum — don't mix the exponents.
- If a question gives u_1 and u_4 to find r, you get $r^3 = u_4/u_1$; on a non-calculator paper the numbers are chosen so the cube root is exact.
- "Find the smallest n such that $S_n > \dots$" needs logs to solve r^n for n — see §1.6.

1.4 Sigma notation

Key facts

- $\sum_{k=1}^{n} f(k)$ means add $f(k)$ for $k = 1, 2, \dots, n$. *Notation, not booklet content.*
- The lower limit need not be 1; the number of terms is $(\text{upper limit} - \text{lower limit} + 1)$.
- A sum of the form $\Sigma\ (ak + b)$ is **arithmetic**; a sum of the form $\Sigma\ c \cdot r^k$ is **geometric** — recognise which, then use the §1.2 / §1.3 sum formula.

Method

1. Write out the first term (substitute the lower limit) and identify the structure.
2. Decide: arithmetic or geometric? Extract u_1, and d or r.
3. Count the number of terms n carefully.
4. Apply the relevant S_n formula.

Worked micro-example Evaluate $\sum_{k=1}^{5} (3k + 2)$. Terms: $k=1 \to 5$, $k=2 \to 8$, ..., $k=5 \to 17$. This is arithmetic with $u_1 = 5$, $d = 3$, $n = 5$. $S_5 = (5/2)(5 + 17) = (5/2)(22) = 55$.

Traps

- **Off-by-one in the term count.** $\sum_{k=3}^{9}$ has $9 - 3 + 1 = 7$ terms, not 6.
- The index variable (k, r, i) is a dummy — don't confuse a sigma index r with a geometric ratio r.
- Σ distributes over addition and pulls out constants: $\Sigma(2k) = 2\Sigma k$, but $\Sigma(k^2)$ is **not** $(\Sigma k)^2$.

1.5 Applications: compound interest, depreciation, financial models

Key facts / formulas

- These are **geometric sequences in disguise.** Compound growth: $\text{value after } n \text{ periods} = P \cdot (1 + i)^n$, where i is the rate per compounding period (as a decimal).
- Depreciation: $\text{value} = P \cdot (1 - i)^n$ — same structure, ratio less than 1.

- The IB booklet gives a financial formula `FV = PV(1 + r/(100k))^(kn)` for compounding `k` times per year. *That compound-interest formula IS in the booklet.*
- Heavy numerical interest problems live on Paper 2 (calculator). On Paper 1 you'll typically **set up** the model or do a clean one- or two-step computation.

Method

1. Identify `P` (initial amount), the rate per period, and whether it's growth `(1 + i)` or decay `(1 − i)`.
2. Adjust the rate and exponent for the compounding frequency if needed.
3. Substitute; leave in exact form unless told to round.

Worked micro-example A machine worth £8000 depreciates by 25% per year. Write an expression for its value after `n` years, and find its value after 2 years. Ratio `= 1 − 0.25 = 0.75 = 3/4`. Value after `n` years `= 8000 · (3/4)^n`. After 2 years: `8000 · (3/4)^2 = 8000 · 9/16 = 4500`. Value is £4500.

Traps

- A 25% *decrease* uses ratio `0.75`, not `−0.25`. The ratio is `1 − i`.
- "After `n` years" vs "during the `n`th year" — depreciation models usually count from `n = 0` at purchase, so after `n` years the exponent is `n`.
- Match the rate to the period: 6% per year compounded monthly means `i = 0.06/12` per month and the exponent counts months.

1.6 Laws of exponents (with rational exponents)

Key facts / formulas — *the index laws are assumed knowledge, not printed in the booklet.*

- `a^m · a^n = a^(m+n)`
- `a^m / a^n = a^(m−n)`
- `(a^m)^n = a^(mn)`
- `a^0 = 1` (for `a ≠ 0`)
- `a^(−n) = 1/a^n`
- `a^(1/n) = ⁿ√a` and `a^(m/n) = (ⁿ√a)^m = ⁿ√(a^m)`

Method (simplify / solve on a non-calculator paper)

1. Rewrite roots and reciprocals as powers — everything becomes `base^exponent`.
2. Express every base as a power of the **same prime** where possible (e.g. `4`, `8`, `1/2` all become powers of `2`).
3. Apply the laws to combine; for an equation, set the exponents equal once the bases match.

Worked micro-example Solve `8^(x+1) = 4^(2x)` for `x`. Write both sides base 2: `8 = 2^3`, `4 = 2^2`. So `2^(3(x+1)) = 2^(2(2x))`, i.e. `2^(3x+3) = 2^(4x)`. Equate exponents: `3x + 3 = 4x`, so `x = 3`.

Traps

- $a^{(m+n)} \neq a^m + a^n$. Powers add only when the bases are *multiplied*.
- $(a^m)^n = a^{(mn)}$ — multiply, don't add. Easy to confuse with the product law.
- $a^{(-1/2)} = 1/\sqrt{a}$, not $-\sqrt{a}$. A negative exponent flips; a fractional exponent roots. They are independent operations.
- $(2x)^3 = 8x^3$ — the coefficient gets the power too.
- Don't lose a solution: if an equation reduces to $y^2 = 9$ after a substitution, $y = \pm 3$.

1.7 Logarithms

Key facts / formulas

- Definition: $\log_a x = y \iff a^y = x$, for $a > 0$, $a \neq 1$, $x > 0$. *The defining relationship — memorise it; it's how you switch between log and exponential form.*
- Laws (all *GIVEN in the booklet*):
- Product: $\log_a(xy) = \log_a x + \log_a y$
- Quotient: $\log_a(x/y) = \log_a x - \log_a y$
- Power: $\log_a(x^m) = m \cdot \log_a x$
- Change of base: $\log_a x = (\log_b x)/(\log_b a)$ — *GIVEN*. The habit of *using* it (e.g. to get a common base, or to switch to $\ln$) must be internalised.
- Natural log: $\ln x = \log_e x$. Useful identities: $\ln(e^x) = x$, $e^{(\ln x)} = x$, $\ln 1 = 0$, $\ln e = 1$.

Method (condense or expand)

1. Use the power law to move exponents in or out as coefficients.
2. Use product/quotient laws to combine into a single log, or split into separate logs.
3. Evaluate any logs that come out to integers ($\log_2 8 = 3$).

Worked micro-example Write $2 \log 3 + \log 5 - \log 9$ as a single logarithm. $2 \log 3 = \log 3^2 = \log 9$. So the expression is $\log 9 + \log 5 - \log 9 = \log 5$.

Traps

- $\log(a + b) \neq \log a + \log b$. The product law is about a **product inside**, not a sum.
- $\log(a)/\log(b) \neq \log(a/b)$. The first is a ratio of two logs (change-of-base shape); the second is the quotient law.
- $(\log x)^2 \neq \log(x^2)$. The power law only moves an exponent that is *inside* the log.
- The argument of a log must be **positive** — check, and reject any solution that makes an argument ≤ 0.
- $\log$ with no base on an IB paper usually means base 10, but read the question; $\ln$ is base e.

1.8 Solving exponential and logarithmic equations

Key facts

- Two non-calculator strategies dominate:
- **Same-base method:** rewrite both sides as a power of one base, then equate exponents (see §1.6).
- **Take logs of both sides:** for `a^x = b` where bases won't match, write `x = log_a b`, or take `ln` of both sides to get `x = (ln b)/(ln a)` — an exact answer.
- For log equations: condense to a single log on each side, then use `log_a M = log_a N ⇒ M = N`, or convert `log_a M = c` to exponential form `M = a^c`.

Method (exponential equation)

1. Try the same-base method first — it gives the cleanest exact answer.
2. If bases can't be matched, take `ln` (or `log`) of both sides.
3. Use the power law to bring the unknown exponent down as a coefficient, then isolate it.
4. Leave the answer as an exact log expression.

Method (logarithmic equation)

1. Use log laws to get a single log on each side (or a single log equal to a constant).
2. Apply "equal logs ⇒ equal arguments", or convert to exponential form.
3. Solve the resulting algebraic equation.
4. **Check every solution** against the domain (arguments must stay positive).

Worked micro-example (exponential) Solve `5^(2x) = 5^(x+3)` — wait, that's same-base, easy. Instead: solve `3^x = 7`, exact form. Take `ln` of both sides: `ln(3^x) = ln 7`, so `x ln 3 = ln 7`, giving `x = (ln 7)/(ln 3)`. That is the exact answer — no decimal needed.

Worked micro-example (logarithmic) Solve `log_2(x) + log_2(x − 2) = 3`. Condense: `log_2(x(x − 2)) = 3`. Convert to exponential form: `x(x − 2) = 2^3 = 8`. So `x^2 − 2x − 8 = 0`, `(x − 4)(x + 2) = 0`, giving `x = 4` or `x = −2`. Check: `x = −2` makes `log_2(x)` undefined, so **reject** it. Answer: `x = 4`.

Traps

- Forgetting to bring the exponent down: `ln(3^x)` becomes `x ln 3`, not `ln 3` or `3 ln x`.
- Losing an exact answer by reaching for a decimal — Paper 1 wants `(ln 7)/(ln 3)`.
- Not checking solutions in a log equation — extraneous negative roots are deliberately built in.
- `log_2(x) + log_2(x − 2)` condenses to `log_2(x(x−2))`, **not** `log_2(x) · log_2(x − 2)` or `log_2(2x − 2)`.

1.9 The binomial theorem

Key facts / formulas

- $(a + b)^n = \sum_{r=0}^{n} \binom{n}{r} a^{n-r} b^r$ — *the expansion is GIVEN in the booklet.*
- $nCr = n! / (r!(n - r)!)$ — *GIVEN.* Also written $C(n, r)$ or ${}_nC_r$.
- The **general term** (the term containing b^r) is $\binom{n}{r} a^{n-r} b^r$. Internalise this — it's how you target one term without expanding everything.
- **Pascal's triangle** gives the nCr coefficients row by row; row n starts $1, n, \ldots$. Fine for small n on a non-calculator paper.
- $0! = 1$, $nC0 = nCn = 1$, $nC1 = n$, and $nCr = nC(n-r)$ (symmetry).

Method (find a specific term or coefficient)

1. Write the general term $\binom{n}{r} a^{n-r} b^r$ for the given binomial.
2. Work out the total power of the variable in that term as a function of r.
3. Set that power equal to the one you want; solve for r.
4. Substitute r back to compute the term, or just its coefficient if that's what's asked.

Worked micro-example Find the coefficient of x^3 in $(2 + x)^5$. General term: $\binom{5}{r} \cdot 2^{5-r} \cdot x^r$. The power of x is r, so set $r = 3$. Coefficient $= \binom{5}{3} \cdot 2^{5-3} = 10 \cdot 2^2 = 10 \cdot 4 = 40$. (The full term is $40x^3$; the **coefficient** is 40.)

Worked micro-example with a trickier variable Find the term in x^2 in $(x - 3/x)^4$. General term: $\binom{4}{r} \cdot x^{4-r} \cdot (-3/x)^r = \binom{4}{r}(-3)^r \cdot x^{4-r} \cdot x^{-r} = \binom{4}{r}(-3)^r \cdot x^{4-2r}$. Set $4 - 2r = 2 \Rightarrow r = 1$. Term $= \binom{4}{1}(-3)^1 \cdot x^2 = 4 \cdot (-3) \cdot x^2 = -12x^2$.

Traps

- **Coefficient vs term.** "Coefficient of x^3" is a number; "the term in x^3" includes the x^3. Answer exactly what's asked.
- **r indexing starts at 0.** The term with b^r is the $(r + 1)$th term. The "4th term" means $r = 3$.
- The sign: in $(a - b)^n$ the b is $-b$, so odd-power terms are negative. Keep the minus inside the bracket.
- When the binomial is $(2 + x)^n$, the constant 2 also gets raised to a power — 2^{n-r} is part of the coefficient, don't drop it.
- For terms like $(x - 3/x)^n$, expand the powers of *both* x factors and combine — the variable power is $n - 2r$ here, not $n - r$.

1.10 Simple proof (direct / deductive)

Key facts

- SL proof is **direct (deductive)**: start from what is given or known, apply valid algebraic/logical steps, and arrive at the required statement. Each line must follow from the previous one.
- **Scope note — read this carefully:** proof by **mathematical induction**, proof by **contradiction**, and the systematic use of **counterexamples** are **HL-only**. SL needs only direct proof. Don't waste revision time on induction for SL Paper 1.

- "Show that" questions are direct proofs in miniature: even if you can't complete the derivation, **state the given result and use it** in later parts to keep follow-through marks.

Method (a direct algebraic proof)

1. Write down what you're given, and what you must show.
2. Represent the objects generally — e.g. an even number as `2k` for integer `k`, an odd number as `2k + 1`, consecutive integers as `n` and `n + 1`.
3. Manipulate algebraically, one justified step per line.
4. Reach the target form and state clearly that it satisfies the definition required. End with a brief concluding line.

Worked micro-example Prove that the sum of any two consecutive integers is odd. Let the integers be `n` and `n + 1`, where `n` is an integer. Their sum is `n + (n + 1) = 2n + 1`. Since `n` is an integer, `2n + 1` is one more than an even number, so it is odd. Therefore the sum of any two consecutive integers is odd. ∎

Traps

- **Don't argue from an example.** "It works for 3 and 4" is not a proof — you must use a general variable like `n`.
- Don't assume what you're proving. Start from the givens and work *towards* the conclusion, not from it.
- Define your variables ("let `n` be an integer") — unstated assumptions lose marks.
- Finish the argument: reaching `2n + 1` is not enough; you must say *why* that makes the statement true.

If you remember nothing else

- **Arithmetic uses `+d`, geometric uses `×r`** — and the nth-term exponent/multiplier is `n − 1`, never `n`. Sum and term formulas are in the booklet; pull them straight off.
- **Always write `|r| < 1` before using `S_∞ = u_1/(1 − r)`**. It's a free reasoning mark you cannot skip.
- **To solve `a^x = b`: same base if you can, otherwise take `ln` of both sides** — and leave the answer in exact log form, never a decimal.
- **Log laws act on products, quotients, and powers — not sums.** `log(a + b)` does not break apart. Always check that log arguments stay positive and reject solutions that don't.
- **Binomial general term: `(nCr) a^(n−r) b^r`, with `r` starting at 0.** Set the variable's power equal to the one you want, solve for `r` — and answer "coefficient" or "term" exactly as asked.
- **SL proof is direct proof only.** Use a general variable, justify each line, and state the conclusion. Induction and contradiction are HL — not your problem on this paper.

Functions

Functions is the largest single topic on AA SL Paper 1 — expect roughly a quarter of the 80 marks to draw on it directly, and more indirectly, because quadratics, transformations and graph-reading thread through calculus, trig and modelling questions too. On the non-calculator paper this topic is almost entirely about exact algebra and clean hand-sketches: completing the square in your head, reading a vertex straight off a form, finding an inverse by swapping and solving, and shifting a curve *and its asymptote* together. Marks here are won by being systematic and lost by small direction errors — a transformation reversed, a root left unjustified, an endpoint inequality the wrong way round. Drill the methods until they are automatic, because P1 gives you no calculator to catch a slip.

---

1. Function concept, domain and range

Key facts. A *function* is a rule that assigns to each input exactly one output. A *relation* is any pairing of inputs and outputs; it is a function only if no input maps to two different outputs — the **vertical line test** on a graph. In a mapping diagram, a function has exactly one arrow leaving each element of the domain. The **domain** is the set of permitted inputs; the **range** is the set of outputs actually produced. Notation: $f(x)$ names the output; $f: x \mapsto 2x - 1$ is the same thing in mapping form. None of this is in the formula booklet — it is definitional and must be internalised.

Method (find domain and range of a given function).

1. Domain: exclude inputs that break the rule — values making a denominator zero, or making the inside of a square root negative, or the argument of a log zero or negative.
2. Range: think about what the function *can* output. For a parabola, find the vertex — the range starts (or stops) there. For $1/x$-type functions, the value the function can never reach is the horizontal asymptote.
3. Write both as inequalities or set notation, e.g. $x \in \mathbb{R}$, $x \neq 3$ or $\{y : y \geq -4\}$.

Worked micro-example. For $f(x) = \sqrt{(x - 5)}$: the inside must be ≥ 0, so $x - 5 \geq 0$, giving domain $x \geq 5$. The square root outputs only non-negative values, so range $f(x) \geq 0$.

Traps.

- Stating a domain endpoint with the wrong inequality — $\sqrt{(x - 5)}$ needs $x \geq 5$ (5 is *allowed*, it gives 0), but $1/(x - 5)$ needs $x \neq 5$ and a strict exclusion.
- Forgetting that range depends on the domain you were *given*; a restricted domain shrinks the range.
- Confusing "relation" and "function" — a sideways parabola $x = y^2$ is a relation, not a function.

---

2. Composite functions

Key facts. The composite $(f \circ g)(x) = f(g(x))$ means "do g first, then feed the result into f." Order matters: in general $f(g(x)) \neq g(f(x))$. The notation is not in the booklet; the concept must be internalised. The domain of $f \circ g$ is the set of x in the domain of g whose outputs $g(x)$ also lie in the domain of f.

Method (form and evaluate a composite).

1. Identify the *inner* function — the one written closest to x.
2. Substitute the entire inner expression into every x of the outer function.
3. Expand and simplify carefully.
4. For the domain: start with the domain of the inner function, then remove any x whose g(x) is outside the domain of the outer function.

Worked micro-example. Let $f(x) = x^2 + 1$ and $g(x) = 2x - 3$. Then $f(g(x)) = (2x - 3)^2 + 1 = 4x^2 - 12x + 9 + 1 = 4x^2 - 12x + 10$. Going the other way, $g(f(x)) = 2(x^2 + 1) - 3 = 2x^2 - 1$ — clearly different. To evaluate $f(g(4))$: $g(4) = 5$, then $f(5) = 26$.

Worked micro-example (domain). Let $f(x) = \sqrt{x}$ and $g(x) = x - 9$. Then $f(g(x)) = \sqrt{(x - 9)}$. The inner g has domain all reals, but f needs a non-negative input, so we require $x - 9 \geq 0$, i.e. domain $x \geq 9$.

Traps.

- Reversing the order — computing $g(f(x))$ when $f(g(x))$ was asked. The inner function is the one *touching* the x.
- Substituting only into the first x of the outer function instead of every occurrence.
- Sign and bracket errors when squaring or expanding the substituted expression — $(2x - 3)^2$ is not $4x^2 - 9$.
- Ignoring the domain restriction the inner output places on the composite.

---

3. Inverse functions

Key facts. The inverse f^{-1} undoes f: $f^{-1}(f(x)) = x$ and $f(f^{-1}(x)) = x$. The graph of f^{-1} is the reflection of the graph of f in the line $y = x$. Consequently the domain of f^{-1} is the range of f, and the range of f^{-1} is the domain of f. An inverse function exists only if f is **one-to-one** (each output comes from exactly one input — a horizontal line test). If f is not one-to-one, you must **restrict the domain** to a piece on which it is. The "swap x and y, then solve" procedure is not in the booklet — internalise it. Note $f^{-1}(x)$ does not mean $1/f(x)$.

Method (find an inverse).

1. Write $y = f(x)$.
2. Swap x and y.
3. Solve the new equation for y — that expression is $f^{-1}(x)$.

4. If a square root appears, decide the correct sign using the original function's range/domain, and *justify the choice in words*.
5. State the domain of f^{-1} (it equals the range of f).

Worked micro-example. Let $f(x) = (3x + 2)/(x - 1)$, $x \neq 1$. Write $y = (3x + 2)/(x - 1)$. Swap: $x = (3y + 2)/(y - 1)$. Multiply: $x(y - 1) = 3y + 2$, so $xy - x = 3y + 2$. Collect y terms: $xy - 3y = x + 2$, so $y(x - 3) = x + 2$, giving $f^{-1}(x) = (x + 2)/(x - 3)$, $x \neq 3$.

Worked micro-example (sign choice). Let $f(x) = x^2 - 4$ with restricted domain $x \geq 0$. Swap: $x = y^2 - 4$, so $y^2 = x + 4$, $y = \pm\sqrt{(x + 4)}$. Because the original domain was $x \geq 0$, the *outputs* of f^{-1} must be ≥ 0, so we take the positive root: $f^{-1}(x) = \sqrt{(x + 4)}$, with domain $x \geq -4$.

Traps.

- Leaving the ± in front of a root with no justification — you must state which sign and why, tying it to the original domain or range.
- Writing $f^{-1}(x) = 1/f(x)$ — that is the reciprocal, a different object entirely.
- Forgetting to restrict the domain of a non-one-to-one function before claiming an inverse exists.
- Not stating the domain of f^{-1}, or stating it as the domain (rather than the range) of f.
- Algebra slips when the variable appears twice — always collect *all* y terms on one side and factor.

---

4. Graphs: key features

Key facts. Every sketch on P1 should show, where relevant: **x- and y-intercepts** (set $y = 0$, then $x = 0$), **vertices or turning points**, **asymptotes** (vertical: where the function is undefined and blows up; horizontal: the value the function approaches as $x \to \pm\infty$), **symmetry** (a parabola is symmetric about its axis; some functions are even — symmetric in the y-axis — or odd — rotationally symmetric about the origin), and **end behaviour** (what happens as $x \to \pm\infty$). These reading skills are not booklet items.

Method (sketch any unfamiliar function by hand).

1. Find the y-intercept: evaluate $f(0)$.
2. Find x-intercepts: solve $f(x) = 0$.
3. Find vertical asymptotes: where the denominator is zero (and numerator is not).
4. Find the horizontal asymptote / end behaviour: consider very large $|x|$.
5. Find turning points if the topic requires them.
6. Plot the features, then join with a smooth curve consistent with the asymptotes and end behaviour. Label every feature with coordinates or an equation.

Worked micro-example. Sketch $y = (x + 1)/(x - 2)$. y-intercept: $f(0) = 1/(-2) = -0.5$. x-intercept: numerator zero at $x = -1$. Vertical asymptote: $x = 2$. Horizontal asymptote: for large $|x|$ the ratio $\to 1$, so $y = 1$. Two branches: one through $(-1, 0)$ and $(0, -0.5)$, one in the opposite quadrant of the asymptote cross.

Traps.

- Sketching a curve that crosses its own horizontal asymptote without reason, or that doesn't actually approach the asymptote.
- Not labelling features — an unlabelled correct shape often scores nothing.
- Reading an asymptote as a number rather than writing it as an equation ($x = 2$, not "2").

---

5. Linear functions

Key facts. Forms of a line: **gradient–intercept** $y = mx + c$; **point–gradient** $y - y_1 = m(x - x_1)$; **general** $ax + by + d = 0$. Gradient between two points: $m = (y_2 - y_1)/(x_2 - x_1)$. **Parallel** lines have equal gradients, $m_1 = m_2$. **Perpendicular** lines have $m_1 \times m_2 = -1$ (each gradient is the negative reciprocal of the other). The **midpoint** of (x_1, y_1) and (x_2, y_2) is $((x_1 + x_2)/2, (y_1 + y_2)/2)$. The **distance** between them is $\sqrt{((x_2 - x_1)^2 + (y_2 - y_1)^2)}$. The distance and midpoint formulas *are* in the AA SL formula booklet; the gradient and perpendicular conditions are not — internalise them.

Method (find the equation of a line perpendicular to a given line through a given point).

1. Find the gradient of the given line (rearrange to $y = mx + c$ if needed).
2. Take the negative reciprocal for the perpendicular gradient.
3. Substitute that gradient and the given point into $y - y_1 = m(x - x_1)$.
4. Rearrange to the form requested.

Worked micro-example. Find the line perpendicular to $2x + 3y = 12$ passing through $(3, -1)$. Rearrange: $y = -(2/3)x + 4$, so $m = -2/3$. Perpendicular gradient: $3/2$. Equation: $y - (-1) = (3/2)(x - 3)$, so $y + 1 = (3/2)x - 9/2$, giving $y = (3/2)x - 11/2$.

Traps.

- Using the reciprocal but forgetting the negative sign (or vice versa) for perpendicular gradients.
- Subtracting coordinates in inconsistent order in the gradient formula — keep "point 2 minus point 1" on both top and bottom.
- Sign errors when distributing $m(x - x_1)$, especially with negative points.
- Adding instead of averaging in the midpoint, or dropping the division by 2.

---

6. Quadratic functions

Key facts — the three forms.

- **Standard form** $y = ax^2 + bx + c$ — reveals the y-intercept c immediately, and the sign of a tells you whether the parabola opens up ($a > 0$) or down ($a < 0$).
- **Vertex / completed-square form** $y = a(x - h)^2 + k$ — reveals the vertex (h, k) and hence the axis of symmetry $x = h$ directly.
- **Factored / intercept form** $y = a(x - p)(x - q)$ — reveals the x-intercepts p and q directly; the axis of symmetry is their midpoint, $x = (p + q)/2$.

The **axis of symmetry** $x = -b/(2a)$ and the **quadratic formula** $x = (-b \pm \sqrt{(b^2 - 4ac)})/(2a)$ *are* in the AA SL formula booklet. The **discriminant** $\Delta = b^2 - 4ac$ is also in the booklet. What each form *reveals*, and the completing-the-square procedure, must be internalised.

The discriminant. With $\Delta = b^2 - 4ac$:

- $\Delta > 0$ — two distinct real roots (graph crosses the x-axis twice).
- $\Delta = 0$ — one repeated real root (graph touches the x-axis; the x-axis is tangent at the vertex).
- $\Delta < 0$ — no real roots (graph entirely above or entirely below the x-axis).

Method (complete the square for $y = ax^2 + bx + c$, $a \neq 1$).

1. Factor a out of the first two terms: $y = a(x^2 + (b/a)x) + c$.
2. Inside the bracket, halve the x-coefficient and square it; add and subtract it: $x^2 + (b/a)x + (b/2a)^2 - (b/2a)^2$.
3. Write the perfect square: $y = a((x + b/2a)^2 - (b/2a)^2) + c$.
4. Distribute a back over the constant and simplify to $a(x - h)^2 + k$. Read off the vertex (h, k).

Worked micro-example (completing the square). $y = 2x^2 - 12x + 7$. Factor: $y = 2(x^2 - 6x) + 7$. Half of -6 is -3, squared is 9: $y = 2(x^2 - 6x + 9 - 9) + 7 = 2((x - 3)^2 - 9) + 7 = 2(x - 3)^2 - 18 + 7 = 2(x - 3)^2 - 11$. Vertex $(3, -11)$, axis $x = 3$, opens upward so range $y \geq -11$.

Worked micro-example (discriminant condition). For what k does $x^2 + kx + 9 = 0$ have two distinct real roots? $\Delta = k^2 - 4(1)(9) = k^2 - 36$. Two distinct real roots needs $\Delta > 0$ strictly: $k^2 - 36 > 0$, so $k^2 > 36$, giving $k < -6$ or $k > 6$.

Worked micro-example (solving by formula). Solve $3x^2 - 5x - 2 = 0$. $\Delta = 25 - 4(3)(-2) = 25 + 24 = 49$. $x = (5 \pm \sqrt{49})/6 = (5 \pm 7)/6$, so $x = 2$ or $x = -1/3$. (Here Δ is a perfect square, so it would also factor as $(3x + 1)(x - 2) = 0$ — always check for factoring first on P1.)

Traps.

- Using "two distinct real roots" with $\Delta \geq 0$ — distinct demands the *strict* inequality $\Delta > 0$; the equal sign belongs to the repeated-root case.
- Forgetting to factor a out before completing the square when $a \neq 1$, or forgetting to multiply the $-(b/2a)^2$ term back by a.
- Sign error in the formula: $-b$ on top means a positive b becomes negative — and 4ac with a negative c becomes +.
- Reading the vertex of $a(x - h)^2 + k$ as $(-h, k)$ — the form has x *minus* h, so $y = 2(x - 3)^2 - 11$ has $h = +3$.
- Not stating the direction of opening when asked for the range.

---

7. Reciprocal and simple rational functions

Key facts. The basic reciprocal $y = 1/x$ has a vertical asymptote $x = 0$ and a horizontal asymptote $y = 0$, sitting in the first and third quadrants, and is an odd function. A simple rational $y = (ax + b)/(cx + d)$ has a **vertical asymptote** where the denominator is zero, $x = -d/c$, and a **horizontal asymptote** found from the behaviour as $x \to \pm\infty$, namely $y = a/c$. Near the vertical asymptote the function shoots off to $+\infty$ on one side and $-\infty$ on the other; far from it the curve flattens towards the horizontal asymptote. These facts are not in the booklet.

Method (sketch $y = (ax + b)/(cx + d)$).

1. Vertical asymptote: set $cx + d = 0$.
2. Horizontal asymptote: take the ratio of leading coefficients, $y = a/c$.
3. Intercepts: y-intercept from $x = 0$; x-intercept from the numerator zero.
4. Decide which side of the vertical asymptote each branch sits using a test point.
5. Draw two smooth branches hugging both asymptotes; label the asymptote equations.

Worked micro-example. $y = (2x - 4)/(x + 1)$. Vertical asymptote: $x + 1 = 0$, so $x = -1$. Horizontal asymptote: $y = 2/1 = 2$. y-intercept: $f(0) = -4$. x-intercept: $2x - 4 = 0$, so $x = 2$. Test $x = 0$ gives -4 (below the asymptote $y = 2$), so the right branch passes through $(0, -4)$ and $(2, 0)$, rising toward $y = 2$; the left branch sits above $y = 2$.

Traps.

- Getting the horizontal asymptote wrong — it is the ratio of leading coefficients, not the constant terms.
- Drawing the curve crossing the vertical asymptote (impossible — it is a break in the domain).
- Forgetting that the curve *may* cross the horizontal asymptote in the middle region, but must settle to it at the extremes.

---

8. Exponential and logarithmic functions

Key facts. The exponential $y = a^x$ (with $a > 0$, $a \neq 1$) passes through $(0, 1)$, has a **horizontal asymptote $y = 0$**, is increasing for $a > 1$ and decreasing for $0 < a < 1$, with range $y > 0$. The logarithm $y = \log_a x$ is its inverse: it passes through $(1, 0)$, has a **vertical asymptote $x = 0$**, domain $x > 0$, and range all reals. The natural pair are $y = e^x$ and $y = \ln x$. Because they are inverses, their graphs are reflections in $y = x$. A *larger base* makes $y = a^x$ rise more steeply (for $x > 0$) and makes $y = \log_a x$ rise more slowly. Crucially, a **vertical translation moves the horizontal asymptote of an exponential**, and a **horizontal translation moves the vertical asymptote of a logarithm**. The graph shapes and asymptote rules must be internalised; the log laws are in the booklet (covered in the Exponents & Logs topic).

Method (sketch a transformed exponential or log).

1. Start from the parent graph ($y = e^x$ or $y = \ln x$) and note its asymptote.
2. Apply each transformation in turn, tracking where the asymptote moves.
3. Recompute the intercept after transforming — the old intercept will have moved.
4. Label the new asymptote *as an equation* and the new intercept.

Worked micro-example. Sketch $y = e^x - 3$. Parent $y = e^x$ has asymptote $y = 0$ and passes through $(0, 1)$. The "− 3" shifts everything down 3: the **asymptote moves to $y = -3$**, and the y-intercept becomes 1

− 3 = −2. The x-intercept now exists: $e^x - 3 = 0$ gives $e^x = 3$, x = ln 3. Curve rises from just above y = −3 on the left, through (0, −2) and (ln 3, 0).

Worked micro-example (log shift). Sketch y = ln(x + 2). Parent y = ln x has vertical asymptote x = 0. Replacing x with x + 2 shifts the graph *left* 2, so the **asymptote moves to x = −2**. The x-intercept: ln(x + 2) = 0 gives x + 2 = 1, x = −1. Domain becomes x > −2.

Traps.

- Leaving the asymptote at y = 0 or x = 0 after a translation — the single most common exponential/log mark-loss.
- Forgetting that a horizontal shift of a log also changes its domain.
- Mixing up which way $\log_a$ steepness goes as the base changes.
- Drawing an exponential that touches or crosses its asymptote.

---

9. Transformations of graphs

Key facts. Given a parent y = f(x):

- **Vertical translation** y = f(x) + k — moves the graph *up* by k (down if k < 0). Acts "as you expect."
- **Horizontal translation** y = f(x − h) — moves the graph *right* by h (left if h < 0). Note it is *inside the bracket and inverted*: x − h shifts right, x + h shifts left.
- **Vertical stretch** y = a·f(x) — stretches by factor a away from the x-axis; if a < 0 it also reflects in the x-axis. Outside the bracket, "as you expect."
- **Horizontal stretch** y = f(bx) — stretches by factor 1/b away from the y-axis; *inside the bracket and inverted*, so f(2x) is a *compression* by factor 1/2, and f(x/3) is a *stretch* by factor 3. If b < 0 it also reflects in the y-axis.
- **Reflection in the x-axis** y = −f(x); **reflection in the y-axis** y = f(−x).

The governing rule to internalise: **anything happening to x (inside the bracket) is horizontal and behaves the opposite way to how it reads; anything happening outside the bracket is vertical and behaves as it reads.** None of this is in the formula booklet.

Order of composed transformations. When several transformations are combined, apply **horizontal changes** (inside the bracket) in the reverse of the usual operation order, and **vertical changes** (outside) in the natural order. A safe practical approach for y = a·f(b(x − h)) + k: deal with the horizontal pair first — factor the bracket so it reads b(x − h), do the stretch by 1/b then the translation by h — then the vertical pair — stretch by a then translate by k.

Method (describe the transformations taking f(x) to a given expression).

1. Identify what is inside the bracket: a coefficient on x is a horizontal stretch, a constant added to x is a horizontal translation (remember to *invert* both).
2. Identify what is outside the bracket: a multiplier is a vertical stretch (and reflection if negative), an added constant is a vertical translation.
3. State each transformation precisely — direction, axis, and factor or distance.

4. If asked, apply them to specific points or to an asymptote and recheck.

Worked micro-example. Describe how $y = f(x)$ maps to $y = 3f(x - 2) - 1$. Inside the bracket: $x - 2$ is a horizontal translation 2 units *right*. Outside: the ×3 is a vertical stretch of factor 3; the −1 is a vertical translation 1 unit *down*. A point (a, b) on f maps to $(a + 2, 3b - 1)$.

Worked micro-example (horizontal stretch). Describe $y = f(2x + 6)$ relative to $y = f(x)$. Factor inside the bracket first: $f(2(x + 3))$. So a horizontal stretch of factor 1/2 (a compression) and a horizontal translation 3 units *left*. The factoring step is essential — reading it as "shift 6 then stretch" gives the wrong result.

Traps.

- Inverting the direction wrongly — $f(x - 2)$ goes *right*, not left; $f(2x)$ *compresses*, it does not stretch.
- Not factoring out the x-coefficient before reading a combined horizontal transformation — $f(2x + 6)$ is not "stretch then shift 6."
- Applying vertical and horizontal transformations in a jumbled order when both stretch and translate are present.
- Forgetting that a negative multiplier (inside or outside) carries a reflection as well as a stretch.
- Transforming the curve but not its asymptote — if the graph shifts, the asymptote shifts with it.

If you remember nothing else

- **The inner function touches the x.** For $f(g(x))$, substitute g into *every* x of f — and never compute $g(f(x))$ by mistake.
- **Inverse = swap then solve**, and if a root appears, *state the sign and justify it* from the original domain or range. $f^{-1}(x) \neq 1/f(x)$.
- **Each quadratic form reveals one thing for free**: standard → y-intercept, vertex form → vertex and axis, factored form → x-intercepts. $\Delta > 0$ strict for *distinct* real roots; $\Delta = 0$ for a repeated root.
- **Inside the bracket is horizontal and inverted**; outside is vertical and as-it-reads. Factor out any x-coefficient before describing a combined horizontal transformation.
- **When a graph moves, its asymptote moves with it** — write the new asymptote as an equation ($x = \ldots$ or $y = \ldots$), never as a bare number.
- **Label every feature** on every hand-sketch — intercepts, vertices, asymptotes — and watch domain endpoints: ≥ when the value is allowed, strict ≠ or > when it is not.

3. Geometry & Trigonometry

Geometry and trigonometry is worth roughly 16% of the AA SL course — a small share of Paper 1, but pound-for-pound the highest trap density on the whole paper. The arithmetic is rarely the problem; the marks leak through degree/radian confusion, forgotten second solutions, identity proofs done the wrong way, and trig equations solved on too narrow a domain. Because Paper 1 is non-calculator, you cannot lean on a GDC to bail you out: you must know the exact-value table cold, work fluently in radians, and manipulate identities by hand. Treat this section as a checklist of disciplined habits, not a list of facts to recognise.

3.1 Three-dimensional geometry

Key facts / formulas. For points $A(x_1, y_1, z_1)$ and $B(x_2, y_2, z_2)$:

- Distance: $AB = \sqrt{[(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2]}$ — the 2D formula with a third term bolted on. In the formula booklet.
- Midpoint: $((x_1+x_2)/2, (y_1+y_2)/2, (z_1+z_2)/2)$. In the booklet.
- Volume and surface area of standard solids (cuboid, cylinder, cone, sphere, pyramid) are in the booklet — do not memorise them, but do know *which* solid you are looking at.

Angles between lines, or between a line and a plane, are **not** found with a special 3D formula at SL. You always drop the problem onto a **right triangle** that you construct inside the solid, then use SOH CAH TOA.

Method (angle in a solid).

1. Sketch the solid; mark the line and the plane (or second line).
2. Find the foot of the perpendicular from the relevant point to the plane.
3. Identify the right triangle whose hypotenuse is the slanted line and whose base lies in the plane.
4. Compute the two known sides (often via Pythagoras on a face first).
5. Apply tan/sin/cos to get the angle.

Worked micro-example. A cuboid has a square base ABCD of side 6 and height 8; the top face is EFGH with E above A. Find the angle the space diagonal AG makes with the base. The base diagonal $AC = \sqrt{(6^2 + 6^2)} = \sqrt{72} = 6\sqrt{2}$. Triangle ACG is right-angled at C, with CG = 8 vertical. So $\tan \theta$ = opposite/adjacent = $8 / (6\sqrt{2}) = 4/(3\sqrt{2}) = 2\sqrt{2}/3$. The angle is $\arctan(2\sqrt{2}/3)$. On a non-calculator paper you leave it in this exact form unless it is a standard angle.

Traps.

- Using the *edge* (length 6) instead of the *base diagonal* ($6\sqrt{2}$) as the adjacent side — always check that your right triangle actually lies flat in the plane.
- Forgetting to rationalise or simplify the surd before quoting the final ratio.
- Mixing up "angle between line and plane" (line vs its projection) with "angle between two lines."

3.2 Right-angled trigonometry

Key facts. SOH CAH TOA: sin = opp/hyp, cos = adj/hyp, tan = opp/adj. Pythagoras: $a^2 + b^2 = c^2$. None of these are "in the booklet" in a useful sense — they must be automatic.

The SL skill being tested is rarely a bare right triangle. It is **spotting a right triangle hidden inside a larger figure** — an altitude dropped in a non-right triangle, a radius meeting a tangent, a diagonal of a rectangle.

Method.

1. Label the right angle and identify hypotenuse, opposite, adjacent relative to the angle you want.
2. Pick the ratio that uses the two quantities you know/want.
3. Solve; rationalise surds.

Worked micro-example. An isosceles triangle has two equal sides of length 10 and an apex angle of 60°. Find its height. Drop the altitude from the apex; it bisects the apex angle into 30° and bisects the base. In the right triangle, the hypotenuse is 10, the angle at the apex is 30°, and the height is the adjacent side: height = $10 \cos 30° = 10 \cdot (\sqrt{3}/2) = 5\sqrt{3}$.

Traps.

- Using the *full* apex angle instead of the *half* angle after dropping the altitude.
- Assuming a triangle is right-angled because it "looks" it — only use SOH CAH TOA when a right angle is given or genuinely constructed.

3.3 The sine rule, cosine rule, and area of a triangle

Key facts / formulas (all in the booklet).

- Sine rule: $a/\sin A = b/\sin B = c/\sin C$ (or the reciprocal form with sines on top).
- Cosine rule: $c^2 = a^2 + b^2 - 2ab \cos C$, rearranged as $\cos C = (a^2 + b^2 - c^2)/(2ab)$.
- Area = $\frac{1}{2} ab \sin C$.

Although the formulas are given, *choosing* the right one is on you.

Method (which rule?).

1. Have a side and its opposite angle as a matched pair? → **sine rule**.
2. Have all three sides (find an angle), or two sides and the included angle (find the third side)? → **cosine rule**.
3. Need area and you have two sides plus the included angle? → **$\frac{1}{2}$ ab sin C**.

Worked micro-example. In triangle ABC, $b = 7$, $c = 5$, and the included angle $A = 60°$. Find a. By the cosine rule, $a^2 = 7^2 + 5^2 - 2\cdot7\cdot5\cdot\cos 60° = 49 + 25 - 70\cdot(\frac{1}{2}) = 74 - 35 = 39$, so $a = \sqrt{39}$. Then the area = $\frac{1}{2} \cdot 7 \cdot 5 \cdot \sin 60° = (35/2)\cdot(\sqrt{3}/2) = 35\sqrt{3}/4$.

Traps.

- Reaching for the sine rule when you do not have a complete side–angle pair — you will have two unknowns and stall.
- Sign error in the cosine rule: the term is *minus* 2ab cos C; if cos C is negative (obtuse angle) the two minuses make the side longer, which is correct — do not "fix" it.
- Using a non-included angle with the area formula. The angle must sit *between* the two sides.

3.4 The ambiguous case of the sine rule

Key facts. When you use the sine rule to find an **angle**, the equation $\sin\theta = k$ has *two* solutions in $0° < \theta < 180°$: the acute value θ from your calculator-style reasoning, and its supplement $\mathbf{180° - \theta}$. Both can give a valid triangle when the data is "two sides and a non-included angle" (SSA).

Method.

1. Apply the sine rule to get $\sin\theta$ = some value.
2. Write down both candidates: θ and $180° - \theta$.
3. Test each: add it to the angle you already know. If the sum is less than 180°, that triangle exists; if not, discard it.
4. If both survive, the problem has two answers — present both.

Worked micro-example. In triangle ABC, $a = 8$, $b = 6$, and angle $B = 40°$. Find angle A. Sine rule: $\sin A / 8 = \sin 40° / 6$, so $\sin A = 8 \sin 40° / 6 = (4/3) \sin 40° \approx 0.857$. Then $A \approx 59.0°$ or $A \approx 121.0°$. Check: $59.0° + 40° = 99° < 180°$ ✓; $121.0° + 40° = 161° < 180°$ ✓. **Both** triangles exist — the answer is $A \approx 59.0°$ or $121.0°$.

Traps.

- The single most common SL trig error: quoting only the acute angle and missing the obtuse one. Whenever the sine rule gives you an *angle* (not a side), reflexively write "or 180° – ...".
- Conversely, over-applying it: if $180° - \theta$ plus the known angle exceeds 180°, that solution is genuinely impossible — discard it, do not force two answers.
- Finding a *side* with the sine rule is never ambiguous; the second-solution check only applies to angles.

3.5 Bearings, angles of elevation and depression

Key facts. A bearing is measured **clockwise from north**, always written as three digits (e.g. 075°, 240°). An angle of elevation is measured upward from the horizontal; an angle of depression is measured downward from the horizontal. By alternate angles, the angle of depression from A to B equals the angle of elevation from B to A.

Method.

1. Draw a clear diagram with a north arrow at every point that has a bearing.
2. Convert the bearing information into the interior angles of a triangle.
3. Apply sine/cosine rule or right-angled trig as appropriate.

Worked micro-example. A ship sails 12 km on a bearing of 050°, then 9 km on a bearing of 140°. How far is it from the start? The two legs turn through 140° – 50° = 90°, so the path is right-angled. Distance = $\sqrt{(12^2 + 9^2)} = \sqrt{(144 + 81)} = \sqrt{225} = 15$ km.

Traps.

- Measuring the bearing anticlockwise, or from the wrong reference line.
- Forgetting that the change of direction between two legs is the *difference* of bearings, and that the interior angle of the triangle may be that difference or its supplement — the diagram tells you which.
- Confusing elevation and depression, or not using the alternate-angle equality.

3.6 Radian measure

Key facts. π radians = 180°. To convert: degrees → radians multiply by π/180; radians → degrees multiply by 180/π. A radian is the angle subtended at the centre of a circle by an arc equal in length to the radius — it is a *pure ratio* (arc/radius), which is exactly why it has no units and why the calculus and circle formulas come out clean.

Radians are the **default** for the rest of AA: the derivatives of sin and cos, the arc-length and sector formulas, and trig equations are all stated assuming radians. Get used to thinking π/6 rather than 30°.

Method. Memorise the anchor conversions so you never reach for π/180 under time pressure: π/6 = 30°, π/4 = 45°, π/3 = 60°, π/2 = 90°, π = 180°, 2π = 360°.

Worked micro-example. Convert 135° to radians: 135 · π/180 = 135π/180 = 3π/4. Convert 5π/6 to degrees: 5π/6 · 180/π = 5·180/6 = 150°.

Traps.

- Leaving the calculator (or your mental "mode") in degrees while a question is in radians, or vice versa — on a non-calculator paper this surfaces as plugging a degree value into a radian formula.
- Writing π/180 the wrong way up; sanity-check that converting 180° gives π, not $180^2/\pi$.

3.7 Arc length and sector area

Key facts / formulas (both in the booklet). For a sector of radius r and central angle θ **in radians**:

- Arc length: $s = r\theta$
- Sector area: $A = \frac{1}{2} r^2 \theta$

These formulas are *radian-only*. There is no separate degree version to memorise — if you are handed degrees, convert first.

Method.

1. Confirm θ is in radians; if not, convert.
2. Substitute into $s = r\theta$ or $A = \frac{1}{2} r^2 \theta$.
3. For a "segment" (the region between a chord and the arc), compute sector area then subtract the triangle area $\frac{1}{2} r^2 \sin\theta$.

Worked micro-example. A sector has radius 6 and central angle π/3. Arc length = $6 \cdot \pi/3 = 2\pi$. Sector area = $\frac{1}{2} \cdot 6^2 \cdot \pi/3 = \frac{1}{2} \cdot 36 \cdot \pi/3 = 6\pi$. The segment cut off by the chord = sector − triangle = $6\pi - \frac{1}{2} \cdot 6^2 \cdot \sin(\pi/3) = 6\pi - 18 \cdot (\sqrt{3}/2) = 6\pi - 9\sqrt{3}$.

Traps.

- The headline trap of the whole topic: using $s = r\theta$ or $A = \frac{1}{2}r^2\theta$ with θ in **degrees**. The formulas silently give nonsense; nothing flags the error. Always write "(radians)" next to θ before substituting.
- Using $\frac{1}{2}r^2\theta$ when you wanted the triangle's $\frac{1}{2}r^2\sin\theta$, or vice versa — θ vs sin θ.
- For a segment, subtracting in the wrong order or forgetting the triangle is also expressible with $\frac{1}{2}ab \sin C$.

3.8 The unit circle and exact values

Key facts. On the unit circle (radius 1, centred at the origin), a point at angle θ has coordinates (cos θ, sin θ); tan θ = sin θ / cos θ is the gradient of the radius. This single picture *defines* trig for all angles, not just acute ones, and it tells you the **sign in each quadrant**:

- Quadrant 1 (0 to π/2): all positive.
- Quadrant 2 (π/2 to π): sin positive, cos and tan negative.
- Quadrant 3 (π to 3π/2): tan positive, sin and cos negative.
- Quadrant 4 (3π/2 to 2π): cos positive, sin and tan negative.

The mnemonic "All Students Take Calculus" runs anticlockwise from Q1 marking which ratio is positive.

The **exact-value table is NOT in the booklet** — you must internalise it:

θ	0	π/6	π/4	π/3	π/2
sin	0	1/2	$\sqrt{2}/2$	$\sqrt{3}/2$	1
cos	1	$\sqrt{3}/2$	$\sqrt{2}/2$	1/2	0
tan	0	$1/\sqrt{3}$	1	$\sqrt{3}$	undefined

Notice sin runs 0, 1/2, $\sqrt{2}/2$, $\sqrt{3}/2$, 1 and cos is the same list reversed — a quick memory hook. For multiples (2π/3, 5π/6, 7π/6, ...), take the related acute angle's value and attach the quadrant sign.

Method (exact value of any standard angle).

1. Locate the angle on the unit circle; note its quadrant.
2. Find the **reference angle** — the acute angle to the nearest part of the x-axis.

3. Read the magnitude from the table for that reference angle.
4. Attach the sign for the quadrant.

Worked micro-example. Find cos(5π/6). It lies in Q2 (between π/2 and π); the reference angle is $\pi - 5\pi/6 = \pi/6$; $\cos(\pi/6) = \sqrt{3}/2$; cosine is negative in Q2, so $\cos(5\pi/6) = -\sqrt{3}/2$.

Traps.

- Computing the reference angle to the y-axis instead of the x-axis.
- Getting the magnitude right but the sign wrong — always do the quadrant-sign step explicitly.
- Forgetting tan(π/2) is undefined (cos = 0), and that tan repeats every π, not every 2π.

3.9 Trigonometric identities

Key facts / formulas (in the booklet).

- Pythagorean identity: $\sin^2\theta + \cos^2\theta = 1$. Useful rearrangements: $\sin^2\theta = 1 - \cos^2\theta$, $\cos^2\theta = 1 - \sin^2\theta$.
- Double angle: $\sin 2\theta = 2 \sin\theta \cos\theta$.
- Double angle for cosine, **three equivalent forms**: $\cos 2\theta = \cos^2\theta - \sin^2\theta = 2\cos^2\theta - 1 = 1 - 2\sin^2\theta$.
- $\tan\theta = \sin\theta / \cos\theta$.

The forms are all in the booklet, but **which form of cos 2θ to use is a judgement call** you make, not something the booklet decides.

Method (choosing the cos 2θ form).

- If the expression/equation also contains **cos θ** (or you want everything in cosine): use $\mathbf{2\cos^2\theta - 1}$.
- If it contains **sin θ**: use $\mathbf{1 - 2\sin^2\theta}$.
- If you want it factorised or both ratios present: use $\mathbf{\cos^2\theta - \sin^2\theta}$.

Method (proving an identity). Work **one side only** — usually the messier side — and transform it step by step until it equals the other side. Never operate on both sides at once (e.g. squaring both sides, or cross-multiplying across the "="); that assumes what you are trying to prove.

Worked micro-example. Show that $(1 - \cos 2\theta) / \sin 2\theta = \tan\theta$. Take the left side. Use $\cos 2\theta = 1 - 2\sin^2\theta$, so $1 - \cos 2\theta = 2\sin^2\theta$. Use $\sin 2\theta = 2 \sin\theta \cos\theta$. Then LHS = $2\sin^2\theta / (2 \sin\theta \cos\theta) = \sin\theta / \cos\theta = \tan\theta$ = RHS. The choice of the $1 - 2\sin^2\theta$ form is what made the numerator collapse cleanly — picking $\cos^2\theta - \sin^2\theta$ would have left more work.

Traps.

- Manipulating both sides of a proof simultaneously — it scores nothing because the logic is circular.
- Choosing the wrong cos 2θ form and grinding into a dead end; if an algebra path is going nowhere after two lines, swap the form.
- Sign slips when rearranging the Pythagorean identity (it is $1 - \cos^2\theta$, not $\cos^2\theta - 1$, for $\sin^2\theta$).

3.10 Trigonometric equations

Key facts. To solve $\sin x = k$ (or cos, tan) over a domain, you need **every** solution in that domain, not just the principal value. Use the unit circle: sin and cos repeat every 2π, tan every π. Each "base" solution generates a family by adding multiples of the period.

Method.

1. Rearrange to a single trig ratio equals a constant. If there are mixed ratios or a squared term, use an identity (e.g. replace $\sin^2 x$ with $1 - \cos^2 x$) to reduce to one ratio, then treat it like a quadratic if needed.
2. Find the principal value, then use the unit circle to find the **second** base solution in one revolution (for $\sin x = k$ it is $\pi - x$; for $\cos x = k$ it is $-x$ or $2\pi - x$; tan has one per π).
3. Add/subtract the period to sweep the whole stated domain.
4. **Domain expansion for sin(nx) etc.**: if the equation is $\sin(2x) = k$ for $0 \le x \le 2\pi$, substitute $u = 2x$. Then u ranges over $0 \le u \le 4\pi$ — solve for *u* across that *expanded* domain, collecting all solutions, then divide each by 2 to get x. Failing to widen the domain for u is how you lose half the answers.

Worked micro-example. Solve $\sin 2x = \sqrt{3}/2$ for $0 \le x \le \pi$. Let $u = 2x$, so $0 \le u \le 2\pi$. $\sin u = \sqrt{3}/2$ gives $u = \pi/3$ and $u = \pi - \pi/3 = 2\pi/3$ as the base pair; both already lie in $[0, 2\pi]$, and adding 2π would exceed it. So $u \in \{\pi/3, 2\pi/3\}$, hence $x = u/2 \in \{\pi/6, \pi/3\}$.

A reduce-with-identity example: solve $2\cos^2 x - \cos x - 1 = 0$ for $0 \le x \le 2\pi$. Factor as $(2\cos x + 1)(\cos x - 1) = 0$, so $\cos x = -1/2$ or $\cos x = 1$. $\cos x = 1$ gives $x = 0$ and $x = 2\pi$. $\cos x = -1/2$ gives $x = 2\pi/3$ and $x = 4\pi/3$. Solutions: $x \in \{0, 2\pi/3, 4\pi/3, 2\pi\}$.

Traps.

- Not expanding the domain for the inner argument of sin(nx)/cos(nx) — the defining error of this sub-topic.
- Quoting only the principal value and missing the supplement (sin) or the reflection (cos).
- Dividing through by cos x or sin x and silently losing the solutions where that ratio is zero — factorise instead.
- Forgetting endpoint solutions when the domain is closed (e.g. $x = 0$ and $x = 2\pi$ both count above).

3.11 Graphs of sin, cos, tan

Key facts. For $y = a \sin(b(x - c)) + d$ (and the same for cos):

- **Amplitude** = $|a|$ — half the distance between max and min.
- **Period** = $2\pi / |b|$ (for tan it is $\pi / |b|$).
- **c** is the horizontal (phase) shift; **d** is the vertical shift, giving the principal axis $y = d$.
- Maximum value = $d + |a|$, minimum = $d - |a|$.
- The tan graph has vertical asymptotes where cos = 0, period π, no amplitude.

Modelling: a real-world periodic quantity (tides, temperature, a Ferris wheel height) is fitted with a sinusoid by reading d from the midline, a from half the range, b from the period, and c from a known peak/trough.

Method (reading or building a sinusoid).

1. d = (max + min)/2.
2. a = (max − min)/2.
3. b = 2π / (period).
4. c from a feature: a maximum of a cos curve sits at x = c; a zero-going-up of a sin curve sits at x = c.

Worked micro-example. A Ferris wheel: lowest point 2 m, highest 18 m, one revolution every 40 seconds, rider starts at the bottom. Then d = (18+2)/2 = 10, a = (18−2)/2 = 8, b = 2π/40 = π/20. Starting at the minimum suits a negative cosine: h(t) = 10 − 8 cos(πt/20). Check t = 0: h = 10 − 8 = 2 ✓; t = 20: h = 10 + 8 = 18 ✓.

Traps.

- Mixing up period and b — the period is 2π/b, so a *larger* b means a *shorter* period.
- Sign of the phase shift: y = sin(x − c) shifts *right* by c.
- Choosing sin when −cos (or +cos) models the start condition far more naturally — pick the parent curve whose natural starting feature matches the given one.
- Treating the tan graph as having an amplitude or a 2π period.

If you remember nothing else

- **Radians everywhere.** s = rθ and A = ½r^2θ are radian-only; trig equations and graphs assume radians. Convert degrees *before* substituting, and write "(rad)" next to θ as a self-check.
- **The sine rule for an angle is ambiguous.** Whenever it gives you an angle, write down both θ and 180° − θ, then test each against the 180° triangle-sum.
- **Memorise the exact-value table and the quadrant signs** — they are not in the booklet, and Paper 1 gives you no calculator to recover them.
- **Prove identities one side at a time**, transforming the messier side until it matches the other; never operate across the equals sign.
- **Pick the cos 2θ form to match what else is present**: $2\cos^2\theta - 1$ with cosines, $1 - 2\sin^2\theta$ with sines, $\cos^2\theta - \sin^2\theta$ to factor.
- **Expand the domain for sin(nx).** Substitute u = nx, widen the interval by the factor n, solve fully for u, then divide back — this is where half-marks vanish.

4. Statistics & Probability

Statistics and probability together account for roughly 17% of the AA SL course, but the two halves behave very differently on Paper 1. The data-handling half — computing standard deviations, drawing cumulative frequency curves, finding a regression equation — is overwhelmingly *Paper 2* work, because it depends on the GDC. What Paper 1 actually tests from this topic is **probability**: sample spaces, the addition rule, conditional probability, tree and Venn diagrams, discrete random variables, expected value, and small binomial calculations you can do by hand. The data-handling sub-topics still appear on P1, but only in their *conceptual* form — interpreting a box plot, reading off a quartile, explaining why a sample is biased, stating what a correlation coefficient does and doesn't mean, reasoning about the effect of a transformation. Treat this section in two registers: know the definitions and interpretations cold (cheap marks on P1), and know the probability machinery well enough to execute it without a calculator.

4.1 Data, sampling and bias

Key facts. Data is *discrete* (counted, takes separate values — number of siblings) or *continuous* (measured, any value in a range — height). Sampling techniques: **random** (every member equally likely), **systematic** (every k-th member from a list), **stratified** (population split into groups, sampled in proportion to group size), **quota** (interviewer fills fixed numbers per group, non-randomly), **convenience** (whoever is easiest to reach). Reliability roughly decreases down that list: random and stratified samples are representative; quota and convenience samples are prone to **bias** — a systematic skew that makes the sample unrepresentative of the population. None of this is in the formula booklet; it is pure recall and explanation.

Method (identify/criticise a sampling method).

1. Name the technique from how members were selected.
2. State whether selection was random or had a systematic skew.
3. Identify the source of bias if any (who is over- or under-represented).
4. Suggest a better method — usually random or stratified.

Worked micro-example. A café surveys the first 30 customers on a Monday morning about how often they visit. This is a convenience sample. It is biased toward *frequent* customers (regulars come early on weekdays), so it will overstate average visit frequency. A stratified sample across different days and times would be more representative.

Traps. Don't confuse systematic (every k-th, *is* a valid random-ish method) with convenience. An **outlier** is a value far from the rest; it can be a genuine extreme or an error, and it drags the mean and range but barely moves the median — that distinction is worth marks.

4.2 Central tendency and dispersion

Key facts / formulas. Mean $\bar{x} = (\Sigma x)/n$, or for a frequency table $\bar{x} = (\Sigma fx)/(\Sigma f)$. Median = middle value when ordered (average of the two middle values if n is even). Mode = most frequent value; for grouped data the **modal class** is the class with highest frequency. Dispersion: range = max − min; interquartile range IQR = $Q_3 - Q_1$; variance σ^2 and standard deviation σ measure spread about the mean. None of the central-tendency or spread formulas are given in the booklet in a form you'd compute by hand on P1 — and you are not expected to compute σ by hand. What P1 *does* test is the **transformation rules**, which you must internalise:

- Add a constant c to every data value → mean increases by c; **standard deviation is unchanged** (the spread doesn't move).
- Multiply every data value by a constant k → mean is multiplied by k; **standard deviation is multiplied by |k|** (and variance by k^2).

Method (transformation question).

1. Identify whether data was shifted (+c), scaled (×k), or both.
2. Apply +c / ×k to the mean.
3. Apply "unchanged" / ×|k| to the standard deviation.

Worked micro-example. A set of test scores has mean 52 and standard deviation 8. The teacher adds 5 marks to everyone, then doubles the result. After +5: mean 57, sd 8. After ×2: mean 114, sd 16. Variance went from 64 to 256.

Traps. Adding a constant does *not* change the standard deviation — a very common slip. When scaling, the sd multiplies by |k| but the variance multiplies by k^2. Don't quote the modal *value* when asked for the modal *class*.

4.3 Frequency tables, cumulative frequency, quartiles and percentiles

Key facts. A cumulative frequency value is the running total of frequencies up to and including a class. Plotting cumulative frequency against the **upper class boundary** gives the cumulative frequency curve (an S-shape, an "ogive"). From it: the **median** is read at cumulative frequency n/2, the **lower quartile Q_1** at n/4, the **upper quartile Q_3** at 3n/4, and a **percentile** P_k at $(k/100)\cdot n$. Not in the booklet — these are positions you must know.

Method (read a value from a cumulative frequency graph).

1. Compute the position on the cumulative frequency (vertical) axis: n/2, n/4, 3n/4, or $k\%\cdot n$.
2. Draw a horizontal line across to the curve.
3. Drop down to the horizontal axis and read the data value.

Worked micro-example. 80 students' homework times are graphed. n/4 = 20, so trace across from cumulative frequency 20 to the curve and down — say it reads 35 minutes, so $Q_1 = 35$. From n/2 = 40 you might read a median of 48, and from 3n/4 = 60 a Q_3 of 65, giving IQR = 30.

Traps. Plot against the *upper* boundary, not the midpoint. Use n/2 for the median position on a cumulative graph (not (n+1)/2 — that's for an ordered list of raw data). Read carefully off the scale; P1 graph questions are about correct *positions*, and the numbers are chosen to be readable.

4.4 Box-and-whisker plots and the outlier rule

Key facts. A box plot displays the **five-number summary**: minimum, Q_1, median, Q_3, maximum. The box spans Q_1 to Q_3 (width = IQR), with a line at the median; whiskers reach to the smallest and largest values that are *not* outliers. The **outlier rule** (must be memorised — not in the booklet): a value is an outlier if it is

- below $Q_1 - 1.5 \times \text{IQR}$, or
- above $Q_3 + 1.5 \times \text{IQR}$.

Method (check for outliers / draw the box plot).

1. Find Q_1, Q_3, and $\text{IQR} = Q_3 - Q_1$.
2. Compute the fences: $Q_1 - 1.5 \cdot \text{IQR}$ and $Q_3 + 1.5 \cdot \text{IQR}$.
3. Any data value outside the fences is an outlier (mark it as a separate point).
4. Whiskers extend to the most extreme values still inside the fences.

Worked micro-example. A data set has $Q_1 = 12$, $Q_3 = 20$, so IQR = 8 and 1.5·IQR = 12. Fences: 12 − 12 = 0 and 20 + 12 = 32. A value of 35 is an outlier; a value of 31 is not. With non-calculator-friendly quartiles like these, this is a very fair P1 question.

Traps. It is 1.5 × IQR, not 1.5 × the range and not 1.5 × the standard deviation. The fence is measured from the *quartiles*, not from the median or the mean. A box plot tells you spread and skew but says nothing about the *number* of data points or their exact distribution within the box.

4.5 Correlation: scatter diagrams and Pearson's r

Key facts. A scatter diagram plots paired data (x, y) to show association. **Pearson's product-moment correlation coefficient r** measures the strength and direction of a *linear* relationship. Its value satisfies $-1 \le r \le 1$: r near +1 is strong positive linear correlation, near −1 strong negative, near 0 little or no *linear* correlation. You compute r on the GDC (Paper 2), so on P1 the examinable part is **interpretation**: describing correlation as weak/moderate/strong and positive/negative, and — crucially — knowing what r does *not* tell you.

- r measures *linear* association only; data on a perfect parabola can have $r \approx 0$.
- **Correlation is not causation.** A strong r between two variables does not mean one causes the other; there may be a hidden third (confounding) variable, or it may be coincidence.

Worked micro-example. Across a town, the number of ice creams sold and the number of swimming-pool visits have $r = 0.94$. This is strong positive correlation, but ice cream does not cause swimming — hot weather (a confounding variable) drives both.

Traps. Never write "x causes y" from r alone. Don't describe a clearly curved scatter as "no relationship" just because r is small — say "no *linear* relationship". r has no units and doesn't depend on which variable is x.

4.6 Linear regression (y on x)

Key facts. The **least-squares regression line of y on x** is the straight line $y = ax + b$ that best predicts y from x. You find a and b on the GDC, so finding the equation is Paper 2; on P1 you reason about its *properties* and *use*:

- The regression line **always passes through the mean point** $(\bar{x}, \bar{y})$. This is an examinable fact.
- Use the line to **predict** a y-value by substituting an x-value.
- Predicting *within* the range of the data is interpolation (reasonably reliable). Predicting *outside* the range is **extrapolation** and is unreliable — the linear pattern may not continue.

Method (use a given regression line).

1. To predict, substitute the x-value into $y = ax + b$.
2. State whether you are interpolating (safe) or extrapolating (caution).
3. Comment on reliability — also weaker if $|r|$ is low.

Worked micro-example. For data on hours studied (x, ranging 1–6) and test score (y), the regression line is $y = 9x + 28$. Predicting for $x = 4$ gives $y = 64$ — interpolation, reasonable. Predicting for $x = 15$ gives $y = 163$, which is nonsense (above any possible score): extrapolation has broken down.

Traps. The line of y on x predicts y from x, not the other way round. Don't trust a prediction far outside the data range. A regression line can be fitted to *any* data, even data with no real correlation — always check r before relying on it.

4.7 Probability basics

Key facts. The **sample space** U is the set of all possible outcomes. For equally likely outcomes, P(A) = (number of outcomes in A)/(total number of outcomes). Probabilities satisfy $0 \le P(A) \le 1$. The **complement** A′ is "A does not happen", and $\mathbf{P(A') = 1 - P(A)}$ — often the fastest route to an answer ("at least one" problems especially). These facts are foundational and not separately listed in the booklet.

Method.

1. List or count the sample space.
2. Count the outcomes in the event.
3. Divide; or, if the event is "at least one" / awkward, compute 1 – P(complement).

Worked micro-example. Two fair coins are tossed. Sample space {HH, HT, TH, TT}. P(at least one head) = 1 – P(TT) = 1 – 1/4 = 3/4.

Traps. Make sure outcomes really are equally likely before counting. "At least one" is the complement of "none" — reach for the complement rather than adding many cases.

4.8 Combined events: the addition rule and mutually exclusive events

Key facts / formulas. The **addition rule** is in the formula booklet:

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

You subtract $P(A \cap B)$ because outcomes in the overlap were counted twice. Events are **mutually exclusive** if they cannot both happen — then $P(A \cap B) = 0$, so the rule simplifies to $P(A \cup B) = P(A) + P(B)$. "Mutually exclusive" is *not* in the booklet as a definition; you must know it.

Method.

1. Identify $P(A)$, $P(B)$, and the overlap $P(A \cap B)$.
2. If told the events are mutually exclusive, set $P(A \cap B) = 0$.
3. Apply $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Worked micro-example. From a standard pack, P(King) = 4/52, P(Heart) = 13/52, P(King and Heart) = 1/52. So P(King or Heart) = 4/52 + 13/52 − 1/52 = 16/52 = 4/13.

Traps. Don't forget to subtract the overlap when events *can* coincide. Conversely, don't subtract an overlap that is genuinely zero. Drawing a a quick Venn diagram prevents both errors.

4.9 Venn diagrams and tree diagrams

Key facts. A **Venn diagram** partitions the sample space into regions: A only, B only, $A \cap B$, and neither ($A' \cap B'$). Fill it from the **inside out** — put the intersection in first, then subtract to get the "only" regions. A **tree diagram** shows a sequence of events; the two rules are:

- **Multiply along the branches** (probabilities for a path of consecutive events).
- **Add across the branches** (separate paths giving the same overall outcome).

Branch probabilities at each node sum to 1.

Method (Venn).

1. Place the intersection count/probability first.
2. Subtract to fill "A only" = $P(A) - P(A \cap B)$, similarly "B only".
3. The remainder of the total goes in "neither".

Method (tree).

1. Draw a branch for each stage; label every branch with its probability.
2. Multiply along each complete path.
3. Add the paths that satisfy the question.

Worked micro-example. A bag has 3 red and 2 blue counters; two are drawn without replacement. P(both red) = (3/5)(2/4) = 6/20 = 3/10. P(one of each) = P(RB) + P(BR) = (3/5)(2/4) + (2/5)(3/4) = 6/20 + 6/20 = 3/5.

Traps. "Without replacement" changes the second-stage probabilities — the denominator and the relevant numerator both drop by 1. Multiplying when you should add (or vice versa) is the classic tree error: multiply *within* a path, add *between* paths.

4.10 Conditional probability

Key facts / formula. Conditional probability — the probability of A *given that* B has happened — is in the formula booklet:

$P(A \mid B) = P(A \cap B) / P(B)$.

The conditioning event B becomes the new "denominator world": you restrict attention to the cases where B occurred.

Method.

1. Identify which event is the *given* (condition) — that goes in the denominator.
2. Find $P(A \cap B)$ — the outcomes where both happen.
3. Find P(B).
4. Divide.

Worked micro-example. A class of 30: 18 study Spanish, 12 study Art, 7 study both. Given a student studies Art, the probability they also study Spanish is $P(S \mid A) = P(S \cap A)/P(A) = (7/30)/(12/30) = 7/12$.

Traps. The **direction trap**: $P(A \mid B)$ and $P(B \mid A)$ are different numbers — read carefully which event is given. In the example, $P(A \mid S) = 7/18 \neq 7/12$. The condition always lands in the denominator. On a tree diagram, $P(A \mid B)$ corresponds to a *second-stage* branch, not a whole-path product.

4.11 Independent events

Key facts / formula. Events A and B are **independent** if one happening does not change the probability of the other. The test/definition: A and B are independent exactly when

$P(A \cap B) = P(A) \times P(B)$

(equivalently $P(A \mid B) = P(A)$). The multiplication form is effectively the working definition you use; it must be internalised, and the conditional-probability formula it follows from is in the booklet.

The single most important distinction in this whole topic:

- **Mutually exclusive** events *cannot happen together*: $P(A \cap B) = 0$.
- **Independent** events *do not affect each other*: $P(A \cap B) = P(A) \cdot P(B)$.

These are almost opposite ideas. In fact, two events with non-zero probabilities that are mutually exclusive *cannot* be independent — if A happens, B definitely didn't, so A drastically changed B's probability.

Method (test for / use independence).

1. To *test*: compute $P(A)\cdot P(B)$ and compare with $P(A \cap B)$. Equal $\Rightarrow$ independent.
2. To *use* (when told they're independent): $P(A \cap B) = P(A)\cdot P(B)$.

Worked micro-example. $P(A) = 0.5$, $P(B) = 0.4$, $P(A \cap B) = 0.2$. Since $0.5 \times 0.4 = 0.2 = P(A \cap B)$, A and B are independent. If instead $P(A \cap B)$ had been 0, they'd be mutually exclusive — and *not* independent.

Traps. Conflating the two concepts is the headline trap of the topic. "Mutually exclusive" lets you *add*; "independent" lets you *multiply* — never swap them. Don't assume independence unless told or unless you've verified $P(A \cap B) = P(A)\cdot P(B)$.

4.12 Discrete random variables and expected value

Key facts / formula. A discrete random variable X takes specific values with given probabilities — its **probability distribution**. The non-negotiable condition (not in the booklet — internalise it): **all the probabilities must sum to 1.** The **expected value** is in the formula booklet:

$E(X) = \Sigma\, x \cdot P(X = x)$.

$E(X)$ is the long-run mean of X; it need not be an attainable value of X. A "fair game" is one where $E(X) = 0$ (or E(net gain) = 0).

Method.

1. If a probability is unknown, use $\Sigma\, P(X = x) = 1$ to find it.
2. Multiply each value x by its probability $P(X = x)$.
3. Sum the products to get $E(X)$.

Worked micro-example. X has $P(X=1) = 0.2$, $P(X=2) = 0.5$, $P(X=3) = k$. Sum to 1 $\Rightarrow$ $k = 0.3$. Then $E(X) = 1(0.2) + 2(0.5) + 3(0.3) = 0.2 + 1.0 + 0.9 = 2.1$.

Traps. Forgetting the sum-to-1 condition — it's usually the first mark and unlocks the rest. If asked for variance, $\mathbf{Var(X) = E(X^2) - [E(X)]^2}$: you compute $E(X^2) = \Sigma\, x^2\cdot P(X=x)$ and subtract the *square* of $E(X)$. The frequent error is subtracting $E(X)$ itself, or squaring after subtracting — square $E(X)$ first, then subtract.

4.13 The binomial distribution

Key facts / formulas. $X \sim B(n, p)$ counts the number of successes in n independent trials, each with success probability p. The probability formula, the mean and the variance are **all in the formula booklet**:

- $P(X = r) = {}^nC_r \cdot p^r \cdot (1 - p)^{n-r}$
- Mean: $E(X) = np$
- Variance: $Var(X) = np(1 - p)$

Conditions for a binomial model: fixed n, two outcomes per trial, constant p, independent trials.

On Paper 1, you won't get large-n binomial probabilities — those need the GDC (Paper 2). What *can* appear: stating mean/variance via np and np(1−p); computing $P(X = r)$ for **small n** where ${}^nC_r$ and the powers are hand-computable (n up to about 4 or 5, friendly p like 1/2, 1/3, 1/4); and "at least one" via the complement, $P(X \geq 1) = 1 - P(X = 0) = 1 - (1-p)^n$.

Method (hand binomial).

1. Identify n, p, and the r required.
2. Compute ${}^nC_r$ (small — e.g. ${}^4C_2 = 6$).
3. Compute p^r and $(1-p)^{n-r}$.
4. Multiply the three together. For "at least one", do $1 - (1-p)^n$.

Worked micro-example. A fair coin is tossed 4 times; X = number of heads, so $X \sim B(4, 1/2)$. $P(X = 2) = {}^4C_2 \cdot (1/2)^2 \cdot (1/2)^2 = 6 \cdot 1/4 \cdot 1/4 = 6/16 = 3/8$. Mean = np = 4 · 1/2 = 2; variance = np(1−p) = 4 · 1/2 · 1/2 = 1.

Traps. Don't forget the ${}^nC_r$ coefficient — $p^r(1-p)^{n-r}$ alone is just *one* arrangement. The exponents must add to n: r and n − r. Variance is np(1 − p), not np; the standard deviation is its square root. Check the binomial *conditions* hold before modelling — non-constant p or dependent trials breaks it.

4.14 The normal distribution

Key facts / formula. A normal distribution $X \sim N(\mu, \sigma^2)$ is the symmetric bell curve, centred on its mean μ, with spread set by σ; mean = median = mode at the centre, and total area under the curve = 1. Symmetry gives $P(X > \mu) = P(X < \mu) = 0.5$. **Standardisation** converts any normal value to the standard normal $Z \sim N(0, 1)$:

$z = (X - \mu) / \sigma$.

The z-score says how many standard deviations X is from the mean.

Paper 1 vs Paper 2. Numerical normal *probabilities* and *inverse-normal* values require the GDC — that is **Paper 2** territory. What can appear on **Paper 1**: the *shape and symmetry* arguments (e.g. deducing $P(X < \mu) = 0.5$, or using symmetry to relate two probabilities), and the **standardisation step itself** — computing a z-value, or reasoning about it. A value's z-score being negative tells you it lies below the mean.

Method (standardise).

1. Identify μ and σ.
2. Substitute into $z = (X - \mu)/\sigma$.
3. Keep the sign: X below μ gives a *negative* z; X above μ gives positive.

Worked micro-example. Heights are $N(\mu = 170, \sigma^2 = 25)$, so $\sigma = 5$. A height of 162 standardises to $z = (162 - 170)/5 = -8/5 = -1.6$ — i.e. 1.6 standard deviations *below* the mean. A height of 178 gives $z = (178 - 170)/5 = +1.6$, the same distance above; by symmetry $P(X > 178) = P(X < 162)$.

Traps. The **sign error**: when $X < \mu$ the numerator $X - \mu$ is negative, so z is negative — don't drop the minus sign. Use σ, not σ^2, in the denominator (here $\sigma = 5$, not 25). And remember the curve's area is a *probability*: it can never exceed 1.

If you remember nothing else

- **Independent ≠ mutually exclusive.** Mutually exclusive (can't co-occur) ⇒ *add*, $P(A \cap B) = 0$. Independent (don't affect each other) ⇒ *multiply*, $P(A \cap B) = P(A) \cdot P(B)$. Two non-trivial events can't be both.
- **Conditional probability has a direction.** $P(A|B) = P(A \cap B)/P(B)$ — the *given* event goes in the denominator, and $P(A|B) \neq P(B|A)$.
- **A probability distribution must sum to 1** — that's usually your first mark and it unlocks $E(X)$. And $\mathbf{Var(X) = E(X^2) - [E(X)]^2}$ — square the expected value, don't just subtract it.
- **Trees: multiply along a path, add between paths.** Venn diagrams: fill the intersection first, then work outwards.
- **The outlier rule is 1.5 × IQR** beyond the quartiles; adding a constant leaves the standard deviation unchanged while scaling multiplies it by $|k|$.
- **Know what's P1 and what's P2.** Heavy data computation, full binomial/normal probabilities, and finding regression equations are calculator work. Paper 1 wants probability machinery by hand, small binomials, the standardisation step, and the *interpretation* of everything else.

5. Calculus

Calculus is the single heaviest topic on Paper 1 — roughly a quarter of the 80 marks, and the part of the paper with the highest density of avoidable errors. It rewards the same handful of disciplines over and over: announce your method in writing, keep the chain rule switched on, never lose the constant of integration, and answer the question that was actually asked. Because Paper 1 is non-calculator, every derivative and integral here is exact-form and by hand; you cannot get a maximum by reading a graph, so you must *show* the second-derivative value, *show* the sign analysis, *show* $F(b) - F(a)$. Marks in this topic are lost in ones and twos to small process failures, not to not knowing the maths. Treat the rest of this section as a list of those failures and how to not commit them.

5.1 Limits and the derivative as a rate of change

Key facts. The derivative $f'(x)$ is the gradient of the tangent to the curve $y = f(x)$ at a point — equivalently, the instantaneous rate of change of f with respect to x. Informally, it is the limit of the gradient of a chord $(f(x+h) - f(x))/h$ as $h \to 0$. AA SL treats this conceptually rather than computationally: you are not required to differentiate from first principles in an exam, but you are expected to *understand* that dy/dx is a limiting gradient. Notation: $f'(x)$, dy/dx, and (for a value at a point) $f'(a)$ all mean the same thing. None of this needs the formula booklet — it is conceptual scaffolding.

Why it matters. Every "rate of change", "how fast", "gradient at", and "instantaneous velocity" question is asking for a derivative evaluated at a point. Recognising the language is half the mark.

Worked micro-example. If the volume of a balloon is $V(t) = 4t^2 + 3t$ (cm^3, t in seconds), the rate at which the volume is increasing at $t = 2$ is $V'(t) = 8t + 3$, so $V'(2) = 19$ cm^3 per second. The phrase "rate at which ... is increasing" *is* the instruction to differentiate.

Traps. Confusing the value of the function with the value of its rate of change — $V(2) = 22$ is the volume, not the rate. Forgetting units carry the "per second" from the differentiation. Reading "average rate of change between $t = 1$ and $t = 3$" as a derivative — that is a chord gradient, $(V(3) - V(1))/2$, not V' of anything.

5.2 Differentiating polynomials — the power rule

Key facts / formulas. The power rule, $d/dx(x^n) = n \cdot x^{n-1}$, is in the formula booklet. With it: the derivative of a constant is 0; constant multiples pull out, $d/dx(k \cdot f) = k \cdot f'$; and sums differentiate term by term, $d/dx(f \pm g) = f' \pm g'$. These last three are not separately printed — they are the basic algebra of differentiation you must simply have.

Method.

1. Write every term as a power of x — including rewriting roots and fractions (see 5.4).

2. Apply the power rule to each term: multiply by the old exponent, subtract one from the exponent.
3. Tidy: a constant term differentiates to 0 and disappears.

Worked micro-example. $f(x) = 5x^3 - 7x^2 + 2x - 9$. Then $f'(x) = 15x^2 - 14x + 2$. The -9 vanishes. To find the gradient at $x = 1$: $f'(1) = 15 - 14 + 2 = 3$.

Traps. Dropping the constant multiple — $d/dx(5x^3)$ is $15x^2$, not $5x^2$. Forgetting that the lone term $2x$ differentiates to 2, not to $2x$ or to 0. Leaving the -9 in as -9 instead of 0.

5.3 Derivatives of further functions

Key facts / formulas. All of these standard derivatives are in the formula booklet:

- $d/dx(\sin x) = \cos x$
- $d/dx(\cos x) = -\sin x$ (note the minus sign)
- $d/dx(\tan x) = 1/\cos^2 x$ (equivalently $\sec^2 x$)
- $d/dx(e^x) = e^x$
- $d/dx(\ln x) = 1/x$

The two logarithm cases must be kept distinct. $d/dx(\ln x) = 1/x$ cleanly. But $d/dx(\log_a x) = 1/(x \ln a)$ — there is an extra $\ln a$ in the denominator because $\log_a x = \ln x / \ln a$, and $\ln a$ is just a constant multiple. The $\log_a x$ form is in the booklet but is the one students misremember.

Worked micro-example. $g(x) = 3 \sin x + 2e^x - \ln x$. Then $g'(x) = 3 \cos x + 2e^x - 1/x$. Compare $h(x) = \log_2 x$, whose derivative is $1/(x \ln 2)$, *not* $1/x$.

Traps. Differentiating $\cos x$ to $+\sin x$ (the missing minus is one of the most common single errors in the whole paper). Writing $d/dx(\log_a x) = 1/x$ by analogy with ln. Forgetting $\tan x$ exists in the derivative list at all and trying to rebuild it mid-question.

5.4 Negative and rational exponents

Key facts. The power rule works for *all* rational exponents, but only after the term is written as a power. So $1/x = x^{-1}$, $1/x^3 = x^{-3}$, $\sqrt{x} = x^{(1/2)}$, $1/\sqrt{x} = x^{(-1/2)}$. This rewriting step is where marks leak — it is not in the booklet because it is just index laws.

Method.

1. Rewrite the term with a single power of x in the numerator. Move denominators up with a negative exponent; turn roots into fractional exponents.
2. Apply the power rule.
3. Optionally rewrite the answer back into root/fraction form if the question's other parts use that form.

Worked micro-example. $f(x) = 1/x^2 + \sqrt{x} = x^{-2} + x^{(1/2)}$. Then $f'(x) = -2x^{-3} + (1/2)x^{(-1/2)} = -2/x^3 + 1/(2\sqrt{x})$. Note both new exponents are *one less* than before: $-2 - 1 = -3$, and $1/2 - 1 = -1/2$.

Traps. Sign errors subtracting one from a negative exponent — from x^{-2} you go to x^{-3}, not x^{-1}. Forgetting to multiply by the (negative) old exponent, so the minus sign never appears. Differentiating $\sqrt{x}$ as if the exponent were 1. Leaving $1/x^2$ un-rewritten and "differentiating the bottom".

5.5 The product, quotient and chain rules

Key facts / formulas. All three are in the formula booklet:

- **Product rule.** $(uv)' = u'v + uv'$.
- **Quotient rule.** $(u/v)' = (u'v - uv')/v^2$. Order matters in the numerator — $u'v$ *first*.
- **Chain rule.** If $y = g(u)$ and $u = f(x)$, then $dy/dx = (dy/du)(du/dx)$. In function form, $d/dx[g(f(x))] = g'(f(x)) \cdot f'(x)$.

The fact you must *internalise*: the chain rule lives *inside* almost every product or quotient question. When you differentiate u or v and that piece is itself a composite — $\sin(3x)$, $e^{(x^2)}$, $(2x+1)^5$, $\ln(5x)$ — the chain rule fires inside the product/quotient rule. Forgetting the inner derivative is the defining calculus trap of Paper 1.

Method (product or quotient with composites).

1. Label u and v explicitly. Write them down.
2. Differentiate each — and if u or v is composite, *apply the chain rule there* and write the inner-derivative factor.
3. Substitute into the product or quotient formula.
4. Simplify only as far as the question needs; do not expand needlessly.

Worked micro-example (product + chain). $y = x^2 \cdot \sin(3x)$. Let $u = x^2$, $v = \sin(3x)$. Then $u' = 2x$, and $v' = 3\cos(3x)$ — the 3 is the chain-rule inner derivative. So $y' = 2x \cdot \sin(3x) + x^2 \cdot 3\cos(3x) = 2x \sin(3x) + 3x^2 \cos(3x)$.

Worked micro-example (quotient + chain). $y = e^{(2x)}/(x + 1)$. Let $u = e^{(2x)}$, $v = x + 1$. Then $u' = 2e^{(2x)}$ (chain rule gives the 2), $v' = 1$. So $y' = (2e^{(2x)}(x+1) - e^{(2x)} \cdot 1)/(x+1)^2 = e^{(2x)}(2x + 1)/(x+1)^2$.

Traps. The chain rule forgotten inside the product — writing $v' = \cos(3x)$ instead of $3\cos(3x)$. Quotient numerator written in the wrong order, $uv' - u'v$, which flips the whole sign. Forgetting to square the denominator. Trying to "simplify" $e^{(2x)}(x+1)$ into something it isn't. Using the quotient rule when the expression could be split into separate terms and power-ruled more cheaply (e.g. $(x^3 + x)/x = x^2 + 1$).

5.6 Tangents and normals

Key facts. At the point where $x = a$ on $y = f(x)$:

- the **tangent** has gradient $m = f'(a)$;

- the **normal** is perpendicular to the tangent, so its gradient is the **negative reciprocal**, $-1/f'(a)$.

Both lines pass through $(a, f(a))$. The "negative reciprocal" rule is not in the booklet — internalise it. The line equation $y - y_1 = m(x - x_1)$ is the tool.

Method.

1. Find the point: compute $y_1 = f(a)$. You need an actual coordinate.
2. Find the gradient: differentiate, evaluate $f'(a)$. This is the tangent gradient.
3. For a normal, take the negative reciprocal: $m\_normal = -1/f'(a)$.
4. Substitute the point and the appropriate gradient into $y - y_1 = m(x - x_1)$.

Worked micro-example. Curve $y = x^2 - 2x$ at $x = 3$. Point: $y(3) = 9 - 6 = 3$, so $(3, 3)$. Gradient: $y' = 2x - 2$, $y'(3) = 4$. Tangent: $y - 3 = 4(x - 3)$, i.e. $y = 4x - 9$. Normal: gradient $-1/4$, so $y - 3 = -\frac{1}{4}(x - 3)$, i.e. $y = -\frac{1}{4}x + 15/4$.

Traps. Using the tangent gradient for the normal (or vice versa) — the single most common tangent/normal error. Forgetting the point and writing $y = mx$ with no intercept. Negative reciprocal done as just the reciprocal (1/4) or just the negative (−4) instead of −1/4. Substituting $x = 3$ into f' but forgetting to also get the y-coordinate from f.

5.7 Increasing and decreasing functions

Key facts. A function is **increasing** on an interval where $f'(x) > 0$, and **decreasing** where $f'(x) < 0$. To find these intervals you solve the inequality, usually by first finding where $f'(x) = 0$ (the stationary points) and testing the sign of f' between them. Not a booklet item — it is the meaning of the first derivative.

Method.

1. Differentiate.
2. Solve $f'(x) = 0$ to find the boundary x-values.
3. Test the sign of f' in each interval between (and outside) those values.
4. State intervals in x: "increasing for $x < \ldots$", "decreasing for $\ldots < x < \ldots$".

Worked micro-example. $f(x) = x^3 - 3x$. $f'(x) = 3x^2 - 3 = 3(x - 1)(x + 1)$. Zeros at $x = \pm 1$. Testing: at $x = -2$, $f' = 9 > 0$; at $x = 0$, $f' = -3 < 0$; at $x = 2$, $f' = 9 > 0$. So f is increasing for $x < -1$ and $x > 1$, decreasing for $-1 < x < 1$.

Traps. Stating the intervals in terms of y instead of x. Mixing up the inequality direction — $f' > 0$ means increasing. Giving only the boundary points and not the intervals. Forgetting to test a region outside the outermost root.

5.8 Stationary points and their classification

Key facts / formulas. A **stationary point** occurs where $f'(x) = 0$. To classify it, the **second derivative test**:

- $f''(x_0) < 0$ → local **maximum**;
- $f''(x_0) > 0$ → local **minimum**;
- $f''(x_0) = 0$ → test is **inconclusive**; fall back to a first-derivative sign analysis.

This is examinable behaviour you must have cold. The mark for classification is a reasoning (R) mark — you only earn it by *writing down the value of $f''(x_0)$ and its sign*, then stating the conclusion. "It's a maximum" with no f'' value shown is a lost mark even if correct.

Method.

1. Differentiate, set $f'(x) = 0$, solve for x.
2. Find the y-coordinate of each stationary point from the original f.
3. Differentiate again to get f''.
4. Evaluate f'' at each stationary x. State the value and sign in writing, then the conclusion.
5. If $f''(x_0) = 0$: test the sign of f' just left and just right of x_0 — change + to – is a max, – to + is a min, no change is a point of inflection.

Worked micro-example. $f(x) = x^3 - 3x^2 + 4$. $f'(x) = 3x^2 - 6x = 3x(x - 2)$, so stationary points at $x = 0$ and $x = 2$. y-values: $f(0) = 4$, $f(2) = 0$. $f''(x) = 6x - 6$. At $x = 0$: $f''(0) = -6 < 0$ → local maximum at (0, 4). At $x = 2$: $f''(2) = 6 > 0$ → local minimum at (2, 0). Note how the f'' value is stated before the conclusion — that is the R mark.

Traps. Classifying without showing the f'' value — the lost R mark, repeatedly. Solving $f' = 0$ but forgetting to find the y-coordinates. Reading $f'' < 0$ as a minimum (it is the *maximum* — the curve is concave down, like a hill). When $f''(x_0) = 0$, jumping to "inflection" without doing the first-derivative sign check. Giving stationary points as x-values only.

5.9 Concavity and points of inflection

Key facts. The second derivative controls concavity: $f''(x) > 0$ means the curve is **concave up** (holds water), $f''(x) < 0$ means **concave down**. A **point of inflection** is where concavity changes — so $f''(x) = 0$ *and* f'' changes sign there. The condition $f'' = 0$ alone is not enough; the sign must actually change. Not a booklet formula — it is the meaning of f''.

Method.

1. Find f''.
2. Solve $f''(x) = 0$ — these are *candidate* inflection points.
3. Check that f'' changes sign across each candidate. If it does, it is a point of inflection.
4. Get the y-coordinate from the original f.

Worked micro-example. $f(x) = x^3 - 3x^2 + 4$ (same function as 5.8). $f''(x) = 6x - 6 = 0$ at $x = 1$. To the left ($x = 0$) $f'' = -6 < 0$; to the right ($x = 2$) $f'' = +6 > 0$ — concavity changes, so there is a point of inflection at $x = 1$, $y = f(1) = 1 - 3 + 4 = 2$. The point is $(1, 2)$.

Traps. Declaring an inflection point from $f'' = 0$ without checking the sign change. Confusing concave up/down ($f'' > 0$ is *up*). Forgetting the y-coordinate. Mixing inflection points up with stationary points — they are different conditions ($f'' = 0$ with sign change vs $f' = 0$).

5.10 Optimisation

Key facts. Optimisation is a maximum/minimum problem dressed in a real-world context with a *constraint*. The four-step method is a discipline you must internalise — there is no booklet entry for it:

1. **Express the quantity to optimise as a function of one variable**, using the constraint equation to eliminate the other variable.
2. **Differentiate and set the derivative to zero**; solve for the critical value.
3. **Justify that it is a maximum (or minimum)** — second-derivative test, or a first-derivative sign check. This is a mark; do not skip it.
4. *Substitute back into the* original *quantity to get the optimum* value*. Also check domain endpoints if the variable is restricted to an interval.

Method. As the four steps above. The two places marks are lost: skipping step 3 (the justification) and stopping at step 2's critical input instead of doing step 4.

Worked micro-example. A rectangle has perimeter 40 cm; maximise its area. Constraint: $2x + 2y = 40$, so $y = 20 - x$. Area $A = x \cdot y = x(20 - x) = 20x - x^2$. Differentiate: $A' = 20 - 2x = 0 \rightarrow x = 10$. Justify: $A'' = -2 < 0$, so this is a maximum. Substitute back: $y = 20 - 10 = 10$, $A = 10 \times 10 = 100\ \text{cm}^2$. The answer to "maximum area" is **100 cm²**, not $x = 10$.

Traps. Giving $x = 10$ as the final answer when the question asked for the maximum *area* — answering with the input, not the value. Skipping the f'' justification (lost R mark). Forgetting to use the constraint, leaving two variables in A. Not checking domain endpoints when the variable is bounded — sometimes the optimum is at an endpoint, not a stationary point. Differentiating the constraint instead of the objective.

5.11 Kinematics

Key facts. For motion in a line, the three quantities form a **derivative ladder**:

- velocity $v(t) = ds/dt$ — the derivative of displacement;
- acceleration $a(t) = dv/dt = d^2s/dt^2$ — the derivative of velocity.

Going the other way, you integrate: $s = \int v\,dt$, $v = \int a\,dt$ (each with a + C fixed by an initial condition). This ladder is not a booklet formula — internalise it. Two distinctions carry marks:

- **Displacement vs distance travelled.** Displacement is the (signed) net change in position; distance travelled is the *total* path length and is never negative. If velocity changes sign during the interval, they differ — distance travelled needs the integral of |v|, computed by splitting the interval at the times where v = 0 and adding the magnitudes.
- **"Momentarily at rest" means v = 0** — not a = 0. Solve v(t) = 0 for those times.

Method (distance travelled when v changes sign).

1. Solve v(t) = 0 to find the turning times within the interval.
2. Integrate v over each sub-interval separately.
3. Take the absolute value of each piece and add them.

Worked micro-example. A particle has velocity $v(t) = t^2 - 4t$ (m/s) for $0 \le t \le 5$. It is momentarily at rest when v = 0: $t(t - 4) = 0$, so t = 0 and t = 4. On $0 \le t \le 4$, $v \le 0$; on $4 \le t \le 5$, $v \ge 0$. Displacement over [0, 5] = $\int_0^5 (t^2 - 4t)dt = [t^3/3 - 2t^2]_0^5 = (125/3 - 50) = -25/3$ m. Distance travelled = $|\int_0^4 v\,dt| + |\int_4^5 v\,dt| = |-32/3| + |7/3| = 39/3 = 13$ m. The two answers genuinely differ.

Traps. Reporting displacement when the question said "distance travelled" (or vice versa) — leaving the answer unsigned/signed wrongly. Setting a = 0 to find when the particle is at rest (that finds where velocity is *stationary*, not zero). Forgetting to split the integral at the sign change, so the negative and positive parts cancel and the distance comes out too small. Mislabelling the ladder — differentiating displacement should give velocity, not acceleration.

5.12 Integration — antiderivatives and the power rule

Key facts / formulas. Integration reverses differentiation. The **power rule for integration** is in the booklet:

$\int x^n\,dx = x^{n+1}/(n + 1) + C$, for $n \ne -1$.

The **+ C** is part of the answer for every *indefinite* integral — it is not optional and not decorative. The $n \ne -1$ exclusion matters: $\int x^{-1}\,dx$ is the special case $\ln|x| + C$, handled in 5.13. Constant multiples pull out and sums integrate term by term, exactly as for differentiation.

Method.

1. Write every term as a power of x (same rewriting as 5.4).
2. For each term, add one to the exponent and divide by the new exponent.
3. Add + C once, at the end.
4. If an initial condition is given (a point on the curve, or F(some value) = something), substitute it to find C.

Worked micro-example. $\int(6x^2 - 4x + 5)dx = 6\cdot x^3/3 - 4\cdot x^2/2 + 5x + C = 2x^3 - 2x^2 + 5x + C$. If additionally the curve passes through (0, 7), then C = 7.

Traps. Omitting the + C — a guaranteed lost mark on an indefinite integral. Dividing by the *old* exponent instead of the new one. Applying the power rule to 1/x — that is the n = −1 case and the rule does not apply. Forgetting to divide by the new exponent at all (just raising the power).

5.13 Integrals of further functions and the linear-composition rule

Key facts / formulas. From the booklet:

- $\int \sin x \, dx = -\cos x + C$
- $\int \cos x \, dx = \sin x + C$
- $\int e^x \, dx = e^x + C$
- $\int (1/x) \, dx = \ln|x| + C$ (the $n = -1$ case; note the absolute value)

The **linear-composition rule**: when the argument is a linear expression $ax + b$, you integrate the outer function and divide by a. So $\int f(ax + b)dx = (1/a)F(ax + b) + C$. Concretely: $\int \cos(3x)dx = (1/3)\sin(3x) + C$, $\int e^{(2x)}dx = (1/2)e^{(2x)} + C$, $\int (2x + 1)^4 \, dx = (2x + 1)^5/(5\cdot 2) + C$. This "divide by a" is the integration counterpart of the chain-rule inner factor — and it only works because the inside is *linear*.

Method.

1. Identify the outer standard integral.
2. Identify the linear inside, $ax + b$, and read off a.
3. Apply the outer integral, then divide the whole thing by a. Add + C.

Worked micro-example. $\int (4\cos(2x) + e^{(-x)})dx = 4\cdot(1/2)\sin(2x) + (1/-1)e^{(-x)} + C = 2\sin(2x) - e^{(-x)} + C$.

Traps. Forgetting to divide by a — writing $\int \cos(3x)dx = \sin(3x) + C$ (missing the 1/3). Getting the sign wrong on $\int \sin x \, dx$ (it is *minus* cos). Multiplying by a instead of dividing. Trying the linear-composition rule when the inside is *not* linear (e.g. $\int \cos(x^2)dx$ — this is not an SL by-hand integral). Dropping the absolute value in $\ln|x|$.

5.14 Integration by substitution

Key facts. Substitution handles the by-hand non-calculator cases where an integrand is a composite times (a multiple of) the derivative of its inside. Let u = (the inside), compute $du = (du/dx)dx$, rewrite the whole integral in u, integrate, then substitute back. The recognisable pattern is $\int f'(x)\cdot g(f(x))dx$ — for instance $\int 2x\cdot(x^2 + 1)^5 \, dx$, where 2x is exactly the derivative of $x^2 + 1$.

Method.

1. Choose u = the inner function. Compute du/dx and hence du.
2. Rewrite the integral *entirely* in terms of u — every x must disappear, including via the dx.
3. Integrate with respect to u.
4. Substitute u back to x. Add + C (for an indefinite integral). For a definite integral, either substitute back first or convert the limits to u-values.

Worked micro-example. $\int 2x(x^2 + 1)^5\,dx$. Let $u = x^2 + 1$, so $du = 2x\,dx$. The integral becomes $\int u^5\,du = u^6/6 + C = (x^2 + 1)^6/6 + C$. Check by differentiating: $(6(x^2 + 1)^5 \cdot 2x)/6 = 2x(x^2 + 1)^5$. ✓

Traps. Choosing u so that the leftover x's do not all cancel — if du does not "mop up" the rest of the integrand, the substitution does not work cleanly and you have likely picked the wrong u. Forgetting to convert dx into du terms. For a definite integral, changing the variable to u but leaving the limits as the old x-values. Forgetting to substitute back to x at the end of an indefinite integral.

5.15 The definite integral

Key facts / formulas. The **fundamental relationship**: if F is an antiderivative of f, then

$\int$ from a to b of $f(x)\,dx = F(b) - F(a)$.

This is in the booklet. A definite integral evaluates to a *number* — there is no + C, because it cancels in the subtraction. The standard layout is $\int_a^b f\,dx = [F(x)]_a^b = F(b) - F(a)$.

Method.

1. Find an antiderivative F (no + C needed).
2. Write it in square brackets with the limits.
3. Substitute the upper limit, then the lower limit, and subtract: $F(b) - F(a)$.
4. Evaluate to an exact number.

Worked micro-example. $\int_1^3 (2x)dx = [x^2]_1^3 = 9 - 1 = 8$.

Traps. Doing $F(a) - F(b)$ — limits the wrong way round flips the sign. Carrying a + C into a definite integral (harmless but signals confusion, and wastes time). Substituting the lower limit into only some of the terms. Arithmetic slips in the subtraction, especially with negative F(a).

5.16 Area under a curve and area between curves

Key facts / formulas. Both formulas are in the booklet:

- Area between a curve and the x-axis, from a to b: $\int_a^b y\,dx$ — *provided $y \geq 0$ across the interval*.
- Area between two curves, $y = f(x)$ above $y = g(x)$ on $[a, b]$: $\int_a^b (f(x) - g(x))\,dx$ — upper minus lower.

The **sign issue**: where a curve dips *below* the x-axis, $\int y\,dx$ is negative. If the question asks for an *area* (which is always positive), you must split the integral at the x-intercepts, integrate each piece, and add the *absolute values*. A single $\int$ across a sign change gives the *net signed* value, where positive and negative regions cancel — that is not the area.

Method (area, curve crosses the axis).

1. Find where the curve meets the x-axis: solve $y = 0$ for the x-intercepts in the interval.
2. Integrate over each sub-interval separately.
3. Take the absolute value of each result and add them.

Method (area between two curves).

1. Find the intersection points: solve $f(x) = g(x)$ for the limits.
2. Decide which function is on top in that interval (test a point if unsure).
3. Integrate (top – bottom) between the intersection x-values.

Worked micro-example (axis crossing). Area between $y = x^2 - 1$ and the x-axis, from $x = 0$ to $x = 2$. Intercept at $x = 1$. On $[0, 1]$, $y \le 0$; on $[1, 2]$, $y \ge 0$. $\int_0^1 (x^2 - 1)dx = [x^3/3 - x]_0^1 = (1/3 - 1) = -2/3$. $\int_1^2 (x^2 - 1)dx = [x^3/3 - x]_1^2 = (8/3 - 2) - (1/3 - 1) = 2/3 - (-2/3) = 4/3$. Area $= |-2/3| + |4/3| = 2/3 + 4/3 = 2$.

Worked micro-example (between curves). Area enclosed by $y = x$ and $y = x^2$. They meet where $x = x^2$, i.e. $x = 0$ and $x = 1$; on $(0, 1)$ the line $y = x$ is above. Area $= \int_0^1 (x - x^2)dx = [x^2/2 - x^3/3]_0^1 = 1/2 - 1/3 = 1/6$.

Traps. Computing one ∫ straight across an axis-crossing and reporting a too-small (or negative) "area" — the defining area trap. Forgetting to find the intercepts/intersections that set the limits. Doing (bottom – top) and getting a negative area between curves. Not checking which curve is on top. Reporting a negative number as an area.

If you remember nothing else

- **The chain rule lives inside almost every product and quotient.** When you differentiate u or v and it is composite — $\sin(3x)$, e^(2x), $(2x+1)^5$, $\ln(5x)$ — write the inner-derivative factor. The most common calculus error in P1 is the missing inner derivative.
- **Classify stationary points by writing the f″ value and its sign**, then the conclusion. $f'' < 0$ is a maximum, $f'' > 0$ a minimum. "It's a max" with no f″ shown is a lost reasoning mark. If $f'' = 0$, fall back to a first-derivative sign check.
- **Normal gradient = negative reciprocal of the tangent gradient**, $-1/f'(a)$. And every tangent/normal needs an actual point, not just a gradient.
- **+ C is not optional** on any indefinite integral, and it is *absent* from every definite integral. A definite integral is a number: $F(b) - F(a)$, upper minus lower.
- **Optimisation is four steps** — eliminate a variable using the constraint, differentiate and set to zero, *justify* the max/min, then substitute back into the *original* to get the optimum value. Answer with the value, not the input. Check domain endpoints.
- **Area is always positive.** Where the curve crosses the x-axis, split at the intercepts and add absolute values; a single integral across a sign change gives net signed value, not area. Between curves: top minus bottom.
- **Kinematics is a derivative ladder** — displacement → velocity → acceleration by differentiating, and back by integrating. "Momentarily at rest" means $v = 0$, not $a = 0$. Distance travelled ≠ displacement when v changes sign — split at $v = 0$ and add magnitudes.
- **Keep the two log derivatives apart**: $d/dx(\ln x) = 1/x$, but $d/dx(\log_a x) = 1/(x \ln a)$. And rewrite every negative/rational power as a power of x *before* applying the rule — that rewriting step is where the sign errors hide.