OCR A A-level Maths Notes

James Chen

2026-02-15

This document provides my full notes for OCR A A-level Mathematics (for the 2026 sitting).
The content covered is directly based on the specification.
This document does not distinguish between AS and A-level.
Anything you do not need to know but may be useful and/or interesting to know is written in red.
Examples are in orange, but the main priority of this document is focusing on the content.
Exam technique strategies are covered throughout, written in blue.

Contents

1 General Notes

1.1 Exam Structure

The exam consists of three papers, each of which being mandatory: Pure Mathematics, Pure Mathematics & Statistics, Pure Mathematics & Mechanics.
Each paper consists of 100 marks.
The Pure Mathematics paper consists of 100 marks of pure questions.
The Pure Mathematics & Statistics paper consists of 50 marks of pure questions, then 50 marks of statistics questions.
The Pure Mathematics & Mechanics paper consists of 50 marks of pure questions, then 50 marks of mechanics questions.

1.2 Command Words

- Exact: unrounded
- Prove: formal mathematical argument in detail, including a concise conclusion
- Show that: sufficiently detailed explanation to cover every step of working
- Determine: justification and working needed
- Verify: a clear substitution of the given value needed
- Find/Solve/Calculate: working may be necessary but no justification is required
- Give/State/Write down: no working nor justification is required
- In this question you must show detailed reasoning: detailed and complete analytical method which allows the argument to be followed (not restricting use of a calculator)
- Hence: next step should be based on what has gone before
- Hence or otherwise: using information from the previous parts may be useful but not necessary, alternate methods may be more time-consuming or complex
- You may use the result: a result that you don't need to know can be used
- Plot: mark points and potentially join with lines
- Sketch: a diagram showing the key features, like turning points and asymptotes
- Draw: draw to a sensible level of accuracy

1.3 What to do when stuck

- Could it be a certain key fact or detail you are missing?
- Scan the formula booklet to see if there any relevant formulae you have forgotten
- Look at other questions in the paper to see if there is anything related
- See if you can obtain the answer immediately using your calculator, and then let the answer guide you
- Scattergun to obtain as many marks as possible
- If you need to show that something is true, work backwards from the final statement to see what you could do to get there in the other direction
- Harvest as many marks as possible from the other parts of the question using what you know (especially if you're stuck on a show that)

1.4 Checking Procedures

Some checking procedures are specific to the content, so this is a list of generic ones:

- Significant figures
- +c
- Simplifying fractions
- Division by 0
- Correct reversal of inequality signs
- Substitute answer into original question
- Sanity checks (do the answers make sense?) (how many solutions should I have?)
- Information given in question (e.g. domain restrictions)
- Input into calculator correctly

2 Proof

There are several methods of proof (or disproof) in the specification:

- Proof by deduction: proceeding from known truths to the conclusion via logical steps

- Proof by exhaustion: checking every possible case The possible cases may be given as a range, e.g. prove something is true for all numbers between 2 and 10. Sometimes, the exhaustion may not be explicit, such as proving something is true for all natural numbers by proving it is true for odd numbers and even numbers separately.

- Disproof by counterexample: giving at least one counterexample to a claim in order to disprove it

- Proof by contradiction: assuming some statement is true and reaching a logical contradiction, implying the statement must have been false

The logical connective $\equiv$ refers to congruence, i.e. identities, and implies the relation is always true (not just for certain cases or values).
The logical connective $\Rightarrow$ refers to implication: $A \Rightarrow B$ means "if A is true then B is true". This does not necessarily mean that if B is true, then A is true.
The logical connective $\iff$, known as 'iff' or 'if and only if', refers to implication in both directions: $A \iff B$ means "if A is true then B is true" and "if B is true then A is true".

Integers are whole numbers and real numbers are any number on the number line.
Rational numbers can be written in the form $\frac{p}{q}$, where p, q are coprime integers (numbers with no common factors apart from 1), and $q \neq 0$. Irrational numbers cannot be written in this form.

There are two standard proofs by contradiction that you have to know about: the proof of the irrationality of $\sqrt{2}$, and the infinitude of the primes.

Assume that $\sqrt{2}$ is rational, so it can be written in the form $\frac{p}{q}$, where p, q are coprime integers (numbers with no common factors apart from 1), and $q \neq 0$.
$\sqrt{2} = \frac{p}{q} \Rightarrow 2 = \frac{p^2}{q^2} \rightarrow 2q^2 = p^2$
Since p^2 is even, p is even, so p can be written as $2k$ for some integer k.
Therefore, $2q^2 = (2k)^2 = 4k^2 \Rightarrow 2k^2 = q^2$, so q is even, but p is also even, so p and q are not coprime, causing a contradiction, so our assumption was false, and therefore $\sqrt{2}$ is irrational.

Assume there are finitely many primes. Let's assume there are n primes, which are $p_1, p_2, p_3, ..., p_n$. Then, let $P = p_1 p_2 p_3 ... p_n + 1$. P is either a prime or a product of primes (as every number is a product of primes). Dividing P by each of the primes in our set gives a remainder of 1, so P is not perfectly divisible by any prime in our set. P is therefore either a prime larger than the primes in our set or a number with at least one prime factor which is not in our set. In both cases, we have found a prime not in our set, contradiction our assumption that we have found all the primes, so there must be infinitely many prime numbers.

Many people use the symbol # to represent a contradiction, but this is not widely recognised, and may risk marks lost if the examiner is not familiar with it. You should always explicitly say that a contradiction has been reached.

3 Algebra and Functions

3.1 Indices

The laws of indices are:

- $x^a x^b = x^{a+b}$
- $\frac{x^a}{x^b} = x^{a-b}$
- $(x^a)^b = x^{ab}$
- $x^{-a} = \frac{1}{x^a}$
- $x^{\frac{m}{n}} = \sqrt[n]{x^m}$
- $x^0 = 1$

3.2 Surds

Fractions of the form $\frac{a}{\sqrt{b}}$ can be rationalised by multiplying by $\frac{\sqrt{b}}{\sqrt{b}}$. Fractions of the form $\frac{a}{b+\sqrt{c}}$ can be rationalised by multiplying by $\frac{b-\sqrt{c}}{b-\sqrt{c}}$.
It is usually best to rationalise the denominator of any fractions with irrational denominators, though questions may not always penalise not doing so.

3.3 Simultaneous equations

Variables can be eliminated from simultaneous equations by adding or subtracting them (typically after scaling one or both of the equations). Rearranging to obtain an expression for one of the variables can allow for substitution into the other equation.

Consider $x + y = 1, x^2 + y^2 = 1$. We can solve as follows:
$y = 1 - x \Rightarrow x^2 + (1-x)^2 = 1 \Rightarrow x^2 + 1 - 2x + x^2 - 1 = 0 \Rightarrow 2x^2 - 2x = 0 \Rightarrow x(x-1) = 0 \Rightarrow x = 0, x = 1$
Plugging these values into the first equation, we get the points $(0, 1)$ and $(1, 0)$.

If the question asks for solutions or points of intersections, never forget to plug your x solutions back in to one of the equations to get the corresponding y solutions.

3.4 Quadratic functions

For a quadratic $ax^2 + bx + c = 0$, the discriminant is $\Delta = b^2 - 4ac$. The notation D can also be used. If $b^2 - 4ac > 0$, there are two real distinct roots. If $b^2 - 4ac = 0$, there is one repeated real root. If $b^2 - 4ac < 0$, the roots are not real.
You should know how to complete the square. Reminder that for a quadratic written in the form $a(x + p)^2 + q$, the turning point is at $(-p, q)$.

We may not necessarily have quadratic functions in x. We could replace x with anything! For example, $x^4 - x^2 - 1$ is a quadratic in x^2, $e^{2x} - e^x - 1$ is a quadratic in e^x, and $\sin^2 x - \sin x - 1$ is a quadratic in $\sin x$. Using the third as an example, we have $\sin x = \frac{1 \pm \sqrt{1^2 - 4(-1)}}{2} = \frac{1 \pm \sqrt{5}}{2}$ but remember that $-1 \leq \sin x \leq 1$ so we only have the solution $x = \arcsin(\frac{1-\sqrt{5}}{2})$ as $\frac{1+\sqrt{5}}{2} > 1$.
To make things easier, it may be worth using a different letter like y to represent the thing the quadratic is in, then substituting x back at the end.

For your interest, here is a derivation of the quadratic formula:
Complete the square on $ax^2 + bx + c = 0$.
$ax^2 + bx + c = a(x^2 + \frac{b}{a}x) + c = a((x + \frac{b}{2a})^2 - \frac{b^2}{4a^2}) + c = a(x + \frac{b}{2a})^2 - \frac{b^2}{4a} + c$
$4a^2(x + \frac{b}{2a})^2 = b^2 - 4ac$
$2a(x + \frac{b}{2a}) = \pm\sqrt{b^2 - 4ac}$
$x + \frac{b}{2a} = \frac{\pm\sqrt{b^2-4ac}}{2a}$
$x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$

and here is an explanation of the discriminant rules:
The two solutions come from the $\pm$ sign.
Notice the $\sqrt{b^2 - 4ac}$ which is 'undefined' if $b^2 - 4ac$ is negative and defined if it is positive. If it is 0, plus or minus 0 are the same, giving one repeated solution.

3.5 Inequalities

You should know how to solve linear and quadratic inequalities. Reminder that to solve quadratic inequalities, you find the critical values (the roots when the inequality sign is replaced with an equals sign), and sketch a graph of the quadratic to identify the region you need.
When plotting inequalities on graphs, we tend to shade the accepted region, and use solid lines for non-strict inequalities and dotted lines for strict inequalities.

There are two types of notation for intervals: interval notation and set notation.

The interval $a < x < b$ is represented in interval notation by (a, b). We use a round bracket for strict inequalities and a square bracket for non-strict inequalities. The interval $a \leq x < b$, for example, is represented by $[a, b)$. If either of the limits are infinity, we tend to use round brackets.

Sets are denoted with curly brackets: { }. We start a set with $\{x :$ which means "x such that" and then write our condition, for example $\{x : 2 \leq x < 3\}$ means the set of all values of x such that x is greater than or equal to 2 and less than 3. We use $\cup$ in between two sets to represent 'or' and $\cap$ to represent 'and'.

3.6 Polynomials

You should know how to expand brackets, collect like terms, and factorise.
You should know what quadratics, cubics, and parabola are.

The factor theorem states that $f(a) = 0$ if and only if $(x - a)$ is a factor of $f(x)$. Also, $f(\frac{b}{a}) = 0$ if and only if $(ax - b)$ is a factor of $f(x)$.
Sometimes, a factor of a polynomial has to be determined by inspection before you can use polynomial division.

You should be able to simplify rational expressions and perform polynomial division.

For your interest, here is a derivation of the factor and remainder theorems:

Dividing anything by anything leaves a quotient and remainder.
This also applies to functions.
Let $f(x) = (x - a)q(x) + r(x)$, where $q(x)$ is the quotient and $r(x)$ is the remainder (this is like saying $67 = 2 \times 33 + 1$).
Substitute $x = a$: $f(a) = (a - a)q(x) + r(x) = r(x)$.
Therefore if $f(a) = r$ then r is the remainder of $f(x)$ divided by $(x - a)$.
Considering $r = 0$ gives the factor theorem.

3.7 The modulus function

The modulus function $|x|$ is equal to x if $x \geq 0$ and is equal to $-x$ if $x < 0$, i.e. it turns x positive no matter its sign (gives the magnitude of x).
For example, $|5| = 5, |-5| = 5, |x - 5| = 3 \Rightarrow x - 5 = 3$ or $x - 5 = -3$.
Many questions can be done by splitting into the different cases depending on whether whatever is in the modulus signs is positive or negative.

3.8 Curve sketching

You should be able to sketch curves of polynomials up to quartics and curves of the form $y = \frac{a}{x}$ or $y = \frac{a}{x^2}$.
To sketch the graph of a function which includes modulus functions, you have to reflect any portions of the graph under the x axis about the x axis.

3.9 Functions

The definition of a function is a mapping from the domain to the range such that for each x in the domain, there is a unique y in the range with $f(x) = y$.
The domain is the set of all possible inputs of the function, and the range is the set of all possible outputs of the function.
The range must be given in terms of $f(x)$ (or f), not x. For example, the range of $f(x) = x^2$ is $f(x) \geq 0$, not $x \geq 0$.
The inverse function of a function $f(x)$ is the function $f^{-1}(x)$ such that $f^{-1}(f(x)) = x$. The inverse function is the reflection of the function in the line $y = x$.
We write fg to refer to $f(g(x))$, i.e. putting $g(x)$ into $f(x)$ to form a composite function.

3.10 Graph transformations

If $y = f(x)$ then $y = f(x + a)$ translates by a units to the left, $y = f(x) + a$ translates by a units up, $y = f(ax)$ stretches parallel to the x axis by scale factor $\frac{1}{a}$, and $y = af(x)$ stretches parallel to the y axis by scale factor a.
When transformations are combined, remember to apply these rules to every instance of x in the function.

3.11 Partial fractions

We can decompose rational fractions into partial fractions.
$\frac{1}{(ax+b)(cx+d)(ex+f)}$ can be written as $\frac{A}{ax+b} + \frac{B}{cx+d} + \frac{C}{ex+f}$ for constants A, B, C.
$\frac{1}{(ax+b)(cx+d)^2}$ can be written as $\frac{A}{ax+b} + \frac{B}{cx+d} + \frac{C}{(cx+d)^2}$

To determine the values of A, B, C, we multiply everything by the denominator of the left hand side, and then substitute any values of x we want as the identity has to hold for all values of x. We tend to choose convenient values of x which makes some of the terms on the right hand side go away. We can obtain simultaneous equations for the constants.

Consider the example $\frac{1}{(1+x)(1-x)^2}$. We wish to decompose this rational function into the partial fractions: $\frac{1}{(1+x)(1-x)^2} = \frac{A}{1+x} + \frac{B}{1-x} + \frac{C}{(1-x)^2}$. Multiplying through by $(1 + x)(1 - x)^2$, we get $1 = A(1 - x)^2 + B(1 + x)(1 - x) + C(1 + x)$. This identity needs to hold for every single value of x, so we can choose any values we want to determine the values of A, B, C. If we choose $x = 1$, we get $1 = 2C \Rightarrow C = \frac{1}{2}$. If we choose $x = -1$, we get $1 = 4A \Rightarrow A = \frac{1}{4}$. Choosing $x = 0$ gives $1 = A + B + C = \frac{1}{4} + B + \frac{1}{2} \Rightarrow B = \frac{1}{4}$. Now that we have each value of A, B, C, we have $\frac{1}{(1+x)(1-x)^2} = \frac{1}{4(1+x)} + \frac{1}{4(1-x)} + \frac{1}{2(1-x)^2}$.

3.12 Models in context

What it says on the tin.

4 Coordinate Geometry in the x-y Plane

4.1 Straight lines

The equation of a line can be written in the forms $y = mx + c, y - y_1 = m(x - x_1), ax + by + c = 0$.
The midpoint of the line segment between the points (x_1, y_1) and (x_2, y_2) is $(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2})$. The distance between these points is $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$.
For two lines with gradients m_1, m_1, they are parallel if $m_1 = m_2$ and they are perpendicular if $m_1 m_2 = -1$.

4.2 Circles

The equation of a circle is $(x - a)^2 + (y - b)^2 = r^2$ where the center is (a, b) and the radius is r.
Some questions give the equation of a circle in the form $ax^2 + by^2 + cx + dy + e = 0$. This can be converted into the standard equation by completing the square on both the x terms and the y terms.

The following circle theorems should be known:

- The angle in a semicircle is a right angle
- The perpendicular from the centre of a circle to a chord bisects the chord
- The radius of a circle at a given point on its circumference is perpendicular to the tangent to the circle at the point

4.3 Parametric equations of curves

We may define a curve by writing x and y in terms of some other parameter, usually t or θ. To convert from parametric to cartesian, we need to find a way to eliminate the parameter, sometimes using known identities (e.g. trig identities).

The most famous example is the parameterisation of a circle using $x = r\cos\theta, y = r\sin\theta$. This can be converted back into cartesian form by squaring and adding: $x^2 + y^2 = r^2\cos^2\theta + r^2\sin^2\theta = r^2(\cos^2\theta + \sin^2\theta) = r^2$ which is exactly the equation of a circle.

Another example of converting from parametric to cartesian:
$x = 1 - \cos t, y = \sin t \sin 2t$. We have $y = 2\sin^2 t\cos t = 2(1 - \cos^2 t)\cos t$ and $\cos t = 1 - x$, hence $y = 2(1 - (1 - x)^2)(1 - x)$ which can be simplified.

4.4 Parametric equations in context

What it says on the tin.

5 Sequences and Series

5.1 Binomial expansion

If n is a natural number, then:
$(a+bx)^n = a^n + \binom{n}{1}bx + \binom{n}{2}(bx)^2 + ... + \binom{n}{r}a^r(bx)^{n-r} + ... + b^n$, which is in the formula booklet.
$n!$, or n factorial, is the product of every natural number from 1 to n, with 0! defined as 1.
${}^nC_r$ or $\binom{n}{r}$ is the number of ways to choose r objects from n objects (without order mattering), and ${}^nC_r = \frac{n!}{r!(n-r)!}$, which is in the formula booklet.
The value of $\binom{n}{r}$ is the rth entry in the nth row of Pascal's triangle.

We can extend the binomial expansion to any rational n using the formula $(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + ...$, which is in the formula booklet. If the term which isn't x is not 1, we can divide the entire expression by some factor to make it 1, e.g. $(2+x)^{0.5} = 2^{0.5}(1+\frac{x}{2})^{0.5}$ and we can proceed using $0.5x$ instead of x.
The range of validity is $|x| < 1$, with x replaced with whatever you used in the expansion.

These expansions can be used for approximations. If x is small, the terms with high powers can be neglected as they are close to 0. We can substitute particular values of x into these expansions to obtain these approximations.

Here is an intuition for the binomial expansion for integer powers:

Consider $(x+y)^4 = (x+y)(x+y)(x+y)(x+y)$.
For the x^4 term, we simply multiply all four x terms (same for y^4.
For the x^3 term, we multiply any three x terms, which leaves room for one more y term. We can choose any three of the x terms, so we need to multiply by $\binom{4}{3}$ which is the same as $\binom{4}{1}$. This gives $4x^3y$
For the x^2 term, we multiply any two x terms, giving $\binom{4}{2}x^2y^2 = 6x^2y^2$.

and here is an intuition for the extended binomial expansion to rational powers:

We know that $(1+x)^n = \sum_{r=0}^{n} \binom{n}{r}x^n$
We could say that $\binom{n}{r}$ is 0 whenever r is greater than n, as it is impossible to choose a larger number of objects from a smaller number of objects.
Therefore, $(1+x)^n = \sum_{r=0}^{\infty} \binom{n}{r}x^n$
We have $\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n(n-1)(n-2)...(n-r+1)}{r!}$
We now have a formula for $\binom{n}{r}$ that does not depend on integer values of n. We just need to also say $\binom{n}{0} = 1$.
We can now expand.

Since we are adding an infinite series whenever n is not an integer, this is analogous to the sum of a GP formula, which indicates the range of validity $|x| < 1$ (which has to be adapted depending on the expansion).

5.2 Sequences

Sequences can be given as a formula for the nth term or a recursion relation of the form $x_{n+1} = f(x_n)$.
Increasing sequences are where every term is greater than or equal to the last.
Decreasing sequences are where every term is less than or equal to the last.
If the terms of a sequence repeat after certain intervals, it is periodic, and the number of terms between each repeat is the period.
A series represents the sum of a certain number of terms of a sequence.
A finite sequence has a finite number of terms and an infinite sequence has an infinite number of terms.

5.3 Sigma notation

The notation $\sum_{r=a}^{b} f(r)$ represents the sum of the values of $f(r)$ from a to b inclusive (every integer in this range).

5.4 Arithmetic sequences

An arithmetic sequence is where the terms increase with a constant difference. If we denote the first term as a and the common difference as d, then the nth term is $a + d(n-1)$, and the sum of the first n terms is $\frac{n}{2}(a+l)$ or $\frac{n}{2}(2a + d(n-1))$, and these sum formulae are in the formula booklet.

Proof:
Consider $a, a+d, a+2d, ..., a+(n-1)d$.
To sum these terms, we could do some clever pairing.
$S_n = (a + a + (n-1)d) + (a + d + a + (n-2)d) + (a + 2d + a + (n-3d)) + ...$
$= (2a + (n-1)d) + (2a + (n-1)d) + (2a + (n-1)d) + ...$
Aka we sum the first and last term, second and second last, etc.
There are n terms so $\frac{n}{2}$ pairs, giving our sum as $S_n = \frac{n}{2}(2a + (n-1)d)$

5.5 Geometric sequences

A geometric sequence is where the terms increase with a common ratio. If we denote the first term as a and the common ratio as r, then the nth term is ar^{n-1} and the sum of the first n terms is $S_n = \frac{a(1-r^n)}{1-r}$, and this sum formula is in the formula booklet. As long as $|r| < 1$, the sum to infinity is $S_\infty = \frac{a}{1-r}$, which is also in the formula booklet. $|r| < 1$ guarantees that the series converges (ends up on a specific value), whereas $|r| > 1$ means the sequence diverges (does not end up on a specific value).

Proof:
Consider $a, ar, ar^2, ..., ar^{n-1}$
$S_n = a + ar + ar^2 + ... + ar^{n-1}$
We multiply each term by r
$rS_n = ar + ar^2 + ar^3 + ... + ar^n = S_n + ar^n - a$
$S_n(r-1) = a(r^n - 1)$
$S_n = \frac{a(1-r^n)}{1-r}$
As n tends to infinity, $S_n \to \frac{a(1-r^\infty)}{1-r}$
If $|r| > 1$ then the progression is clearly divergent
If $|r| < 1$ then the progression is convergent
$|r| < 1 \Rightarrow \lim_{n\to\infty} r^n = 0$
Therefore $S_\infty = \frac{a}{1-r}$

5.6 Modelling

What it says on the tin.

6 Trigonometry

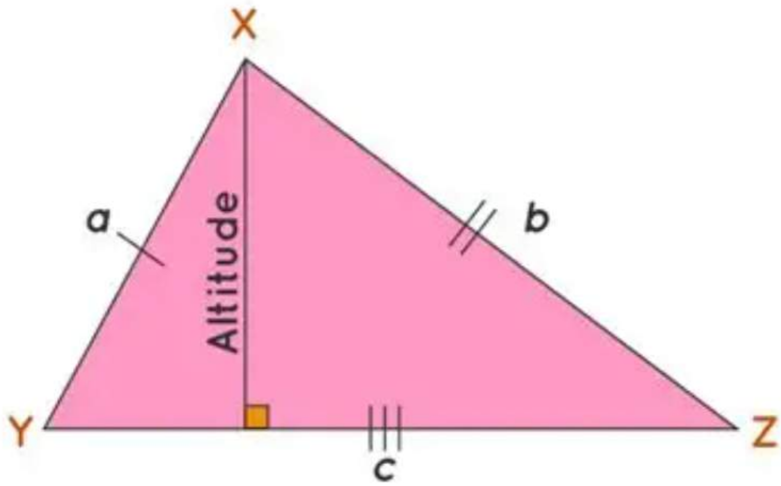
6.1 sin, cos and tan for all arguments

You should know the definitions of sin, cos, and tan.

6.2 Sine and cosine rules

You should know the sine and cosine rules.

Proof of the sine rule:
Consider a triangle and drop an altitude (a perpendicular height).



Let's call this altitude h. Clearly, $h = a \sin Y = b \sin Z$.
Therefore, $a \sin Y = b \sin Z \Rightarrow \frac{a}{\sin Z} = \frac{b}{\sin Y}$
The rest of the sine rule can be derived similarly using the other altitudes.

Proof of the cosine rule:
For the same triangle as above, let O be the intersection of the altitude and the line YZ, i.e. the point directly below X on the base. Now let $OY = c_1, OZ = c_2$ such that we have $c_1 + c_2 = c$.
$c_1 = a \cos Y, h = a \sin Y$
$c_2 = c - a \cos Y$
$c_2^2 + h^2 = b^2 \Rightarrow (c - a \cos Y)^2 + a^2 \sin^2 Y = b^2$
$a^2 \sin^2 Y + a^2 \cos^2 Y - 2ac \cos Y + c^2 = b^2$
$b^2 = a^2 + c^2 - 2ac \cos Y$
Which can be transformed into the more familiar form by relabelling

The area of a triangle is given by $A = \frac{1}{2} ab \sin C$, where C is the included angle between sides a, b.

Proof:
We know $A = \frac{1}{2} bh$.
In the triangle above, $h = a \sin Y$, so $A = \frac{1}{2} ac \sin Y$.
This can be relabelled as desired.

Because $\sin x = a$ has two solutions in the range $0 < x < 180$ for $0 < a < 1$, there are often two possible triangles for certain given conditions, leading to an ambiguous case.

6.3 Radians

Multiplying an angle in degrees by $\frac{\pi}{180}$ converts it to radians (and vice versa). If θ is the angle subtended by a sector of a circle with radius r, the arc length is $r\theta$ and the area is $\frac{1}{2}r^2\theta$.

6.4 Small angle approximations

When x is small and measured in radians, $\sin x \approx x$, $\cos x \approx 1 - \frac{1}{2}x^2$, $\tan x \approx x$. These are given in the formula booklet.

There are several ways of deriving the small angle approximations that we use in OCR. I will give a brief overview.
Note that some content from Further Maths is used.

We know that $\lim_{x\to 0} \frac{\sin x}{x} = 1$, which can be proven as follows:
A result is that $\sin x \leq x \leq \tan x$.
We can divide through by $\sin x$, which is positive when x is small and positive, so the inequality sign is unchanged.
$1 \leq \frac{x}{\sin x} \leq \frac{1}{\cos x}$
Taking the limit as x tends to 0, $1 \leq \frac{x}{\sin x} \leq 1$.
By the Squeeze Theorem, $\lim_{x\to 0} \frac{\sin x}{x} = 1$.
Therefore, $\sin x \approx x$ for small x.
Note: we could have used L'Hôpital, but this involves differentiating $\sin x$, which requires the use of the limit we are trying to find to differentiate $\sin x$ by first principles, so this method is technically circular reasoning.

We know, using Maclaurin Series, that:
$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + ...$
$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + ...$
When x is small, the terms with higher powers are very close to 0, so can be neglected, therefore $\sin x \approx x, \cos x \approx 1 - \frac{x^2}{2}$.

Let's consider $\tan x$ in a different way. Consider a sector of a circle with a really small angle θ.
The arc length is $r\theta$.
When θ is small, the sector looks similar to a right-angled triangle.
The opposite side is therefore approximately $r \tan \theta$.
Since the arc length and opposite side are approximately equal, dividing by r gives $\tan \theta \approx \theta$.

6.5 Graphs of the basic trigonometric functions

You should the know the graphs, symmetries, and periodicities of sin, cos, and tan.

6.6 Exact values of trigonometric functions

A new thing is recognising the exact values for angles given in radians, which may take some practice.

6.7 Inverse and reciprocal trigonometric functions

We define $\csc\theta = \frac{1}{\sin\theta}, \sec\theta = \frac{1}{\cos\theta}, \cot\theta = \frac{1}{\tan\theta}$. We also define $\arcsin\theta, \arccos\theta, \arctan\theta$ as the inverse functions of $\sin\theta, \cos\theta, \tan\theta$ respectively, giving principal values between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$. The domains of these inverse functions are the ranges of their respective trig functions. Their graphs are the reflections of their respective trig functions in the line $y = x$ (for the appropriate domain).
The reciprocal trig functions have the same domain as their respective trig functions except for where the denominator is 0. The ranges of $\csc x$ and $\sec x$ is all reals apart from between -1 and 1. The range of $\cot x$ is all reals.

6.8 Trigonometric identities

$\tan\theta = \frac{\sin\theta}{\cos\theta}$
$\sin^2\theta + \cos^2\theta = 1$
$1 + \tan^2\theta = \sec^2\theta$
$1 + \cot^2\theta = \csc^2\theta$

Here is a quickfire list of 'proofs' of the trig identities to be known.

- $\tan x = \frac{\sin x}{\cos x}$
  Considering SOHCAHTOA, we know that $\sin x = o/h, \cos x = a/h, \tan x = o/a$, from which the identity follows.

- $\sin^2 x + \cos^2 x = 1$
  Again considering SOHCAHTOA, $\sin^2 x + \cos^2 x = \frac{o^2+a^2}{h^2}$. From Pythagoras, $o^2 + a^2 = h^2$, from which the identity follows.

- $\sec^2 x = 1 + \tan^2 x$
  Divide the second identity through by $\cos^2 x$.

- $\csc^2 x = 1 + \cot^2 x$
  Divide the second identity through by $\sin^2 x$.

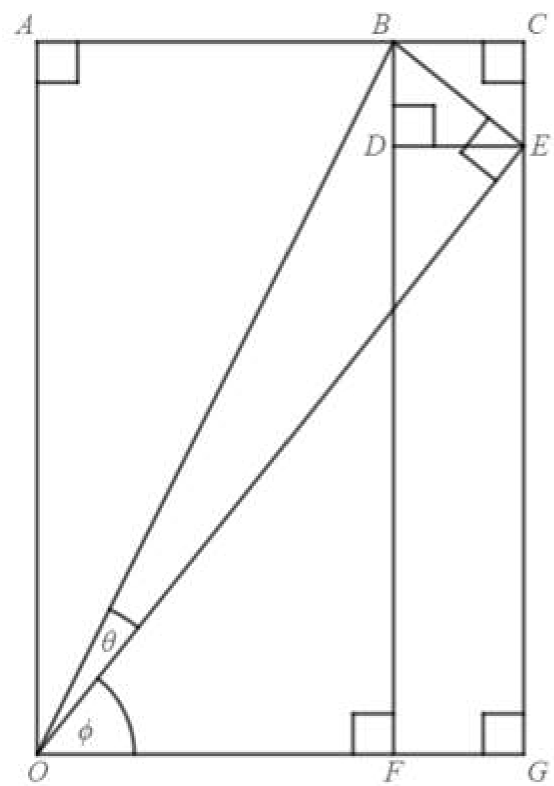
6.9 Further trigonometric identities

These are given in the formula booklet: $\sin(x+y) = \sin x \cos y + \sin y \cos x$
$\cos(x+y) = \cos x \cos y - \sin x \sin y$
$\tan(x+y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}$

Replacing y with x gives important double angle formulae (which aren't in the formula booklet):
$\sin 2x = 2 \sin x \cos x$
$\cos 2x = 2\cos^2 x - 1 = 1 - 2\sin^2 x = \cos^2 x - \sin^2 x$
$\tan 2x = \frac{2 \tan x}{1 - \tan^2 x}$

You do not need to know but need to understand the geometric proofs of the compound angle formulae:

Consider the following diagram.



Why should we draw such a diagram? We want to consider $\cos(\theta + \phi)$, and

this could be heavily related to right-angled triangles. We could therefore consider the separate right-angled triangles with angles $\theta, \phi, \theta + \phi$.
$\cos(\theta + \phi) = OF/OB$
$\cos\theta = OE/OB$
$\cos\phi = OG/OE$
$OG = OB\cos\theta\cos\phi =$
$OF = OG - FG = OG - BC = OG - BE\sin\phi = OG - OB\sin\theta\sin\phi$
$\cos(\theta + \phi) = \frac{OG-BC}{OB} = \frac{OB\cos\theta\cos\phi - OB\sin\theta\sin\phi}{OB} = \cos\theta\cos\phi - \sin\theta\sin\phi$

Now consider $\sin(\theta + \phi) = \sqrt{1 - \cos^2(\theta + \phi)} = \sqrt{1 - (\cos\theta\cos\phi - \sin\theta\sin\phi)^2}$
$= \sqrt{\sin^2\theta + \cos^2\theta - \sin^2\theta\sin^2\phi - \cos^2\theta\cos^2\phi + 2\sin\theta\sin\phi\cos\theta\cos\phi}$
$= \sqrt{\sin^2\theta(1 - \sin^2\phi) + \cos^2\theta(1 - \cos^2\phi) + 2\sin\theta\sin\phi\cos\theta\cos\phi}$
$= \sqrt{\sin^2\theta\cos^2\phi + \sin^2\phi\cos^2\theta + 2\sin\theta\sin\phi\cos\theta\cos\phi} = \sin\theta\cos\phi + \sin\phi\cos\theta$

We can divide sin by cos to get tan:
$\tan(\theta + \phi) = \frac{\sin\theta\cos\phi + \sin\phi\cos\theta}{\cos\theta\cos\phi - \sin\theta\sin\phi}$
Let's divide through by $\sin\theta\cos\phi$.
$\tan(\theta + \phi) = \frac{1 + \tan\phi\cot\theta}{\cot\theta - \tan\phi}$
And let's now multiply by $\tan\theta$.
$\tan(\theta + \phi) = \frac{\tan\theta + \tan\phi}{1 - \tan\theta\tan\phi}$

We can write $a\sin x + b\cos x$ in any of the forms $R\sin(x+\alpha), R\sin(x-\alpha), R\cos(x+\alpha), R\cos(x-\alpha)$. We have, in all cases, $R^2 = a^2 + b^2$. We can determine α by expanding the required form using a compound angle formula and by comparing coefficients of $\sin x, \cos x$ on both sides to determine the value fo $\tan\alpha$. By convention, $R > 0$ and $0 \leq \alpha \leq \frac{\pi}{2}$.
This makes graphs of $a\sin x + b\cos x$ easier to sketch, max and min values are easier to find, and equations can be solved easier.

6.10 Trigonometric equations

Nothing much to say.
The question will give a range for the solution. If this needs to be changed via a transformation of the variable, it is worth explicitly changing the required range.

6.11 Proof involving trigonometric functions

What it says on the tin.

6.12 Trigonometric functions in context

What it says on the tin.

7 Exponentials and Logarithms

7.1 Properties of the exponential function

You should be familiar with the graph of a^x when a is positive, including the example e^x.

7.2 Gradient of e^{kx}

By the chain rule, the gradient of e^{kx} is ke^{kx}, so exponential curves are proportional to their gradient, which is a common thing in the real world and it makes the exponential model suitable for many applications.

7.3 Properties of the logarithm

$\log_a x$ is the inverse function of a^x (where x and a are positive).
$a = b^c \Rightarrow c = \log_b a$
$\log_a a = 1, \log_a 1 = 0$
$\ln x$, the natural logarithm, is the inverse function of e^x and its graph is the reflection of e^x in the line $y = x$.

7.4 Laws of logarithms

$\log_a x + \log_a y = \log_a xy$
$\log_a x - \log_a y = \log_a \frac{x}{y}$
$\log_a(b^c) = c \log_a b$

Proof:
Let $a^A = x, a^B = y \Rightarrow a^{A+B} = xy \Rightarrow A + B = \log_a xy$
But $A = \log_a x, B = \log_a y, \Rightarrow \log_a xy = \log_a x + \log_a y$
We may prove $\log_a \frac{x}{y} = \log_a x - \log_a y$ similarly, or by using the power law with $\frac{1}{y} = y^{-1}$.

Let $A = \log_a x \Rightarrow a^A = x \Rightarrow a^{AB} = x^B \Rightarrow AB = \log_a x^B = B \log_a x$

I shall also prove the change of base rule even though it is not on the A-level specification.
Let $A = \log_a x, B = \log_a y, C = \log_y x, \Rightarrow a^A = x, a^B = y, y^C = x$
$a^{BC} = a^A \Rightarrow A = BC \Rightarrow \log_a x = \log_a y \log_y x \Rightarrow \log_y x = \frac{\log_a x}{\log_a y}$

7.5 Equations of exponentials

Equations of the form $a^x = b$ can be solved by taking logarithms of both sides. Any base may be used, but the base a may be most useful here.

7.6 Reduction to linear form

For relationships of the form $y = ax^n$. We may take the logarithm of both sides (with any base) to obtain a linear form. If we use the natural logarithm, we have $\ln y = \ln(ax^n) = \ln(a) + n \ln x$ (using the laws of logarithms). This means that a graph of $\ln y$ against $\ln x$ has gradient n and y-intercept $\ln(a)$.

7.7 Modelling using exponential functions

What it says on the tin.

8 Differentiation

8.1 Gradients

The derivative of $f(x)$ represents the gradient of the tangent to $y = f(x)$ at the point $(x, f(x))$. This is the rate of change of y with respect to x. If the derivative is at $x = a$, this is the limit of the gradient of a chord as x tends to a. We tend to denote the derivative of y as $\frac{dy}{dx}$ and the derivative of $f(x)$ as $f'(x)$. We represent second derivatives as $f''(x)$ and $\frac{d^2y}{dx^2}$. The second derivative is the rate of change of the gradient. If it is positive over a certain interval, the function is convex in that interval. If it is negative, the function is concave. If it is zero and the curve changes from concave to convex or vice versa, there is a point of inflection.

8.2 Differentiation from first principles

The derivative is defined as $f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$ (which is in the formula booklet). Using this formula to differentiate functions is known as differentiation by first principles. We substitute in our function of x into the formula, expand, simplify, divide by h, and let h tend to 0.

The derivatives of $\sin x$ and $\cos x$ by first principles need to be known:
$\frac{d}{dx} \sin x = \lim_{h\to 0} \frac{\sin(x+h)-\sin x}{h} = \lim_{h\to 0} \frac{\sin x \cos h+\sin h \cos x-\sin x}{h}$
$= \lim_{h\to 0}(\frac{\cos h-1}{h} \sin x + \frac{\sin h}{h} \cos x)$.
We now need to use some standard limits (which I shall leave unproven): $\lim_{h\to 0} \frac{\sin h}{h} = 1, \lim_{h\to 0} \frac{\cos h-1}{h} = 0$. This gives us $\frac{d}{dx} \sin x = \cos x$.

$\frac{d}{dx} \cos x = \lim_{h\to 0} \frac{\cos(x+h)-\cos x}{h} = \lim_{h\to 0} \frac{\cos x \cos h-\sin x \sin h-\cos x}{h}$
$= \lim_{h\to 0}(\frac{\cos h-1}{h} \cos x - \frac{\sin h}{h} \sin x) = -\sin x$, using the same standard limits as before.

The first of these two limit results can be proven by the Squeeze Theorem (see Small Angle Approximations) and the second is based on the following: $\frac{\cos h-1}{h} = \frac{-2\sin^2 \frac{h}{2}}{h} = -\sin \frac{h}{2} \frac{\sin \frac{h}{2}}{\frac{h}{2}}$.

8.3 Differentiation of standard functions

In general, $\frac{d}{dx}(f(x)+g(x)) = \frac{d}{dx}f(x) + \frac{d}{dx}g(x)$, $\frac{d}{dx}(f(x)-g(x)) = \frac{d}{dx}f(x) - \frac{d}{dx}g(x)$, $\frac{d}{dx}(af(x)) = a\frac{d}{dx}f(x)$ for a constant a.

$\frac{d}{dx}x^n = nx^{n-1}$
$\frac{d}{dx}e^x = e^x$
$\frac{d}{dx}a^x = a^x \ln a$
$\frac{d}{dx}\sin x = \cos x$
$\frac{d}{dx}\cos x = -\sin x$
$\frac{d}{dx}\tan x = \sec^2 x$ (this is in the formula booklet)
$\frac{d}{dx}\ln x = \frac{1}{x}$

8.4 Tangents, normals, stationary points, increasing and decreasing functions

Differentiating the equation of a curve and plugging in a to the gradient function gives the gradient of the tangent at the point $x = a$. Taking the negative reciprocal gives the gradient of the normal at this point. The point can be used to find the equation of the tangent or normal.
Stationary points occur when $\frac{dy}{dx} = 0$. A stationary point is a local maximum when $\frac{d^2y}{dx^2} < 0$ and a local minimum when $\frac{d^2y}{dx^2} > 0$. If $\frac{d^2y}{dx^2} = 0$ and there is a sign change in the second derivative either side close to this point, there is a point of inflection. If $\frac{dy}{dx} = 0$ is also true at that point, it is a stationary point of inflection.
A function is increasing when $\frac{dy}{dx} \geq 0$ and decreasing when $\frac{dy}{dx} \leq 0$.

8.5 Techniques of differentiation

The product rule states that $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$
The quotient rule states that $\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x)-f(x)g'(x)}{(g(x))^2}$, which is in the formula booklet.
The chain rule states that $\frac{d}{dx}f(g(x)) = \left(\frac{d}{dx}g(x)\right)\frac{d}{dg(x)}f(x)$.
We also have the following results: $\frac{dy}{dx} = 1 \div \frac{dx}{dy}$, $\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$.

Proof of the chain rule:
We need to accept a few results first:
If a function is differentiable, then it is continuous.
If $g(x)$ is differentiable and continuous, then $g(x+h) - g(x)$ tends to 0 as h tends to 0.
The limit of a product is the product of the limits

Let $y = f(g(x))$. $y' = \lim_{h\to 0} \frac{f(g(x+h))-f(g(x))}{h}$.
Let's let $\Delta g(x) = g(x+h) - g(x)$

$y' = \lim_{h\to 0} \frac{f(\Delta g(x)+g(x))-f(g(x))}{h} = \lim_{h\to 0}(\frac{f(\Delta g(x)+g(x))-f(g(x))}{\Delta g(x)} \times \frac{\Delta g(x)}{h}) = \lim_{h\to 0} \frac{f(\Delta g(x)+g(x))-f(g(x))}{\Delta g(x)} \times \lim_{h\to 0} \frac{\Delta g(x)}{h} = \lim_{h\to 0} \frac{f(\Delta g(x)+g(x))-f(g(x))}{\Delta g(x)} \times \lim_{h\to 0} \frac{g(x+h)-g(x)}{h}$
Since $\Delta g(x) \to 0$ as $h \to 0$, the first limit is the definition of $f'(g(x))$.
The second limit is the definition of $g'(x)$.
Therefore, $y = f(g(x)) \Rightarrow y' = g'(x)f'(g(x))$.

Proof of the product rule:
$(f(x)g(x))' = \lim_{h\to 0} \frac{f(x+h)g(x+h)-f(x)g(x)}{h} = \lim_{h\to 0} \frac{f(x+h)g(x+h)+f(x+h)g(x)-f(x+h)g(x)-f(x)g(x)}{h}$
$= \lim_{h\to 0} \frac{f(x+h)(g(x+h)-g(x))+g(x)(f(x+h)-f(x))}{h} = f(x)g'(x) + f'(x)g(x)$

Proof of the quotient rule:
We take the chain and product rules to be proven.
$\left(\frac{f(x)}{g(x)}\right)' = (f(x)\frac{1}{g(x)})' = \frac{f'(x)}{g(x)} + f(x)(-\frac{g'(x)}{g(x)^2}) = \frac{f'(x)g(x)}{g(x)^2} + f(x)(-\frac{g'(x)}{g(x)^2}) = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$

8.6 Parametric and implicit differentiation

We can differentiate implicit relations of x and y. We can differentiate both sides with respect to x by differentiating any term in x as normal, and by differentiating any term in y with respect to y, and multiplying these by $\frac{dy}{dx}$.

We can also differentiate parametric relations. If y and x are defined in terms of t, we can find $\frac{dy}{dt}, \frac{dx}{dt}$ and use the result $\frac{dy}{dx} = \frac{dy}{dt} \times \frac{dt}{dx}$.

8.7 Constructing differential equations

What it says on the tin.

9 Integration

9.1 Fundamental theorem of calculus

The fundamental theorem of calculus states that differentiation and integration are essentially the reverse of each other. If $\frac{d}{dx}f(x) = f'(x)$ then $\int f'(x)dx = f(x) + c$ for some arbitrary constant c.

9.2 Indefinite integrals

Similar to differentiation, the following results apply: $\int f(x)+g(x)dx = \int f(x)dx + \int g(x)dx$, $\int f(x) - g(x)dx = \int f(x)dx - \int g(x)dx$, $\int af(x)dx = a\int f(x)dx$ for a constant a.
$\int x^n dx = \frac{1}{n}x^{n+1} + c$ as long as $n \neq 1$
$\int e^x dx = e^x + c$
$\int \frac{1}{x}dx = \ln x + c$

$\int \sin x dx = -\cos x + c$
$\int \cos x dx = \sin x + c$

These are all indefinite integrals as the integral signs have no limits, and we need to add an arbitrary constant. NEVER forget to add an arbitrary constant for indefinite integrals.

9.3 Definite integrals and areas

The integral $\int_a^b f(x)dx$ represents the area between the curve $y = f(x)$ and the lines $x = a, x = b$. This integral is definite as there are limits. We do not need an arbitrary constant. The integral returns the signed area (positive if the curve is above the x axis, negative if the curve is below the x axis). If a question asks for the area, not the value of an integral, the area must be positive. We may therefore have to split up the integral into several integrals based on where the curve is above or below the x axis.

The area between the two curves $y = f(x), y = g(x)$ from $x = a$ to $x = b$ is $|\int_a^b f(x) - g(x)dx|$.

We can also integrate parametric relations. We can either use x or y bounds as our limits, but we have to be integrating with respect to the same variable. We can use $\frac{dy}{dt}$ or $\frac{dx}{dt}$ to change our variables. e.g. $\int y dx = \int y \frac{dx}{dt} dt$.

9.4 Integration as the limit of a sum

Integration can be viewed as the approximation of the area under a curve with a bunch of rectangles, where the limit of the sum of the areas of the rectangles is taken as the width of the rectangles tends to 0.

9.5 Integration by substitution

We may wish to transform an integral in the variable x to an integral in a different variable, such as u. If $u = f(x)$ and we have $I = \int_{x=a}^{x=b} g(x)dx$, then $I = \int_{u=f(a)}^{u=f(b)} g(f^{-1}(u)) \frac{dx}{du} du$. This is essentially a reverse form of the chain rule, and there is a particularly important result to recognise (which is in the formula booklet): $\int \frac{f'(x)}{f(x)} dx = \ln(f(x)) + c$.
Integrals of the form $\int f'(x)(f(x))^n dx = \frac{1}{n+1}(f(x))^{n+1} + c$ can typically be done by inspection or by an explicit substitution if you wish. This result is in the formula booklet.

9.6 Integration by parts

We may integrate products of functions using integration by parts: $\int u'v dx = [uv] - \int uv' dx$. This is the reverse version of the product rule. Sometimes, the

method has to be used more than once. Consider the product rule: $(uv)' = u'v + uv' \Rightarrow u'v = (uv)' - uv'$
Integrate both sides: $\int u'vdx = [uv] - \int uv'dx$

To integrate $\ln x$, you can use parts with $1 \times \ln x$ to obtain $\int 1 \times \ln x dx = (\int 1dx) \ln x - \int (\int 1dx) \frac{d}{dx} \ln x dx = x \ln x - \int \frac{x}{x} dx = x \ln x - x + c$.

9.7 Use of partial fractions in integration

Some rational functions such as $\frac{1}{(ax+b)(cx+d)}$ are hard to integrate. We can use partial fractions to decompose them into several fractions, which are easier to integrate as they are scalar multiples and translations of $\frac{1}{x}$, which integrates to $\ln x + c$.

9.8 Differential equations with separable variables

For differential equations relating $\frac{dy}{dx}, y, x$ we have to use the method of the separation of variables. We need all y terms on the same side as $\frac{dy}{dx}$ and all x terms on the other side. We can then integrate both sides with respect to x, so we get rid of the $\frac{dy}{dx}$ and integrate its remaining side with respect to y, and integrate the other side with respect to x. We add an arbitrary constant, which could be evaluated using any given boundary conditions.
The arbitrary constant must be added at the first integration. Every time it changes into a different arbitrary constant, it is worth using a different variable to represent it.
A common trick is for the question to describe an object at rest at a certain time, which is a boundary condition as the derivative of its displacement with respect to time is 0.

Consider $\frac{dy}{dx} = xy \Rightarrow \frac{1}{y}\frac{dy}{dx} = x \Rightarrow \int \frac{1}{y} dy = \int x dx \Rightarrow \ln y = \frac{x^2}{2} + c \Rightarrow y = Ae^{\frac{x^2}{2}}$ where $A = e^c$. Given a boundary condition, such as $y = 1$ when $x = 0$, we can evaluate A.

9.9 Interpreting the solution of a differential equation

What it says on the tin.

10 Numerical Methods

10.1 Sign change methods

As long as $f(x)$ is continuous on a certain interval, if $f(a) > 0, f(b) < 0$ or $f(a) < 0, f(b) > 0$, then $f(x)$ has a root between a and b. For change of sign questions, you should always explicitly say that there is a change of sign after plugging in the values.

This method can be used to determine roots to a certain accuracy by considering upper and lower bounds for what rounds to the root for a given number of significant figures.

Change of sign methods can fail when the function touches the x axis or has a vertical asymptote.

10.2 Formal iterative methods

If we can rearrange an equation $f(x) = 0$ into the form $x = g(x)$, we can use the iterative formula $x_{n+1} = g(x_n)$ with a chosen value of x_1 to look for roots. We can draw cobweb and staircase diagrams to represent this method. We draw the function curve and the line $y = x$ on a graph. We take an x_n, draw a line up to the function curve, draw a horizontal line to $y = x$, and repeat. This can in result in what looks a staircase if the function is convex (its second derivative is greater than 0) and it looks like a cobweb if the function is concave (its second derivative is less than 0).
The iterative formula $x_{n+1} = g(x_n)$ only converges to a root $x = \alpha$ if $|g'(\alpha)| < 1$ and if x_1 is sufficiently close to α.

The Newton-Raphson is a special iterative method where to estimate roots of $f(x) = 0$, we use $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ for a given value of x_0, which is in the formula booklet.

Consider a function $f(x)$ and the tangent at $x = x_n$.
The gradient of the tangent is $f'(x_n)$
We use the equation of a line: $y - f(x_n) = f'(x_n)(x - x_n)$
Let the x-intercept be x_{n+1}, so $-f(x_n) = f'(x_n)(x_{n+1} - x_n)$
$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_{n+1})}$

Why did we take this approach? Well, we can approximate the curve at any point using it's tangent. It's not a great approximation, but it's the best possible linear approximation. We use this linear approximation to estimate the root by finding the x-intercept of the tangent. We do this iteratively to improve on our approximation until it becomes exact.

10.3 Numerical integration

We may wish to estimate the area under a curve $y = f(x)$ from $x = a$ to $x = b$ using trapezia. If we use n trapezia of equal width then the width of each one, h, is $\frac{b-a}{n}$. Let's let the x values $x_0, x_1, x_2, ..., x_n$ denote $a, a+h, a+2h, ..., b$ respectively. We can then use the trapezium rule (which is given in the formula booklet): $A \approx \frac{h}{2}(f(x_0) + f(x_n) + 2(f(x_1) + f(x_2) + ... + f(x_{n-1})))$. This gives an underestimate of the area when the curve is concave and an overestimate when the curve is convex.
We could also use rectangles to approximate integrals, and give upper and lower

bounds for areas.

The area of a trapezium is $\frac{h}{2}(a+b)$.
If we approximate the area under a curve with trapezia, we need to sum each of their areas.
Notice that if you use the formula for the area of a trapezium for each one, all of their widths are the same, and each vertical side is counted twice, apart from the two at the ends, giving:
$A \approx \frac{h}{2}(y_0 + y_n + 2(y_1 + y_2 + ... + y_{n-1})$

10.4 Use numerical methods in context

What it says on the tin.

11 Vectors

11.1 Vectors

A scalar has a magnitude and no direction while a vector has a magnitude and a direction. A two dimensional vector can be represented as $x\mathbf{i} + y\mathbf{j}$ or $\binom{x}{y}$. The vector notations $\vec{AB}$ or $\mathbf{a}$ or $\underline{a}$ could also be used.

Vectors in three dimensions can be represented as $x\mathbf{i} + y\mathbf{j} + \mathbf{k}$ or $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

11.2 Magnitude and direction of vectors

The magnitude of the vector $\mathbf{a}$ is denoted as $|\mathbf{a}|$. If $\mathbf{a} = x\mathbf{i} + y\mathbf{j}$ then $|\mathbf{a}| = \sqrt{x^2 + y^2}$. Similarly, if $\mathbf{a} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ then $|\mathbf{a}| = \sqrt{x^2 + y^2 + z^2}$.

By convention, we define the direction of a vector as the angle it makes with the positive x axis, which should lie in the interval $[0, 360)$.
The direction can be calculated using $\tan^{-1} \frac{y}{x}$, but may not always be exactly equal to this, as this does not always return the angle desired. It may be the case that this value has to be subtracted from 180. It is worth drawing a quick diagram to make sure you have the correct angle.

We define the modulus as the distance of the vector from the origin. That's why $|x|$ gives x but positive, since that represents the distance of x from 0 on the number line. By Pythagoras, the modulus of a vector of n dimensions $(x_1, x_2, ..., x_n)$ is $\sqrt{x_1^2 + x_2^2 + ... + x_n^2}$.

11.3 Basic operations on vectors

We can add vectors by adding the components, or add them diagrammatically by drawing them tip to tail and drawing the resultant vector from the start to the end (to form a triangle of vectors).
We can multiply vectors by scalars by multiplying each component by the scalar (this has the geometric effect of enlarging the vector).

11.4 Position vectors

A position vector is a vector which starts at the origin and points outwards to a point. We tend to use lowercase letters, such as **a** to denote position vectors.
A resultant vector can be split into its horizontal and vertical component vectors.
Two vectors are parallels if and only if they are scalar multiples of each other, i.e. you can multiply one by a constant to get the other.
Two vectors are equal if and only if their components are equal.
A unit vector has magnitude 1.

11.5 Distance between points

The distance between the points represented by position vectors $a\mathbf{i} + b\mathbf{j}$ and $c\mathbf{i} + d\mathbf{j}$ is $\sqrt{(c-a)^2 + (d-b)^2}$.

11.6 Problem solving using vectors

What it says on the tin.

12 Statistical Sampling

The population refers to the entire set of people or objects for some context.
A sample is a subset of this population. Different samples can lead to different conclusions about the population.

A census involves recording data from every single member of the population,

Simple random sampling refers to giving a unique number to each member of the population, choosing a set of these numbers using a random number generator, and only recording data for the people represented by these numbers.
Advantages of simple random sampling include that it is less time consuming and easier than a census, and gives each member an equal chance of being considered, but disadvantages include it not necessarily representing the entire population.
For the exam, you may assume the population is large enough to sample without replacement unless told otherwise.

Opportunity sampling is only using the first members you come across. Advantages include the convenience, whereas disadvantages include the potential bias.

Systematic sampling is essentially sampling with intervals (e.g. taking every 3rd person).

Stratified sampling is where we divide the population into strata based on some category (e.g. age or gender) and then use a different sampling method to sample within these strata. This may represent differences between different demographics if necessary, but is reliant on the advantages and disadvantages of other sampling methods.

Cluster sampling is where the population is split into clusters, such as by location, where each cluster represents the entire population. A sample of clusters is chosen and every member of these clusters contribute (sometimes sampling can be done within these clusters). It may be hard to ensure each cluster reflects the entire population, and there may be bias caused by the choice of particular clusters.

Quota sampling is where members of the population are chosen such that there are enough to reach pre-decided minimum numbers for certain quota (such as age or location).

13 Data Presentation and Interpretation

13.1 Single variable data

You should be familiar with vertical line charts, dot plots, and bar charts.

Reminder that stem and leaf diagrams consist of a stem (usually representing the first digit) with leafs representing all the data. There must be a key.

Reminder that box-and-whisker-plots consist of vertical lines at the minimum and maximum values, and the median, lower quartile, and upper quartile, with the two quartiles forming a box, and the min and max values connected to this box via whiskers.

Reminder that cumulative frequency diagrams show the total of a variable as some other variable changes, and must be strictly increasing.

Reminder that a histogram has bars of unequal width, with frequency density on the y axis and the classes on the x axis. The area represents frequency.

Each data presentation method has advantages and disadvantages, such as whether or not they show all the data, how clearly they present data, whether or not they present the key features of the data (e.g. median), etc.

13.2 Bivariate data

You should be able to interpret scatter diagrams. The regression line is essentially the line of best fit. Scatter diagrams may include distinct sections of the population.
Correlation is essentially where one variable changes linearly as a different variable changes. They may either change in the same direction (positive correlation) or change in opposite directions (negative correlation) or may not impact each other at all (zero correlation). No matter what, correlation does not necessarily imply causation.

13.3 Measures of average and spread

You should know how to determine the mean, median, and mode of a dataset.
The ath percentile is the item which lies at the a% position when the elements are in order.
Reminder that the interquartile range is the upper quartile minus the lower quartile.

The standard deviation is the root mean square deviation from the mean. It is a measure of spread calculated by squaring the differences between each item and the mean, taking the average of these squared differences, and square rooting.
The variance is the square of the standard deviation.

13.4 Calculations of mean and standard deviation

We have $\bar{x} = \frac{\sum x}{n}$.
For variance, we have $\sigma^2 = \frac{\sum(x-\bar{x})^2}{n} = \frac{\sum x^2}{n} - \bar{x}^2$ (and the standard deviation is the square root of this). These are in the formula booklet.
If we are dealing with grouped frequency, then n is the sum of the frequencies, and we need to multiply $(x - \bar{x})^2$ and x^2 with their frequencies (which is also in the formula booklet).

One standard deviation formula is as follows:
$\sigma = \sqrt{\frac{(x_1-\mu)^2+(x_2-\mu)^2+\ldots+(x_n-\mu)^2}{n}}$
This is the root mean square formula applied to the differences between each data element and the mean. The root mean square is just another type of average like the arithmetic mean, but you square each item, add them together, divide by the number of items, and square root.
We could have taken the average of the differences between each item and the mean to get some idea of the spread, but then we don't know which be positive or negative (depending on whether they are greater than or less than the

mean), so some may cancel out with each other. We need the magnitude of the differences.
In theory we could use the modulus function, but that can be tricky to work with, so we square instead to ensure each difference is non-negative.

We can obtain another formula as follows:
$\sigma^2 = \frac{1}{n}\sum_{r=1}^{n}(x_r - \mu)^2 = \frac{1}{n}\sum_{r=1}^{n} x_r^2 - 2x_r\mu + \mu^2$
Notice that $\frac{1}{n}\sum_{r=1}^{n} x_r$ is simply μ
$\sigma^2 = \frac{1}{n}\sum_{r=1}^{n} x_r^2 - 2\mu^2 + \frac{n\mu^2}{n} = \frac{1}{n}\sum_{r=1}^{n} x_r^2 - \mu^2$

13.5 Outliers and cleaning data

There are two definitions for outliers in the spec: more than 1.5 times the interquartile range away from the nearer quartile, or more than 2 times the standard deviation from the mean. Sometimes, these two definitions may disagree on whether an element is an outlier or not. In that case, it is inconclusive.

14 Probability

14.1 Mutually exclusive and independent events

Reminder that mutually exclusive events cannot occur at the same time, and independent events do not affect the probabilities of each other occurring.
Reminder that $P(A)$ is the probability of the event A occurring, and $P(A') = 1 - P(A)$ is the probability that A does not occur (the complement of A).
$P(X = x)$ is the probability that the random variable X takes the value x.

14.2 Probability

You should be familiar with tree diagrams, sample space diagrams, and Venn diagrams.
We have $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, $P(A|B) = \frac{P(A \cap B)}{P(B)}$ (probability of A given B), and the special case $P(A \cap B) = P(A)P(B)$ when A and B are independent.
When A and B are independent, $P(A|B) = P(A)$.

We have $P(A|B) = \frac{P(A \cap B)}{P(B)}$.
If A and B are independent, then $P(A|B) = P(A)$, since the outcome of B has no effect on A (which is what independence is).
Therefore, $P(A \cap B) = P(A)P(B)$.

Now imagine a Venn diagram with circles for A and B. If we wish to find $P(A \cup B)$, we need the area bounded by the two circles. If we simply add $P(A) + P(B)$ as expected, we double count the area shared between A and B, so we need to subtract off $P(A \cap B)$.

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

We have the formula $P(A|B) = \frac{P(A \cap B)}{P(B)}$.
We can build intuition for this formula as follows:
We want the probability of A given B, i.e. the probability A is true if B is true. Well, since we know B is true, then our entire set of possible outcomes is just B.
If A also needs to be true, the successful outcomes are when A and B are true. Since probability is the number of successful outcomes divided by the number of total outcomes (as long as they are equally likely), we can divide $P(A \cap B)$ by $P(B)$.

14.3 Modelling with probability

What it says on the tin.

15 Statistical Distributions

15.1 Discrete probability distributions

Discrete probability distributions can be defined in a table or by a formula, which must specify which values the random variable can take.

The conditions for a binomial distribution are: independent repetitions of a trial, two mutually exclusive and exhaustive events, and constant probabilities of success and failure.
Conditions must be satisfied for a model to hold, and assumptions are where we assume the conditions are met so that the model can be seen as suitable.
Assumptions must be made in context, e.g. for the number of MCQ questions answered correctly, you have to say 'MCQ questions are answered independently of one another' and not 'they are independent' or 'MCQ questions are independent'.

If $X \sim B(n, p)$ then there are n trials with probability of success p, and $P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$. This is in the formula booklet.

We may approximate the binomial distribution using a normal distribution when n is large and p is close to 0.5, in which case we need to use $\mu = np, \sigma^2 = np(1-p)$, which are in the formula booklet.
I am aware that Edexcel expects knowledge of continuity corrections, but this is not on the OCR specification. I'll mention it anyway:
We may need to use a continuity correction due to rounding. For example, since any value between 5 and 5.5 rounds to 5, if we're finding $P(X \leq 5)$ for a binomial distribution, we need to find $P(X \leq 5.5)$ if we're using the normal approximation.

If we are given an event with n trials and probability p of success, and this event satisfies the conditions for a binomial distribution, we can consider the probability of each possible number of successes occurring.
Let's call this random variable X and say it takes integer values x.
The probability $P(X = x)$ is the probability the event succeeds x times and the event fails $n - x$ times, i.e. the probability it succeeds x times but fails every other time.
However, we haven't accounted for which trials these successes and fails occur at. The situations SSFFS and SFSFS both represent the same numbers of successes and fails, but in different orders.
The number of ways we can position the x successes in n trials is n choose x. The remaining spots must all be failures.
Therefore, we have $P(X = x) = \binom{n}{x}p^x(1-p)^{n-x}$

We shall now find the expectation and variance of a binomial distribution (which gives us our values we would use if we wished to approximate it with a normal distribution).
We use the formulae for expectation and variance from Further Maths.
$E(X) = \sum_{r=0}^{n} rP(X = r) = \sum_{r=0}^{n} r\binom{n}{r}p^r q^{n-r}$
We may wish to find an expression for $r\binom{n}{r}$. We know that $\binom{n-1}{r-1} = \frac{(n-1)!}{(r-1)!(n-r)!}$.
$r\binom{n}{r} = \frac{rn!}{r!(n-r)!} = \frac{n(n-1)!}{(r-1)!(n-r)!} = n\binom{n-1}{r-1}$.
Therefore, $E(X) = \sum_{r=0}^{n} n\binom{n-1}{r-1}p^r q^{n-r}$
n is just a constant so we can take that out. What we are left with looks suspiciously similar to a binomial expansion, although some of our terms are shifted by 1.
$E(X) = n\sum_{r=0}^{n} \binom{n-1}{r-1}pp^{r-1}q^{n-r}$.
Taking our p out, and noticing the binomial expansion left sums to 1 since total probability is 1, we get:
$E(X) = np$.

I cannot be bothered to do the entire variance proof. It can be found here: `https://www.probabilisticworld.com/binomial-distribution-mean-variance-formulas-proof/`

15.2 The normal distribution

The normal distribution is denoted as $X \sim N(\mu, \sigma^2)$, where μ is the mean and σ^2 is the variance. Probabilities can only be evaluated with your calculator.

The standard normal distribution Z is distributed $Z \sim N(0, 1)$ and $Z = \frac{X-\mu}{\sigma}$.
If we need to find values of μ and σ given known probabilities, we have to transform to the standard normal.

A normal distribution resembles a roughly bell-shaped unimodal symmetric his-

togram which tapers at the extremes.
If you asked for reasons why a distribution may be normal, the main reasons to give are symmetric and unimodal. Tapering at the ends doesn't tend to be in the mark scheme.

About two thirds of the data lie within one standard deviation of the mean. About 95% lie within two standard deviations. Almost all lie within three standard deviations. The points of inflection of a normal curve lie at approximately one standard deviation from the mean.

15.3 Sample Mean

We use expectation and variance results from Further Maths.
Consider n identically distributed variables $X_1, X_2, ..., X_n$, each with mean μ and variance σ^2.
Let $\bar{X}$ be the sample mean, i.e. $\bar{X} = \frac{X_1+...+X_n}{n}$
We have $E(\bar{X}) = \frac{1}{n}E(X_1 + ... + X_n) = \frac{1}{n}n\mu = \mu$.
And $Var(\bar{X}) = \frac{1}{n^2}Var(X_1 + ... + X_n) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$
These give the results we use when considering the sample mean for normal distribution hypothesis tests.

15.4 Selecting an appropriate distribution

A binomial distribution with large n can be approximated with a normal distribution.

16 Statistical Hypothesis Testing

16.1 The language of hypothesis testing

The null hypothesis, H_0, tends to be what would happen if our hypothesis were not true, and nothing has changed. The alternative hypothesis, H_1 tends to be what would happen if our hypothesis were true, and things have changed.
The significance level is the pre-decided maximum probability which would lead to rejection of the null hypothesis.
The test statistic is the distribution we use for the test (for our spec we only cover normal, binomial, PMCC).
A 1-tail test is where we only test either if something is lower than it should be, or higher than it should be.
A 2-tail test is where we test for both. We tend to split the significance level in half in this case.
The critical value is the value such that anything more extreme leads to rejection of the null hypothesis.

The critical region is the range of values which leads to rejection of the null hypothesis.
The acceptance region is the range of values which leads to acceptance of the null hypothesis.
A p value (or the significance level) is the probability the null hypothesis is rejected if it is true.

This is the general recipe for a hypothesis test:

H_0: [insert null hypothesis]
H_1: [insert alternative hypothesis]
Define the parameter in context (you typically have to say the parameter represents the whole population)
Divide the significance level by 2 if the test is two-tailed.
State the test statistic / distribution
Determine the critical region and determine whether the data given is in it, OR determine the probability of the data occurring and compare to the significance level.
Conclude: "there is sufficient/insufficient evidence to reject H_0 at the [significance level]: [insert thing] is likely to [alternative hypothesis]".

16.2 Hypothesis test for the proportion in a binomial distribution

We can use the above recipe to perform hypothesis tests for the proportion in a binomial distribution.
For a given significance level, the probability of rejecting the null hypothesis will be less than or equal to this level.

16.3 Hypothesis test for the mean of a normal distribution

We can consider a sample mean $\bar{X}$ as a random variable, and we have that if $X \sim N(\mu, \sigma^2)$ then $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, where n is the sample size. This is in the formula booklet.
Critical values can be calculated using calculator functions or the percentage point table in the formula booklet.

Exam question from 2022:

In the past the masses of new-born babies in a certain country were normally distributed with mean 3300 g. Last year a publicity campaign was held to encourage pregnant women to improve their diet. Following this campaign, it is required to test whether the mean mass of new-born babies has increased. A random sample of 200 new-born babies is chosen, and it is found that their mean mass is 3360 g. It is given that the standard deviation of the masses of new-born babies is 450 g. Carry out the test at the 2.5% significance level. [7]

11			**See the exemplars at the end of the MS**			**NB. Use of a "continuity correction" loses 1st A1 only**
			Hypotheses:			
			H_0: $\mu = 3300$ H_1: $\mu > 3300$ where μ = (population) mean mass	**B1** **B1**	**1.1** **2.5**	Allow other letter (including X) only if clearly defined Subtract B1 for each error eg: 2-tail B1B0 Undefined μ B1B0 not in terms of parameter B1B0 Not include 3300 B0B0 $\overline{X}$ stated or implied B0B0 $H_0 = 3300$ etc: B0B0 μ = sample mean implied B0 & (B1or B0)
			Calculation and comparison $\overline{X} \sim N(3300, \frac{450^2}{200})$ or N(3300, 1012.5) oe and $\overline{X} > 3360$	**M1***	**3.3**	Correct distribution and $\overline{X}$ (allow 3359.5, 3360.5, 3659) stated or implied eg by 0.0297 or 0.0307 or 0.0286 even if within incorrect statement eg $P(X = 3360) = 0.0297$ Allow $450^2 \div \sqrt{200}$ or $450^2 \div 200^2$ Not 0.0297 from $\mu = 3360$, $P(\overline{X} < 3300)$
			$P(\overline{X} > 3360) = 0.0297$ **(NB 3 sf)**	**A1**	**3.4**	**BC cao**

20

H240/02 Mark Scheme June 2022

Question			Answer	Mark	AO	Guidance
			$0.0297 > 0.025$	**A1**	**1.1**	Explicit comparison Allow compare (**any value $\le$ 0.35**) with 0.025
11	**ctd**		**Conclusion** Not reject H_0	**M1 dep**	**1.1**	Allow Reject H_1 or Accept H_0 Dep M1A1A1 or M1A0A1 or M1A1A0 Dep also on comparing like with like, eg not 0.970 > 0.025 May be implied by their conclusion, if M1 criterion is met ft with opposite conclusion if, eg $P(\overline{X} > 3360) = 0.024$
			There is insufficient evidence that (mean) mass (or mean) is > 3300 (g) or has increased	**A1f**	**2.2b**	Dep M1A1A1 or M1A0A1 or M1A1A0 In context, not definite; eg not "Mean mass not > 3300 g" Not "There is evidence that mean mass hasn't increased" Not "This suggests that mean mass hasn't increased" Not "Unlikely (or unsure) that mean has increased" ft with opposite conclusion if, eg $P(\overline{X} > 3360) = 0.024$
						Alternative scheme for incorrect method using 2-tails: Hypotheses **B1B0** Calculation: as above **M1A1** Compare 0.0125 oe **A1** Conclusion **M0A0**
				[7]		

16.4 Hypothesis test using Pearson's correlation coefficient

The PMCC, Pearson's Product Moment Correlation Coefficient, is a number from -1 to 1 which represents how correlated data is. 1 represents perfect positive correlation, -1 represents perfect negative correlation, and 0 represents no correlation. It can be used in hypothesis test using the table in the formula booklet.

This is fairly complicated, and explored in detail in Further Maths, so I will provide a very brief overview.
We define $s_{xx} = Var(x), s_{yy} = Var(y), s_{xy} = Cov(x, y)$ where $Cov(x, y)$ is the covariance between x and y.
Covariance is essentially a measure of how the two variables vary together.
The formula for the PMCC is $r = \frac{S_{xy}}{\sqrt{S_{xx}Syy}}$.
Some very shaky considerations based on intuition:
If two variables have no correlation, their covariance can be considered as 0, and therefore the PMCC is 0.
If two variables are perfectly positively correlated, they might as well be identical and have identical variances and covariance, giving a PMCC of 1.

17 The Large Data Set

The large data set is pre-release material specific to OCR (and is the same for every year). Some of the data may be used in statistics questions, but you are never expected to memorise any of its data, however familiarity with the data set tends to be an advantage. The following information is useful to be aware of, but you do not necessarily need to memorise it.
The LDS can be accessed here in 'Pre-release Materials':
`https://www.ocr.org.uk/qualifications/as-and-a-level/mathematics-a-h230-h240-from-2017/assessment/`

The data is split by local authority district (LAD) and unitary authority (UA). You do not need to know the difference, but it is important to know that the data is split by location in the UK.
There are two topics whose data is presented: method of travel and age structure. There are two tables for each of these topics, one for 2001 and one for 2011.
The method of travel is the method of transport used for the part with the longest duration in the commute to work (only applicable to people in work a week before the census). The tables for 2001 and 2011 give the number of people, by local authority, using each type of transport, as well as the number of unemployed people.
In the age structure tables, the number of people for each age range by local authority is given, with infants under the age of 1 treated as age 0, and seniors over the age of 115 treated as invalid. The mean and median age are included.

18 Quantities and Units in Mechanics

18.1 SI units

Length (m), time (s), and mass (kg), are the three base quantities in the S.I. system and are mutually independent. Other quantities, such as velocity and force, are derived from these base quantities.
The unit for moment is Nm.

19 Kinematics

19.1 Language of kinematics

The displacement of an object is a vector quantity representing the magnitude and direction of its position from a given origin. (Distance is a scalar, the magnitude of displacement, and overall distance travelled has to be determined from the overall path of the object, and cannot be determined from the displacement alone).
Velocity is the vector version of speed, representing the magnitude and direction of the speed.
Acceleration is the rate of change of velocity, also a vector.

19.2 Graphical representation

The gradient of a displacement-time graph is the velocity. The gradient of a velocity-time graph is acceleration. The area between an acceleration-time graph and the x axis is velocity. The area between a velocity-time graph and the x axis is displacement.

19.3 Constant acceleration

The following SUVAT equations are given in the formula booklet (and only apply for constant acceleration):
$v = u + at, s = ut + \frac{1}{2}at^2, s = \frac{1}{2}(u + v)t, v^2 = u^2 + 2as, s = vt - \frac{1}{2}at^2$
These equations can be extended to 2D by replacing u, v, a, s with their respective vectors (these are also in the formula booklet). We exclude $v^2 = u^2 + 2as$ from this.

There are a variety of ways to derive these formulae, for example:

- $v = \int adt = at + c, t = 0 \Rightarrow v = u$ which gives $v = u + at$
- $s = \int vdt = \int u + atdt = ut + \frac{1}{2}at^2 + c, t = 0 \Rightarrow s = 0$ which gives $s = ut + \frac{1}{2}at^2$
- Finding formulae for areas under graphs using triangles
- Finding formulae for gradients

- Substituting $v = u + at$ into $s = \frac{1}{2}(u+v)t$ to get $s = \frac{1}{2}(2u+at)t = ut + \frac{1}{2}at^2$

Acceleration is simply defined as the rate of change of velocity, much like how velocity is defined as the rate of change of displacement.
Therefore, $a = \frac{v-u}{t} \Rightarrow v = u + at$.
This also tells us that acceleration is the derivative of velocity with respect to time, and velocity is the derivative of displacement with respect to time (these are really just definitions).

We can integrate both sides of $v = u + at$ with respect to time to get $s = ut + \frac{1}{2}at^2$.

Since $u = v - at$, we can substitute to get $s = (v-at)t + \frac{1}{2}at^2 = vt - at^2 + \frac{1}{2}at^2 = vt - \frac{1}{2}at^2$

Average speed is $\frac{u+v}{2}$ but average speed is also $\frac{s}{t}$, therefore $\frac{u+v}{2} = \frac{s}{t}$.

$v + u = \frac{2s}{t}, v - u = at$. Therefore, $(v+u)(v-u) = v^2 - u^2 = 2as \Rightarrow v^2 = u^2 + 2as$.

19.4 Non uniform acceleration

For non-constant acceleration, we have $v = \frac{ds}{dt}, a = \frac{dv}{dt} = \frac{d^2s}{dt^2}, v = \int adt, s = \int vdt$.
We can extend this to vectors in two dimensions by differentiating or integrating the horizontal and vertical components separately.

19.5 Gravity

A projectile moves only under the influence of gravity, i.e. $\mathbf{a} = \binom{0}{-g} = -g\mathbf{j}$, where $g = 9.8\text{ms}^{-2}$ for this spec. There may be limitations of modelling motion with projectile motion, such as the influence of air resistance or the non-uniformity of the object.
Given initial velocity u at angle θ to the horizontal, the horizontal component of the velocity is $u\cos\theta$ and the vertical component is $u\sin\theta$.
The total duration of the motion is given by $s = ut + \frac{1}{2}at^2 \Rightarrow 0 = ut\sin\theta - \frac{1}{2}gt^2 \Rightarrow t = \frac{2u\sin\theta}{g}$.
The horizontal distance travelled for this duration is $\frac{2u\sin\theta}{g}u\cos\theta$.
The maximum height is given by $v^2 = u^2 + 2as \Rightarrow 0 = u^2\sin^2\theta - 2gh \Rightarrow h = \frac{u^2\sin^2\theta}{2g}$.
We can derive a cartesian equation for the projectile's motion. The vertical height at a certain time is given by $y = ut\sin\theta - \frac{g}{2}t^2$ and the horizontal displacement at a certain time is $x = ut\cos\theta \Rightarrow t = \frac{x}{u\cos\theta}$, so substituting gives $y = \frac{x}{u\cos\theta}u\sin\theta - \frac{g}{2}(\frac{x}{u\cos\theta})^2 = x\tan\theta - \frac{gx^2}{2u^2}\sec^2\theta = \tan\theta - \frac{gx^2}{2u^2}(1+\tan^2\theta)$.

20 Forces and Newton's Laws

20.1 Newton's first law

A force is a vector and changes the velocity of an object with mass.
Newton's first law states that an object at rest or moving with a constant velocity remains at this constant velocity until acted upon by an external force.
When the resultant force is 0, an object is in equilibrium, and all the forces on it form a closed shape.

20.2 Newton's second law

Newton's second law states that $F = ma$ for motion in a straight line of bodies with a constant mass moving under the action of a constant force. Force and acceleration may be considered as a two-dimensional vector here.
Forces may need to be resolved, in which case we can multiply by $\cos\theta$ or $\sin\theta$ to get the horizontal or vertical components of the force (depending on what θ is).

20.3 Weight

We have $W = mg$ for a body moving under gravity. Our spec states $g = 9.8\text{ms}^{-2}$ for all questions, but you need to be aware that g is not a constant, and depends on location in the universe.

20.4 Newton's third law

Newton's third law states that every action has an equal and opposite reaction.
A system in which none of its components has any relative motion may be modelled as a single particle.
When an object is resting on a horizontal surface, the normal reaction force is equal and opposite to the weight.
Contact is lost when the reaction force is 0.

Surfaces can be modelled as smooth, i.e. without friction, which may be inaccurate.

When particles are connected, every component has the same acceleration.
A particle is in equilibrium if and only if the sum of the resolved parts of all the forces in any given direction is 0. Forces may be resolved horizontally and vertically, parallel and perpendicular to a surface, or in any chosen direction.

20.5 Applications of vectors in a plane

Vector addition can be used to find resultant forces of two or more forces acting on an object.

The velocity vector gives the direction of motion and the acceleration vector gives the direction of resultant force.

20.6 Frictional forces

Frictional forces act in the opposite direction to a driving force.
The components of the contact force between two rough surfaces are the normal force and the friction force.
We use the model $F \leq \mu R$, where F is the frictional force, R is the normal contact force, and μ is the coefficient of friction.
Limiting friction describes the maximum friction before slipping occurs (before the system no longer remains in equilibrium) and occurs when $F = \mu R$.

21 Moments

21.1 Statics

The moment of a force about a pivot is the product of the magnitude of the force and the perpendicular distance from the line of action of the force to the pivot.
A rigid is body is in equilibrium if and only if the resultant force and resultant moment are 0.
For a uniform rod, the weight acts at the midpoint of the rod.
For a non-uniform rod, the weight acts at a given specific point or is to be determined by moments.
For a rectangular lamina, the weight acts at the point of symmetry.