MATHEMATICAL SUPPLEMENTS

For H2 Math Students

MATHEMATICAL SUPPLEMENTS

For H2 Math Students

First Edition

Yang Xu

A LETTER TO THE READER

Hello, and thank you for picking up my book! I am Yang Xu from the 2021 VJC cohort, and I am really excited to share with you my journey with Mathematics! Math wasn't always my strongest subject, and I really struggled with it back in secondary school. I understand the frustration when concepts do not click despite the effort, and the disappointment at the marks lost to careless mistakes. But one thing Mathematics has taught me is that with the right mindset, appropriate guidance, and some practice, you can really start to see improvement, and even start enjoying math! Having gone through this journey myself, I firmly believe that everyone is capable of doing well in Mathematics, provided the effort and appropriate understanding of the subject.

Mathematics is a subject of immense beauty, but only if you allow yourself to see it. In my humble opinion, the structure of the examined syllabus and the competitive nature of exams often overshadow the true elegance of math, making it easy to miss the profound connections between topics. Through this book, I hope to help you discover new insights and develop a deeper appreciation for math, especially for the examinable topics. But even if you don't come away with that appreciation right away, I still encourage you to keep an open mind and give your best effort. Remember, the true beauty of mathematics goes far beyond what I can express here. I urge you to continue exploring on your own if you're curious; there is truly so much more to discover!

In this book, I have put together a non-exhaustive summary of the key ideas and conventional solving methods from the H2 Math syllabus. I've also included some personal tips, shortcuts, and insights that I found helpful as a student. These sections are designed to help refresh and review the H2 content while offering some new perspectives on familiar concepts. I have also dedicated a part towards the end of this book to exploring some interesting topics in the syllabus, though not explicitly taught. These concepts have personally helped me a lot in my math journey, but they're only the tip of the iceberg. I encourage you to dive deeper and explore more beyond what's presented here.

Please note that this book is only supposed to serve as a supplement, not a

replacement of your school notes. Certain parts also require H2 knowledge that is not covered explicitly in this book.

On a side note, I would also like to mention that Math concepts beyond the syllabus can also develop your critical thinking and problem-solving skills in H2 Math. After all, that is what you are being examined for. These topics either broaden your way of thinking, or provide valuable insights into H2 Math topics. For instance, learning about 'Linear Algebra' has helped me to appreciate the interconnections between vectors, matrices, and systems of linear equations. I would occasionally share some of the insights I have gained throughout the span of this book. Meanwhile, I also encourage you to read beyond the syllabus, to see the underlying interconnection between different topics, and appreciate the beauty of Mathematics.

As a disclaimer, some of the phrasings in this book are layman paraphrases for the sake of simplicity and ease of comprehension. They may not be accepted in working presentations, hence I would urge you to clarify the presentation aspects with your tutors. The content here is also non-exhaustive, hence don't expect to find everything here. This just serves as a brief recap, an extension, and some hopefully-useful tips.

Moreover, if you face any difficulties in your studying, please do not hesitate to contact your subject tutors, or me via my VJC email (yang.xu.2021@vjc.edu.sg). With that, I wish you all the best for all your Math tests and exams; I really really believe that all of you are able to do well in them!

Best Wishes,
Yang Xu
Class of 21S51

CONTENTS

Part I

Content Topics

CHAPTER 1

GENERAL TIPS

1 Analysing the Question

Read ALL questions carefully.

A good habit would be to have a pen or pencil in hand when reading the question. Underline or circle whatever is important and annotate beside it.

For example, 'maximum point at $x = 2$' implies $\frac{d^2y}{dx^2} < 0$ and $\frac{dy}{dx} = 0$ at $x = 2$.

If you are like my past self, you might think: "I already have them in my head, why do I need to write it down and waste my time?" To this, I would say that annotating can help to direct your thought process in the right direction, while helping you to store the information safely. At times, questions contain so much information that it is very difficult to remember everything. This also leads me to the next point.

Try to practice the habit of copying the question.

Even when you are stumped and completely clueless upon reading the question, just copy it down first as your first line of working. Sometimes (and I can say that it is not uncommon), copying the question can help you to remember certain concepts related to it, which will be the appropriate formulae or concepts required to solve the question.

Behaves like muscle memory, works like magic. I do this all the time and I can guarantee it helps!

Take on different viewpoints when analysing the question.

Acknowledge that there are many solutions to the same question.

For example, when you see $ax^2 + bx + c = 0$, how will you find the number of real roots? You could draw out the graph and see if there are intercepts, or you can make use of the discriminant value to determine the number of real solutions.

Some questions can allow you to have multiple perspectives to view it. An expansion of $(1+x)^n$ can be seen as a Binomial Expansion, a Maclaurin Expansion, a polynomial expansion, or even a combinatorics problem where the number of ways to combine one term from each of the 'n' brackets are used to find the terms.

Taking on multiple viewpoints can help you find faster and more efficient methods to solve questions. Some questions even require you to combine different perspectives in order to solve it, like finding the number of intersections between an ellipse and a parabola.

When you are stuck, try taking on a different approach. Even if it does not provide you with a solution, it may offer you valuable insights that are useful in solving the question.

Start with the end in mind

Arguably, the most important element to solving a problem is to have a clear direction. Navigating a maze with a planned route is exponentially easier than without one. As much as possible, try to mark out different "checkpoints" along the way, to ensure that you are on the right path to solving the problem. Here are some useful questions to help plan this route:

- What information is required? Where is the end-point here?
- What information do I currently have? What tools can I use with this information?
- Are there any information I must have in order to reach the end-point?

As a hiker myself, this is analogous to planning a hike. Imagine planning a hike up a snowy cliff: We will have to pass by steep hills and snowy terrain, rest overnight at a mountain lodge before climbing up a twenty meter cliff to reach the summit. For the steep hills we will require walking sticks; crampons and an ice axe for the snow; a harness with ropes and camming devices for the rock climbing. All these have to be planned out before setting off for the hike itself.

Likewise, you need to plan out a road map before embarking on each problem, with areas you must pass by and tools you will require before reaching the destination. This might sound vague, so we illustrate this with an example.

Example : 2023 SMO Q8

Two planes $x + y + 3z = 4$ and $2x - z = 6$ intersect at the line

$$\vec{r} \times \begin{pmatrix} -1 \\ a \\ b \end{pmatrix} = \begin{pmatrix} -2 \\ c \\ d \end{pmatrix}$$

Find $|a + b + c + d|$. (Don't worry, this is solvable with H2 Maths.)

We start off with a sanity check: We know that two planes indeed intersect on a line, which aligns with the information provided by the question, so we are on the right track.

Planning: Given the equations of the two planes, we are able to determine the line of intersection. We are also given the equation (in a manipulated form) of the same line, which contains the unknowns a, b, c, d; so we are likely expected to find the equation of the line ourselves, then compare it to the unknowns given. The end-point is where we find the sum of all the unknowns.

We break it into two sub-problems: Finding the equation of this line, then comparing it with the second equation (with unknowns) to find the unknowns.

Sub-problem 1: Find the equation of the line of intersection of the two planes.

(Planning the approach) The line must lie on both planes, and so this line must be perpendicular to the normal of both planes.

(Tool 1) The cross-product of two vectors conveniently gives a third vector which is perpendicular to the first two, so we will need that. This gives the direction of the line, but is insufficient since a line is defined by a point and a direction vector. So we will need to find a point that lies on both planes.

(Tool 2) That should not be too complicated, since a point lies on a plane if it satisfies the equation of the plane; we can make intelligent guesses and verify easily whether a point lies on both planes. With the point and the direction vector, we can successfully assemble the equation of the line, and sub-problem 1 is completed.

Sub-problem 2: Use the known equation to find the unknowns.

(Planning) We know that $\vec{r}$ can be rewritten as a column vector with components x, y, z so we can compute the cross product on the left-hand side, then equate it to the column vector on the right.

(Tool 1) From the previous part we have already found an entire line of points (x, y, z) that lie on both planes. We can simply pick and choose some of these points (and sub them in to the cross-product equation) to form equations in a, b, c, d, and the question is solved.

Though it looks like a whole lot, the above planning is much easier than it looks. Planning is usually done in a matter of seconds, and comes even quicker with practice. Practice this in your day-to-day problems, and soon it will become second-nature to you.

That said, the problem is actually rather simple after we laid out the road map.

Solution:

$$\vec{n_1} \times \vec{n_2} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \times \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 7 \\ -2 \end{pmatrix}$$

(3, 1, 0) lies on both planes, since it satisfies both plane equations. So equation of l:

$$\vec{r} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 7 \\ -2 \end{pmatrix}$$

Computing the given cross-product yields:

$$\begin{pmatrix} by - az \\ -z - bx \\ ax + y \end{pmatrix} = \begin{pmatrix} -2 \\ c \\ d \end{pmatrix}$$

So using $x = 3$, $y = 1$, $z = 0$ from the above point (3, 1, 0):

$$b(1) - a(0) = -2 \quad \Rightarrow \quad b = -2$$

$$-(0) - (-2)(3) = c \quad \Rightarrow \quad c = 6$$

We use another point (2, 8, -2), as generated from the equation of the line we found:

$$(-2)(8) - a(-2) = -2 \quad \Rightarrow \quad a = 7$$

$$7(2) + 8 = d \quad \Rightarrow \quad d = 22$$

Therefore they sum to give 33. ■

At times, the road ahead is foggy and a clear direction cannot be seen. In times like these, you would have to plan as much as possible, before diving into the problem head-first. Do not forget to continue planning along the way. Some problems require you to play around with the expressions given, before you can see the clear path, like the integral of $(1-e^{-x})^{-1}$.

Your plan at the start may also be incorrect at times, akin to a turn down a wrong path. These situations require you to re-evaluate your direction and amend some of your existing plans.

Most problems you face in the syllabus would likely be less unorthodox than the one above, and so (hopefully) easier to plan out. From my experience, most questions in the H2 syllabus are rather standard, with only a handful of question types, each requiring similar routes. This is why practicing setting out these plans provide much value: As you are likely to encounter it again. Practice (in the right manner) makes perfect!

2 Solving the Question

Mathematics is all about solving problems. Having certain good habits can make the difference between solving the question quickly and struggling to even start.

Be Systematic

The idea of being systematic (an organised method to do something) is always valued in mathematics. It provides efficiency and prevents unnecessary mistakes.

How would you expand the following expression?

$$(3x^2 - 2 + 5x^{-1})(6x^3 - 4 + 7x)$$

Without an organised system, one would jump in directly and multiply using the "rainbow method", pairing each term in the first bracket with a term in the second bracket.

$$18x^5 - 12x^2 + 21x^3 - 12x^3 + 8 - 14x + 30x^2 - 20x^{-1} + 35$$

After which terms with the same "x power" is combined.

Consider an alternate method, where we start from the lowest possible x power, and work our way up. The lowest possible x power here is -1, so we find all ways to create a term with x^{-1}, that is, one way: picking $5x^{-1}$ from the first bracket and -4 from the second. So the coefficient of x^{-1} is -20.

Next we repeat the same for x power 0. Each time, ask yourself, "which two terms multiple to give me a term of this power?". Repeat this procedure for each x power until we reach the maximum possible power of 5.

$$(5)(-4)x^{-1} + [(-2)(-4) + (5)(7)] + (-2)(7)x + [(5)(6) + (3)(-4)]x^2$$

$$+[(-2)(6) + (3)(7)]x^3 + (3)(6)x^5$$

This method is more systematic and makes it easier to keep track of your terms, especially as the expression gets more complex.

I would advise you to take out a pen and a piece of paper to try it yourself, to see that the latter method is indeed much more organised.

A systematic method could also be breaking down a complex problem into smaller and simpler problems, then tackling them one at a time:

- To find the equation of a straight line, we can break it into three steps: First finding its gradient, next finding a point on the line, then assembling them to form the equation of the line.

- To find a complicated complex number like $\frac{z^5}{(w^*)^2}$, we can break it into three steps: First finding its modulus, then finding its argument, lastly assembling them to find the complex number.

- To find the projection of vector $\vec{a}$ on line L, we can break it into three steps: First find the length of the projection, next find the unit vector of the direction of L, lastly assemble them to find the projection vector.

The idea of being systematic extends beyond the listed examples. Systematic approaches exist for every single topic, though a method may be systematic to one person but not another; what makes a method systematic is whether it appears organised to you. In general, having a systematic approach to problems will always be more efficient and reduce the likelihood of careless mistakes.

Pattern Recognition

The ability to recognise patterns is required to solve certain questions, and can also provide computational shortcuts in some cases. Take this compound interest question, for example:

> ***Example : Recognising the Pattern***
> \$1000 is deposited into a bank at the start of month 1. At the end of every month, the total amount grows by 1%, whereas at the start of every month, \$1000 more is deposited into the bank. Find the total amount at the end of n months.
>
> Since the question is interested in the amount at the *end* of the month, we let $\$T_n$ denote the balance at the end of n months. We list the first few terms and generalise using the pattern.
>
> $$T_1 = 1.01(1000), \quad T_2 = 1.01(1000 + 1.01(1000)) = 1.01(1000) + 1.01^2(1000)$$
>
> $$T_3 = 1.01(1000 + 1.01(1000 + 1.01(1000))) = 1.01(1000) + 1.01^2(1000) + 1.01^3(1000)$$
>
> Notice the pattern: All T_n start with 1.01, increasing their power by 1 each term, ending at 1.01^n. Thus we generalise: $T_n = 1000 \sum_{r=1}^{n} 1.01^r$ which represents a GP. ■

We consider another problem utilising patterns as a shortcut.

Example : Generalising with Patterns
Given two terms of a GP: $a_{17} = 27$ and $a_{20} = 729$, find a_n.

The slightly more obvious method would be to find the common ratio:

$$r^3 = \frac{a_{20}}{a_{17}} = 27 \quad \Rightarrow \quad r = 3$$

Then the first term:

$$a_1 = \frac{a_{17}}{r^{16}} = \frac{1}{3^{13}}$$

Lastly plugging it into the formula for a_n:

$$a_n = a_1 r^{n-1} = \frac{1}{3^{13}}(3^{n-1}) = 3^{n-14}$$

However, we can consider a shortcut: Notice that $a_{17} = 3^3$ and $a_{20} = 3^6$. The subscript and the index always has a difference of 14: $17-3 = 14$ and $20-6 = 14$. Thus, $a_n = 3^{n-14}$.

Important Note: This shortcut must be used with care, especially in considering the case where r might be equal to -3. Suppose we were instead given $a_{17} = 27$ and $a_{19} = 243$. Using the same shortcut (without considering that r can be negative) will give us $a_n = 3^{n-14}$, when in reality the answer is $a_n = (\pm 3)^{n-14}$. ■

Fortunately in most other contexts, this shortcut can be used rather freely without such corner cases. The above just serves to highlight that it should not be applied blindly.

Simplify your Expression

Learning to simplify your expressions can save you considerable time and make your work much neater. This is beneficial to you since disorganised work tends to invite careless mistakes. It is also helpful when checking your work later on, and much more pleasant for the person marking your script. Simplifying your expressions can be done in a few ways:

1. **Cancelling redundant terms:** Always remove the terms if they are not necessary. Do not waste your time rewriting it over and over again. Contrast the two presentations:

Example : Cancelling redundant terms in inequalities
Find the range of x for which the following inequality is satisfied.

$$\frac{(x^2-4)^2}{x^3-6x^2+11x-6} > 0$$

Inefficient presentation:

$$\frac{(x^2-4)^2}{x^3-6x^2+11x-6} > 0$$

Using factor theorem we simplify the denominator:

$$\frac{(x^2-4)^2}{(x-1)(x^2-5x+6)} > 0$$

$$\frac{((x+2)(x-2))^2}{(x-1)(x-2)(x-3)} > 0$$

$$\frac{(x+2)^2(x-2)}{(x-1)(x-3)} > 0$$

(Continue solving using test point method with critical points -2, 1, 2, 3)

The above method is evidently inefficient. Since the numerator is never negative, we only need to consider the denominator (find the x values for which denominator is positive). Contrast the above with the efficient presentation; much more time is required to write out the inefficient method.

Efficient presentation:

$$\frac{(x^2-4)^2}{x^3-6x^2+11x-6} > 0$$

Using factor theorem we simplify the denominator. We also cancel the numerator since it is never negative.

$$\frac{1}{(x-1)(x^2-5x+6)} > 0$$

$$\frac{1}{(x-1)(x-2)(x-3)} > 0$$

Then test-point method can then be conducted using critical points 1, 2, 3. ■

Such inefficiencies scale with the complexity of the question. While the above example may not illustrate a big difference, the contrast is definitely greater for tedious problems. It is hence advisable to cancel the redundant terms at the first opportunity.

2. **Multiplying across by a non-zero factor:** Whenever dealing with multiple fractions, multiply across by the denominator if possible. It is much neater that way.

Example : Multiplying across
Find the equation of the tangent to the curve $y = \frac{4}{7}x^2$ at $x = \frac{1}{3}$.

$$\frac{dy}{dx} = \frac{8}{7}x$$

At $x = \frac{1}{3}$,

$$y = \frac{4}{63}, \quad \text{Gradient} = \frac{8}{21}$$

So equation of line is:

$$y - \frac{4}{63} = \frac{8}{21}\left(x - \frac{1}{3}\right)$$

$$y = \frac{8}{21}x - \frac{8}{63} + \frac{4}{63} = \frac{8}{21}x - \frac{4}{63}$$

Instead of shifting $\frac{4}{63}$ to the right hand side first, it is advisable to multiply across by 63 to get rid of the fractions. This is also neater should you require this equation again in a later part.

$$63y - 4 = 24x - 8 \quad \Rightarrow \quad 63y = 24x - 4$$

Multiplying across by a constant (other than zero) is always permissible. Do note, however, that in order to multiply across by an unknown (like x or $f(x)$), you first have to ensure that it does not, or will not equal to zero.

3. **Factorising:** Factorise out any common factors whenever you have the chance to. This reduces the amount of things you have to write.

Example : Factorising
Given $y = \frac{2x^2}{\sqrt{3x^3-7}}$, find $\frac{dy}{dx}$ and use it to find the x-coordinates of the stationary points.

We expand using the product rule:

$$\frac{dy}{dx} = \frac{4x}{\sqrt{3x^3-7}} + 2x^2 \cdot \left(-\frac{1}{2}\right) \cdot \frac{(9x^2)}{(3x^3-7)^{3/2}}$$

$$= \frac{1}{(3x^3-7)^{3/2}}(4x(3x^3-7) - 9x^4)$$

$$= \frac{12x^4 - 9x^4 - 28x}{(3x^3-7)^{3/2}}$$

The numerator can then be set to 0 to find the stationary points.

4. **Avoiding unnecessary expansions and simplifications:** Sometimes, simplifications are not required to solve for the answer. Unnecessary expansions can also make

the problem more complicated. Keep the expression in its factored form until it's absolutely necessary to expand it (when there is no other way to solve).

> ***Example : Avoiding unnecessary expansions***
> Find the derivative of $y = (x^2+1)^3(2x^2-1)$ at $x = 1$.
>
> Instead of expanding this expression, we differentiate using the product rule:
>
> $$\frac{dy}{dx} = 3(x^2+1)^2(2x^2-1) + (x^2+1)^3(4x)$$
>
> Now instead of simplifying the expression above, we can simply substitute $x = 1$ into the right-hand-side to find the value of $\frac{dy}{dx}$, since that is all that the question requires.
>
> Realise also that expanding first would have required additional steps and led to more opportunities for mistakes. ■

Be Sensitive to the Nuances

It's important to be sensitive to the subtleties in the question. Pay attention to conditions such as whether a variable is positive or negative, whether it is an integer or a real number, or the presence of certain constraints like specific domains or principle ranges. These nuances may require us to exclude certain solutions.

For example when working with functions, special attention must be given to the domain and range. Suppose we want to find the inverse function of $y = x^2$ for $x < 0$. When we take the inverse of this function, we obtain $x = \pm\sqrt{y}$. However, because we are restricted to $x < 0$, we must choose the negative root, i.e. $x = -\sqrt{y}$, to satisfy the condition.

Another common situation arises with modulus. For example, suppose k is an unknown negative constant, then $|k| = -k$ rather than k. This is crucial when solving inequalities or equations involving absolute values, since the incorrect interpretation of $|k|$ will definitely lead to the wrong solution.

3 Checking Your Work

Once you've completed a problem, it is essential to verify your solution. Checking your work can prevent careless mistakes, improve accuracy, and help you understand any steps where you might have gone wrong. Here are some methods you can use to check your solutions effectively.

Substitute the Solution Back into the Original Problem

The most straightforward way to check if your solution is correct is to substitute it back into the original equation or inequality. This ensures that the solution satisfies the problem's conditions.

> ***Example : Substituting Back***
> Solve for the intersection of the lines $y = 2x + 3$ and $2y = 3x + 7$.
>
> $$2(2x + 3) = 3x + 7 \quad \Rightarrow \quad x = 1 \quad \Rightarrow \quad y = 5$$
>
> So our answer is (1, 5). To check, substitute $x = 1$ and $y = 5$ back into the original equations:
>
> $$2(1) + 3 = 5$$
> $$2(5) = 3(1) + 7$$
>
> Since both sides of the equations are equal, the solution is verified as correct. ■

Check for Special Cases

In problems involving inequalities or domains, ensure that your solution holds true for special cases or boundary conditions. For example, in an inequality problem, try testing values near critical points to see if they satisfy the inequality.

> ***Example : Checking Boundary Conditions***
> Consider the inequality $\frac{x-2}{x+3} > 0$. The critical points are $x = 2$ and $x = -3$. Test values near these points:
>
> $$x = 0 : \quad \frac{0-2}{0+3} = \frac{-2}{3} \quad \text{(Not greater than 0)}$$
>
> $$x = 3 : \quad \frac{3-2}{3+3} = \frac{1}{6} \quad \text{(Greater than 0)}$$
>
> Testing values around the critical points can help to catch careless mistakes should there be one. ■

Solving Using a Different Method

If time permits, try solving the problem using a different approach. For example if you have already solved for the intersection of two parabolas algebraically, you can then graph it out using your GC to check whether your solutions align with the graph.

Estimate the Result

If the problem involves numerical answers, try estimating the answer to see if it seems reasonable. Using these approximations can catch careless mistakes if your solution happens to be far off the sensible range. This can be done either before or after solving; the key is to have a rough gauge of where your answer should lie.

Example : Estimation for Reasonableness
Consider the integral $\int_0^1 (3x^2+2)\,dx$. Without solving it exactly, we can approximate:

$$3x^2+2 \quad \text{is approximately} \quad 3(0.5^2)+2 = 2.75 \quad \text{at} \quad x = 0.5.$$

So the integral should be around $2.75 \times 1 = 2.75$. Solving it exactly gives:

$$\int_0^1 (3x^2+2)\,dx = \left[x^3+2x\right]_0^1 = 1+2 = 3.$$

The estimate was close, verifying that our exact solution is reasonable. ■

CHAPTER 2

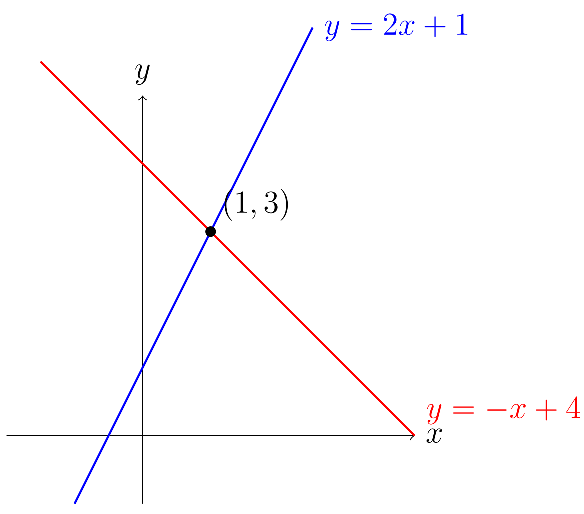
SYSTEM OF LINEAR EQUATIONS

1 Simultaneous Equations

A system of linear equations is just as its name suggests, a collection of one or more linear equations, containing the same set of variables (e.g. a, b, and c). It is commonly known as 'simultaneous equations'.

One common application is finding the point of intersection between two lines, where each line can be represented by a linear equation (i.e. $y = mx + c$). When determining the point where two lines intersect, the coordinates of this intersection point must satisfy both line equations simultaneously, thus generating the simultaneous equations.

Consider the following example, where we find the intersection point of two lines:



The equations of the two lines are given by:

$$y = 2x + 1$$
$$y = -x + 4$$

We can solve this system by equating the two equations since they both represent y:

$$2x + 1 = -x + 4$$
$$3x = 3$$
$$x = 1$$

Substitute $x = 1$ into one of the original equations to find y:

$$y = 2(1) + 1 = 3$$

*Why does it suffice to substitute x into only one equation? It is because both lines have the same y value at that x value.

Thus, the point of intersection is $(1, 3)$.

As shown in the diagram, the lines $y = 2x + 1$ and $y = -x + 4$ intersect at the point $(1, 3)$. Once again, this method is valid since the point of intersection must simultaneously satisfy the equations of both lines. The point to be made is that a system of linear equations is useful where multiple conditions have to be satisfied simultaneously.

2 Solving Approaches

1. **Simultaneous Equations:** A concept you may be more familiar with, 'simultaneous equations', is one of the common methods to solve a system of linear equations.

   One of the variables is made the subject, and substituted into another equation. The process is repeated until one variable is solved for. The other two can then be found accordingly.

2. **Using a Graphing Calculator (GC):**

   On → Apps → 4: PlySmlt2 → 2: Simultaneous Eqn Solver → Choose the number of unknowns (a.k.a. variables) and the number of equations. Input coefficients of variables in the GC matrix, lastly press 'F5: Graph' to solve.

3 Tips and Tricks

- **Identify all the relevant information from the question before listing down your equations.**

> ***Example : Identifying the equations***
> When the polynomial $f(x) = x^3 + ax^2 + bx + c$ is divided by $(x-4)$, $(x-1)$, and $(x+2)$, the remainders are 10, 15, and -5 respectively. Find the values of a, b, and c.
>
> Here, we have 3 unknowns, 3 divisors and 3 remainders. Using the remainder theorem would give us 3 variables and 3 equations, which can be solved by GC. ■

- **Most of the time, the number of equations equals the number of variables.**

Like the above example, most questions have the same number of equations as variables, which generally means they can be solved such that each variable has a specific value.

However, do not spend too long trying to find equations if you cannot. Some questions require you to use logical deductions to determine the last variable despite having fewer equations than variables. These questions usually involve finding the 'minimum' or 'maximum' value for a specific variable.

> ***Example : More variables than equations***
> Given that the real numbers x, y and z satisfy:
>
> $$5x - y + 3z = 15$$
>
> $$2x + 3y + 5z = 19$$
>
> Find the maximum value of x, provided that z is non-negative.
>
> In this problem, we have 3 unknowns but only 2 equations. Solving requires expressing x and y in terms of z. For example:
>
> $$x = -\frac{14}{17}z + \frac{64}{17}, \quad y = -\frac{19}{17}z + \frac{65}{17}$$
>
> Then, we can find the appropriate z value to maximise x. ■

- **Assign identities to the variables you are using.**

This is just a simple statement like "Let a, b, and c represent the number of apples, bananas, and cherries bought, respectively". This is not necessary when the question has already given the names to the variables.

- **Be extremely careful and look out for tiny 'additional' details.**

At times, questions tend to contain too much information, such that some small details are overlooked. Be extra cautious with these details, such as the "fixed income" in the question below.

John and Sam are drivers for a ride-hailing company and are paid \$$p$ and \$$q$ per trip on weekdays, respectively. On weekends, they are paid 30% more per trip. Additionally, John and Sam receive a fixed monthly salary, \$$m$ and \$$n$, respectively. The table below shows the number of trips made by each driver in the first four months of 2023 and the combined monthly income of both drivers. Solve for p, q, m, and n.

Month	No. of Weekday Trips		No. of Weekend Trips		Income
	John	Sam	John	Sam	Total Payout
January	15	20	5	6	6426
February	12	18	4	8	6238
March	20	22	7	5	6837.5
April	18	16	6	7	6519.5

In addition to solving for p and q, you also have to account for the monthly salaries, m and n, which are additional values given by the question. So your equations should look something like this:

$$15p + 20q + 5(1.30)p + 6(1.30)q + m + n = 6426$$

CHAPTER 3

INEQUALITIES

1 Overview

Inequalities are statements of relationships (greater than '$>$', greater than or equal to '$\geq$', and less than '$<$') between two or more numbers, or more commonly algebraic expressions.

Some Common Knowledge (Non-exhaustive):

- $x^n \geq 0$ for all real and even n.
- $|x| \geq 0$ for all real x.
- $|x| > k \Rightarrow x > k,$ or $x < -k$.
- $|x| < k \Rightarrow x < k,$ and $x > -k \Rightarrow -k < x < k$.
  *The use of 'or' and 'and' have very different meanings. This important concept will be covered in more detail on the next page.
- $e^x \geq 0$ for all real x.
- $-1 \leq \sin x, \cos x \leq 1$.
- $-\frac{\pi}{2} \leq \sin^{-1} x \leq \frac{\pi}{2}$, and $0 \leq \cos^{-1} x \leq \pi$.
- For increasing functions (i.e. $f'(x) > 0$), if $a > b$, then $f(a) > f(b)$.

2 Solving Approaches

1. **Graphing and Observation**

   Provided that the question allows, you can graph the function using your GC or by hand, and observe where the graph lies (i.e. above or below the x-axis) for each segment.

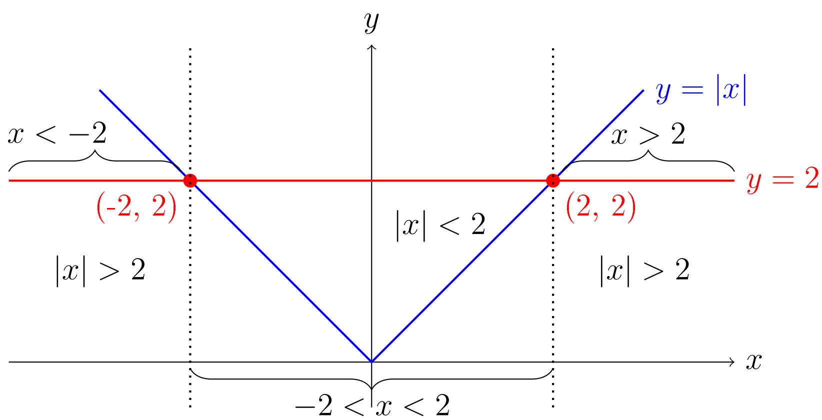
2. **Algebraic Simplification and Test-point Method**

   Most of the time, you would be required to perform algebraic manipulations to the inequality given. After reaching a simplified inequality, the test-point method can be used to determine the solutions.

3 Modulus

Modulus is a concept that stumps many and is rather confusing to understand by just looking at the algebraic expressions. As such, a visual representation would help in the comprehension of the modulus function in inequalities.

Consider the following example:

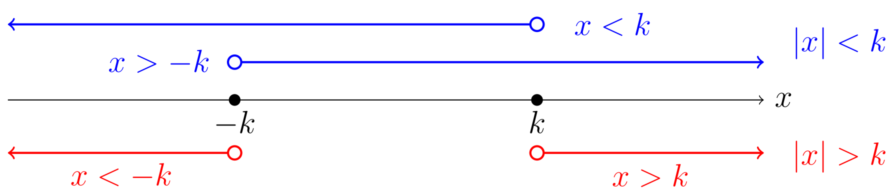


With reference to the above graph of $y = |x|$ and $y = 2$, the segment where $|x| < 2$ (below the red line) is continuous, while the segment where $|x| > 2$ (above the red line) is discontinuous. This is because it is possible for x to be greater than -2 and smaller than 2 simultaneously, while x cannot be greater than 2 while being less than -2 at the same time.

Hence, this explains the above inequalities:

$$|x| > k \Rightarrow x > k, \text{ or } x < -k$$
$$|x| < k \Rightarrow x < k, \text{ and } x > -k \Rightarrow -k < x < k$$

In a more mathematical sense, 'or' implies union, while 'and' implies intersection. Another method to visualize this is through the number line:

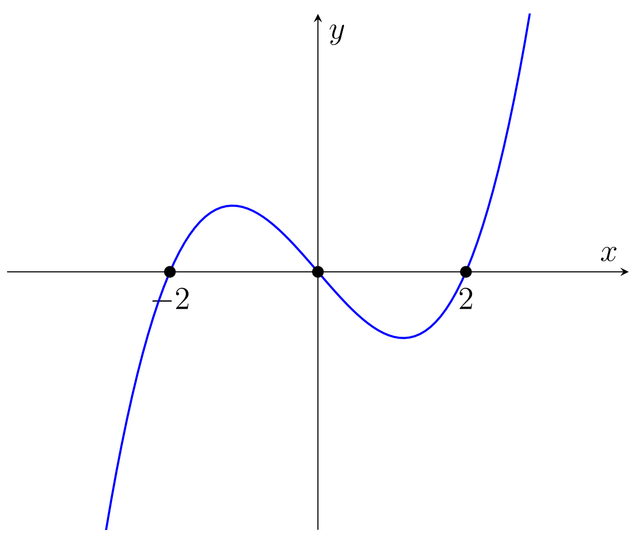
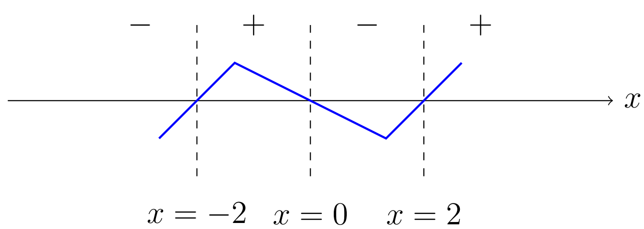


Notice how we take the union of the disjoint red lines, and the intersection of the overlapping blue lines. This illustrates the above point.

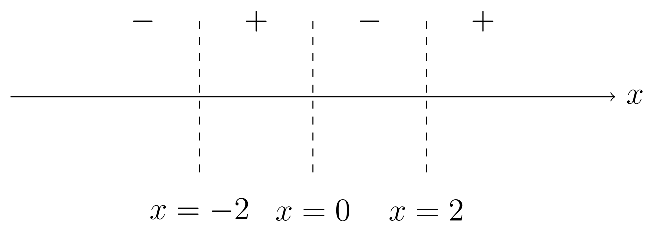
4 Tips and Tricks

- **The number line can be thought of as a graph, but with only the x-axis.**

  The only y component is the sign (i.e. $+$ or $-$) to indicate whether the function is positive, negative, or zero in the respective regions of x.

> ***Example : Abstracting away the y-axis***
> Consider the graph of $y = x^3 - 4x$.
>
> 
>
>
> The graph can then be "compressed" along the y-axis, to give a number line:
>
> 
>

Abstracting away the graph in blue, we get the number line:



This visualization aids in understanding the relationship between a cartesian graph and the number line. It also shows why the test-point method works.

- **Always shift all terms to one side before solving.**

  This reduces careless mistakes. For example:

$$\frac{2}{x+2} < \frac{3}{3-x} \Rightarrow \frac{2}{x+2} - \frac{3}{3-x} < 0$$

- **Never multiply across by $f(x)$.**

  There is a high chance that you will lose a solution. For example, do NOT do this:

$$\frac{2x^2-3x+1}{x+2} < 0$$

$$\frac{2x^2-3x+1}{x+2} \times (x+2) < 0 \times (x+2) \qquad \Longrightarrow \qquad 2x^2-3x+1 < 0$$

  By doing the above, you will lose the condition of $x \neq -2$, hence losing a solution.

  $f(x)$ refers to any term containing x. However, constants can be multiplied across as long as they are not zero.

- **Be sensitive to "impossible" values.**

  Always check for values that make the expression undefined. For example, in the inequality $\frac{2x^2-3x+1}{x+2} < 0$, $x = -2$ is an "impossible" value since it makes the denominator zero.

  Mark these values on the number line with a hollow dot.

- **Be mindful of changing domains when substituting.**

  Always consider the domain of the original term and the new term after substitution.

  > ***Example : Being careful with domains***
  > Find the set of values of x for which $\frac{6}{x+5} > x$. Hence, solve for $\frac{6}{\ln x+5}$.
  >
  > From the first part of the question, the set of x values is:
  >
  > $$\{x \in \mathbb{R} : -5 < x < 1 \text{ or } x < -6\}$$
  >
  > When substituting $\ln x$, we must remember that $\ln x$ is only defined for $x > 0$, so the solution set becomes $\{x \in \mathbb{R} : e^{-5} < x < e \text{ or } 0 < x < e^{-6}\}$. ■

- **Be very careful with what the question is asking for, specifically for *'determine the set'*.**

  Set refers to the set builder notation, which is written as: $\{x \in \mathbb{R} : a < x < b\}$.

  Explanation of the components:
  - The curly braces $\{\}$ is how we represent a set of values.
  - The first x represents all possible numbers in the set.
  - The symbol $\in$ means "is an element of."
  - $\mathbb{R}$ represents the set of real numbers.
  - The colon (:) means "such that."
  - $a < x < b$ is the range of x that was found through solving the inequality.

CHAPTER 4

CURVE SKETCHING

1 Overview

Unsurprisingly, the topic of curve sketching essentially revolves around sketching a curve. The behaviors of graphs (e.g. asymptotes, shapes, etc.) with respect to the algebraic functions are also studied in this topic.

Some Common Knowledge (Asymptotes)

- **Exponential** (e^x)
  Asymptote: $y = 0$

- **Tangent** ($\tan x$)
  Asymptote: $x = \frac{(2n-1)\pi}{2},\ n \in \mathbb{Z}$

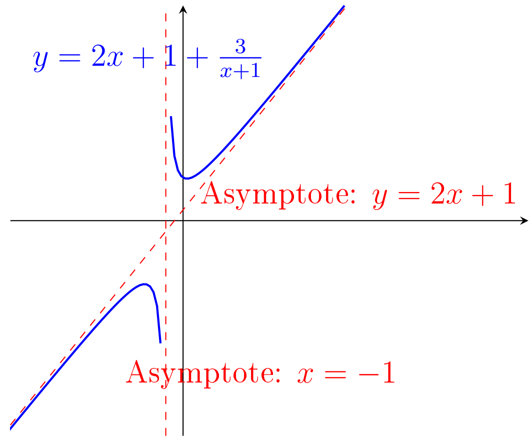
  *$2n - 1$ is a general representation of an odd number.

- **Logarithmic** ($\ln x$)
  Asymptote: $x = 0$

- **Rational function** $\left(y = ax + b + \frac{c}{g(x)}\right)$
  Asymptotes: $y = ax + b,\ g(x) = 0$

> ***Example :*** $y = 2x + 1 + \frac{3}{x+1}$

Vertical Asymptote: $x = -1$ Oblique Asymptote: $y = 2x + 1$



2 Solving Approaches

1. **Plot graph on GC and copy.**

   This is the go-to if you are provided with the equation with no unknowns, or are unsure of how to plot the graph yourself.

2. **Plot by hand (for functions with unknowns, etc.).**

   Before rushing to plot the graph, make sure you:
   - Determine all the asymptotes (if there are any).
   - Calculate the axial intercepts.
   - Deduce the general shape of the graph if possible.
   - Be clear about the domain of the function.

   Then, and only then, do you proceed to sketch the graph.

   When plotting a graph, be sure to label the following (if applicable):
   - Axes
   - Coordinates of axial intercepts
   - Coordinates of turning points
   - Coordinates of start and end points
   - Coordinates of vertices

- Coordinates of intersections
- Equation of the function (except for parametric equations)
- Equations of asymptotes

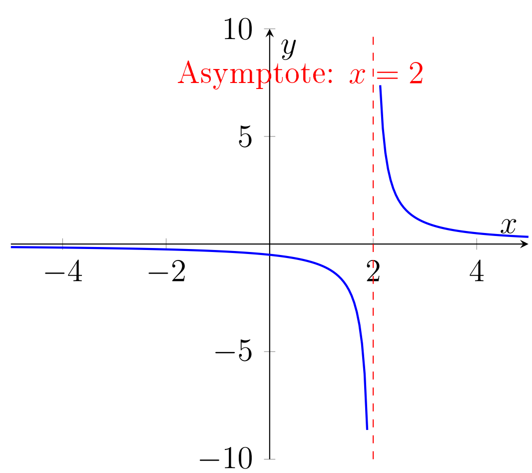
3 Visualizing Asymptotes

- **Vertical Asymptotes:**

Vertical asymptotes are lines where the function becomes undefined and the value of $f(x)$ grows without bound (positive or negative infinity). These generally occur when the denominator of a function equals zero.

Tip: Look for the points where the denominator equals zero. At these points, the function shoots up or down toward infinity.

Example : $y = \frac{1}{x-2}$
In the function $y = \frac{1}{x-2}$, the vertical asymptote occurs at $x = 2$, where the denominator becomes zero and the function explodes to infinity.



As x approaches 2 from the left, y tends toward $-\infty$, and from the right, y tends toward ∞.

- **Horizontal Asymptotes:**

Horizontal asymptotes describe how a function behaves as x gets larger and approaches infinity or negative infinity. If the function approaches a horizontal line and flattens out, this horizontal line is known as the horizontal asymptote.

Tip: Evaluate the function as $x \to \infty$ and $x \to -\infty$. For rational functions:

- If the degree of the denominator is greater than the degree of the numerator, the horizontal asymptote is $y = 0$.

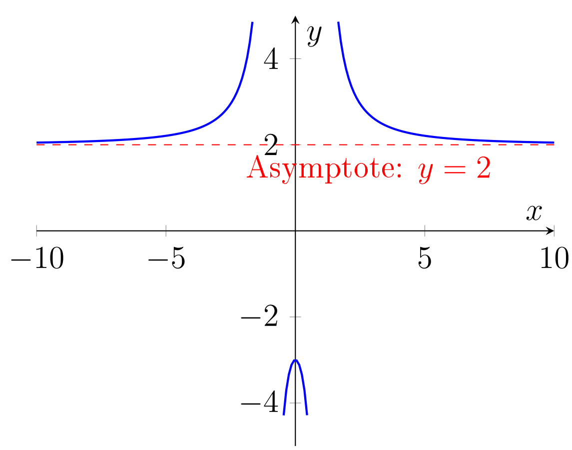
- If the degrees are equal, the horizontal asymptote is the ratio of the leading coefficients.

Usually we are only concerned with the leading terms (terms of the highest x power) since it will be significantly larger than all other subsequent terms when x gets large.

For example when $x = 1000$, $3x^3 + 2x = 3000002000$ which is almost the same as $3x^3 = 3000000000$ (in fact, the percentage error is 0.000067% which is almost negligible!). This error will become smaller as x gets larger, thus we can effectively ignore all smaller powers when evaluating large x.

The converse is true when we look at very very small x: we ignore high powers since they shrink much quicker than lower powers.

Example : $y = \frac{2x^2+3}{x^2-1}$
In this function, as $x \to \infty$, the horizontal asymptote is determined by the ratio of the leading coefficients, which gives $y = 2$.



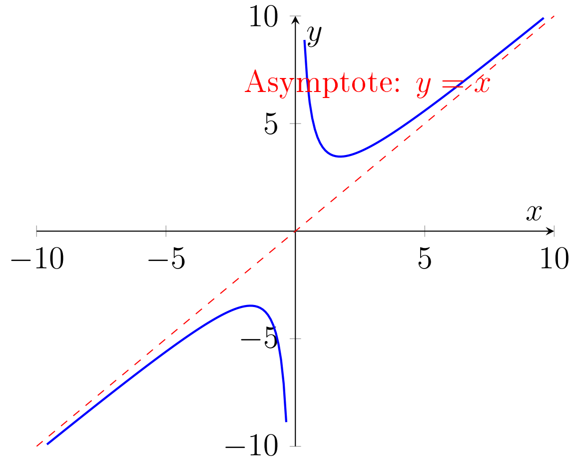
As $x \to \infty$, the function flattens out to the horizontal line $y = 2$.

- **Oblique Asymptotes:**

Oblique asymptotes, also known as slant asymptotes, occur when the degree of the numerator is exactly one more than the degree of the denominator. In this case, the function approaches a straight-line path as x becomes large, rather than a horizontal asymptote.

Tip: Perform polynomial long division to find the equation of the oblique asymptote.

Example : $y = \frac{x^2+3}{x}$
Performing long division on $\frac{x^2+3}{x}$, we get $y = x + \frac{3}{x}$. So the oblique asymptote is $y = x$.



As $x \to \infty$, the function behaves more and more like $y = x$. This is because the remaining term $\frac{3}{x}$ becomes insignificant in comparison to x.

4 Tips and Tricks

- **Do not trust the GC entirely.**

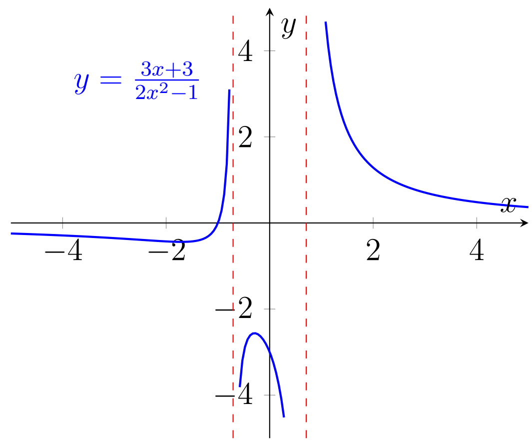
  Always read your GC with caution. The GC only provides an approximate value, but it may not be exact, accurate, or precise.

  Use your mathematical intuition when approaching a graph (e.g. general shape of the graph, presence of asymptotes and intercepts, value of $f(x)$ as $x \to \infty$, etc.). You can develop this by plotting more graphs by hand. ("Desmos" is a good graphing site for practice).

- **Zoom in zoom out.**

  To be safe, zoom out to see the general shape of the graph, then zoom in to observe the tiny details

Example : $y = \frac{3x+3}{2x^2-1}$
It may be difficult to notice the existence of 2 turning points without zooming in.



There is a minimum point near $x = -1$.

Note: Some rational graphs (like the above) have two turning points; always zoom in to confirm.

- **Read the question carefully and take note of the domain.**

  Especially for functions with specified domains, do not draw the graph beyond the given domain.

  For parametric equations, input your domain into your GC using the 'Windows' function.

- **Scale your axes properly as much as possible.**

  Ideally, the two axes should be proportionally scaled. A poor scale leads to a higher chance of misinterpretation and careless mistakes, hence it is advisable to avoid.

- **Observe your graph carefully, it might direct you to solve subsequent parts.**

CHAPTER 5

FUNCTIONS

1 Overview

Functions are mathematical relations that map each element in one set (called the domain) to exactly one element in another set (called the range).

2 Key Concepts and Common Mistakes

- **Functions vs Images**

  f is a function while $f(x)$ is the value of the function at x. It is important to note that while f can be a one-to-one function, $f(x)$ itself cannot be one-to-one. The distinction is that f refers to the entire function, while $f(x)$ is simply a value or output of the function.

- **Graph of a Function**

  A graph of $y = f(x)$ is the set of all images $f(x)$. Therefore, we can have a graph of $y = f(x)$, but not of $y = f$, because f represents the function itself, not its output values.

- **Inverse Functions and the Horizontal Line Test**

  Only one-to-one functions can have an inverse function. The existence of an inverse is tested using the horizontal line test. (Why does the horizontal line test work?)

- **Reflections of Functions and Their Inverses**

The graphs of $y =$ and $y = f^{-1}(x)$ are reflections of each other in the line $y = x$. This symmetry occurs because the inverse function f^{-1} essentially reverses the mapping of the original function f. So $f^{-1}(f(x)) = x$.

- **Composite Functions**

A composite function, denoted as $fg(x)$, means that the function g operates first, and then the function f operates on the result of $g(x)$. This order is crucial, since it dictates the few following results on composites.

Because g operates first, the domain of the composite function fg is the domain of g.

- **Existence of Composite Functions**

The existence of the composite function fg requires that the range of g (R_g) is a subset of the domain of f (D_f). This is because g operates first, passing on its output to f afterwards, hence f must be able to accommodate whatever g outputs. This ensures that f can operate on the output of g.

- **Range of Composite Functions**

The range of the composite function fg is determined by limiting the domain of f to the range of g. Therefore, the range of fg depends on both the functions f and g.

- **Note on Function Composition**

It is important to note that ff^{-1} may not always be the same as $f^{-1}f$. This is because the composition of functions is not necessarily commutative, meaning that the order in which functions are composed matters. This will be covered in more detail later on.

- **Learn to Sketch**

Whenever you are stuck, sketch! Even a simple sketch can provide insights, and chances are that you might gain some inspiration and be able to continue afterward.

3 Domain and Range

The importance of domain and range cannot be overemphasized in the topic of functions. Here are some cases where the domain (D) and range (R) are crucial:

1. **Drawing any graph:**

   When drawing the graph of a function, knowing the domain and range helps you understand the limits of the function and where to plot points. For instance, if the function is not defined for certain values, you can avoid plotting beyond those points.

2. **Determining the inverse of a square function:**

   For a function such as $y = (x + 1)^2 + 3$, the inverse function only exists for restricted domains. For example, if $y \leq 1$, the inverse of this function exists within that restricted range, ensuring the function remains one-to-one.

   Example : Inverse of $y = (x + 1)^2 + 3$ ***for*** $x \leq -1$

   $$y - 3 = (x + 1)^2 \quad \Rightarrow \quad x = \pm\sqrt{y - 3} - 1$$

   To find the inverse function required, we have to consider the original domain $x \leq -1$

   Since $\sqrt{y - 3} \geq 0$, the only way $x \leq -1$ can be satisfied is if we take

   $$x = -\sqrt{y - 3} - 1$$

   Thus

   $$f^{-1}(y) = -\sqrt{y - 3} - 1 \;\Rightarrow\; f^{-1}(x) = -\sqrt{x - 3} - 1$$

   ***Important side note:** Only change the variable from y to x at the final step to avoid confusion. The original domain restriction applies to x, and changing the variable too early can cause misunderstandings about whether the restriction applies to the original function or the inverse.

3. **Composite functions:**

   The domain of a composite function $fg(x)$ depends on both the domain of $g(x)$ and the range of $g(x)$. If $g(x)$ maps outside the domain of $f(x)$, the composite function cannot exist at those points.

4. **Graphing ff^{-1} and $f^{-1}f$:**

   It's important to note that ff^{-1} and $f^{-1}f$ are not always the same. The graph of ff^{-1} typically maps back to the identity function on the range of $f(x)$, while $f^{-1}f$ maps back to the identity function on the domain of $f(x)$.

   $R_{f^{-1}f} = D_f$ since f operates first.
   $R_{ff^{-1}} = D_{f^{-1}}$ since f^{-1} operates first.

   So if $D_f \neq D_{f^{-1}}$ then $R_{f^{-1}f} \neq R_{ff^{-1}}$.

   > ***Example : Graphing ff^{-1} and $f^{-1}f$***
   > Consider the function $f(x) = \ln x$ for $x > 5$, so $f^{-1}(x) = e^x$ for $x > \ln 5$. Contrast the two:
   > $$f(f^{-1}(x)) = x \quad \text{for } x > \ln 5$$
   > $$f^{-1}(f(x)) = x \quad \text{for } x > 5$$
   > Notice that the domain of the two composites are different, although both give the same rule of x. This is due to the difference in domain of the "inner functions". ■

4 Why does reflecting a function give its Inverse?

In the H2 syllabus, you are only required to know that reflecting a function gives its inverse; you need not know how it works. But if you are interested, this section offers a unique perspective to answer this question. I have tried my best to express the ideas coherently; read it carefully, but feel free to skip ahead if you do not understand my points.

Inverse functions are functions that reverse the mapping of a given function. We know that the graph of the inverse function $y = f^{-1}(x)$ can be obtained by reflecting the graph of $y = f(x)$ in the line $y = x$, but why does reflection work?

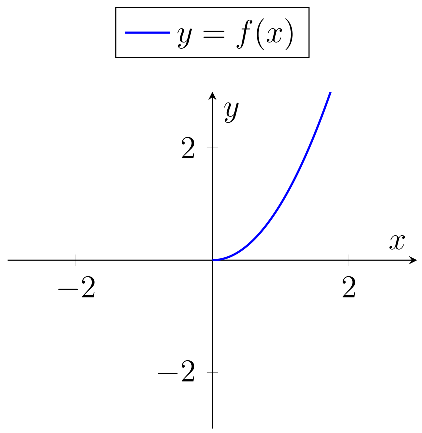
- **Swapping of Axes:**

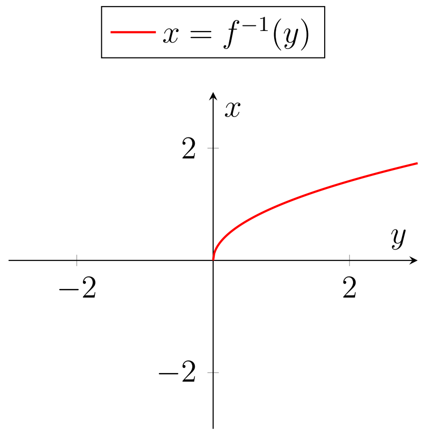
  When we talk about inverse functions, we are swapping the roles of the input (domain) and output (range). If a function $f(x)$ takes an input x and maps it to an output y, the inverse function $f^{-1}(x)$ does the opposite: it takes y and maps it back to x. In this sense, we are switching the roles of x and y.

  This swapping can be visualised as a change in axes, where x and y swap places with each other. Pay close attention to this idea of swapping.

- **Taking on Two Different Perspectives:**

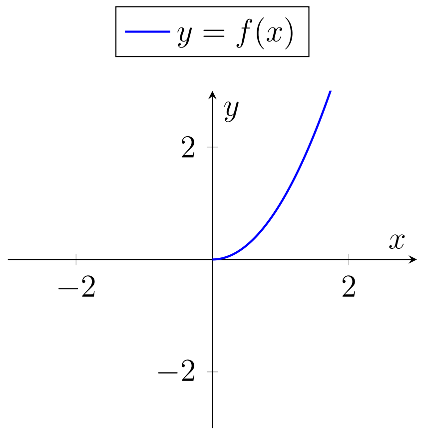
Consider the following two graphs. One represents the usual graph of $y = f(x)$, and the other is the same graph but with the axes swapped; we are essentially looking at the same graph from two different perspectives. Both graphs are mathematically identical, but the labeling of the axes is different.

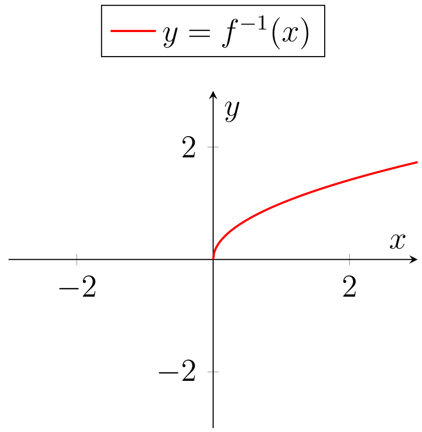




In the left graph, we see the function $y = f(x) = x^2$ for $x \geq 0$. On the right, we have swapped the axes so that y is on the horizontal axis and x is on the vertical axis. This corresponds to the equation $x = f^{-1}(y) = \sqrt{y}$; we have simply shifted f from the left hand side to the right hand side. Though they are written differently, both still represent the same equation mathematically.

Notice that in the red graph, if we were to swap the positions of x and y in both the equation and on the axes, we will get the exact graph of $y = f^{-1}(x)$ (illustrated below). This is exactly the inverse function that we set out to find. Once again, the above pair are both the same function, while the bottom pair are inverses of each other.





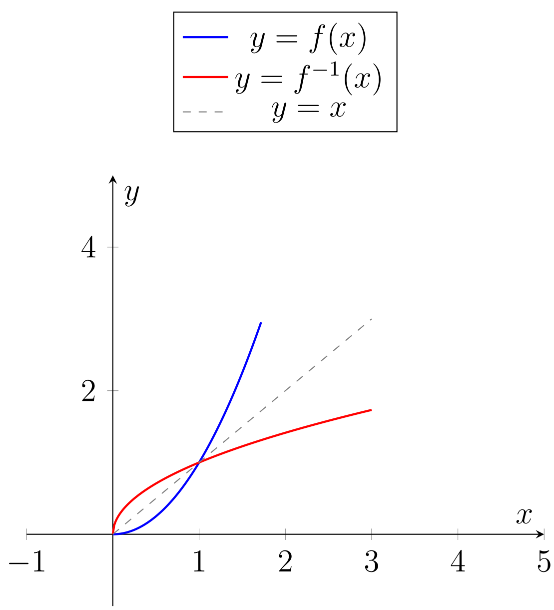
- **Why Reflections Work:**

  The reflection came in the first step when we took on two different perspectives at the same function, after which we changed the position of the axes. This had the overall effect of reflecting the graph of $y = f(x)$ in the line $y = x$. While this is a shortcut we often use, I see value in understanding the mechanism behind this concept: the swapping of x and y.

  An alternate perspective is that when we reflect in the line $y = x$, the coordinates of each point are swapped. For example, a point (a, b) on the graph of $y = f(x)$ corresponds to the point (b, a) on the graph of $y = f^{-1}(x)$. Again, this leads us back to the concept of swapping x and y, the very essence of inverse functions.

  Example : Inversing the coordinates
  Let $f(x) = x^2$ for $x \geq 0$. The inverse function is $f^{-1}(x) = \sqrt{x}$.

  In the original graph of $f(x)$, the point $(2, 4)$ means that $f(2) = 4$. On the inverse graph, the point $(4, 2)$ corresponds to $f^{-1}(4) = 2$. Swapping these points reflects the reversal of the input-output relationship between the function and its inverse.



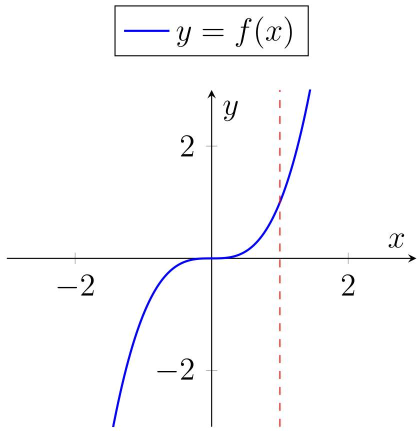
  The point $(2, 4)$ on $y = f(x)$ reflects to the point $(4, 2)$ on $y = f^{-1}(x)$.

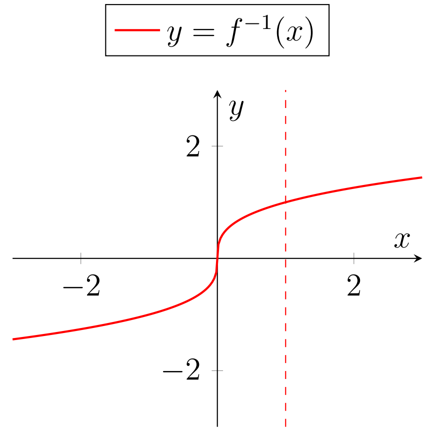
5 An Interesting Perspective on the Horizontal Line Test

We know that the horizontal line test checks if a function is one-one, and we know that only one-one functions can have inverses. This is a perfectly valid explanation of why the horizontal line test works. However, in this section I would like to take on a different perspective to "discover" the horizontal line test.

We know from previous results that some functions can be reflected in the line $y = x$ to produce their inverse. However, we need to verify if the new, reflected graph is a valid function. To do this, we will reflect $y = f(x)$ and test whether the resulting graph is a proper function by running the vertical line test.

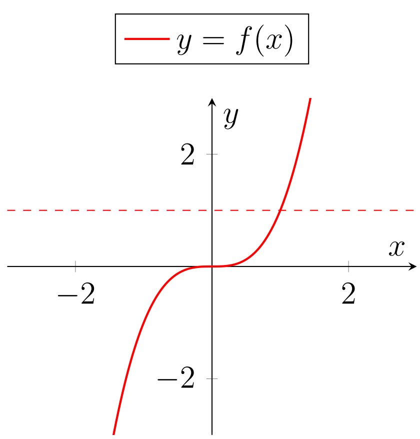
Consider the graph of a one-to-one function. Below is the graph of $y = f(x)$, and next to it is the graph of $y = f^{-1}(x)$, obtained by reflecting $y = f(x)$ in the line $y = x$.



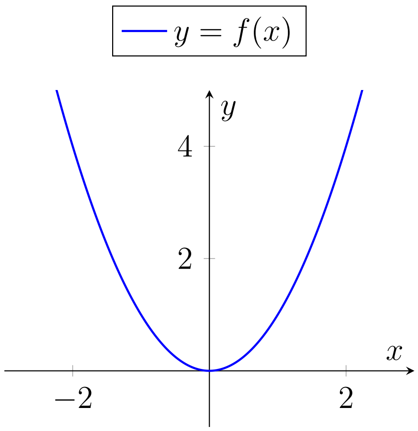


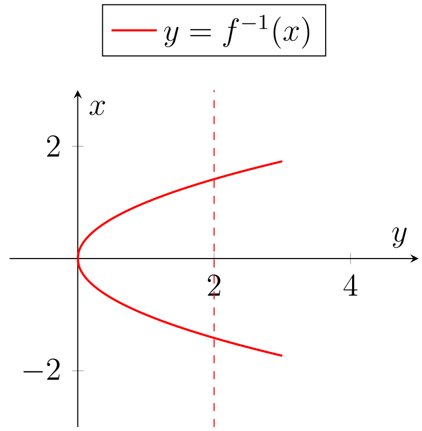
In this case, the vertical line test is satisfied in both the original and reflected graphs, meaning both the function f as well as the inverse f^{-1} exist, and are both valid functions.

Now notice that when we reflect the red graph along with the vertical line back to its original position as illustrated below, the graph now represents $y = f(x)$ again, while the vertical line now becomes a horizontal line. This is what we call the horizontal line test, since it tests if the inverse function is valid (whether it exists).



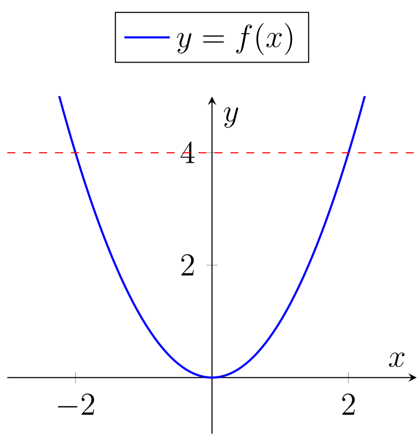
Now, we consider a case where the horizontal line test fails. Depicted below is the graph of $y = x^2$; it can be reflected in the line $y = x$ to give the graph of the supposed "inverse" function.





However, we see that the vertical line test fails on the reflected graph (in red) indicating that $f^{-1}(x)$ is not a valid function. Thus, the inverse of $f(x) = x^2$ does not exist.

This failed vertical line test can be reflected back to show a failed horizontal line test as shown below. Once again, we see that the horizontal line test is simply a reflection of the vertical line test in the line $y = x$.



Summing up the above: By reflecting a function in the line $y = x$, we obtain a candidate for its inverse. To check if this inverse is a valid function, we use the vertical line test on the reflected graph. When we reflect the graph back to its original form, the vertical line transforms into a horizontal line on the original graph. This gives us the horizontal line test.

Important Tip: The horizontal line is conducted on the graph of $y = f(x)$, and not $y = f^{-1}(x)$.

CHAPTER 6

GRAPH TRANSFORMATIONS

1 Transformations

Graph transformations deal primarily with three types of changes: translation, stretching, and reflection. As their names suggest, they transform the original graph into a similar shape, even though the graph may be re-positioned or scaled.

Transformations are different from deductions (next chapter) whereby the latter derives a different graph from the original graph. These two graphs may not even look alike.

Translation: This involves shifting the graph, usually associated with the addition or subtraction of a constant. It changes the position of the graph without altering its shape.

Stretching: This distorts the graph, usually associated with multiplication or division by a constant. Stretching can either expand or compress the graph along the x- or y-axis.

Reflection: This involves flipping the graph about an axis, usually associated with a change in polarity (a negative sign). Reflections occur over the x-axis or y-axis.

2 Tips, Tricks, and Ideas

- **Understanding through axes relocation**

  While not directly important for solving problems, you can think of transformations as changing the axes or relocating them. This aids in understanding but is not required content-wise. More details on this idea will be covered in the next section.

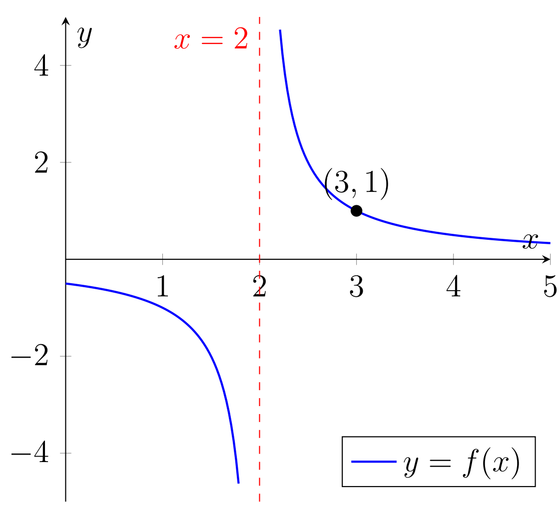
- **Changes in coordinates**

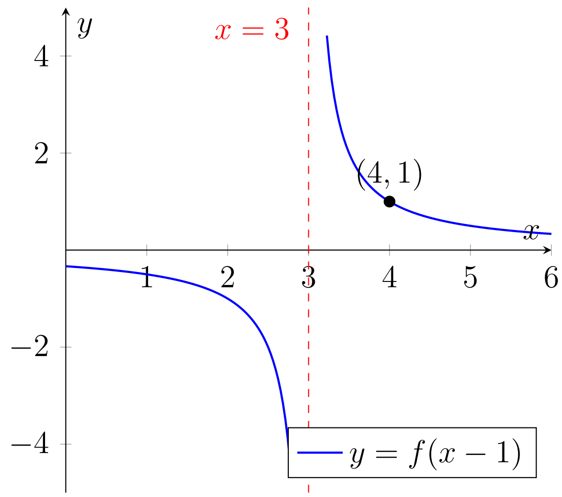
Another useful way to think about transformations is by focusing on the changes in the coordinates of points, rather than the overall shape of the graph. This is just a thought process—avoid presenting it this way during exams.

> ***Example : Transformation from*** $y = f(x)$ ***to*** $y - 4 = f(x)$
> Consider points $A(0,4)$, $B(-3,0)$, and $C(4,0)$ on the graph of $y = f(x)$. To describe the transformation to $y - 4 = f(x)$, we can analyze the coordinate changes.
>
> Initially, each point is $(x, y) = (x, f(x))$ since $y = f(x)$. After the transformation, the final points are $(x, y) = (x, f(x)+4)$, since $y = f(x)+4$ after the transformation. This means we add 4 to each y-coordinate.
>
> Thus, the transformed points are:
>
> $$A'(0,8), \quad B'(-3,4), \quad C'(4,4)$$

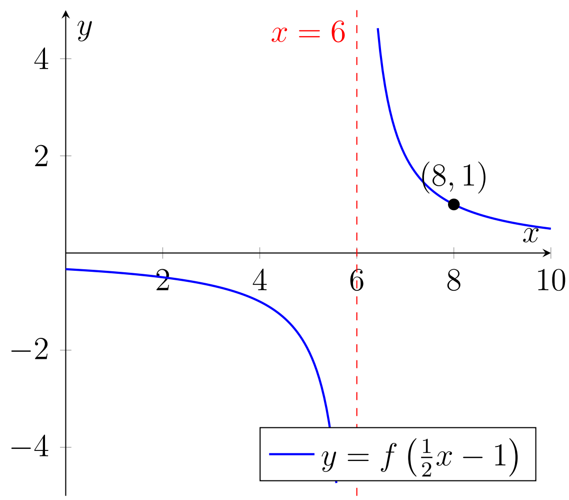
- **One step at a time**

Always go step-wise with transformations, especially when multiple transformations are involved. There's no need to rush unless you're extremely confident.

> ***Example : Transformation from*** $y = f(x)$ ***to*** $y = f\left(\frac{1}{2}x - 1\right)$
> Let $f(x)$ be a rational function with a vertical asymptote. We perform the transformation in two steps:
>
> 
>
>
> **Step 1:** Replace x with $x - 1$, which translates the graph horizontally.



Step 2: Replace x with $\frac{1}{2}x$, which stretches the graph horizontally by a factor of 2.



Note that the order of transformations matter: It would be incorrect to first replace x with $\frac{1}{2}x$ then replace x with $x - 1$.

- **Use completing the square for quadratic transformations**

For transformations involving quadratic functions, completing the square simplifies the process.

Example : Transformation from $y = x^2$ ***to*** $y = x^2 + 2x - 5$
We first complete the square:

$$y = x^2 + 2x - 5 \quad \Rightarrow \quad y = (x+1)^2 - 6$$

Now the transformation becomes simpler. First, replace x with $x + 1$ (translation), followed by replacing y with $y+6$ (translation). ■

- **Separate y-transformations from x-transformations**

  It's a good practice to keep the transformations of y and $f(x)$ separate. Manipulate the equation until you have $f(ax+b)$ on the right-hand side, with all other terms on the left. This makes it easier to identify the transformations.

  Example : Transformation from $y = \frac{1}{2}f(x+1) - 6$
  First, manipulate the equation:

  $$2y + 12 = f(x+1)$$

  Now the transformations are clear:
  1. Replace x with $x+1$ (translation).
  2. Replace y with $y+12$ (translation).
  3. Stretch the graph vertically by a factor of $\frac{1}{2}$.

  ■

- **For replacement of x, only replace x, not the entire expression**

  When replacing x, ensure that you are only replacing the x-term itself, and not the entire expression $(ax+b)$.

  Example : Transformation involving $f(x+1)$
  If you replace x with $2x$, $f(x+1)$ becomes $f(2x+1)$, not $f(2x+2)$. ■

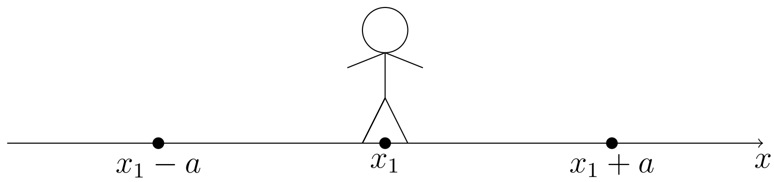
3 Visualising Transformations

As a student, one concept I struggled quite a bit with was the counter-intuitive nature of graph transformations: Why does replacing x with $x-a$ shift the graph in the positive direction? Intuitively, if we are subtracting a, shouldn't we be moving in the negative direction instead?

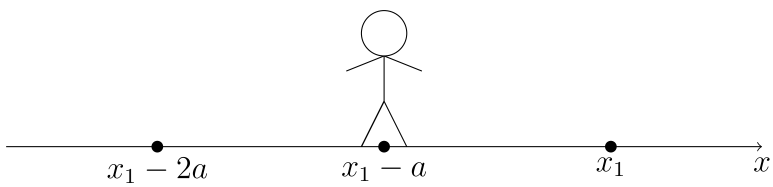
To break this down, let's look at how the transformation works and why this happens:

Replacement on the x-axis

Imagine you are standing on a value x_1 on the x-axis. The point $x_1 - a$ is "a" units behind you, and $x_1 + a$ is "a" units in front of you.



Now we execute the replacement by replacing x with $x - a$. This also means replacing every x-coordinate x_i by $x_i - a$: So the point $x_1 - a$ becomes $x_1 - 2a$, x_1 becomes $x_1 - a$, and $x_1 + a$ becomes x_1.



Notice that even though you have not moved, you are now standing on $x_1 - a$; meaning every point on the axis has effectively shifted right by "a" units.

This means that replacing x with $x - a$ causes your graph to be translated in the positive x-direction by "a" units.

Understanding it as a Delay

Think of $f(x - a)$ as a function where every input x is "delayed" by a. You now need to move a units further along the x-axis to achieve the same output $f(x)$ had at x.

The confusion arises from thinking that $x - a$ implies moving left. But remember, the transformation happens to the input, not the output. Replacing x with $x - a$ delays the function's response by a units, causing the graph to shift in the positive direction to make up for this delay.

These visualisations can also be applied to other transformations like scaling or reflection. Always think of transformations as moving the points relative to where they used to be. A graph is, after all, just a collection of many points.

CHAPTER 7

GRAPH DEDUCTION

Though similar to graph transformation, graph deduction is not exactly the same. The former performs changes on the original graph, while the latter creates a new graph by deducing certain features from the original graph.

In the H2 syllabus, there are only three types of graphs you need to know how to deduce: Modulus graphs (i.e. $y = |f(x)|$ and $y = f(|x|)$), reciprocal graphs (i.e. $y = \frac{1}{f(x)}$), and the derivative graph (i.e. $y = f'(x)$).

The steps required to produce the desired graph will not be covered extensively, as those are already detailed in your H2 notes. Instead of flooding you with graphs and examples like the previous chapter, here we will explore the further intuition that some students may not have fully developed.

True understanding would help you to deduce any combinations of the above without memorising, with the flexibility to accommodate additional transformations if the question requires.

1 Modulus Graphs

The two main types of modulus graphs are $y = |f(x)|$ and $y = f(|x|)$, which behave very differently. Recall from functions that given an equation like $y = f(x)$, x is the input while $f(x)$ is the output.

For $y = |f(x)|$, the modulus is applied on the output (y-coordinates), while for $y = f(|x|)$ the modulus is applied on the input (x-coordinates).

For $y = |f(x)|$:

As mentioned, here we are computing the function with its usual domain, however the range has changed due to the added modulus outside $f(x)$. So where the original function may have had negative outputs, these outputs now become positive due to this outer modulus.

In other words, any negative y-values of the original function will be reflected into positive values while keeping all positive values the same. The modulus literally "reflects" the negative portion of the graph upwards into the positive region.

For $y = f(|x|)$:

Since the modulus is now applied inside the bracket, the modulus actually operates on x before the function f can even access x. In other words, this is a composite function like fg where the above case would be gf.

Firstly notice that when the modulus is applied to positive x values, nothing happens. f takes in the same (positive x) inputs and produce the same outputs, and business is as usual. So we should expect the graph for the positive x values (rightward of the y-axis) to be untouched.

What happens to the negative x values then? They first undergo the modulus operation to become positive, and then are passed into f as positive inputs. So, they produce the exact same outputs as their positive counterparts. For instance, $x = -1$ produces the same output $f(1)$ as $x = 1$.

In summary, the expression $f(|x|)$ implies that the function treats both the positive and negative values of x the same way since the modulus of x is always non-negative.
(The modulus is not "always positive" because it can equal zero, which is not positive.)

Visually, the portion of $y = f(x)$ where $x \geq 0$ is mirrored onto the negative x-axis.

2 Reciprocal Graphs

The reciprocal of something, in layman terms, is to take "1 over something". For instance the reciprocal of 3 is simply $\frac{1}{3}$; and the reciprocal of 0 is undefined for obvious reasons.

To deduce the graph of $y = \frac{1}{f(x)}$ from $y = f(x)$, it is helpful to realise that the reciprocal function operates on the outputs, and not the inputs. This is again, a composite function. The latter case would be $y = f(\frac{1}{x})$ which is rather difficult to deduce, hence excluded from the syllabus.

It is also useful to recognise some key behaviors of the reciprocal function.

The Special Number 1

1 is the only number invariant under the reciprocal function, as $\frac{1}{1} = 1$. Of course, -1 as well, but for simplicity we summarise both as its magnitude, which is 1.

Note: For the rest of this section, "larger" or "smaller" would refer to the comparison of magnitudes, so 3 is smaller than 5 and -3 is smaller than -5. Take note that this is not the case usually, but only when referring to magnitudes; under normal circumstances, -3 is said to be larger than -5 since it is less negative.

Under the reciprocal function, all numbers larger than 1 become smaller than 1, while all numbers smaller than 1 become larger than 1. To illustrate, 2 is larger than 1 but its reciprocal $\frac{1}{2}$ is smaller than 1.

When deducing the graph of $y = \frac{1}{f(x)}$, an imaginary horizontal line $y = 1$ can be plotted in your head: All points above this line will now go below, while all points below go above. Since the reciprocal function reverses the order of magnitudes, the graphical implications are as such: All maxima become minima, while all increasing portions become decreasing; vice-versa.

Yet there remains one thing unchanged under the reciprocal function, that is, the polarity. So we should expect all positive portions (above the x-axis) to stay above in the new graph.

The Reciprocal of Zero

What happens at the x-intercepts of the original graph? We take $\frac{1}{0}$ which is not defined so we represent it using a vertical asymptote. In the same vein, all vertical asymptotes then become x-intercepts.

How about those points near the x-intercept? For very small values of $f(x)$, such as $f(x) = 0.001$, taking its reciprocal gives a large number $\frac{1}{f(x)} = 1000$; as $f(x)$ gets smaller and smaller, it is no surprise that $\frac{1}{f(x)}$ explodes to infinity.

Special care must be taken with negative values, since the reciprocal of small negative numbers like $f(x) = -0.001$ gives large negative values like $\frac{1}{f(x)} = -1000$. So as the graph approach the x-axis from the below the axis, the reciprocal graph then explodes to negative infinity.

3 Derivative Graphs

The graph of the derivative, $y = f'(x)$, describes the gradient of the original graph $f(x)$ at every point. There isn't much more intuition than what you already know:

Stationary points of $y = f(x)$ correspond to $f'(x) = 0$, so the new graph would have an x-intercept wherever $f(x)$ is stationary. If $f(x)$ is increasing, $f'(x)$ will be positive, and if $f(x)$ is decreasing, $f'(x)$ will be negative. The steeper the slope of $f(x)$, the larger the magnitude of $f'(x)$.

4 Two Practical Tips

Partition the Regions: One way to prevent careless mistakes is to section the regions on your graph according to their behaviours, so we group regions that are alike. Partitions are typically vertical lines that separate the regions according to x value. This depends mainly on the type of graphs we are looking at.

For example for derivative deductions, we should partition the regions where the graph is strictly increasing, then when the graph is strictly decreasing. Similar partitions can be applied for reciprocals.

This breaks the original problem down into smaller problems, allowing us to isolate each region and sketch the required graph, part by part.

Derivative Graphs: One of my favourite tricks for drawing derivative graphs is to use a ruler as a "tangent line". I would place my ruler tangent to the leftmost point of the graph, and begin tracing my ruler along the graph.

Throughout, the ruler remains tangent to every point; this traces out the behaviour of the gradient. If the tangent points upwards, we know the derivative graph must be positive. If the tangent gets steeper, we assign a larger value to the derivative. Lastly, horizontal tangents indicate an x-intercept as mentioned.

CHAPTER 8

SERIES AND SEQUENCES

1 Overview

A sequence is a set of ordered terms, usually denoted as $u_1, u_2, u_3, \ldots$. Each term u_n corresponds to the n-th element of the sequence.

A series is the sum of terms of a sequence, expressed as $u_1 + u_2 + u_3 + \cdots$. More formally, we are summing terms from a sequence to form a new value.

Important Note: The convergence of a sequence, defined as $\lim_{n\to\infty} u_n$, is not the same as the convergence of a series, which is defined as $\lim_{n\to\infty} \sum u_n$.

2 Important Results

- The relationship between a term u_n of a sequence and its sum:

  $$u_n = S_n - S_{n-1}$$

  where S_n is the sum of the first n terms of the series.

- The existence of a finite limit $\lim_{n\to\infty} u_n$ implies that the sequence converges. Otherwise, the sequence diverges.

- Convergence of series implies convergence of sequence. However convergence of sequence may not necessarily imply convergence of series.

3 Tips for Exams

Presentation is key in this chapter. How you lay out your work and present the limits can make a difference in communicating your understanding of convergence and divergence.

- **Convergence of Sequences**

  When checking for convergence, always begin with the phrase "As $n \to \infty$," followed by how the term is affected as n increases. Finally, state the limit of the entire expression if it exists.

  Example : Convergence of $S_n = 1 - \frac{1}{(n+1)!}$
  As $n \to \infty$, the factorial $(n+1)!$ grows extremely large, so the reciprocal $\frac{1}{(n+1)!}$ approaches zero.

  Therefore, S_n approaches 1 as $n \to \infty$.

  $$\lim_{n\to\infty} S_n = 1$$

- **Finding Limits: Intuition and Rigour**

  By intuition: When dealing with limits of sequences, a general rule of thumb is that if the denominator grows faster than the numerator, the limit of the ratio will tend towards zero.

  Mathematically, this would be when the degree of the numerator is larger than that of the denominator. So when n gets larger and larger, the denominator would grow much faster.

  Also, only the leading terms (largest power) have to be considered. This is because as n gets very large, the subsequent terms are much smaller than the leading term; we can hence neglect them due to their insignificant size.

  Example : Intuitive Approach
  Consider the following limits as $n \to \infty$:

  $$\lim_{n\to\infty} \frac{3n+1}{2n^2-9} = 0$$

  $$\lim_{n\to\infty} \frac{3n+1}{2n-9} = \frac{3}{2}$$

By rigour: To find the limit rigorously, divide both the numerator and the denominator by the highest power of n in the denominator.

This simplifies the expression by leaving a select few constant terms, with the rest being reciprocals in n (i.e. powers of $\frac{1}{n}$) meaning they tend to zero as $n \to \infty$.

Example : Rigorous Approach
Find the limit of $\frac{3n^2+2n}{5n^2+n+1}$ as $n \to \infty$.

To find the limit as $n \to \infty$, we divide both the numerator and denominator by n^2, as it is the highest power in the denominator:

$$\frac{3n^2 + 2n}{5n^2 + n + 1} = \frac{3 + \frac{2}{n}}{5 + \frac{1}{n} + \frac{1}{n^2}}$$

As $n \to \infty$, the terms $\frac{2}{n}$, $\frac{1}{n}$, and $\frac{1}{n^2}$ approach zero, leaving:

$$\lim_{n\to\infty} \frac{3 + (0)}{5 + (0) + (0)} = \frac{3}{5}$$

CHAPTER 9

ARITHMETIC AND GEOMETRIC PROGRESSIONS

1 Important Results

Arithmetic and Geometric Progressions are two special types of sequences. It is important to remember that they are sequences and not series, unless the question explicitly refers to an "arithmetic series" or "geometric series".

Arithmetic Progression (AP):

An AP is defined by:

$$u_n - u_{n-1} = d$$

where d is a constant known as the common difference. The sum of the first n terms of an AP is given by:

$$S_n = \frac{n}{2}\left[2a + (n-1)d\right]$$

where a is the first term and d is the common difference.

Geometric Progression (GP):

A GP is defined by:

$$\frac{u_n}{u_{n-1}} = r$$

where r is a constant known as the common ratio. The sum of the first n terms of a GP is given by:

$$S_n = \frac{a(r^n - 1)}{r - 1} \quad \text{for } r \neq 1$$

or equivalently,

$$S_n = \frac{a(1 - r^n)}{1 - r} \quad \text{for } r \neq 1$$

The sum to infinity for a converging geometric progression (where $|r| < 1$) is given by:

$$S_\infty = \lim_{n \to \infty} S_n = \frac{a}{1 - r}$$

2 Derivations

Arithmetic Series

The sum of the first n natural numbers is a classic example of an arithmetic progression where the first term $a = 1$, the common difference $d = 1$, and the number of terms n.

Let's represent this sum:

$$S_n = 1 + 2 + 3 + \cdots + n$$

One way to find the sum is by pairing the terms. Let's write the sum forwards and backwards, then add the two equations:

$$S_n = 1 + 2 + 3 + \cdots + n$$

$$S_n = n + (n - 1) + (n - 2) + \cdots + 1$$

Adding these two:

$$2S_n = (1 + n) + (2 + (n - 1)) + (3 + (n - 2)) + \cdots + (n + 1)$$

Each of these pairs sums to $n + 1$, and there are n pairs. Therefore:

$$2S_n = n(n + 1)$$

$$S_n = \frac{n(n + 1)}{2}$$

This idea of reversing the order and pairing can be generalised to derive the formula $S_n = \frac{n}{2}[2a + (n - 1)d]$ as well.

Geometric Series

A geometric progression has the first term a, common ratio r, and number of terms n. The sum of the first n terms is:

$$S_n = a + ar + ar^2 + \cdots + ar^{n-1}$$

To derive the formula, multiply both sides by r:

$$rS_n = ar + ar^2 + ar^3 + \cdots + ar^n$$

Now subtract these two equations:

$$S_n - rS_n = a - ar^n$$
$$S_n(1-r) = a(1-r^n)$$

Thus, the sum is:

$$S_n = \frac{a(1-r^n)}{1-r}, \quad \text{for } r \neq 1$$

Sum to Infinity for GP

When $|r| < 1$, the sum of an infinite geometric series can be calculated by taking the limit as $n \to \infty$:

$$S_\infty = \lim_{n\to\infty} \frac{a(1-r^n)}{1-r}$$

Since $r^n \to 0$ as $n \to \infty$ for $|r| < 1$, we get:

$$S_\infty = \frac{a}{1-r}, \quad \text{for } |r| < 1$$

This is the formula for the sum to infinity of a converging geometric progression.

3 Tips for Exams

- **"Pure" APGP questions**

  "Pure" APGP questions here refer to questions regarding the terms of an AP or a GP. They typically do not have contexts or stories attached to them.

  These are often just simultaneous equations in disguise. You need to carefully extract the information from the question and use it to form the relevant equations.

  Some questions require 3 unknowns, while some only require 2. Most of the time, you are required to solve them by hand, since they may not be a system of linear equations (which can be solved by GC).

Example : Complicated AP-GP Question
The 4th, 6th, and 9th terms of an AP are also consecutive terms of a GP. The sum of the 2nd and 5th terms of the AP is 20. Find the first term and common difference of the AP, and the common ratio of the GP.

Step 1: Form equations for the AP.

Let the first term and common difference of the AP be a and d, respectively.

$$T_4 = a + 3d, \quad T_6 = a + 5d, \quad T_9 = a + 8d$$

Since these terms are consecutive terms of a GP, we use the property of the geometric progression:

$$(T_6)^2 = T_4 \times T_9$$
$$(a + 5d)^2 = (a + 3d)(a + 8d)$$
$$a^2 + 10ad + 25d^2 = a^2 + 11ad + 24d^2$$
$$-ad + d^2 = 0 \implies d(a - d) = 0$$

Since $d \neq 0$, we conclude that $a = d$.

Step 2: Use the sum of the 2nd and 5th terms of the AP.

From the question, we know that the sum of the 2nd and 5th terms is 20:

$$T_2 + T_5 = 20$$

Using $a = d$, the terms become:

$$T_2 = a + d = 2a, \quad T_5 = a + 4d = 5a$$

Thus:

$$2a + 5a = 20 \implies 7a = 20 \implies a = \frac{20}{7}$$

Since $a = d$, the common difference is also $d = \frac{20}{7}$.

Step 3: Find the common ratio of the GP.

The common ratio r of the GP is given by:

$$r = \frac{T_6}{T_4} = \frac{a + 5d}{a + 3d}$$

Substituting $a = d = \frac{20}{7}$:

$$r = \frac{\frac{20}{7} + 5\left(\frac{20}{7}\right)}{\frac{20}{7} + 3\left(\frac{20}{7}\right)} = \frac{\frac{120}{7}}{\frac{80}{7}} = \frac{120}{80} = \frac{3}{2}$$

Thus, the common ratio of the GP is $r = \frac{3}{2}$.

Final Answer: The first term and common difference of the AP are both $\frac{20}{7}$, and the common ratio of the GP is $\frac{3}{2}$. ■

- **Compound interest questions**

Carefully identify all given conditions before writing down anything. Pay attention to details like the starting date, whether the interest is compound or simple, and the rate. It's important to capture these details correctly to ensure that the equations you form are accurate.

Example : Interest Questions
\$1000 is deposited into a bank at the start of month 1. At the end of every month, the total amount grows by 1%, whereas at the start of every month, \$1000 more is deposited into the bank. Find the total amount at the end of n months.

Since the question is interested in the amount at the *end* of the month, we let $\$T_n$ denote the balance at the end of n months. We list the first few terms and generalise using the pattern.

$$T_1 = 1.01(1000), \quad T_2 = 1.01(1000 + 1.01(1000)) = 1.01(1000) + 1.01^2(1000)$$

$$T_3 = 1.01(1000 + 1.01(1000 + 1.01(1000))) = 1.01(1000) + 1.01^2(1000) + 1.01^3(1000)$$

Thus $T_n = 1000 \sum_{r=1}^{n} 1.01^r$. ■

If we had set T_n to be the amount at the start of the month, the method would have been much messier, possibly inviting careless mistakes.

- **Draw tables**

To be safe, especially in interest rate or problems involving many things changing simultaneously, drawing tables can help keep track of information in an organised manner. If there is a time constraint, the listing method (used in the previous example) can be used as well; it is much faster than drawing a table.

- **Presentation**

Always copy the exact phrasing from the question when it refers to concepts like "arithmetic sum" or "sum to infinity of geometric series". This ensures that you are properly fulfilling the question's requirements.

CHAPTER 10

SUMMATIONS

1 Overview

Summations primarily deal with adding up the terms of a specific sequence. A summation is usually denoted by the summation symbol $\sum$, which represents the sum of the terms from a sequence according to specific bounds. For example, $\sum_{r=1}^{n} u_r$ means summing up the terms $u_1, u_2, \ldots, u_n$.

2 Solving Approaches

This topic generally revolves around manipulating your expression into an appropriate form where known formulae can be applied to find the sum.

Common Results:

$$\sum_{r=1}^{n} ku_r = k\sum_{r=1}^{n} u_r$$

$$\sum_{r=1}^{n}(u_r + v_r) = \sum_{r=1}^{n} u_r + \sum_{r=1}^{n} v_r$$

Tip: This is especially important when your sequence is the sum of an AP and GP; this formula helps to split it into two separate sums that are much easier to compute. For example, $\sum \ln(3(5)^r)$ is easier to compute when split into $\ln 3 + r\ln(5)$.

$$\sum_{r=m}^{n} u_r = \sum_{r=1}^{n} u_r - \sum_{r=1}^{m-1} u_r$$

Tip: You are doing this to make both lower limits $r = 1$, so that known formulae can be applied.

Extra Results (Good to memorise but not necessary):

$$\sum_{r=1}^{n} r = \frac{n(n+1)}{2}$$

The above formula can be derived from sum of an AP.

$$\sum_{r=1}^{n} r^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\sum_{r=1}^{n} r^3 = \frac{n^2(n+1)^2}{4}$$

3 Tips and Tricks

- **When using a previously proven result**

  Make sure the lower limits of the summations are the same. This is usually $r = 1$, but it may not always be the case.

Example : Handling Different Lower Limits
Find $\sum_{r=8}^{13} \frac{1}{\sqrt{r}+\sqrt{r-1}}$ given the formula:

$$\sum_{r=6}^{n} \frac{1}{\sqrt{r}+\sqrt{r-1}} = \sqrt{n} - \sqrt{5}$$

In order to apply the given formula, we must express our sums such that the lower limit is always $r = 6$.

$$\sum_{r=8}^{13} \frac{1}{\sqrt{r}+\sqrt{r-1}} = \sum_{r=6}^{13} \frac{1}{\sqrt{r}+\sqrt{r-1}} - \sum_{r=6}^{7} \frac{1}{\sqrt{r}+\sqrt{r-1}}$$

Now applying the formula we get

$$(\sqrt{13} - \sqrt{5}) - (\sqrt{7} - \sqrt{5}) = \sqrt{13} - \sqrt{7}$$

■

Notice that whenever we have a summation formula (e.g. $\sum_{r=1}^{n} r = \frac{n(n+1)}{2}$), the expression on the right-hand side changes along with the upper limit (i.e. n) of the left-hand side.

So while we can apply the formula with a different upper limit (e.g. $\sum_{r=1}^{10} r = \frac{10(10+1)}{2}$), we cannot apply the formula directly when the lower limit is different (e.g. $\sum_{r=5}^{10} r$). In these cases, the sum must be rewritten into two different sums with lower limits 1:

$$\sum_{r=5}^{10} r = \sum_{r=1}^{10} r - \sum_{r=1}^{4} r$$

The lower limits must match that of the formula in order to apply it.

- **Try to rewrite u_r so that known formulae can be applied.**

> ***Example : Manipulating u_r***
> Consider $\sum_{r=1}^{n}(ar^2 + br + c)$. This should be rewritten as:
>
> $$\sum_{r=1}^{n}(ar^2 + br + c) = a\sum_{r=1}^{n} r^2 + b\sum_{r=1}^{n} r + c\sum_{r=1}^{n} 1$$
>
> Splitting them up allows for the known formulae $\sum r^2$, $\sum r$, and $\sum 1$ to be applied.

- **Look out for expressions u_r containing AP, GP, or both.**

In these cases, the AP and GP formulae can be applied to solve the sum more easily.

> ***Example : AP and GP in a term***
> Consider the following summation. The terms contain both an AP and GP:
>
> $$\sum_{r=1}^{n} \ln(3^r 5^{6^r}) = \sum_{r=1}^{n}(r \ln 3) + \sum_{r=1}^{n}(6^r \ln 5)$$
>
> Breaking them up makes it much easier to compute the sum.

- **Present your observations and manipulations appropriately.**

Once again, you should always aim to simplify the expression into a form where known results can be applied.

Be precise with terminology: Are you dealing with an "Arithmetic Progression", or "Arithmetic Series"? One is a sequence while the other is a sum.

One way to stay safe is to copy the exact phrasing of the question, such as "infinite sum" or "limit of infinite series", to ensure that you are answering the question's requirement correctly.

4 Method of Differences (MOD)

- **Look out for partial fractions and sigma notation.**

When you encounter partial fractions and summations expressed in sigma notation, it's a strong indicator that the method of differences might be useful.

Denominators containing multiple factors can usually be simplified using partial fractions. For example $\sum \frac{1}{(r)(r+1)}$. It may be handy to memorise a couple of common partial fraction decomposition so you do not need to derive it every time, such as $\frac{1}{(r)(r+1)} = \frac{1}{r} - \frac{1}{r+1}$.

Meanwhile, denominators containing surds should be rationalised to reveal the cancellations. For instance $\sum \frac{1}{\sqrt{r}+\sqrt{r-1}}$ which appeared in one of the earlier examples.

- **Be sensitive to repeating terms.**

Pay close attention to repeating terms in the sequence, especially those involving consecutive integers like $(r-1, r, r+1)$ or $(r-2, r-1, r)$, etc. These often signal that terms will cancel out.

- **Write out the first few terms.**

It's often helpful to explicitly write out the first few terms of the summation. This allows you to observe a symmetry or cancellation pattern. Usually, the number of uncancelled terms at the beginning will equal the number of uncancelled terms at the end.

> ***Example : Using the Method of Differences***
> Consider the sum:
> $$\sum_{r=1}^{n}\left(\frac{1}{r}-\frac{1}{r+1}\right)$$
> Writing out the first few terms:
> $$\left(\frac{1}{1}-\frac{1}{2}\right)$$
> $$+\left(\frac{1}{2}-\frac{1}{3}\right)$$
> $$+\quad\cdots$$

$$+\left(\frac{1}{n}-\frac{1}{n+1}\right)$$

Notice that $-\frac{1}{2}$ cancels with $+\frac{1}{2}$.

Most other terms cancel in a similar manner, leaving only $\frac{1}{1}$ at the top. By symmetry we know that only one term is left at the bottom, that is, $\frac{1}{n+1}$.

Therefore, the sum is:

$$S_n = 1 - \frac{1}{n+1}$$

Write out enough terms such that the cancellations happen. Similar cancellations can be done throughout

CHAPTER 11

DIFFERENTIATION

1 What Exactly is Differentiation?

While many students think differentiation means finding the gradient at a point, it is actually far more than that. Broadly speaking, differentiation studies relative rates of change, that is, how much one variable changes compared to another variable.

The Limit of a Ratio

We look at one of the most familiar expressions, $\frac{dy}{dx}$. It is fundamental yet important to know that $\frac{dy}{dx}$ is not a ratio itself, rather the limit of a ratio.

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

Since $\frac{dy}{dx}$ is a limit, we cannot treat it as a fraction and "take it apart". For example, $\frac{dy}{dx}$ cannot be rewritten as $dy = 2dx$. Practically however, it behaves exactly like a fraction, which is why many techniques involving it work similarly. Take the chain rule, for example:

$$\frac{dy}{dx} = \frac{dy}{dt} \times \frac{dt}{dx}$$

Though the dt in the numerator and denominator seem to cancel out, we cannot write it that way since $\frac{dy}{dx}$ is not a fraction hence cannot cancel.

Differentiation as a Comparison of Change

Again we look at the above formula:

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

We see that the right hand side resembles the gradient formula,

$$\frac{y - y_1}{x - x_1} = \frac{\Delta y}{\Delta x}$$

This is because the gradient compares the change in y (the rise) to the change in x (the run), which is exactly what differentiation does: to compare changes.

Meanwhile, the limit as $\Delta x \to 0$ resembles bringing the two points of interest closer and closer to each other, until they almost overlap. This produces the supposed "gradient at a point". In fact it is comparing a small change in x to a small change in y.

Differentiation as an Operator

Just a side note, that differentiation is not a function but an operator. A function takes in an input value and returns an output value. An operator takes in an input function and returns an output function. The two are somewhat analogous, but it would be inaccurate to refer to $\frac{d}{dx}$ as a function.

2 Visualising the Chain Rule

Consider the following scenario.

3 runners, A, B and C take part in a race; A goes against B in the first round, while B goes against C in the second round. For the sake of discussion we assume that each runner maintains a constant pace. A runs twice as fast as B, while B runs three times as fast as C, so how much faster is A compared to C?

The solution is simple and intuitive:

$$\frac{Speed\ of\ A}{Speed\ of\ C} = \frac{Speed\ of\ A}{Speed\ of\ B} \times \frac{Speed\ of\ B}{Speed\ of\ C} = 2 \times 3 = 6$$

So A is 6 times as fast as C.

This is analogous to the chain rule, where we compare the changes of two variables using a series of similar comparisons. For instance:

$$\frac{dy}{dz} = \frac{dy}{dx} \times \frac{dx}{dz} \quad \text{or} \quad \frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}}$$

In the pursuit of computation, the meaning behind the chain rule is frequently overlooked. The chain rule simply helps to compare changes, which is the exact purpose of derivatives.

3 Stationary Points

A stationary point occurs when the derivative of a function is zero, i.e. $\frac{dy}{dx} = 0$. This implies that the slope of the tangent line to the curve at that point is horizontal.

Stationary points can be classified into three categories:

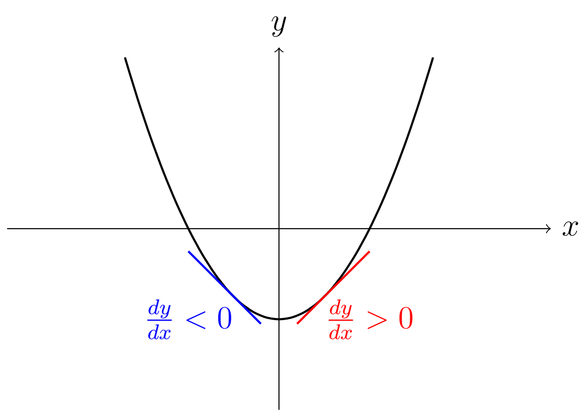
- **Minimum point**: The second derivative $\frac{d^2y}{dx^2} > 0$.

- **Maximum point**: The second derivative $\frac{d^2y}{dx^2} < 0$.

- **Point of inflection**: The second derivative $\frac{d^2y}{dx^2} = 0$.

We are familiar with the above conditions, but do we know why they are as such?

Consider the graph with a minimum point. We know that the second derivative $\frac{d^2y}{dx^2} > 0$ at the minimum point.



Observe how $\frac{dy}{dx}$ changes along with x. As x increases from left to right, the gradient $(\frac{dy}{dx})$ changes from negative (blue) to zero (at the stationary point) then to positive (red). This means that $\frac{dy}{dx}$ increases as x increases, implying $\frac{d^2y}{dx^2} > 0$.

4 Common Problem Types

The H2 syllabus primarily tests a few types of questions: differentiation techniques, tangents and normals, rate of change, and optimisation (stationary points).

Differentiation Techniques

Questions of this form are the most straightforward, typically requiring you to differentiate an expression. To score well you need to know the rules of differentiation well: power rule, product rule, chain rule, quotient rule, as well as the derivatives of terms like $e^x, a^x, \ln x, \sin x$ and so on.

Tangents and Normals

Generally asking for the equation of a tangent or normal, the gradient of the tangent can be easily found by differentiation.

You may find the following coordinate geometry information helpful:

- The equation of any straight line can be found using two pieces of information: the gradient and the coordinates of any one point on the line.

- For a straight line with gradient m passing through a point (x_1, y_1), the equation of the line is given by $(y - y_1) = m(x - x_1)$. This is simply the gradient formula rewritten: $m = \frac{y-y_1}{x-x_1}$.

- The product of gradients of two perpendicular lines is -1 (i.e. $m_{normal} = \frac{-1}{m_{tangent}}$).

- Gradient $= \tan\theta$. If you do not already know this, refer to the chapter on Trigonometry.

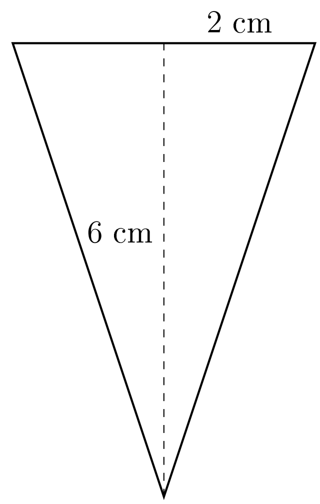
Rate of Change

These questions typically require an understanding of a certain context given, after which you may be required to form your own equations, which will then be differentiated.

To differentiate these equations, there are generally two approaches: chain rule or implicit differentiation. The better method is the one you are more comfortable with. For implicit differentiation, do not forget to apply the chain rule. For instance when you apply $\frac{d}{dx}$ on y^3, you have to multiply $3y^2$ by $\frac{dy}{dx}$ due to the chain rule.

Tip: You cannot have an equation with more than two variables in the same equation, since you cannot differentiate it. If you do, you should express one of them in terms of the other two. There will be a way.

Example : Volume of a cone
Given the cross sectional diagram of a cone of height 6 cm and radius 2 cm, find a general formula for V in terms of r.



Using the formula for volume we have:

$$V = \frac{1}{3}\pi r^2 h$$

However this cannot be differentiated yet since there are 3 variables. So we have to express one in terms of the other two.

From the diagram we see the relationship (similar triangles) between r and h, that $\frac{h}{r} = 3$ so we can write $h = 3r$.

Thus,

$$V = \frac{1}{3}\pi r^2(3r) = r^3$$

Optimisation (Maximising/Minimising)

Questions of this form would always require setting the derivative to zero. That is where some computational tricks come in handy.

Tip: Identify which factors in the derivative cannot equal zero; you can cancel these terms. Derivatives in these questions are often rather long and complicated, so cancelling the redundant terms early on can save you significant time.

Example : Cancelling non-zero terms
Given $y = 21^{1-\sec(2x+\frac{\pi}{7})}$, find $\frac{dy}{dx}$ and use it to find the x-coordinates of the stationary points.

First, differentiate the function:

$$\frac{dy}{dx} = 21^{1-\sec(2x+\frac{\pi}{7})} \cdot \ln 21 \cdot \left(-2\sec(2x+\frac{\pi}{7}) \cdot \tan(2x+\frac{\pi}{7})\right)$$

Since $21^{1-\sec(2x+\frac{\pi}{7})}$, $\sec(2x+\frac{\pi}{7})$ and $\ln 21$ are non-zero, we can cancel these terms, leaving:

$$\tan(2x+\frac{\pi}{7}) = 0$$

Solving $\tan(2x+\frac{\pi}{7}) = 0$ gives:

$$x = \frac{n\pi}{2} - \frac{\pi}{14} \quad \text{for integers } n$$

CHAPTER 12

INTEGRATION TECHNIQUES

The syllabus deals with three types of integration: integration by inspection, integration by parts, and integration by substitution.

When approaching an expression to integrate, always try inspection before moving on to integration by parts. If the question requires you to integrate by substitution (like $u = x^2$), the substitution will definitely be provided in the problem.

1 Integration by Inspection

Since integration is essentially the inverse operation of differentiation, expressions can often be manipulated into forms that are known derivatives, allowing them to be integrated directly. For example:

$$\frac{d}{dx}\ln(f(x)) = \frac{f'(x)}{f(x)} \quad \Rightarrow \quad \int \frac{f'(x)}{f(x)}\,dx = \ln|f(x)| + C$$

Note that the C is a mandatory arbitrary constant for indefinite integrals, as all constants vanish under differentiation (indefinite integrals are integrals without upper and lower bounds).

Tips and Techniques

- **Use trial and error intelligently:** Integration by inspection often involves trying different possibilities. Make smart guesses based on common patterns and known derivatives.

- **Train your eyes for standard forms:** Common integrals include chain rule expressions like $nf'(x)f(x)^{n-1}$, or natural log forms like $\int \frac{f'(x)}{f(x)}\,dx$. Recognizing these forms is key to using inspection successfully.

- **Play around with the expression:** Isolate terms and find their derivatives to see if you can match them with known forms. For example, by isolating a term like $f(x)$ and differentiating it, you can identify if the expression follows the chain rule pattern $nf'(x)f(x)^{n-1}$.

> ***Example : Playing around with terms***
> Consider the integral $\int \frac{\cot x}{\ln \sin x}\, dx$. By differentiating the denominator, we find:
>
> $$\frac{d}{dx} \ln \sin x = \cot x$$
>
> which is exactly the numerator of the given expression. This shows that the integrand is of the form $\frac{f'(x)}{f(x)}$, and we can directly apply the standard result:
>
> $$\int \frac{f'(x)}{f(x)}\, dx = \ln |f(x)| + C$$
>
> Thus, $\int \frac{\cot x}{\ln \sin x}\, dx = \ln |\ln \sin x| + C$. ■

- **Always try inspection before moving to integration by parts:** Many integrals that look complicated at first glance can often be handled quickly with inspection. Some may require subtle manipulations like splitting or rewriting the expression in a similar form.

> ***Example : Try solving these integrals by inspection!***
>
> $$\int \frac{1}{1 - e^{-x}}\, dx$$
>
> $$\int \frac{2x + 1}{(x + 1)^2}\, dx$$
>
> $$\int \tan \theta \, d\theta$$
>
> ■

- **Beware of the "inverse chain rule":** When we differentiate, byproducts often result from the chain rule. These byproducts must be present during integration as an "inverse chain rule".

> ***Example : Inverse Chain Rule***
> Consider $\frac{d}{dx} \sin(x^2) = \cos(x^2)(2x) = 2x \cos(x^2)$.
>
> The factor $2x$ appears because of the chain rule. Therefore, $\int \cos(x^2)\, dx \neq \sin(x^2) +$

> C, as this neglects the chain rule. The $2x$ must be present to account for the "inverse chain rule", that is, $\int 2x\cos(x^2)\,dx = \sin(x^2) + C$.

- **Be cautious with higher-order terms:** Standard forms can usually be applied safely to first-order terms (i.e. where the power of x is 1), for instance $\int \sin(2x)\,dx$. Extra care is required when dealing with higher-order terms, like $\int \sin(x^2)\,dx$, which cannot be integrated directly by inspection.

> ***Example : A Wrong Approach!***
> Please note that you cannot do this:
>
> $$\int \cos(x^2)\,dx = \frac{1}{2x}\int 2x\cos(x^2)\,dx = \frac{1}{2x}\sin(x^2) + C$$
>
> While this method is valid for constants, you cannot use it to generate an x term within the integrand, because x is not a constant in this context.

2 Integration by Parts

The formula for integration by parts is:

$$\int u\frac{dv}{dx}\,dx = uv - \int v\frac{du}{dx}\,dx$$

This formula is derived from the product rule of differentiation. Integration by parts involves selecting appropriate terms for u and $\frac{dv}{dx}$, then finding $\frac{du}{dx}$ and v. Finally, we apply the above formula to solve the integral.

Tips and Techniques

- **Simplification through by parts:** The purpose of integration by parts is to simplify the expression. If the result becomes more complicated than the original expression, you're likely applying the technique incorrectly.

- **Choosing u:** Usually, we choose u to be the term that simplifies when differentiated. For example, in $\int x^2\cos(2x)\,dx$, we set $u = x^2$ because differentiating x^2 reduces its order (making it simpler). Meanwhile, $\cos(2x)$ remains a similar form after integration, as it gives a term with $\sin(2x)$.

  A bit of foresight is needed, as you should notice that picking $u = x^2$ ultimately reduces to $du = 2dx$ after two steps. In contrast, $\cos(2x)$ is "unable to be simplified" since repeated operations on it result in expressions of similar form.

- **The LIATE Acronym:** The rule for choosing u can be summarized by the acronym

LIATE, which stands for:

1. **L:** Logarithmic functions ($\ln x$)
2. **I:** Inverse trigonometric functions ($\tan^{-1} x$, $\sin^{-1} x$, etc.)
3. **A:** Algebraic functions (x^n, x^2, x, etc.)
4. **T:** Trigonometric functions ($\sin x$, $\cos x$, etc.)
5. **E:** Exponential functions (e^x, 2^x, etc.)

1 should be picked first while 5 last. This serves as a general guide to select the correct u for integration by parts.

- **Reappearance of the original expression:** In some cases, applying by parts leads to the original integral reappearing. This can be useful, as it allows you to solve the equation by isolating the original expression and solving algebraically.

> ***Example : Reappearing Expression***
> Consider the integral $\int \frac{(\ln x)^2}{x}\,dx$.
>
> We set:
> $$u = (\ln x)^2 \quad \text{and} \quad \frac{dv}{dx} = \frac{1}{x}$$
> Then:
> $$\frac{du}{dx} = \frac{2\ln x}{x} \quad \text{and} \quad v = \ln x$$
> Applying the integration by parts formula:
> $$\int \frac{(\ln x)^2}{x}\,dx = (\ln x)^3 - 2\int \frac{(\ln x)^2}{x}\,dx$$
> The original expression $\int \frac{(\ln x)^2}{x}\,dx$ reappears, and we solve by shifting terms:
> $$\int \frac{(\ln x)^2}{x}\,dx = \frac{1}{3}(\ln x)^3 + C$$
> Letting the original expression reappear allows us to simplify the problem and solve by dividing.
>
> (Did you notice a quicker approach?) ■

3 Integration by Substitution

Integration by substitution involves expressing the integral in terms of another variable, which is provided in the H2 syllabus. This new expression is generally easier to integrate directly (often by inspection), which is the purpose of substitution.

For example, when our original integral is in terms of x, and a substitution such as $t = x^2$ is provided, we must change every single part of the integral to be in terms of t. There should not be any x terms left. This includes the upper and lower limits, the integrand (the expression to be integrated), and the dt term (instead of dx).

After rewriting, integrate the new expression, and finally substitute back to the original variable. In the above example, since we began with x, the final result must be expressed in terms of x, not t.

A Systematic Approach to Substitution

- **Identify the substitution:** Look for a part of the expression that can simplify through substitution. In all H2 problems, the substitution is provided (e.g. $t = x^2$).

- **Rewrite all parts of the integral:** Express the entire integral, including limits (for definite integrals), the integrand, and the differential term (dx) in terms of t. To express dx in terms of t, this means finding $dx = \frac{dx}{dt}\, dt$ because $\frac{dx}{dt}$ is in terms of t, and so is dt.

- **Perform the integration:** Solve the new integral, which should now be simpler.

- **Substitute back:** After solving the integral, revert back to the original variable if necessary (i.e. express the final answer in terms of x instead of t if this is an indefinite integral. If it is a definite integral, your answer will be a constant).

It is best to express all substitutions (upper and lower limits, integrand, and differential terms) before rewriting the integral in the new variable.

> ***Example : Integration by Substitution***
> Evaluate the integral $\int_{-1}^{1} \sqrt{1 - x^2}\, dx$ using the substitution $x = \sin t$.
>
> **Step 1:** Identify the substitution. We are given $x = \sin t$, so $\frac{dx}{dx} = \cos t$.
>
> **Step 2:** Change the limits of integration. When $x = -1$, $\sin t = -1$ implies $t = -\frac{\pi}{2}$. Similarly, when $x = 1$, $\sin t = 1$ implies $t = \frac{\pi}{2}$.
>
> **Step 3:** Rewrite the integrand. The integrand $\sqrt{1 - x^2}$ becomes:
>
> $$\sqrt{1 - \sin^2 t} = \cos t$$

So the integral becomes:

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{1-\sin^2 t}\cos t\,dt = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 t\,dt$$

Step 4: Perform the integration. Use the identity $\cos^2 t = \frac{1+\cos 2t}{2}$ to rewrite the integrand:

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 t\,dt = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1+\cos 2t}{2}\,dt$$

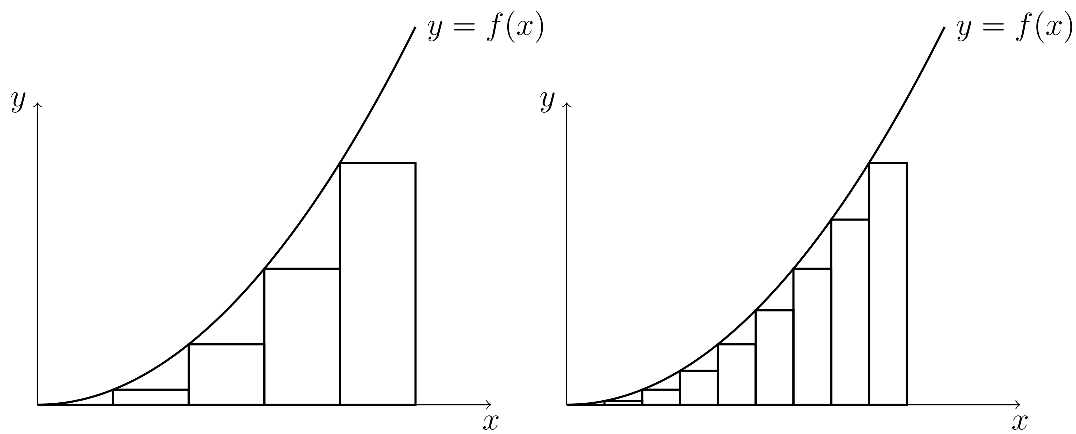
Thus, $\int_{-1}^{1} \sqrt{1-x^2}\,dx = \frac{\pi}{2}$.

CHAPTER 13

DEFINITE INTEGRALS

1 Integration as a Sum

You may be familiar with the idea of approximating the area under a curve by summing rectangles.



As the rectangles get thinner, our approximation becomes more accurate, and the sum of their areas gets closer to the true area under the curve. By taking the limit of this sum as the width of the rectangles approaches zero, we arrive at the exact area under the graph.

This is one of the applications of integration. However, beyond simply finding the area, integration should be seen as a sum. Like how differentiation is the limit of a ratio, likewise integration is the limit of a sum. In fact, the integration symbol is simply an elongated "S", which represents "sum". Recognising integrals as sums can shed light on some of its workings.

The mathematical definition of the integral can be written as:

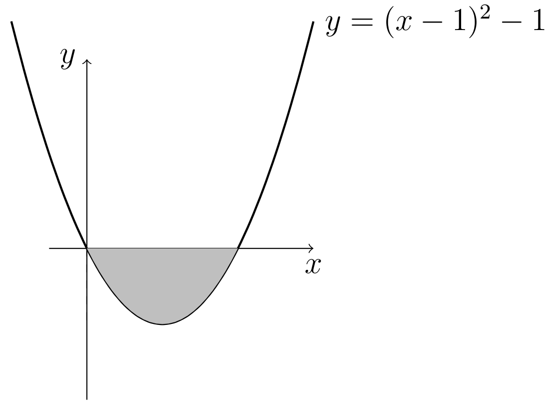
$$\int_{x_1}^{x_2} f(x)\,dx = \lim_{\Delta x \to 0} \sum f(x_i)\Delta x$$

On the right-hand side, $f(x_i)$ represents the height of a rectangle at $x = x_i$, while Δx represents the constant width of the rectangles. Multiplying these gives the area of one rectangle, and summing them gives the total area of all the rectangles, which approximates the area under the curve.

This notion of summing reflects the true nature of the integral. In fact, when moving on to higher level math, integrals are no longer used merely for finding area. They are used to sum up very small portions of items of interest, such as the work done in Physics. Learning to view integrals as sums would be helpful if you are interested in pursuing related fields in the future.

2 Negative Areas

When we integrate a curve that lies below the x-axis, why do we get a negative result? Consider the function $y = (x-1)^2 - 1$ which lies below the x-axis between $x = 0$ and $x = 2$.



As the curve dips below the x-axis, the area calculated between $x = 0$ and $x = 2$ appears to be negative.

$$\int_0^2 \left((x-1)^2 - 1\right) dx = -\frac{4}{3}$$

But why does this integral produce a negative result? How can an area be negative?

To understand this better, we return to the formula:

$$\int_{x_1}^{x_2} f(x)\, dx = \lim_{\Delta x \to 0} \sum f(x_i)\Delta x$$

Here, $f(x_i)$ represents the height, and Δx is the width. If the curve is below the x-axis, $f(x_i)$ will be negative, since it is simply the output of the function when we input x_i. In simpler terms, the heights of the rectangles are "negative" because the rectangles are below the x-axis. Since Δx is always positive (width is always positive), each rectangle will have a "negative area", as $f(x_i)\Delta x$ is negative, thus producing an integral of negative value.

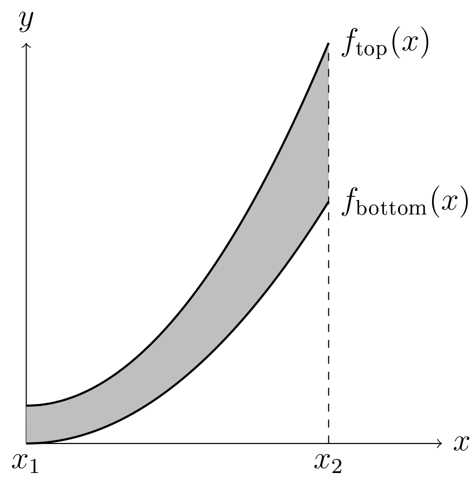
Important Note: In reality, quantities like height and area cannot be negative, so you will definitely be penalised if your solution contains a negative height or area. Always convert them to positive values, using modulus or otherwise.

3 Area Between Two Curves

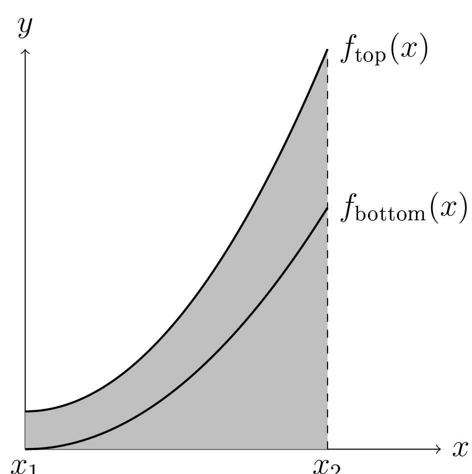
The general rule for finding the area between two curves is to integrate the function on top minus the function below:

$$\text{Area} = \int_{x_1}^{x_2} f_{\text{top}}(x) - f_{\text{bottom}}(x)\, dx$$

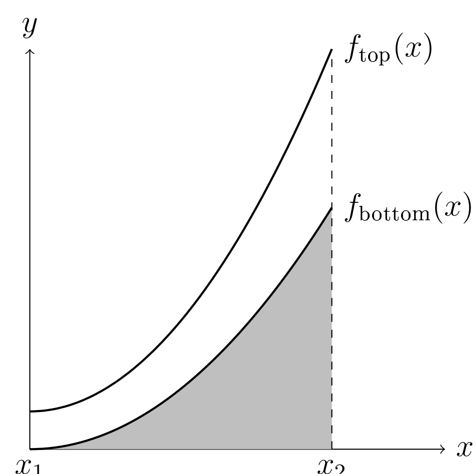
Let's consider the case where both curves are above the x-axis.



Here, area A represents $\int_{x_1}^{x_2} f_{\text{top}}(x)\, dx$, while area B represents $\int_{x_1}^{x_2} f_{\text{bottom}}(x)\, dx$, as illustrated below.



Area A



Area B

It can be observed visually that taking subtracting area B from area A gives the area between the two curves. Thus verifying the relationship:

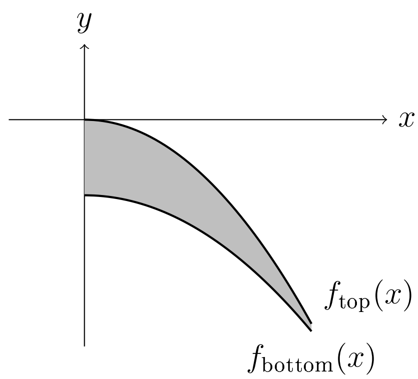
$$\text{Area between two curves} = \int_{x_1}^{x_2} f_{\text{top}}(x) - f_{\text{bottom}}(x)\,dx$$

Now let's consider the case where both curves are below the x-axis. The integral works in the same way as before, where the result of subtracting the two areas will still give the correct area between the curves. The function above is still chosen as f_{top}.

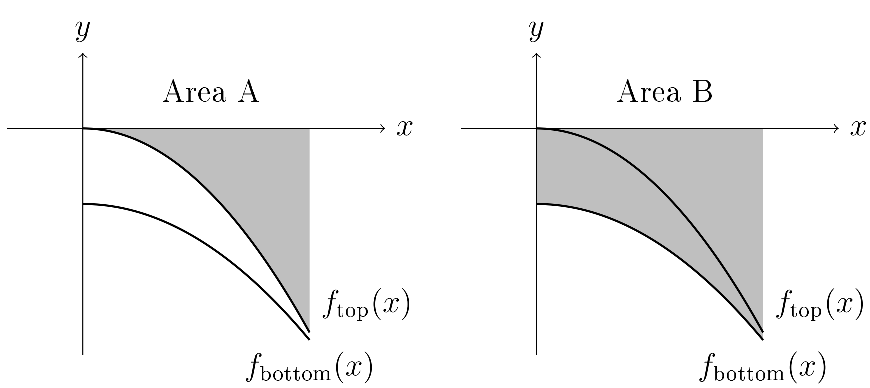
The general formula remains:

$$\text{Area} = \int_{x_1}^{x_2} f_{\text{top}}(x) - f_{\text{bottom}}(x)\,dx$$

Let's illustrate this with an example where both curves are below the x-axis:



In this case, both curves are below the x-axis, so the areas A and B are negative. However, subtracting B from A still gives the positive area between the two curves, because while $\int_{x_1}^{x_2} f_{\text{top}}(x)\,dx$ is negative, it is less negative than $\int_{x_1}^{x_2} f_{\text{bottom}}(x)\,dx$. A is larger as it is less negative, and B is smaller since it is more negative, just like how -1 is larger than -10.



Thus, the principle remains:

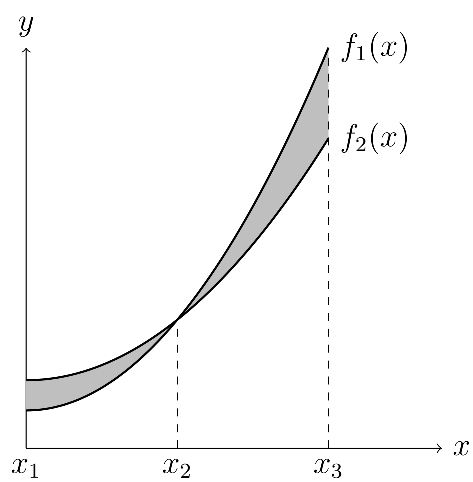
$$\text{Area between two curves} = \int_{x_1}^{x_2} f_{\text{top}}(x) - f_{\text{bottom}}(x)\,dx$$

Even when both curves are below the x-axis, the subtraction still results in a positive area. Subtracting a larger negative number from a smaller one results in a positive value, just like how $(-1) - (-10) = 9$.

Things get more complicated when the curves cross the x-axis, but the principle remains the same, and so the above formula still holds true.

Crossing Curves

However, things have to be handled differently when the two curves cross. We split the integral into two or more intervals to ensure that one curve is on top of the other throughout the entire interval.



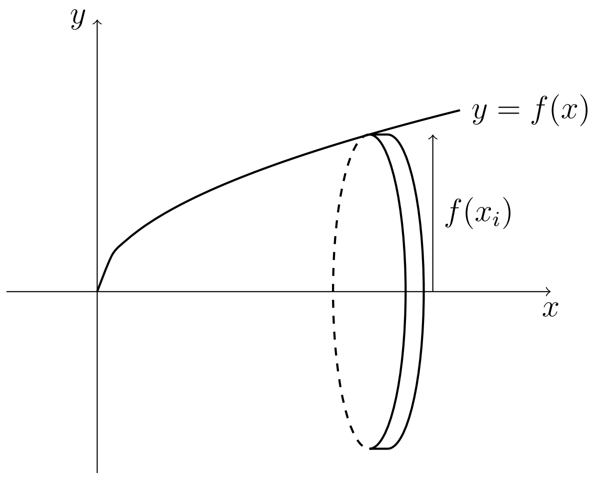
For example, if $f_2(x)$ is above $f_1(x)$ from x_1 to x_2, but below $f_1(x)$ from x_2 to x_3, you should split the integral:

$$\text{Total area} = \int_{x_1}^{x_2} f_2(x) - f_1(x)\,dx + \int_{x_2}^{x_3} f_1(x) - f_2(x)\,dx$$

This ensures that you are always integrating the top curve minus the bottom curve within each interval.

Evaluating one integral across the entire interval would cause one portion to be positive while the other negative, so summing them cancels each other out, giving us an area lesser than the actual.

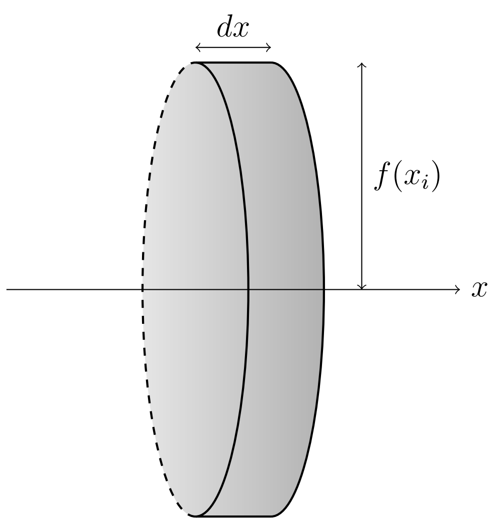
4 Volume of Revolution



When asked to find the volume of a solid generated by revolving a curve around the x-axis, the volume can be found using the following formula:

$$V = \int_a^b \pi f(x)^2 \, dx$$

This formula treats the solid of revolution as a collection of many thin circular disks stacked along the x-axis. The thickness of each disk is dx, and the radius of the disk at any point x_i is given by $f(x_i)$ (i.e. the height of the curve at x_i).



- $f(x_i)$ is the radius of the circular disk at any point x_i.
- The area of the circular disk at point x_i is $\pi[f(x_i)]^2$.
- The volume of this small disk (a slice of the solid) is $(\pi[f(x_i)]^2 \, dx)$.
- Lastly, the integration is applied to sum up all volumes of discs.

Reiterating the notion of summing, the definite integral from a to b adds up all these small volumes, giving the total volume of the solid generated by the revolution.

The same method can be applied to more complex solids, where the curve might have multiple inflection points or intersections with the x-axis. In these cases, the integral should be broken down into separate intervals to account for any changes in the curve's shape.

Note that this method can only be applied when revolving around the x-axis (or similarly for the y-axis by integrating with respect to y instead). For curves being revolved around lines other than the axes, the method can still be applied, but additional adjustments must be made depending on the context (such as the subtraction of any extra volume).

5 Additional Tips and Techniques

- **Be careful with negative areas**

  When integrating below the x-axis, remember that the integral gives a negative value. If you are asked for the total area, take the absolute value of these negative integrals.

  This also means that you should draw out your curve if you suspect some regions to be negative. This helps you to recognise and partition the regions appropriately.

- **Check for symmetry**

  If a function or region is symmetric, you can simplify the problem by integrating only one part of the graph and multiplying by an appropriate factor. For example,

  $$\int_0^4 |x-2|\,dx = 2\int_0^2 2-x\,dx$$

  Notice also that we took the negative part of the modulus since the interval exists where $x \leq 2$. If unsure why this is true, feel free to plot it out, or refer to the chapter on "Modulus" to learn more.

- **Watch out for discontinuities**

  If the function has a discontinuity in the interval of integration, split the integral at the discontinuity points to avoid errors.

  Because of the way the integral is defined, it cannot accommodate discontinuous curves. Adaptations have to be made to apply the integration.

CHAPTER 14

SERIES EXPANSION

1 Maclaurin Series

In essence, the Maclaurin series aims to approximate functions as an infinite power series by creating a polynomial whose each derivative matches that of the function at $x = 0$.

It is very important to note that the Maclaurin series is only a fair approximation near $x = 0$, in other words when x is small. This is because the "matching" is done exactly at $x = 0$. We generally refer to "small" as $|x| < 1$, so numbers like -0.75 or $\frac{1}{\sqrt{2}}$ work, but not $x = 1$.

Why Do We Use This?

Firstly, recognise that polynomials are nice functions that are very easy to deal with; we can integrate, differentiate, and sketch them with ease, unlike most other functions like $\ln(x)$ or $\cot(x)$, whose differentiation or integration becomes complicated after a few cycles.

Oftentimes in real-world applications, we do not require the exact value but rather one that is simply "close enough". For instance, if you need to estimate the value of $\sin(0.1)$ to 3 decimal places, using a series expansion can give a very quick approximation. This is where the Maclaurin series comes in handy, since an approximate polynomial can be computed easily.

How Does This Work?

Polynomials are very flexible functions, so we can "build our own polynomials", customising each derivative to match the original function. This is analogous to how an artist chooses specific colours to apply to different regions of a painting; we choose the values for each of our derivatives.

Suppose we have an infinite polynomial:

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

and we want to customise its third derivative at $x = 0$ to be 12. We differentiate this polynomial three times, giving:

$$f'''(x) = 3!a_3 + 4(3)(2)a_4x + \dots$$

Notice that the 3! in front of a_3 comes from differentiating three times. Each differentiation brings down a power from x^n, multiplying as we differentiate further.

Since we want $f'''(0) = 12$, we substitute $x = 0$, and all the subsequent terms containing x disappear, leaving:

$$3!a_3 = 12 \quad \Rightarrow \quad a_3 = 2$$

Here, we have customised a_3 to accommodate the condition $f'''(0) = 12$. This illustrates the idea of "customisation" of the polynomial coefficients.

We generalise this concept to write the Maclaurin series as:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)x^2}{2!} + \dots + \frac{f^{(n)}(0)x^n}{n!} + \dots$$

This represents matching every single derivative of the polynomial at $x = 0$ to that of the function. The key point is that the derivatives must be evaluated at $x = 0$, unlike the more general Taylor series (out of the syllabus), which can be used for any point $x = a$.

2 Exam Tips and Techniques

Fortunately, many Maclaurin series for different functions have already been computed, so we do not need to waste our time computing them again. In the H2 syllabus the relevant ones are present in the MF26 formula list as the "Standard Series". Knowledge of these is crucial to do well in this topic.

Know How to Apply Standard Series from the MF26

This includes being aware of the range of validities, should they exist. The Maclaurin series approximates functions near $x = 0$, so the approximation is only valid for small x, i.e. $|x| < 1$ for some cases.

There are a few special cases where the Maclaurin series exactly equals the function for all x-values, like $\sin x$, $\cos x$, and e^x. These are exceptions, and you should not expect other series to behave in the same manner.

For composite terms like $e^x \sin x$, simply expand them out separately, then multiply them together. For example, expanding up to x^3:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots, \quad \sin x = x - \frac{x^3}{3!} + \dots$$

Multiplying these gives the combined expansion:

$$e^x \sin x = \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots\right)\left(x - \frac{x^3}{3!} + \dots\right) = x + x^2 + \frac{x^3}{2!} - \frac{x^3}{3!} + \dots$$

We only expand to the required powers and omit higher terms.

Be Sensitive Toward the Standard Formula

Some standard formulae require a very specific form before you can apply them. For instance, the $(1+x)^n$ expansion in the MF26 requires the constant term to be 1, while x can be replaced with other expressions. The same applies to the $\ln(1+x)$ expansion. Neglecting this condition would result in an incorrect series.

Learn to be flexible to accommodate these conditions. For example, suppose we want to expand $(a + bx)^{-1/2}$. We need the constant term to be 1. Simply factor out a:

$$(a + bx)^{-1/2} = a^{-1/2}\left(1 + \frac{bx}{a}\right)^{-1/2}$$

Do not forget that whatever is factored out must carry the same index. In this case, a is raised to the power $-1/2$.

The Order of Expansion

Consider a composite function $\sin(e^x)$, which we need to expand. A common question is: should we apply the Maclaurin series for $\sin x$ first, or for e^x first?

The short answer is: it does not matter. Try and prove this by yourself!

Be Sharp About Where Series Can Be Applied

Consider an expression like $\frac{3x-1}{2+x}$, and the question asks you to expand it in increasing powers up to and including x^3. This may seem challenging until you realise that the expression can be rewritten as $(3x - 1)(2 + x)^{-1}$, where the series can then be applied easily.

Sometimes, the question does not directly suggest a method. You need to be sharp to distinguish where and how the series expansion can be applied.

Only Write Out What Is Necessary

Suppose we need to expand $e^x \sin x$ up to and including x^3. We only need to expand each term up to x^3 in the brackets:

$$\left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots\right)\left(x - \frac{x^3}{3!} + \dots\right)$$

Since higher powers do not multiply to give lower powers, special care is needed when one bracket contains negative powers, and the other contains positive powers. Be aware of how different powers can combine, and write out only what is necessary.

Descending Powers

We are used to expanding in increasing powers, but sometimes the question asks us to expand in decreasing powers. How is this possible? Do we start from infinity?

Consider $(1+x)^{-1}$ expanded in decreasing powers. We factor out x from the bracket first:

$$(1+x)^{-1} = x^{-1}\left(1 + \frac{1}{x}\right)^{-1}$$

Then, expand the bracket and multiply with the x^{-1} factor in front:

$$\frac{1}{x}\left(1 - \frac{1}{x} + \frac{1}{x^2} + \dots\right) = \frac{1}{x} - \frac{1}{x^2} + \frac{1}{x^3} + \dots$$

Thus, we successfully expand in decreasing powers up to x^{-3}.

A Brief Mention of Notation

Notation is crucial in mathematics, and presenting your solutions with clear and accurate notation is key to communicating effectively.

The ellipsis "…" is used to indicate that the pattern in the sequence or expression continues indefinitely. For instance:

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

This notation shows that the expansion continues following the same pattern. It can also mean that there are more terms afterwards, that have not been explicitly written out.

The approximation "≈" is used to indicate that a value is approximately equal to another. It signals that the result is close but not exact. For instance:

$$\sin x \approx x \quad \text{for small values of } x$$

This means that for small x, the value of $\sin x$ is very close to x, but not exactly equal.

Poor usage of such symbols will definitely result in penalisation.

3 A Common Question

One of the most common questions involves two parts. First, you are asked to differentiate an expression multiple times to prove an equation. In the second part, you are asked to find the Maclaurin series of the function. These questions often require a simplification step before differentiating to avoid a messy process.

Let's go through an example where this technique is applied. We will consider the function $y = \ln(\cos x)$, simplify it, and then differentiate implicitly.

Example : Simplification and Implicit Differentiation
Given $y = \ln(\cos x)$, prove that:

$$\frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^2 + 1 = 0$$

We start by simplifying the expression to make the differentiation process cleaner:

$$y = \ln(\cos x) \quad \Rightarrow \quad e^y = \cos x$$

Now, instead of differentiating $\ln(\cos x)$ directly, we differentiate implicitly:

$$\frac{d}{dx}\left(e^y\right) = \frac{d}{dx}\left(\cos x\right)$$

This gives us:

$$e^y \frac{dy}{dx} = -\sin x$$

Differentiating both sides with respect to x again:

$$\left(e^y \frac{dy}{dx}\right)\left(\frac{dy}{dx}\right) + e^y \frac{d^2y}{dx^2} = -\cos x$$

Lastly we substitute back $e^y = \cos x$:

$$\cos x \left(\frac{dy}{dx}\right)^2 + \cos x \frac{d^2y}{dx^2} = -\cos x$$

Dividing both sides by $\cos x$ yields the required differential equation.

At this point, you would be asked to derive the Maclaurin series for y. It is advisable to begin with "At $x = 0, \ldots$" to prompt yourself to stay on track. Then, you just need to find the corresponding values of y, y', y'' and so on systematically. It is best to start from y and move upwards from there.

Notice how we did not jump directly into differentiating directly, which would result in a mess. Always keep an eye out for any available simplifications.

CHAPTER 15

DIFFERENTIAL EQUATIONS

Differential equations are equations that relate the derivatives of a function. In the H2 syllabus, there are two broad categories of questions: Formulating differential equations, and solving them.

1 Formulating

In problems requiring you to formulate a differential equation, the question will typically provide a story or context where you must translate the given information into mathematical terms. This is analogous to translating a phrase from one language to another; here we are translating from English to Math. For example, "The initial rate of change of Area is 1 cm^2 per second" translates to "At $t = 0$, $\frac{dA}{dt} = 1$".

To do well in formulating differential equations, it is important to clearly understand the meaning of derivatives and how they are applied in context. Below are a few common keywords and their mathematical implications:

- "Rate of change of x", "x increases at a rate of": Translates to $\frac{dx}{dt}$, the derivative of x with respect to time.

- "Initially", "at the start", "starting at the reference point": Translates to "at $t = 0$".

- "Reaches a stationary value", "momentarily at rest", "maintained at": Translates to $\frac{dx}{dt} = 0$, indicating a stationary point.

- "A proportional to B": Translates to $A = kB$ for some constant k.

- "A inversely proportional to B": Translates to $A = \frac{k}{B}$ for some constant k.

- "At time t, the distance is x", "when $t = \ldots$, $x = \ldots$": Refers to a specific value of t, which can be used to find constants after solving. This can be thought of as a specific point on the x-t graph.

It is useful to make a habit of using arbitrary constants (like k) to represent proportionality. These constants may not be given to you at the start of the question, but their values must be found by the end of the question through substitution of initial conditions or specific values.

Usually, when formulating the differential equation, you do not need to worry about the information like "initially $x = 5$" at the start. These are points provided to you, $(x, t) = (5, 0)$ in this case, and typically only used afterward to find the value of certain constants (like c, the constant of integration, or k, the constant of proportionality) via substitution. These problems are called "initial value problems."

Example : Initial values being used only at the end
Solve for y given that $\frac{dy}{dx} = \sec y$, and the point $(0, \frac{\pi}{2})$

So first we can ignore the point $(0, \frac{\pi}{2})$, since it is only relevant at the end. We solve the integral:

$$\int \cos y \, \frac{dy}{dx} \, dx = \int 1 \, dx \quad \Rightarrow \quad \sin y = x + c$$

We can substitute in the initial conditions now, using $(0, \frac{\pi}{2})$:

$$\sin\left(\frac{\pi}{2}\right) = (0) + c \quad \Rightarrow \quad c = 1$$

Thus our equation is $\sin y = x + 1$ which should be rewritten as $y = \sin^{-1}(x + 1)$. ■

Of course, you can choose to rewrite the equation as $y = \sin^{-1}(x + c)$ first, then use the point to find c; I prefer the above presented way, but this is entirely up to personal preference. Point is, the initial values presented is simply a point to help determine unknown constants, and are only relevant towards the end. Do not let them confuse you at the start of the question.

2 Solving

Once the differential equation is formulated, solving it typically requires standard integration techniques, which you should already be familiar with by now.

Sometimes, certain simplifications may be required, such as "separation of variables", which may be a technique new to some. It involves rearranging the equation so that all terms containing y are on one side and all terms containing x are on the other side, allowing you to integrate both sides separately.

> ***Example : Separation of variables***
> Solve the differential equation $\frac{dy}{dt} = yt$ provided that $y = 2$ intially.
>
> Notice that this is an unusual equation that cannot be integrated directly. So ee move all terms involving y to one side and terms involving t to the other:
>
> $$y\,\frac{dy}{dt} = t$$
>
> Integrate both sides with respect to t, and the dt terms in the numerator and denominator can be thought to cancel out (do not actually cancel them):
>
> $$\int \frac{1}{y}\,\frac{dy}{dt}\,dt = \int t\,dt$$
>
> This gives:
>
> $$\ln|y| = \frac{t^2}{2} + C$$
>
> First, exponentiate both sides to get rid of the logarithm:
>
> $$|y| = e^{\frac{t^2}{2}+C} = e^{\frac{t^2}{2}}e^C$$
>
> To remove the modulus, simply add $\pm$ to the right-hand side:
>
> $$y = \pm e^C e^{\frac{t^2}{2}}$$
>
> Since $\pm e^C$ is just a constant, we can write it as A, so:
>
> $$y = Ae^{\frac{t^2}{2}}$$
>
> Some students get confused over why A is used instead of $\pm A$. The rationale will be explained soon.

Use the initial condition $y = 2$ when $t = 0$ to find A:

$$2 = Ae^0 = A$$

So $A = 2$, and the solution is:

$$y = 2e^{\frac{t^2}{2}}$$

3 Arbitrary Constants

We revisit the following portion of the above example:

Since $\pm e^C$ is just a constant, we can write it as A, so:

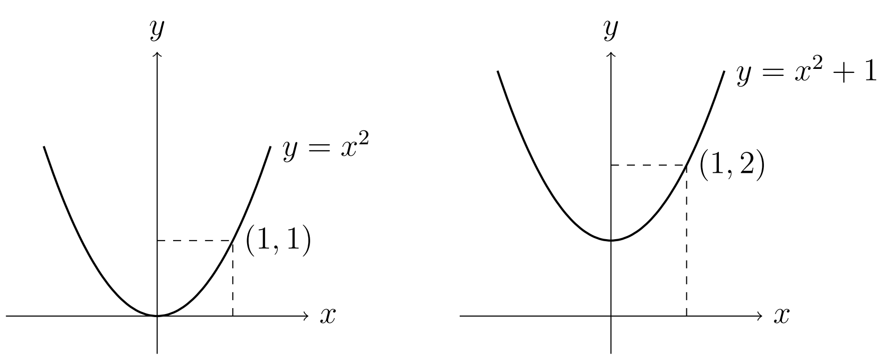
$$y = Ae^{\frac{t^2}{2}}$$

As mentioned earlier, there is often confusion surrounding the arbitrary constant; so we dedicate a section to it.

First it would be helpful to understand why an arbitrary constant is necessary. When can we leave the answer with an arbitrary constant (no specific value), and when must we find the value of this constant?

Vertical Translation

To answer this question, we consider the graphical meaning behind a simple differential equation $\frac{dy}{dx} = 2x$. This tells us that when at any point on the graph, the gradient of the curve is twice as large as its x-coordinate. Notice how we made no mention of the y-coordinate, since the y-coordinate does not affect the gradient. We observe this graphically:



The graphs of $y = x^2$ and $y = x^2 + 1$ both have the same gradient of 2 at $x = 1$, even though they differ by a vertical translation (i.e. the $+1$ constant). In fact, a vertical translation by any amount will produce a function that still satisfies this differential equation! Imagine shifting the graphs up and down; while the y-coordinates change, the gradients at each x value does not. Therefore, all functions of the form $y = x^2 + c$ satisfies this differential equation. All such functions are said to belong to the same "family".

Of course, we all know that when we integrate $\frac{dy}{dx} = 2x$, an arbitrary constant must be added. This, is the graphical meaning behind the arbitrary constant.

General Solutions vs Specific Solutions

In the context of differential equations, a function satisfying the differential equation is called a solution; and there are two types of solutions: General solutions and specific solutions.

As its name suggest, general solutions are general: You must give the general form of all possible solutions, hence requiring an arbitrary constant c to denote that all curves in this "family" satisfy the differential equation.

Likewise, specific solutions are specific: There is only one such function satisfying this differential equation. This is often specified by a point that the curve must pass through, which generally restricts the set of solutions to one curve only. To illustrate, the only $y = x^2 + c$ that passes through the point $(1, 1)$ is the curve $y = x^2$. Prove it yourself!

Now we come back to the previous problem, where because $\pm e^C$ is just a constant, we can write it as A, so:

$$y = Ae^{\frac{t^2}{2}}$$

This question asks us for a specific solution, as inferred from the point (initial condition) given. So instead of leaving the constant $\pm e^C$ or A in the final answer, we must find its value. Furthermore, because it is a specific solution (one solution only), this constant can only have one value, meaning $\pm e^C$ is either positive or negative; it cannot accommodate both possibilities.

This justifies replacing it with an unknown constant A which has only one value (be it positive or negative). This makes the specific solution easier to find since A is definitely easier to work with than $\pm e^C$, even though both refer to the same shift.

4 One Important Tip

It is essential to maintain consistent units throughout the problem, especially when two different units measuring the same quantity appear in the question. Mixing units can lead to incorrect results, so always ensure that all quantities are expressed in the same units before performing any calculations.

For example, if $\frac{dh}{dt}$ (the rate of change of height) is given in cm per second, then h (height) should not be measured in meters, but in centimeters.

Example : The Importance of Consistent Units
The height h of a liquid in a rectangular tank is increasing at a rate of $\frac{dh}{dt} = 5\,\text{cm/s}$. Find the time taken for the height to increase by 1 meter.

Incorrect Approach: We mistakenly assume the units are consistent and proceed directly with the equation $\frac{dh}{dt} = 5$:

$$t = \frac{h}{\frac{dh}{dt}} = \frac{1}{5} = 0.2\,\text{seconds}$$

However, the answer is incorrect because $h = 1\,\text{m}$ and $\frac{dh}{dt}$ is given in cm/s. The units do not match.

Correct Approach: We must convert the height $h = 1\,\text{m}$ into centimeters, so $h = 100\,\text{cm}$:

$$t = \frac{h}{\frac{dh}{dt}} = \frac{100}{5} = 20\,\text{seconds}$$

By ensuring that the units are consistent, we now obtain the correct answer. ■

CHAPTER 16

COMPLEX NUMBERS

1 Complex Number Representation

There are 3 ways to represent a complex number, namely the "cartesian", "trigonometric", and "polar" forms

Cartesian form:

$$z = x + iy$$

Where x is the real part of the number, and iy is the imaginary part. We represent this using $\mathrm{Re}(z) = x$ and $\mathrm{Im}(z) = y$.

Trigonometric form:

$$z = r(\cos\theta + i\sin\theta)$$

Where r is the modulus and θ is the argument.

Polar form:

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

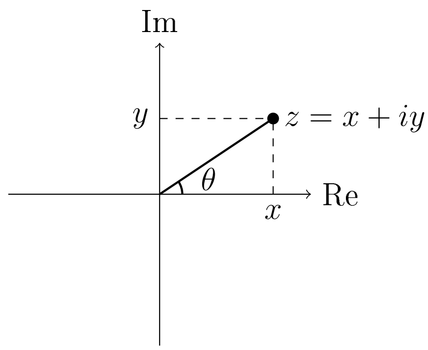
Where r is the modulus and θ is the argument.

These forms can be converted from one to another; different forms have different purposes and provide us with different information. For instance the cartesian form gives us the real and imaginary parts, while the polar form gives us the modulus and argument. The trigonometric form is usually an intermediate form when converting from cartesian to polar, since the cartesian and polar forms are mainly used.

2 Modulus and Argument

The modulus of a complex number $z = x + iy$ is its distance (length) from the origin in the complex plane. It is denoted by $|z|$ and calculated by the Pythagoras' Theorem:

$$|z| = \sqrt{x^2 + y^2}$$



The argument of a complex number is the angle θ from the positive real axis to the line representing z. The argument is denoted by $\arg(z)$ and is measured in radians. By definition, the argument must fall within the range $-\pi < arg(z) < \pi$.

There are two important results for modulus and argument, and they provide many computational shortcuts. For complex numbers z and w:

$$|zw| = |z||w|$$

$$arg(zw) = arg(z) + arg(w)$$

All other results can be derived from these two, such as $arg(z^n) = n\ arg(z)$.

It is fundamental, yet crucial to realise that a complex number is defined uniquely by its modulus and argument. For instance, a complex number with modulus $\sqrt{2}$ and argument $\frac{\pi}{4}$ must be equal to $\sqrt{2}e^{i\frac{\pi}{4}}$, or $1 + i$. These two forms represent the same number.

Knowing this unique correspondence also means that we can build a complex number if we know its modulus and argument. This is really helpful when finding complicated complex numbers, since we can break it into two simpler problems of finding the modulus, then finding the argument.

A typical H2 question goes like this:

> ***Example : Building a Complex Number using the Modulus and Argument***
> $z = 1 + i$ and $w = 2\sqrt{3} + 2i$. Find $\frac{z^5}{w^2}$.

Let $\alpha = \frac{z^5}{w^2}$.

$$|\alpha| = \left|\frac{z^5}{w^2}\right| = \frac{|\sqrt{2}|^5}{|4|^2} = \frac{\sqrt{2}}{4}$$

$$arg(\alpha) = arg(\frac{z^5}{w^2}) = 5arg(z) - 2arg(w) = \frac{11\pi}{12}$$

Thus, $\alpha = \frac{\sqrt{2}}{4}e^{i\frac{11\pi}{12}}$ is the answer. ■

Evidently, the concept of decomposing a complex number into its modulus and argument serves as a handy shortcut when it comes to complicated numbers.

3 Important Theorems

Fundamental Theorem of Algebra: A polynomial of degree n has exactly n roots, which may be complex or real, identical or distinct.

Conjugate Root Theorem: If the coefficients of a polynomial are all real, then for any complex root $z = a + ib$, the conjugate $z^* = a - ib$ must also be a root.

It is crucial to ensure that the said conditions are met before applying the Conjugate Root Theorem.

4 Tips and Techniques

Understand when different techniques are used

Rationalising: Rationalising is typically used for fractions containing a complex number in the denominator. It helps to eliminate imaginary components in the denominator, so as to isolate the real and imaginary parts.

> ***Example : Rationalising***
> Find the real and imaginary parts of $\frac{1+i}{1-i}$.
>
> To rationalise, multiply both the numerator and denominator by the conjugate of the denominator $1+i$:
>
> $$\frac{1+i}{1-i} \times \frac{1+i}{1+i} = \frac{(1+i)^2}{(1-i)(1+i)} = \frac{1+2i-1}{1^2-(-i)^2} = \frac{2i}{2} = i$$
>
> Thus, $\frac{1+i}{1-i} = i$. So the real and imaginary parts are 0 and 1 respectively. ■

Mod-Arg Decomposition: This method is helpful for determining complicated complex numbers, especially those made by multiplying or dividing multiple complex numbers. The modulus and argument are found one at a time, then put together to "build" the complex number. This has been covered in the previous section.

Let $z = a + ib$: This technique is typically used when there is only one complex number z, and works by expanding out z and determining a and b by comparing real and imaginary parts. If there are two or more complex numbers, such substitution is not advised since there would be too many variables.

> ***Example : Let*** $z = a + ib$
> Given that $z + \frac{1}{z} = 2i$, find z.
>
> Let $z = a + ib$. Then:
>
> $$z + \frac{1}{z} = a + ib + \frac{1}{a+ib}$$
>
> Rationalise the denominator:
>
> $$\frac{1}{a+ib} = \frac{a-ib}{a^2+b^2}$$
>
> Substitute this into the equation:
>
> $$a + ib + \frac{a-ib}{a^2+b^2} = 2i$$

Now, separate real and imaginary parts:

$$\text{Real part:} \quad a + \frac{a}{a^2 + b^2} = 0$$

$$\text{Imaginary part:} \quad b - \frac{b}{a^2 + b^2} = 2$$

Solve these equations to find a and b.

Quadratic Formula: Akin to its usual usage, quadratic equations with complex coefficients can also be solved using the quadratic formula. To illustrate the point, the above example can also be solved using the quadratic formula.

Again, though these cover many techniques in the syllabus, it is not exhaustive. The point here is that a firm understanding of how and why each technique works is essential to deploy them appropriately.

Double check that your argument falls within range

By definition, the argument of any complex number must fall in the interval $(-\pi, \pi]$. This is to ensure that no angle is "repeated": For instance $\frac{\pi}{5}$ and $\frac{11\pi}{5}$ describe the same angle on the argand diagram.

If your argument is out of range, add or subtract 2π repeatedly until the argument falls within range.

If unsure, draw the Argand Diagram

When dealing with modulus and arguments, drawing the argand diagram can help to visualise the modulus and argument much clearer, making them easier to be determined.

For example, it might be challenging to figure out the argument of $z = -3 - 5i$. The best method here would be to draw out the argand diagram, then use ASTC (the four quadrant method from A Math) to solve for the argument.

5 Conclusion

Checkpoint to Test Understanding

This is a checkpoint to see if you have any gaps in your understanding. Listed below are a few facts; if you do not understand why they are true, consider reading some of the above parts in further detail.

- $|e^{i\theta}| = 1$ for any θ
- $|\cos\theta + i\sin\theta| = 1$ for any θ
- $z^* = \frac{1}{z}$ when $|z| = 1$
- $\arg(zw) = \arg(z) + \arg(w)$
- $|z^n| = |z|^n$
- If $z = re^{i\theta}$, then the conjugate $z^* = re^{-i\theta}$
- $\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2}$ (Not required, but you should be able to derive it)
- $zz^* = |z|^2$

Links and Extensions beyond the Syllabus

Complex numbers have strong ties to Vectors, Geometry, and Trigonometry, and should not be viewed as an isolated topic.

Complex numbers are very useful to describe 2D rotations, which is not surprising, considering that multiplying two complex numbers sum their arguments. Complex numbers can even be extended to describe 3D rotations, using a type of number called "quaternions". This is definitely beyond the syllabus, but I encourage you to explore further if it interests you!

CHAPTER 17

VECTORS

1 Two Types of Vectors

Vectors can be thought of as a path between two points in space. Unlike scalars, which only have magnitude (e.g. speed, distance), vectors have both magnitude and direction (e.g. velocity, displacement). Vectors come in two main types: position vectors and free vectors.

Position vectors start from the origin and point towards a specific location in space. A position vector describes the location of a point relative to the origin, so we only look at them when the coordinates of a point is concerned.

Free vectors do not have a specific starting point. They describe a direction and magnitude of travel, but they can be placed anywhere in space. The direction vector of a line, normal vector of a plane, or the projection of one vector one another; these are all free vectors. Where we put them does not matter.

2 Magnitude and Direction

All vectors can be broken down into two components: their magnitude and their direction. We can depict this using the trivial formula:

$$\mathbf{v} = |\mathbf{v}|\ \hat{\mathbf{v}}$$

The magnitude of a vector is its length and is usually denoted by $|\mathbf{v}|$. Since it represents length, it can never be negative.

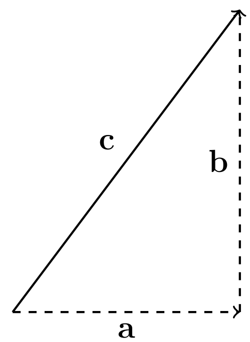
The direction of a vector is described by a unit vector that points in the same direction as the original vector but has a magnitude of 1. The notion of a unit vector is very important, and will be covered in further detail soon.

The unit vector of $\mathbf{v}$ is denoted by $\hat{\mathbf{v}}$, and it is obtained by dividing the vector by its magnitude:

$$\hat{\mathbf{v}} = \frac{1}{|\mathbf{v}|}\mathbf{v}$$

A unique vector can be described by its magnitude and unit vector. For example, a vector of magnitude 5 with a unit vector $\left\langle \frac{3}{5}, \frac{4}{5}, 0 \right\rangle$ must result in the vector $\mathbf{v} = \langle 3, 4, 0 \rangle$.

3 Resolving a Vector



Like how we can add two perpendicular vectors **a** and **b** tip-to-tail to give a vector **c**, likewise we can decompose **c** into **a** and **b**. We call this resolving a vector, where **a** and **b** are called the components of **c**. Each component is itself a vector.

We often use resolving as we are usually more concerned of one component but not the other. For example in Physics, where we resolve a force (vector) **F** and into its horizontal and vertical components to analyse them separately.

The horizontal component is given by:

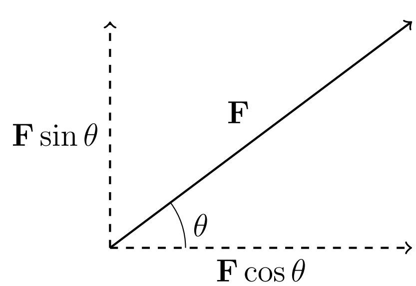
$$\mathbf{F}\cos\theta$$

The vertical component is given by:

$$\mathbf{F}\sin\theta$$

Notice we intentionally omitted the modulus, since the components are vectors and not lengths.

These components form the sides of a right-angled triangle with **F** as the hypotenuse.



This concept of resolving will help greatly in understanding all subsequent results.

4 Dot Product

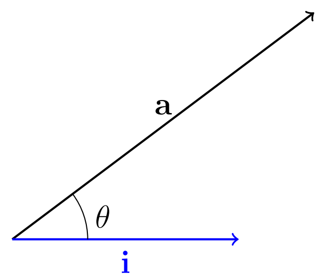
We are all too familiar with the dot product. Given two vectors **a** and **b** separated by an angle θ:

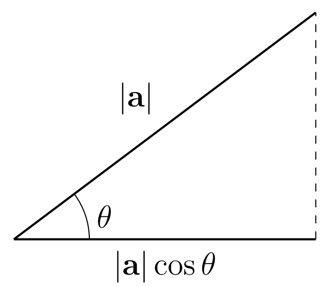
$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$$

But besides knowing how to compute a dot product, do we know what it is measuring?

Consider the formula for the Body Mass Index (BMI $= \frac{m}{h^2}$). Beyond knowing how to compute it, we also know what it is measuring: a rough index to indicate how fit one is.

Likewise, we also ought to know what the dot product is measuring. To make things clear we use a simple example, taking the dot product of the vector **a** and the unit vector **i**.





Computationally:

$$\mathbf{a} \cdot \mathbf{i} = |\mathbf{a}||\mathbf{i}| \cos\theta = |\mathbf{a}| \cos\theta$$

From the diagram we can clearly see that the result of the dot product, $|\mathbf{a}| \cos\theta$, is exactly the component of **a** that is in the same direction as **i**.

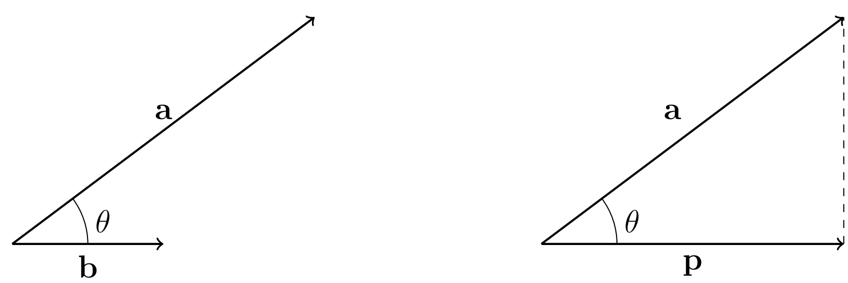
We may start to develop an intuition that the dot product represents how aligned two vectors are. So now we look at the dot product formula again: $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$.

On the right-hand side, $|\mathbf{b}|$ simply represents the length of vector **b**, while $|\mathbf{a}| \cos\theta$ represents the length of the component of vector **a** that is in line with **b**. Multiplying them gives the product of the length of these two components.

The more aligned **a** is with **b**, the larger $|\mathbf{a}| \cos\theta$ is, hence a larger dot product. Meanwhile if **a** and **b** are perpendicular, no component of **a** aligns with **b**, rendering a zero dot product. Thus, dot products should be thought of as how closely two vectors align.

5 Projections

Projections can be understood as the "shadow" one vector casts onto another. Using a similar example as before, we project **a** onto **b**, naming this projection vector **p**:



Notice that the length of **b** did not affect the projection **p** in any way. **b** could have been longer or shorter, but **p** will remain exactly the same. Only the direction of **b** is involved in the projection of **a** onto **b**.

Mathematically, the projection of vector **u** onto vector **v** is given by:

$$\text{Proj}_{\mathbf{v}}(\mathbf{u}) = (\mathbf{u} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}}$$

Let's try to break down the formula to interpret the meaning of each part.

- Firstly we look at the terms: $(\mathbf{u} \cdot \hat{\mathbf{v}})$ is a scalar, while $\hat{\mathbf{v}}$ is a vector.

- The vector $\hat{\mathbf{v}}$ should be easier to understand, since we should expect the projection of **u** onto **v** to have the same direction of **v**. This is the direction of the projection vector.

- Next, the dot product $(\mathbf{u} \cdot \hat{\mathbf{v}})$, like we saw earlier, simply gives the length of **u** that is in line with **v**. This is the exact magnitude of the projection vector.

In section 2 we mentioned how a vector is composed of a magnitude and direction, so we put them together to find:

$$\text{Proj}_{\mathbf{v}}(\mathbf{u}) = (\mathbf{u} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}}$$

I want to make a special emphasis on why we use $\hat{\mathbf{v}}$ for the direction, and not any other vector like **v**. While both of them have the same direction, $\hat{\mathbf{v}}$ is the only such vector that has a magnitude of 1, so multiplying it with $(\mathbf{u} \cdot \hat{\mathbf{v}})$ does not alter the magnitude of the projection vector. Any other vector in the same direction will change the magnitude.

Once you grasp the concept of projections, many other parts of vectors come naturally. In the H2 syllabus, projections are used to find many things, especially shortest distances between two items, be they planes, lines or points.

6 Cross Product

The cross product is another operation we can perform on two vectors, but unlike the dot product, the result of a cross product is a vector rather than a scalar.

By definition:

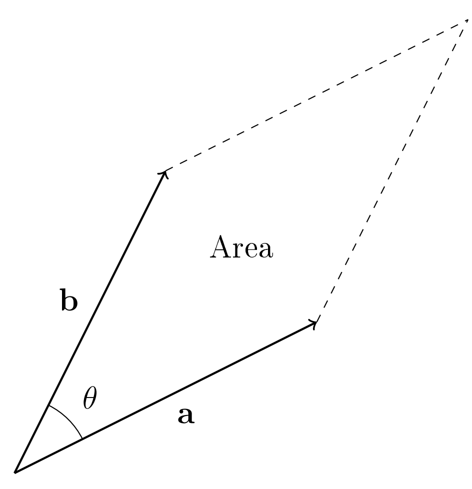
$$\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}| \sin\theta \, \hat{\mathbf{n}}$$

where:

- $|\mathbf{a}|$ and $|\mathbf{b}|$ are the magnitudes of the vectors,
- θ is the angle between them, and
- $\hat{\mathbf{n}}$ is a unit vector perpendicular to both $\mathbf{a}$ and $\mathbf{b}$, following the right-hand rule.

The magnitude of the cross product $|\mathbf{a} \times \mathbf{b}|$ represents the area of the parallelogram formed by the two vectors $\mathbf{a}$ and $\mathbf{b}$.

The right-hand rule helps determine the direction of the resulting vector: curl your right hand's fingers from $\mathbf{a}$ to $\mathbf{b}$, and your thumb will point in the direction of the resulting vector.



The cross product measures two things:

- The magnitude $|\mathbf{a}||\mathbf{b}| \sin\theta$ gives the area of the parallelogram formed by the two vectors. This can be thought of as two congruent triangles, each with areas $\frac{1}{2}|\mathbf{a}||\mathbf{b}| \sin\theta$.

- The direction is given by the vector perpendicular to both $\mathbf{a}$ and $\mathbf{b}$.

The cross product can be thought of as measuring the perpendicularity between two vectors; the more perpendicular two vectors are, the larger the resultant area.

The fact that the resultant vector perpendicular to both $\mathbf{a}$ and $\mathbf{b}$ is also extremely helpful, for example finding the normal vector to a plane, given two (non-parallel) vectors that lie on the plane.

The exact mechanics of how cross products work is not covered in the syllabus, but it is rather elegant. So I will just make a brief mention of it, for those who may be interested. An alternate way to define the cross product is through the determinant of a matrix.

For example, given vectors $\mathbf{a} = \langle 1, 0, 0 \rangle$ and $\mathbf{b} = \langle 0, 1, 0 \rangle$:

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix}$$

It just so happens that the determinant of a matrix represents the area scale factor of the linear transformation described by the matrix. Is it a coincidence that both the determinant and the cross product deal with area? Probably not. Vectors and matrices do not just look alike; they are intrinsically interwoven. If you are interested, 3Blue1Brown has an amazing series covering this on YouTube! (Disclaimer: Most of it is not in the syllabus)

7 Two Types of Questions

The H2 syllabus primarily has two types of vectors questions: One that I will call "algebraic", while the other is "application". A firm understanding in all vector operations (why they are used, how they are used, what they measure, and so on) is crucial for both question types; however there are some key differences in their nature.

Application Questions

This type of question deals mainly with the applications of vectors in real-world contexts, hence they appear together with a story most of the time.

> ***Example : Application Question***
> **Question:** A ship is located at point $A(2, 3, 5)$ and must travel to the line $r = \mathbf{a} + \lambda\mathbf{b}$, where $\mathbf{a} = \begin{pmatrix} 4 \\ 4 \\ 0 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$. What is the length of the shortest path that the ship can take from point A to the line?

Such questions provide you with a handful of information, like the coordinates of certain points or equations of certain lines, then has a few details you need to find. This could be the shortest path from some point to another, or the coordinate of another point.

The tools required here have been covered to some degree in the preceding sections, especially that of projections, a concept key to this question type. Needless to say, being able to utilise projections appropriately is often key to solving these questions.

Algebraic Questions

These questions usually appear like algebra questions and require you to manipulate them like an algebraic equation. Usually, there is no story attached to it; and the operations are much cleaner than the "application" questions, in that less column vectors have to be computed, etc.

> ***Example : Algebraic Question***
> Given two non-parallel vectors $\mathbf{p}$ and $\mathbf{q}$ are related by $(\mathbf{p} \times \mathbf{q}) \cdot (\mathbf{p} \times \mathbf{q} + \mathbf{q}) = 9$. Find $|\mathbf{p} \times \mathbf{q}|$.

Questions like these focus more on the algebraic manipulation of vectors, rather than how they are applied in the real world. So, vector-algebra operations are key to solving these problems.

These concepts include, but are not limited to:

- All of the above operations like the dot and cross
- Commutative property of dot products, but $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
- Associativity of dot products
- Distributive property of dot and cross product
- $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$
- $\mathbf{a} \times \mathbf{a} = \mathbf{0}$ (Note that this is not zero, but the zero vector)
- $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = 0$ (Why is this true? Is 0 the zero vector or the scalar zero?)

CHAPTER 18

PERMUTATIONS AND COMBINATIONS

1 Overview

Permutations and combinations are fundamental concepts in counting and probability, allowing us to determine how many ways objects can be arranged or grouped. The key difference between them is whether order matters: order matters in Permutations, but not in Combinations. This point will be expanded in later sections.

We will first build the intuition from the basic principles of counting and then explore various forms of permutations and combinations.

2 Addition and Multiplication Principles

Addition Principle: If we have two tasks, and one can be done in m ways and the other in n ways, and they are mutually exclusive (i.e. you can only do one or the other), then the total number of ways to do either task is $m + n$.

> ***Example : Addition Principle***
> To go to work, John can choose between taking 2 trains or 3 different buses. How many different routes can he take?
>
> The number of choices for trains = 2.
> The number of choices for buses = 3.
> Since the choices are mutually exclusive, we sum them. Thus, the total number of routes is $3 + 2 = 5$. ■

Multiplication Principle: If one task can be done in m ways and another task can be done in n ways, and both tasks need to be done in succession, then the total number of ways to perform both tasks is $m \times n$.

Example : Multiplication Principle
Max can choose either a red, blue, or green shirt, and then pair it with a black or blue pair of jeans.

The number of choices for shirts = 3.
The number of choices for jeans = 2.
Since the shirt and jeans are chosen in succession, we multiply them. Thus, the total number of outfits is $3 \times 2 = 6$. ■

3 Permutations

A permutation is an arrangement of objects where the order matters. An example of a permutation would be a queue, where each person in line has a specific position; the arrangement of the queue would be changed if two people were to swap positions.

In the syllabus, there are only a handful of question types; we will go through them below.

Permutations with No Restrictions

If we have 6 (distinct) people and we want to arrange them in a queue, there are 6 ways to assign the first position, 5 ways to assign the second position (since one person is already assigned), and so on. The total number of ways to arrange these 6 people is:

$$6 \times 5 \times 4 \times 3 \times 2 \times 1 = 6! = 720$$

Now, consider the case where we don't use all of the people. Suppose we want to arrange 4 people from a group of 6 people in a queue. Using the same logic as above, we have:

$$6 \times 5 \times 4 \times 3 = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{2 \times 1} = \frac{6!}{(6-4)!} = {}^6P_4 = 360$$

Thus, ${}^nP_r$ is the number of ways to arrange r objects from a set of n distinct objects and is given by:

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

Permutations with Identical Objects

When objects are not all distinct, we would have to make slight changes to the formula.

Let's consider 6 markers: 3 red, 1 blue, 1 green and 1 yellow. To illustrate the change, we first assume the red markers are labeled as distinct: R_1, R_2, and R_3.

Since the red markers are distinct, we can arrange the 6 markers in 6! ways. Let's focus on some permutations where only the red markers R_1, R_2, and R_3 swap places among themselves. Here are 6 different permutations:

$$R_1, R_2, R_3, B, G, Y \qquad R_1, R_3, R_2, B, G, Y$$

$$R_2, R_1, R_3, B, G, Y \qquad R_2, R_3, R_1, B, G, Y$$

$$R_3, R_1, R_2, B, G, Y \qquad R_3, R_2, R_1, B, G, Y$$

Each of these arrangements is distinct because the red markers are labeled R_1, R_2, and R_3. So, permuting them among themselves results in different permutations.

In reality, the red markers are *identical*, labeled simply as R, so the above 6 permutations are actually the same arrangement R, R, R, B, G, Y. In other words, we are counting the same arrangement multiple times because swapping the red markers among themselves does not create a new arrangement.

We have *over-counted* the permutations 6 times, as there are $3! = 6$ ways to arrange the red markers among themselves. To correct this over-counting, we divide by the number of ways the identical red markers can be arranged among themselves, which is 3!. This gives the correct number of distinct arrangements:

$$\frac{6!}{3!} = \frac{720}{6} = 120$$

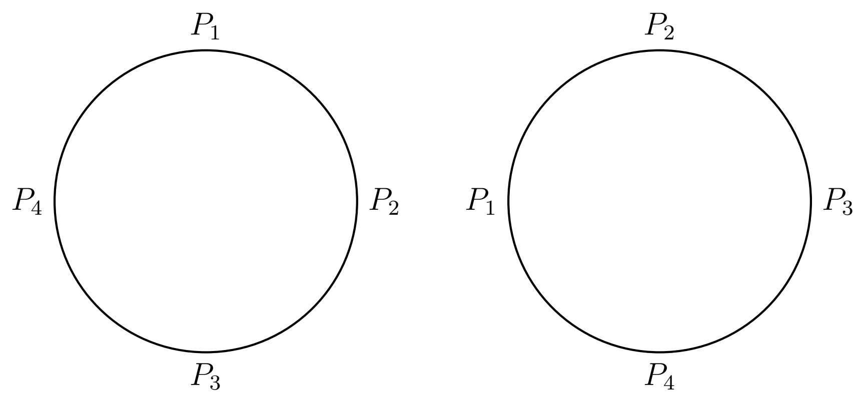
This idea generalises to account for any number of identical objects.

> ***Example : Permutations with Identical Objects***
> Given 3 green, 6 blue and 2 red pens, how many ways are there to arrange them in a row?
>
> Total no. of objects $= 3 + 6 + 2 = 11$
> No. of ways $= \frac{11!}{3!6!2!}$ ■

Permutations in a Circle

For circular arrangements, the arrangement is considered the same even when rotated.

Consider 4 people P_1, P_2, P_3, and P_4 seated around a circle. Rotating the circle does not produce a new arrangement, as everyone simply shifted but retain the same relative positions.



Original Arrangement Rotated Arrangement

As we can see, the second arrangement is just a rotation of the first one. Since everyone maintained the same relative positions (e.g. P_1 is still to the right of P_2), the two arrangements are considered identical.

This suggests that the original permutation formula ${}^nP_r$ has to be modified slightly again to prevent over-counting rotations of the same arrangement. There are two perspectives to understand this change.

Perspective 1: Suppose we arrange 6 people in a circle. We first permute them without accounting for the rotation of the table, effectively arranging them in a line, which means there are 6! arrangements.

Now accounting for rotation, each rotation of the circle results in the same arrangement. This means we have to divide by the number of times we can rotate (since that is the number of times we have over-counted rotations of the same arrangement). This is equal to the number of seats, since everyone can be moved down by this number of times before moving back to the starting arrangement. Hence, the number of distinct circular permutations is:

$$\frac{6!}{6} = 5!$$

In general, the number of ways to arrange n distinct objects in a circle of r spots is $\frac{{}^nP_r}{r}$.

Perspective 2: There is only 1 way to assign the seat to the first person, since all seats are identical when nobody is seated yet.

After the first person sits down, the seats now become distinct, since they differ in relative position to the first person. Thus we permute everyone else normally, where number of ways is 5!.

In general, the number of ways to arrange n distinct objects in a circle is $(n-1)!$.

What happens when we only require r objects from n to be arranged? We would have to choose r objects first, then permute them. The general formula is ${}^nC_r \times (r-1)!$ which is mathematically identical to the formula in Perspective 1. The idea of choosing is covered later on.

Permutations with Restrictions

Restrictions are essentially added conditions that forbid certain permutations, so we have to find a systematic approach to exclude these forbidden arrangements. Restrictions can come in any form; some restrictions include:

- Two objects cannot be adjacent.
- Two objects must be adjacent.
- Certain objects must take on certain positions.
- Two objects cannot appear in the same arrangement.
- At least x number of objects of a type must appear.

In most cases, restrictions will appear in conjunction with conditions like identical objects and arrangement in a circle. To effectively tackle these questions, we must equip ourselves with appropriate strategies. These techniques will be covered in a later section.

4 Combinations

A combination is a selection of objects where the order does not matter. This is different from permutations because we do not care about the arrangement, only the group itself. An example of a combination would be a bag of marbles; even if two marbles swapped position in the bag, it is still the same bag of marbles.

Like Permutations, the syllabus only has a handful of question types on Combinations.

Combinations with No Restrictions

To understand Combinations, we first consider Permutations. Suppose we want to know how many ways we can group 3 people from a set of 6 people.

We first calculate the permutations (as if the order mattered):

$$6 \times 5 \times 4 = {}^6P_3 = 120$$

For illustration, consider the 6 following permutations:

$$\begin{array}{ll} P_1, P_2, P_3 & P_1, P_3, P_2 \\ P_2, P_1, P_3 & P_2, P_3, P_1 \\ P_3, P_1, P_2 & P_3, P_2, P_1 \end{array}$$

Notice that even though these are 6 different permutations, they are in fact the same combination since the members in the group are the same. This means that for every 1 combination we have 6 permutations. More generally, for every group of r distinct members we have $r!$ permutations, since there are $r!$ number of ways to permute these r distinct members within the group.

Hence, we must divide by the number of ways to arrange these 3 people among themselves (which is 3!):

$$\frac{{}^6P_3}{3!} = \frac{120}{6} = 20$$

By the above explanation we have that:

$${}^nP_r = r! \times {}^nC_r$$

since for every 1 combination of r distinct objects we have exactly $r!$ permutations.

It follows that:

$${}^nC_r = \frac{{}^nP_r}{r!} = \frac{n!}{r!(n-r)!}$$

This formula gives the number of ways to choose r objects from a set of n distinct objects where the order does not matter.

Important Note: There are two situations where combinations can be applied:

1. Choosing r objects out of n distinct objects to form a group.
2. In a row of n objects, choose r objects to be of one type, while all remaining objects are another type by default. There must be exactly two types.

> ***Example : Choosing in a Row***
> In a row of 12 red and blue balls, how many arrangements contain exactly 7 blue balls?
>
> No. of arrangements = ${}^{12}C_7$
> This is because out of 12 balls, we choose 7 to be blue while the rest are red by default.
>
> Alternatively, no. of arrangements = ${}^{12}C_5$
> Since we can also choose 5 balls out of 12 to be red, while the rest are blue by default.
>
> Notice that both methods produce the same result: they are mathematically identical,

hence the result:

$$^nC_r = {}^nC_{n-r}$$

Choosing r objects to include in a group is the same as choosing $(n-r)$ objects to exclude from the group.

Forming Groups

Combinations often involve arranging objects into different groups. Special care must be taken when considering if the groups are distinct or identical.

Consider the grouping of 9 distinct books.

Example : Groups of 2, 3, and 4
Number of ways = $^9C_2 \times {}^7C_3 \times {}^4C_4$
Because after choosing 2 books for the first group, we are left with 7, and so on.

Also, notice that the order of grouping does not matter (i.e. we can pick the group of 4 first as well):
Number of ways = $^9C_4 \times {}^5C_2 \times {}^3C_3$
This yields the same result.

Chronology does not matter when grouping in this manner.

In the above case, the groups are clearly distinct, since a group of 4 is different from a group of 2. Now we consider identical groups.

Example : Three groups of 3
Number of ways = $({}^9C_3 \times {}^6C_3 \times {}^3C_3) \div (3!)$

Let's see why the $\div(3!)$ is necessary:
Treating all 3 groups as distinct initially, there would be $({}^9C_3 \times {}^6C_3 \times {}^3C_3)$ ways to group the 9 books into 3 distinct groups. However, the groups are identical in reality, so we have actually over-counted the number of ways 3! times, which is the number of ways the 3 groups can be permuted among themselves. Thus we divide by 3!.

This is the same way we accounted for identical objects. In some sense, groups are also objects that should be accounted for when they are identical.

Choosing in Succession

A binary number is where each digit is either a 0 or 1; how many binary numbers of 16 digits can we form?

Perspective 1: We take a walk down the line, starting from the first digit. At each digit, we have exactly 2 choices, either a '0' or '1'. So for the first digit we have 2 choices, for the second digit we also have 2 choices. This repeats till the last digit where we have 2 choices.

These 16 rounds of choosing occurred in succession, so we have $2 \times 2 \times \cdots \times 2 = 2^{16}$ such binary numbers.

Perspective 2: We consider mutually exclusive cases, that is, where there is exactly zero '1's, where there is exactly one '1', where there is exactly 2 '1's, and so on.

For the case of zero '1's, no. of ways $= {}^{16}C_0$
For the case of one '1', no. of ways $= {}^{16}C_1$
For the case of two '1's, no. of ways $= {}^{16}C_2$
This is repeated for all 17 cases, afterwards we sum to get:
Total number of ways $= {}^{16}C_0 + {}^{16}C_1 + {}^{16}C_2 + \cdots + {}^{16}C_{16} = 2^{16}$

Notice the previous line where the binomial coefficients summed up to yield a neat result. It is no coincidence.

$${}^{n}C_0 + {}^{n}C_1 + \cdots + {}^{n}C_n = \binom{n}{0}(1)^0 + \binom{n}{1}(1)^1 + \cdots + \binom{n}{n}(1)^n = (1+1)^n$$

In fact, the principles that define combinations is the very backbone of the Binomial Theorem.

Combinations with Restrictions

Restrictions are essentially added conditions that forbid certain groupings, so we have to find a systematic approach to exclude these forbidden combinations. Restrictions can come in any form; some restrictions include:

- Two objects cannot appear in the same group.
- Two objects must appear in the same group.
- At least x number of objects of a type must appear in the group.

In most cases, restrictions will appear in conjunction with conditions like forming groups or choosing in succession, or even together with any of the conditions from Permutations. To effectively tackle these questions, we must equip ourselves with appropriate strategies. These techniques will be covered very soon.

5 Solving Techniques

When faced with more complex problems, especially those involving restrictions, the following techniques can be used to simplify the problem and arrive at a solution more efficiently. These methods allow us to break down problems into more manageable parts.

Complements

Complements is an extremely useful technique, especially when it is easier to find the exact opposite (or complement) of a situation. Instead of calculating the total number of ways directly, we calculate the number of ways for the opposite scenario, and subtract it from the total possible number of outcomes.

> ***Example : Complements***
> Suppose we want to count how many 5-digit numbers can be formed using the digits 1 through 9, where 7 appears in at least one position. Instead of directly calculating the valid numbers, we can find the total number of possible 5-digit numbers without any restrictions and then subtract the cases where the digit 7 does not appear.
>
> Total number of 5-digit numbers without restrictions:
> $9^5 = 59049$
>
> Number of 5-digit numbers where 7 is not used:
> $8^5 = 32768$
>
> Thus, the number of 5-digit numbers where 7 appears at least once is:
>
> $$59049 - 32768 = 26281$$
>
> ■

The complement approach is especially useful when the complement (the opposite case) is easier to compute than the desired case itself.

Grouping

Grouping is a powerful technique to tackle questions with restrictions where certain objects must remain together. For example, if two or more objects must stay adjacent, we treat these objects as a single "group".

> ***Example : Grouping***
> Consider a problem where 6 people need to be arranged in a row, but person A and person B must sit next to each other. Instead of treating them as individuals, we group A and B together as one group, so now we are arranging 5 objects: 4 individuals and one group

(containing A and B). The number of ways to arrange these 5 objects is:

$$5! = 120$$

Within the group, the 2 people can swap positions, so we multiply the number of arrangements by 2! (since there are 2 ways to arrange them within the group):

$$120 \times 2! = 240$$

Thus, there are 240 ways to arrange the people such that the two specific individuals are adjacent. ■

Notice that by grouping them, we are guaranteed to satisfy the condition of the question where they must stay together. At the same time, do not forget that objects within a group can also be permuted.

Slotting

Slotting is a technique useful for problems where two or more objects must not be adjacent. We first arrange one group (usually the non-restricted group) and then "slot" the second group into available spaces to ensure that the restriction is satisfied.

Example : Slotting
Suppose we need to arrange 4 boys and 3 girls in a row, but the boys and girls must alternate. First, we arrange the 4 boys, which gives us:

$$4! = 24 \text{ ways}$$

Next, we slot the 3 girls into the spaces between the boys (there are exactly 3 slots available between 4 boys). The number of ways to arrange the girls in these slots is:

$$3! = 6 \text{ ways}$$

Thus, the total number of ways to arrange the boys and girls alternately is:

$$24 \times 6 = 144\ ways$$

■

Slotting is particularly useful when you need to prevent certain objects from being adjacent, since you manually separate the objects.

Cases

When the number of ways to achieve a solution isn't consistent, it's helpful to split the problem into cases. We break down the problem into mutually exclusive cases and then sum the number of ways for each scenario.

> ***Example : Cases***
> Suppose we need to select a committee of 3 people from a group of 8 men and 5 women, but the committee must include at least 1 woman. To solve this, we split the problem into cases based on the number of women selected:
>
> Case 1: 1 woman and 2 men
>
> $$\binom{5}{1} \times \binom{8}{2} = 5 \times 28 = 140$$
>
> Case 2: 2 women and 1 man
>
> $$\binom{5}{2} \times \binom{8}{1} = 10 \times 8 = 80$$
>
> Case 3: 3 women
>
> $$\binom{5}{3} = 10$$
>
> Total number of ways to select the committee:
>
> $$140 + 80 + 10 = 230$$

Using cases allows us to break down the problem into manageable parts where the conditions are consistent within each case.

Combining Techniques

Often, these techniques are used in conjunction with one another to solve more complex problems. For instance, we may use grouping alongside slotting, or we may use the complement technique in conjunction with cases to exhaust all possibilities.

> ***Example : Combining Techniques***
> 8 people are to be seated in a row, but A and B must sit next to each other, while X and Y cannot sit next to each other.
>
> Step 1: Use *grouping* to treat A and B as a single group, so now we are arranging 7 objects (6 individuals and the group).

Step 2: Use *slotting* to ensure that X and Y do not sit next to each other. First, arrange the remaining objects, and then slot the restricted individuals to satisfy the condition.

By combining these techniques, we ensure that the restrictions are met. ■

More Advanced Techniques (That the syllabus does not exactly require)

- **Symmetry:** Useful for simplifying problems by recognizing symmetrical cases that can be grouped together.

- **Recursion:** Some problems can be broken down into smaller sub-problems that follow the same structure. This is often useful in advanced counting problems.

- **Inclusion-Exclusion Principle:** In situations where objects have overlapping restrictions, this principle helps to calculate the total number of ways by adding and subtracting the cases appropriately.

These techniques (complements, grouping, slotting, and cases) cover a large portion of solving strategies for permutations and combinations problems. Do also remember that most problems have multiple approaches, but we want to find the most efficient one. Problem-solving often requires flexibility and creativity.

With these core techniques in hand, you should be well-prepared to tackle most combinatorial problems in your syllabus.

6 Tips and Ideas

- **Understand Each and Every Operation**

  It is essential to understand both *how* to apply the formulae, and *why* we apply them in a particular way. For instance, why do we multiply in certain cases and add in others? Why do we choose in some situations and permute in others?

  Each step in a permutations or combinations problem involves selecting objects from a set and arranging or grouping them in a particular way. At each step, you need to ask yourself: *How many objects am I picking? How many options am I picking from?*

  Memorizing formulas is not enough—many problems require adapting the logic behind a formula to a new context. For example, a problem might require selecting items from a group but involve a condition that modifies how they are arranged. Knowing the *why* behind each formula will allow you to adapt your approach in novel situations.

- **Approach Questions Like You Are Telling a Story**

One way to ensure that your approach is logical is to approach each problem as if you are telling a story. Each step of the story may be associated with a number of ways.

> ***Example : Selecting and Arranging People for a Movie***
> From a group of 5 men and 7 women, 2 men and 4 women are selected to watch a movie. The two men must not sit next to each other. How many ways can this be done?
>
> First, we select 2 men from 5 and 4 women from 7:
>
> $$^5C_2 = \frac{5 \times 4}{2 \times 1} = 10, \quad ^7C_4 = \frac{7 \times 6 \times 5 \times 4}{4 \times 3 \times 2 \times 1} = 35$$
>
> Total number of ways to select 2 men and 4 women = $10 \times 35 = 350$.
>
> We now arrange the 4 women in a row. Since no restrictions are placed on their seating:
>
> $$4! = 24$$
>
> Total number of ways to arrange the women = 24.
>
> The 4 women create 5 available slots (one before the first woman, between the women, and one after the last woman). Since the two men must be separated, we choose 2 of these 5 slots to place the men:
>
> $$^5C_2 = \frac{5 \times 4}{2 \times 1} = 10$$
>
> Now, we multiply all the steps together:
>
> $$350 \times 24 \times 10 = 84000$$
>
> Thus, there are 84,000 ways to select and arrange the men and women such that the two men are not sitting next to each other. ■

If the story makes sense and flows logically, it is likely that your method is correct. If something feels off, there may be a logical gap in your reasoning. This storytelling approach helps to break down complex problems into simpler, logical steps and ensures that no steps are skipped or done out of order.

- **Be Sensitive Toward the Nuances**

Many problems in permutations and combinations hinge on subtle distinctions. For example, are the objects distinct or identical? Does the order of selection matter? Is this a case where we are choosing or permuting?

> ***Example : Choosing Representatives***
> From a class of 6 boys and 4 girls, a representative has to be chosen for Maths, English, and Science. How many ways can the representatives be chosen such that there is at least one boy and one girl?
>
> What are the nuances here? What is distinct and what is identical? What requires choosing and what requires permuting?

Before even starting to solve a problem, identify these conditions clearly. Misunderstanding a single condition—such as whether the objects are distinct—can completely change the outcome. Techniques like grouping, slotting, or using cases rely on these nuances, so ensuring that you understand the precise nature of the problem is crucial.

- **Do Not Isolate Permutations from Combinations**

Permutations and combinations are often intertwined. Many problems require both in a single solution. For instance, a problem may first require you to select objects (a combination), and then arrange those selected objects (a permutation).

Don't think of them as separate topics but as different tools to solve counting problems. The key is knowing when to choose and when to arrange. Misapplying a permutation when a combination is needed—or vice versa—can lead to incorrect answers, so always double-check what is being asked: *Am I selecting? Or am I arranging?*

- **Always Verify with Smaller Numbers**

If you ever feel unsure about a solution, verify it with smaller numbers. For example, if you're working on a problem with a large number of objects, try solving a simpler version with just 2 or 3 objects. This helps you see whether the logic you've applied holds in the simplified version. If it does, you're likely on the right track.

Simplifying the numbers makes it easier to visualize or count the arrangements manually, which can help build confidence in the approach.

- **Break Down Large Problems into Sub-Problems**

  Sometimes a single problem can seem overwhelming due to the number of objects or conditions. In such cases, break the problem into smaller sub-problems.

  For example, if you're arranging objects with multiple restrictions, deal with one restriction at a time, and use techniques like grouping or slotting. Then, gradually build up to the full solution.

- **Look for Symmetry and Redundancy**

  Some problems can be made easier by identifying symmetrical cases or redundant computations. For instance, when arranging identical objects, be mindful that different permutations may result in the same arrangement. This insight helps you avoid over-counting.

  Similarly, problems involving circular arrangements often have rotational symmetry, meaning that different rotations of the same arrangement are identical. Use this to reduce the complexity of your calculations.

- **Pay Attention to Special and Boundary Cases**

  Always consider special or boundary cases that can change the outcome of a problem. For example, when choosing or arranging objects, ask yourself: What happens when the number of objects chosen is 0 or the maximum number possible?

  > ***Example : Where a Special Case is to be Neglected***
  > How many non-empty subsets of the set $\{1, 2, 3, 4, 5\}$?
  >
  > Here the answer is $2^5 - 1$ where 1 is subtracted since the subset cannot be empty. ■

  By thinking about the extremes, you ensure that your solution covers all possibilities.

- **Check for Double-Counts**

  A common mistake in permutations and combinations is accidentally counting the same arrangement multiple times. This happens frequently when there are overlapping conditions or when the problem involves identical objects. To avoid this, it's crucial to ensure that each combination or arrangement is counted only once.

Let's consider the following example:

Example : Checking for Double-Counting
A committee of 7 people is to be formed from a group of 4 women and 5 men. The requirement is that the committee must have at least 3 women and at least 3 men.

Incorrect Approach:

One might attempt to solve this by:
1. Choosing 3 women from the 4 women:

$$^4C_3 = 4$$

2. Choosing 3 men from the 5 men:

$$^5C_3 = 10$$

3. Then choosing the 7th member from the remaining 3 people (1 woman and 2 men):

$$^3C_1 = 3$$

Multiplying these together:

$$4 \times 10 \times 3 = 120$$

While this looks correct at first glance, it is actually over-counting the scenarios.

Let's take a specific example:
Suppose the 3 women chosen are W_1, W_2, W_3.
The 3 men chosen are M_1, M_2, M_3.
Now, choosing W_4 as the 7th member will give one valid committee.

However, if you had first chosen W_1, W_2, W_4 as the 3 women and M_1, M_2, M_3 as the men, and then chosen W_3 as the 7th member, you would have ended up with the exact same committee.

This means the same committee is being counted more than once, which leads to an over-count.

Correct Approach:

Instead of the above, we can break the problem into two cases:
1. Case 1: Choose 3 women and 4 men.

$$^4C_3 \times ^5 C_4 = 4 \times 5 = 20$$

2. Case 2: Choose 4 women and 3 men.

$$^{4}C_{4} \times ^{5} C_{3} = 1 \times 10 = 10$$

Adding the two cases together:

$$20 + 10 = 30$$

Thus, the total number of ways to form the committee is 30.

By considering both cases separately and avoiding the extra step of choosing one more member from the remaining group, we prevent the problem of double-counting.

Being able to generate examples like the above $W_1, W_2, W_3, M_1, M_2, M_3, W_4$ is helpful in picking out faults in your method. This is a skill that can help you visualise the validity of your logic.

Part II

Connecting The Dots

CHAPTER 19

POLYNOMIALS

1 Overview

Algebra and polynomials are foundational to the H2 syllabus, yet they are not covered explicitly since they do not hold much content in themselves, that are relevant to the H2 syllabus. Thus I dedicated a chapter here to run through all the important bits that you and I use every so often.

I personally love polynomials, as they are one of the easiest functions to deal with. They are easy to differentiate and integrate, we can easily find the output of a specific input, and we know its behaviour at extreme values.

Consider a polynomial $f(x) = 4x^3 + 5x - 2$, and another function like $g(x) = \csc(x)$. Which is easier to find? $f(3)$ or $g(3)$? $f'(x)$ or $g'(x)$? $\int f(x)\,dx$ or $\int g(x)\,dx$? The answer is clear.

In all, polynomials are just easy to manipulate and predict.

This is also the exact purpose of the Maclaurin Series, or more generally the Taylor Series. That is, to express complicated functions as polynomials, simply because polynomials are much easier to work with.

Small angle approximations also exploit the proximity of trigonometric functions with lower-order polynomials at small values to get approximations precise to a few decimal places.

You get the idea. Polynomials encompass great potential and applications, thus we definitely need to look at it in more detail.

2 Factorising

Factorising to polynomials is like a pen to a piece of paper, whatever that may mean. Without a pen, a piece of paper can only do so much.

Factorising is crucial in determining roots, inequalities, cancelling unwanted terms, simplifying expressions, comparing factors, and so much more. It just makes expressions nice to look at. The question should be "how", not "why" we factorise. Below I summarise some tools to help you factorise your expressions.

Common Factorisations

These are the ones ought to be committed to memory, and should be recognised at the first opportunity.

$$\begin{gathered} a^2 + 2ab + b^2 = (a+b)^2 \\ a^2 - 2ab + b^2 = (a-b)^2 \\ a^2 - b^2 = (a+b)(a-b) \\ a^3 + b^3 = (a+b)(a^2 - ab + b^2) \\ a^3 - b^3 = (a-b)(a^2 + ab + b^2) \end{gathered}$$

Grouping

Look for terms that share common factors or structures. Group terms together, factor out the common factors, and simplify.

> ***Example : Factor by Grouping***
> Consider the expression $x^3 - x^2 + x - 1$. We can group the terms:
>
> $$(x^3 - x^2) + (x - 1) = x^2(x-1) + 1(x-1)$$
>
> Then, factor out $(x-1)$:
>
> $$(x-1)(x^2+1)$$

Completing the Square

Completing the square is a method used to simplify quadratic expressions and is especially useful for solving quadratic equations. We can also use it to determine the factors of a quadratic equation.

For example, to complete the square for $x^2 + 6x + 5$:

$$x^2 + 6x + 5 = (x + 3)^2 - 9 + 5 = (x + 3)^2 - 4$$

Example : Solving by Completing the Square
Solve $x^2 + 6x + 5 = 0$ by completing the square.

$$x^2 + 6x + 5 = (x + 3)^2 - 4 = 0$$

$$(x + 3)^2 = 4$$

$$x + 3 = \pm 2$$

$$x = -3 \pm 2$$

$$\therefore x = -1 \text{ or } x = -5$$

Alternatively, we can apply the $a^2 - b^2$ formula to get the factorisation $(x + 5)(x + 1)$.

To use this efficiently, please be clear on how "completing the square" works. You are literally creating a square by inserting (adding or subtracting) a term required to complete the square. Of course, you must make up for this elsewhere (subtracting or adding).

This is important so I use one more example. We all know that $(x + 1)^2 = x^2 + 2x + 1$.

Consider the expression $x^2 + 2x - 3$. We want to create the term $x^2 + 2x + 1$ because it is a square. So we add 1 (and to make up for it we need to subtract 1).

$$x^2 + 2x - 3 = (x^2 + 2x + 1) - 1 - 3 = (x + 1)^2 - 4$$

Only complete the square by adding and subtracting constants! While you can technically complete the square with variables, the final expression often serves you no use.

$$x^2 + 2x - 3 = (x^2 + 2x + 2x + 4) - 2x - 4 - 3 = (x + 2)^2 - 2x - 7$$

Evidently, this expression has only brought us further from the solution, since we cannot determine any factors from it. We do not want variables outside the square, since they are very messy to deal with.

Quadratic Formula

The quadratic formula is a method for solving any quadratic equation $ax^2 + bx + c = 0$.

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

This formula is a byproduct of completing the square; a shortcut derived from it.

Example : Solving by Quadratic Formula
Solve $2x^2 + 3x - 5 = 0$ using the quadratic formula.

Here, $a = 2$, $b = 3$, and $c = -5$.

$$x = \frac{-3 \pm \sqrt{3^2 - 4(2)(-5)}}{2(2)} = \frac{-3 \pm \sqrt{9+40}}{4} = \frac{-3 \pm \sqrt{49}}{4}$$

$$x = \frac{-3 \pm 7}{4}$$

$$\therefore x = \frac{4}{4} = 1 \quad \text{or} \quad x = \frac{-10}{4} = -2.5$$

The discriminant, $\Delta = b^2 - 4ac$, plays a crucial role in determining the nature of the roots:

- If $\Delta > 0$, the equation has two distinct real roots.
- If $\Delta = 0$, the equation has one real root (a repeated root).
- If $\Delta < 0$, the equation has no real roots.

The above conditions are often taken for granted. Do we know why they are true?

The discriminant appearing within the square root is no coincidence. Notice that if the discriminant is negative, the square root is in trouble; it is unable to produce any real values to x, rendering there no real solutions.

When the discriminant is zero, the square root disappears, leaving only $x = -\frac{b}{2a}$ as the only solution. The last case is trivial.

Polynomial Long Division

Polynomial long division works exactly like the long division we are all used to. It is critical when dealing with complex algebraic fractions, providing a quick method of simplification when dividing one polynomial by another.

For example, dividing:

$$\frac{x^3 - 6x^2 + 11x - 6}{x - 1}$$

yields a quotient of:

$$x^2 - 5x + 6$$

with no remainder, which can be done using long division.

Long division is often used in conjunction with the factor theorem to determine the roots of a cubic polynomial.

Remainder and Factor Theorems

The Remainder Theorem states that if a polynomial $f(x)$ is divided by $x-c$, the remainder is $f(c)$. The Factor Theorem builds on this, stating that if $f(c)=0$, then $x-c$ is a factor of $f(x)$.

These theorems can be used to quickly factor polynomials and find roots. For example, consider the polynomial:

$$f(x)=x^3-6x^2+11x-6$$

By evaluating $f(1)$, we find:

$$f(1)=1^3-6\cdot 1^2+11\cdot 1-6=0$$

So, $x-1$ is a factor, and the polynomial can be factored as $(x-1)(x^2-5x+6)$.

The factor theorem is often utilised by trial-and-error with simple integer values like $x=0,\pm 1,\pm 2$, or sometimes simple fractions like $\frac{1}{2}$. For more complex values, the problem would provide relevant hints in the context.

Using Substitution

At times, the equation at hand might seem tricky to factorise. An appropriate substitution can simplify polynomials, or even non-polynomial terms. For example, solving $x^4+4x^2-5=0$ can be simplified by letting $y=x^2$, reducing the polynomial to a quadratic in y.

To illustrate that substitution can be utilised beyond polynomials, consider $e^{4x}+6e^{2x}+9$ using the substitution $y=e^{2x}$. Or an equation in the 2022 SMO, $\sqrt{x+2-2\sqrt{x+1}}+\sqrt{4x^2+5-4\sqrt{x+1}}$ using the substitution $y=\sqrt{x+1}$.

3 Symmetry

The symmetries in polynomials are frequently overlooked. In the H2 syllabus, there is one symmetry that can frequently provide shortcuts; that is, the reflection symmetry of all quadratics.

For those interested, all cubics also possess rotational symmetry about their point of inflexion, though this is rarely touched on in the syllabus. This section focuses more on the former symmetry.

Reflection Symmetry in Quadratics

For a quadratic function $f(x) = ax^2 + bx + c$, the vertex of the parabola occurs at $x = -\frac{b}{2a}$. Why is that?

We begin with the parent function $y = x^2$ which is symmetrical in the line $x = 0$. By completing the square, $ax^2 + bx + c$ can be rewritten as:

$$a\left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a}$$

Looking at the transformation on x, only, this is a shift by $\frac{b}{2a}$ units in the negative x direction. y transformations can be neglected, since they only shift or stretch the graph vertically, which does not change the line of symmetry.

This symmetry allows us to reduce the amount of calculation needed when solving problems involving quadratics.

An Alternate Way to find Roots

One common application of symmetry is when solving quadratic equations or inequalities. If we know one root of the equation, we can use the symmetry to immediately find the other.

> ***Example : Using Symmetry to Find Roots***
> Consider the quadratic equation $f(x) = x^2 - 6x + 8 = 0$. We solve for one root using factor theorem:
>
> $$f(2) = 2^2 - 6(2) + 8 = 0$$
>
> We now know that $x = 2$ is a root. Since the vertex of the quadratic is located at:
>
> $$x_{\text{vertex}} = -\frac{-6}{2(1)} = 3$$
>
> the other root must be symmetrically located on the other side of the vertex, i.e., at $x = 4$. Therefore, the roots are $x = 2$ and $x = 4$, without needing to factor the quadratic explicitly. ■

This method is usually more helpful in verifying roots, since we have easier ways to determine the roots of quadratics. The above relationship implies that the midpoint of the two roots always lies on the line of symmetry. This provides a quick way to check if your roots are correct.

Shortcut in finding Maxima and Minima

We know that all quadratics only have one vertex, and by symmetry this vertex must lie on the line of symmetry. This means we can find the coordinates of the vertex (maxima or minima) simply by inputting the x-coordinates of the line of symmetry into the function.

> ***Example : Finding the Minimum Value***
> Consider the quadratic function $f(x) = 2x^2 - 8x + 5$. We are asked to find the minimum value of the function.
>
> Since the quadratic opens upwards (because $a = 2 > 0$), the minimum occurs at the vertex. The vertex is located at:
>
> $$x_{\text{vertex}} = -\frac{-8}{2(2)} = 2$$
>
> To find the minimum value, we substitute $x = 2$ into the function:
>
> $$f(2) = 2(2)^2 - 8(2) + 5 = 8 - 16 + 5 = -3$$
>
> Thus, the minimum value of the function is -3, without needing to complete the square or use differentiation.

It is also worth noting that the polarity of the leading coefficient (coefficient of the highest x power) determines if the quadratic has a maxima or minima.

4 The Fundamental Theorem of Algebra

We are familiar with the Fundamental Theorem of Algebra, yet we often overlook its implications.

The Fundamental Theorem states that every non-constant polynomial of degree n has exactly n roots (real or complex), counting multiplicities.

This means that for any polynomial, such as $2x^5 - 4x^3 + 7x - 9$, we can rewrite it as:

$$k(x - x_1)(x - x_2) \cdots (x - x_5)$$

where $x_1, x_2, \ldots, x_5$ are the roots of the polynomial, and k is a constant. These roots may be real or complex. This factorised form allows us to explore several properties of the roots themselves.

Sum of Roots and Product of Roots

For example, most of us know the following formulae for a quadratic polynomial:

$$\text{Sum of roots} = -\frac{b}{a}, \quad \text{Product of roots} = \frac{c}{a}$$

for a quadratic polynomial $ax^2 + bx + c$. These formulae can be derived from a simple case of the theorem, by writing the quadratic as:

$$ax^2 + bx + c = a(x - x_1)(x - x_2)$$

Expanding this, we get:

$$a(x^2 - (x_1 + x_2)x + x_1 x_2)$$

From here, by comparing coefficients, we see that:

$$\text{Sum of roots} = -(x_1 + x_2) = -\frac{b}{a}, \quad \text{Product of roots} = x_1 x_2 = \frac{c}{a}$$

This is only one of the many cases out there. In fact, this notion is further generalised into the "Vieta's Formula", a concept widely used in math competitions. It is still somewhat useful for the H2 syllabus nonetheless, providing a shortcut or two at times.

Though this idea is more commonly seen in mathematical Olympiads and less so in the H2 syllabus, understanding the implications of the Fundamental Theorem of Algebra provides deeper insight. It helps clarify certain links, such as why polynomials of higher degrees also follow these properties in more complex forms, extending beyond quadratics.

5 The Behavior of Polynomials

End Behavior

The end behaviour of a graph refers to how it behaves as $x \to \infty$ or $x \to -\infty$. For a polynomial, this is dictated by the leading term, since it dominates all other terms as x becomes very large (in both the positive and negative direction).

> ***Example : End Behavior of a Cubic Polynomial***
> Consider the cubic polynomial:
>
> $$f(x) = x^3 - 6x^2 + 11x - 6$$
>
> As $x \to \infty$, the x^3 term dominates, so the graph goes to $+\infty$. As $x \to -\infty$, the graph goes to $-\infty$, indicating that the polynomial has opposite end behavior. ■

Stationary Points

A polynomial of degree n can have at most $n - 1$ stationary points (where the derivative is zero). This can be easily deduced since a polynomial of degree n would have a derivative of degree $n - 1$.

> ***Example : Turning Points of a Quartic Polynomial***

Consider the quartic polynomial:

$$f(x) = x^4 - 4x^3 + 6x^2 - 4x + 1$$

This polynomial has 3 turning points, which can be found by solving $f'(x) = 0$. ■

6 A Common Pitfall

When it comes to factorising polynomials, I see that many students embrace the shortcut of the "polynomial solver" function in their calculators. While this is definitely helpful and efficient in finding and checking your roots, I have seen many students overlook one step, thereby losing unnecessary marks.

I will first put out the shortcut, then address the said pitfall. As a side note, I personally do not utilise this shortcut, since I see value in factorising polynomials independently. This shortcut is still useful to most, nonetheless.

Say we have a polynomial to be factorised, like $x^3-6x^2+11x-6$. While the traditional method is to trial-and-error with integer roots, divide with long-division, then use the quadratic formula to factorise; the shortcut provides us with the roots instantly.

For older calculators like the Casio fx-96SG PLUS, this has famously been named "mode-3-3", referring to the buttons you press to access the "polynomial root finder"; contemporary calculators share the same feature.

Simply typing in your coefficients and hitting "=", you will get the 3 roots of this equation, $x = 1, 2, 3$.

What most students will do here is to immediately plug each root into a different bracket, in a form like this:

$$x^3 - 6x^2 + 11x - 6 = (x-1)(x-2)(x-3)$$

Or if the root is a fraction like $\frac{3}{2}$, they will write the factor as $(2x - 3)$.

Most of the time, this happens to be correct. This is merely a coincidence, though one of great probability.

Now we consider another polynomial $5x^3 - 30x^2 + 55x - 30$ which gives $x = 1, 2, 3$ again using the polynomial-finder. At this juncture many students will immediately write:

$$5x^3 - 30x^2 + 55x - 30 = (x-1)(x-2)(x-3)$$

Is this really correct though?

If we take the time to multiply out the right-hand side, we will see that it gives us $x^3 - 6x^2 + 11x - 6$. We are actually short by a factor of 5.

This is because the root-finder simply gives you the roots; it does not give you the factorisation directly. Thus we always have to ensure that the coefficients are equal on both sides. The easiest way is to check the coefficient of the leading term (term with the highest power).

Consider again $5x^3 - 30x^2 + 55x - 30 = (x-1)(x-2)(x-3)$, the incorrect equation. We see that the leading term on the left has a coefficient of 5, while that on the left is 1. So we just multiply the left side by 5 to account for this. This, is the step that many students miss out.

CHAPTER 20

TRIGONOMETRY

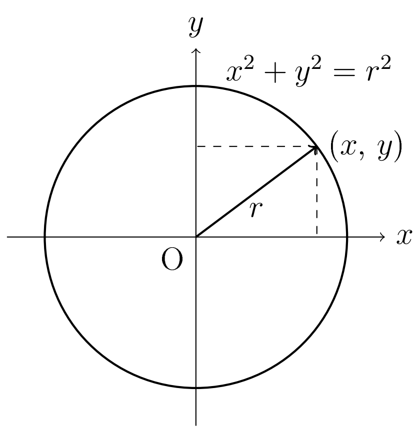
1 The Pythagorean Identity

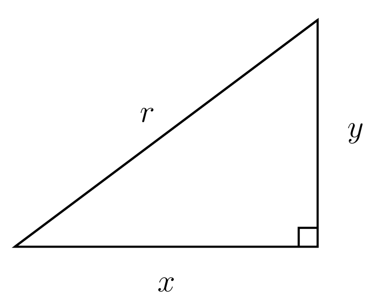
Trigonometry is intrinsically related to geometry. In this chapter we explore some of these links.

Consider the equation of a circle with radius r and center at the origin, given by:

$$x^2 + y^2 = r^2$$

This equation describes the relationship between the coordinates of any point (x, y) on the circle and the radius r.





Does this equation look familiar? In fact, it is a direct application of the Pythagorean Theorem! If we consider a right triangle with sides x and y, and the hypotenuse as the radius r, we see that:

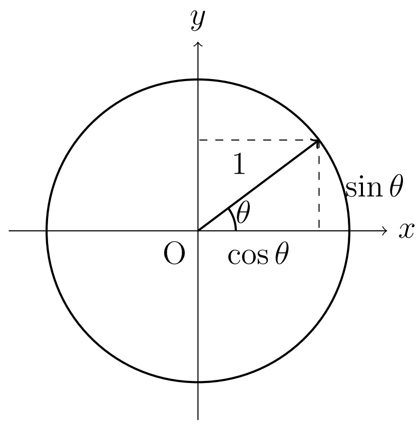
$$x^2 + y^2 = r^2$$

This relationship defines the circle. Imagine one end of the radius r stuck to the origin, while the other end is free to rotate. At any instance during this rotation, there is a right triangle with width x and height y, meaning the other end of the radius must lie at $(x,\ y)$, hence defining the coordinates of the circle using Pythagoras' Theorem.

A special case of the circle is the unit circle, which has a radius $r = 1$. The unit circle is centered at the origin and has the equation:

$$x^2 + y^2 = 1$$

This circle is fundamental in trigonometry because it simplifies many relationships. We construct an arbitrary radius that forms an angle θ with the positive x-axis.



On the unit circle, the coordinates of any point (x, y) can be interpreted as $(\cos\theta, \sin\theta)$ using the trigonometric ratios, where θ is the angle formed with the positive x-axis.

Visually, $\cos\theta$ is the width while $\sin\theta$ is the height of the right triangle at any point on the unit circle. This illustrates the periodicity of sin and cos, since the angle $(2\pi + \theta$ produces the same right triangle as θ, since it takes 2π radians to go around the circle once.

This also gives us the famous Pythagorean identity:

$$\sin^2\theta + \cos^2\theta = 1$$

This relationship holds for all angles θ, making it one of the key identities in trigonometry.

Now notice what happens when we divide throughout by $\cos^2\theta$:

$$\frac{\sin^2\theta}{\cos^2\theta} + \frac{\cos^2\theta}{\cos^2\theta} = \frac{1}{\cos^2\theta}$$

$$\tan^2\theta + 1 = \sec^2\theta$$

Repeating for $\sin^2\theta$:

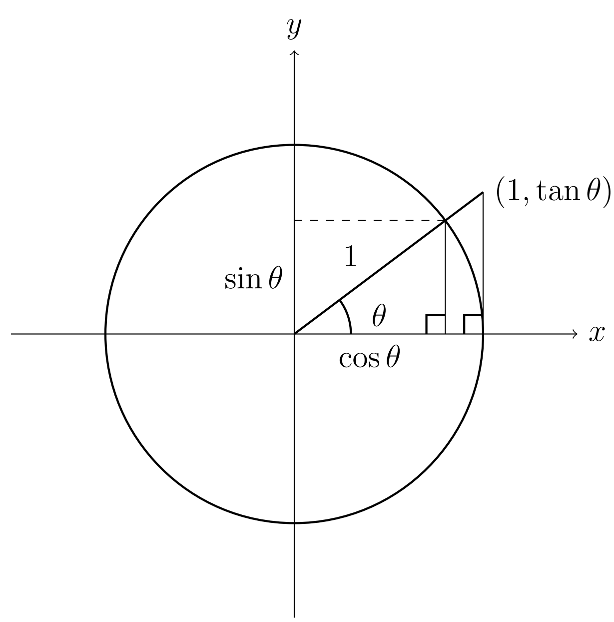
$$\frac{\sin^2\theta}{\sin^2\theta}+\frac{\cos^2\theta}{\sin^2\theta}=\frac{1}{\sin^2\theta}$$
$$1+\cot^2\theta=\csc^2\theta$$

Two more equations can be derived from the Pythagorean Identity, thus it suffices to remember that $\sin^2\theta+\cos^2\theta=1$.

2 Visualizing the Tangent

By definition, the tangent of an angle θ is the ratio of the opposite and adjacent sides of a right triangle, which means that using the definitions of sin and cos, we have that $\tan\theta=\frac{\sin\theta}{\cos\theta}$. But the tangent is also defined a straight line that only touches a curve at one point. Why do they share the same name?

Again we construct a unit circle, this time extending the radius slightly beyond the circle, intersecting the tangent at a point, labelled $(1, \tan\theta)$. The x-coordinate here is 1 since it has the same length as the radius, while the y-coordinate is $\tan\theta$ by the definition of the tangent function (i.e. ratio of opposite side to adjacent side).



Notice the pair of similar right triangles above. The smaller has a width of $\cos\theta$, while the larger has a width of 1. The smaller has a height of $\sin\theta$, while the larger has a height of $\tan\theta$. Comparing the ratio of sides of similar triangles:

$$\frac{\tan\theta}{1}=\frac{\sin\theta}{\cos\theta}$$

Thus we arrive at $\tan\theta = \frac{\sin\theta}{\cos\theta}$.

This beautifully illustrates how $\tan\theta$ is indeed the tangent of a circle. Yet this is only the tip of the iceberg; similar illustrations and reasoning exist for sec, cot, and csc! For starters, the word "secant" is defined as a straight line that cuts a curve at two points. Feel free to look up similar visualisations!

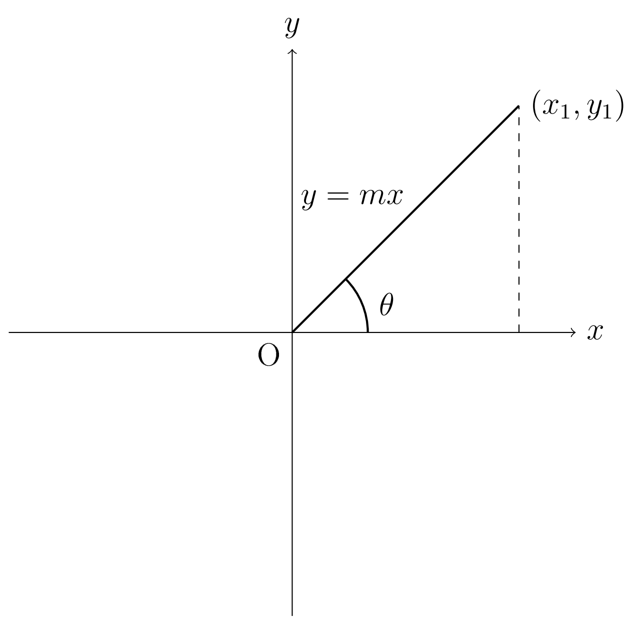
3 Tangent as a Gradient

One of the most fundamental yet often overlooked properties of the tangent function is its direct relationship to the gradient of a straight line passing through the origin. Specifically, the gradient of such a line is simply $\tan\theta$, where θ is the angle the line makes with the positive x-axis.

Consider a straight line passing through the origin at an angle θ with the positive x-axis. The equation of such a line can be written as:

$$y = mx$$

where m is the gradient of the line, and c is neglected since the y-intercept is 0. The relationship $m = \tan\theta$ then holds.



We are of course familiar with the definition of gradient as rise over run. Considering the line segment from the origin (0, 0) to (x_1, y_1), the rise is $(y_1 - 0)$ while the run is $(x_1 - 0)$, thus giving:

$$m = \frac{y_1}{x_1}$$

Now we turn out attention to θ, the angle formed between the straight line and the positive x-axis. By definition, the tangent of an angle is the ratio of the opposite side to the adjacent side in a right triangle. Here, the opposite side is y_1, and the adjacent side is x_1. So we get:

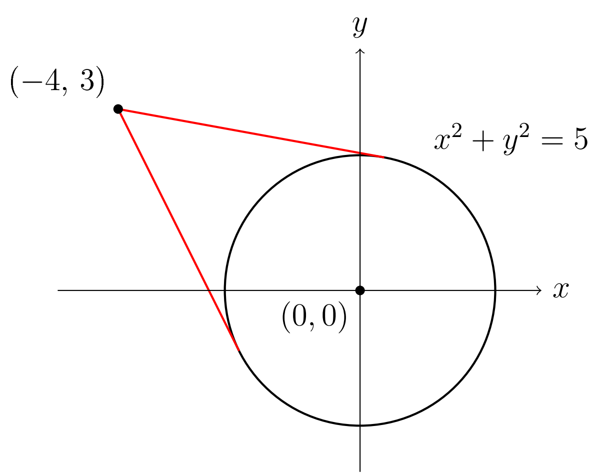
Thus, the gradient $\tan\theta$ is:

$$\tan\theta = \frac{y_1}{x_1}$$

We see that the two above equations imply their equality:

$$m = \tan\theta$$

It should be noted that this method is not limited to straight lines passing through the origin. In general, $m = \tan\theta$ is true for every straight line, where θ is the angle with the horizontal (any arbitrary straight line parallel to the x-axis). At times, this helps to solve tedious questions really quickly.

> ***Example : Find the equation of tangents to the circle*** $x^2 + y^2 = 5$***, where the tangents pass through the point*** $(-4, 3)$***.***
>
> 
>
>
> **Method 1: Brute Force**
>
> The most obvious method would be to label one point of tangency as $(a, \sqrt{5 - a^2})$ and find the distance to the point $(-4, 3)$ in terms of a. Utilising the properties of tangents to circles would give us that both red lines have length $2\sqrt{5}$, which should be equated to the expression in a above. The two solutions for a can be found after intensive algebraic manipulation. With these two coordinates we can then find the two tangent lines using the gradient and their coordinates.

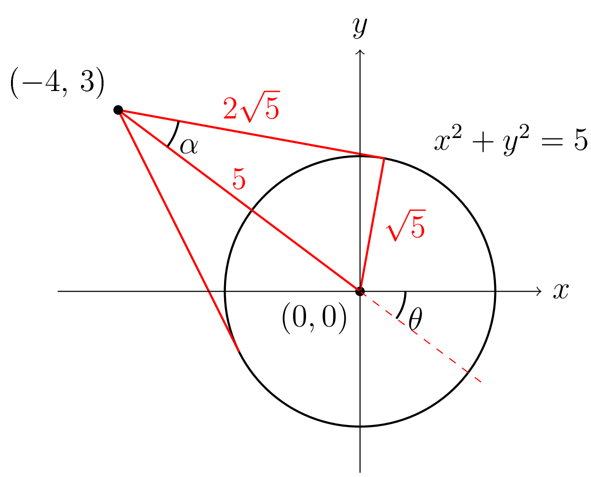
Method 2: Variation of Gradient

We know that the line must be in the form $y-3=m(x+4)$ because of the given condition, so we manipulate it slightly such that $y=m(x+4)+3$. This can then be substituted into the equation of the circle to give a quadratic in x.

In order for there to be exactly one point of tangency, the discriminant must equal to 0, which gives a quartic in m. This can be solved to give exactly two possible values of m. Compared to the above, this method is slightly less tedious, but still algebraically intensive nonetheless.

Method 3: Taking Advantage of Trigonometry

Firstly we can find the distance between (0, 0) and $(-4, 3)$ to be 5, and by Pythagoras' Theorem find the sides of the right triangle, as labelled in red.



We then label the two angles of interest, θ and α. Notice that the two tangents differ from the center line by $\pm\alpha$.

$$\theta=\tan^{-1}\left(\frac{3}{-4}\right),\ \alpha=\tan^{-1}\left(\frac{\sqrt{5}}{2\sqrt{5}}\right)$$

It follows that the two gradients of the two tangents are:

$$m=\tan(\theta\pm\alpha)=\tan\left(\tan^{-1}\left(\frac{3}{-4}\right)\pm\tan^{-1}\left(\frac{\sqrt{5}}{2\sqrt{5}}\right)\right)$$

$$=\frac{(-\frac{3}{4})\pm(\frac{1}{2})}{1\mp(-\frac{3}{4})(\frac{1}{2})}=-\frac{2}{11}\ \ or\ \ -2$$

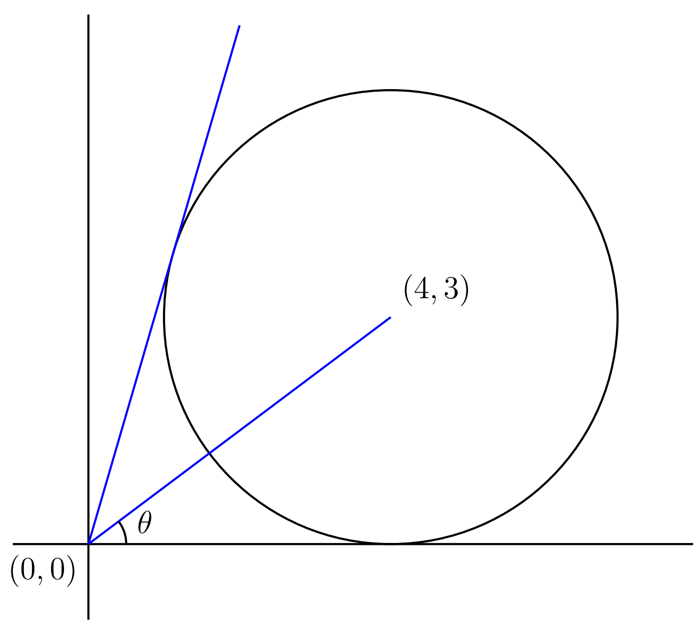
Thus the two equations are:

$$y - 3 = -\frac{2}{11}(x + 4)$$
$$y - 3 = -2(x + 4)$$

Evidently, this method is way quicker and neater than the previous two. This illustrates potential shortcuts by exploiting trigonometry in coordinate geometry. ∎

Next we consider a quick and simple solution to a question in the 2023 Singapore Mathematical Olympiad.

Example : Q14 SMO 2023
Consider the set of all possible pairs (x, y) of real numbers that satisfy $(x-4)^2+(y-3)^2 = 9$. If S is the largest possible $\frac{y}{x}$, find the value of $\lfloor 7S \rfloor$.



This question is basically asking us to find the gradient of the steepest tangent to the circle that passes through the origin, given by

$$S = \frac{y - 0}{x - 0}$$

so we draw in this tangent in blue, along with another line connecting the origin to the center of the circle.

The angle θ can be found using $\tan\theta = \frac{3}{4}$ since $\frac{3}{4}$ is the gradient.

Using the property of tangents at an external point, we can conclude that the gradient of the blue tangent is given by

$$S = \tan 2\theta = \frac{2\tan\theta}{1-\tan^2\theta} = \frac{24}{7}$$

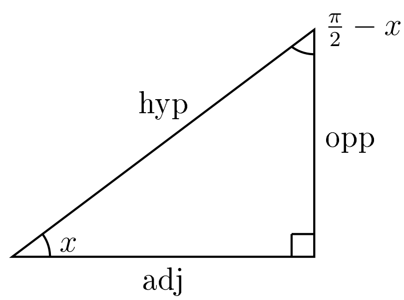
Thus, $\lfloor 7S \rfloor = \lfloor 7(\frac{24}{7}) \rfloor = 24$.

4 Interconnectedness in Trigonometry

The topic of trigonometry is elegantly interconnected within itself in many ways. In this section, we explore how various trigonometric relations can be deduced from geometric constructions and graphical representations. The aim here is not to be exhaustive but to highlight the general essence of deriving these relations using geometry and graphs.

Geometrical Deduction of $\sin x = \cos(\frac{\pi}{2} - x)$

Considering an arbitrary right-angled triangle.



According to the definitions of trigonometric ratios:

$$\sin x = \frac{\text{opp}}{\text{hyp}} \quad \text{and} \quad \cos\left(\frac{\pi}{2} - x\right) = \frac{\text{opp}}{\text{hyp}}$$

Thus,

$$\sin x = \cos\left(\frac{\pi}{2} - x\right)$$

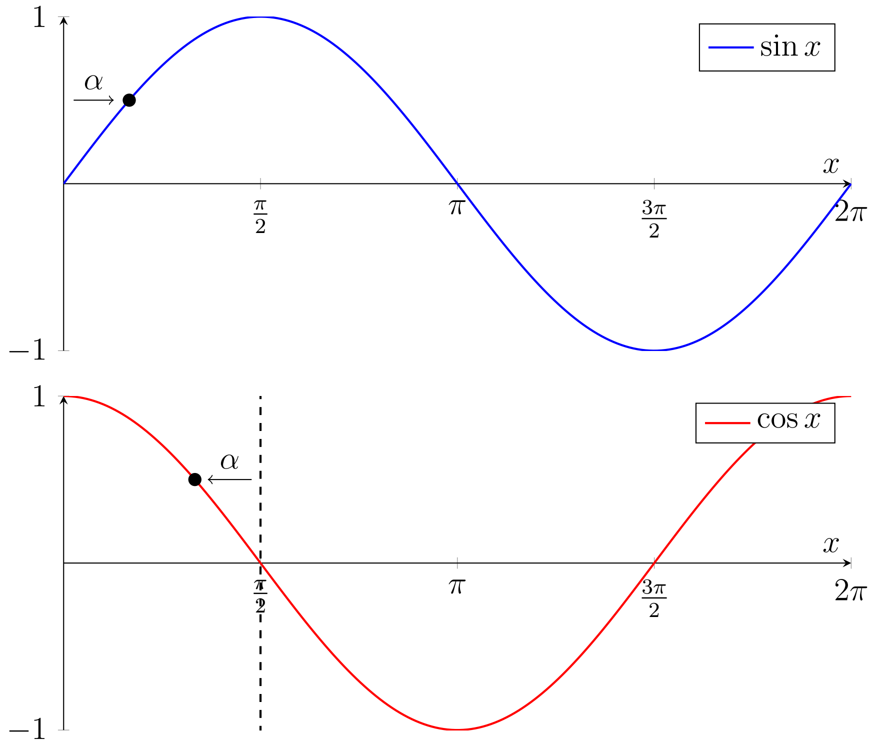
Similar methods can be used to derive:

$$\tan x = \cot\left(\frac{\pi}{2} - x\right) \quad \text{and} \quad \sec x = \csc\left(\frac{\pi}{2} - x\right)$$

Notice the special naming: “sine and cosine”, “tangent and cotangent”, “secant and cosecant”. Here, “co” can be thought of as “complementary”, since their angles complement each other.

Graphical Deductions

We can also derive this relationship from the graphs of $\sin x$ and $\cos x$.



Notice that moving α steps to the right on the sine graph produces the same result (y-value) as moving α steps to the left on the cosine graph, starting from the dashed line $x = \frac{\pi}{2}$. This visually confirms the relationship $\sin x = \cos(\frac{\pi}{2} - x)$.

Similarly, many other relationships, such as $-\sin x = \sin(\pi + x)$, can be deduced using the same method: by considering the output of a function as taking a number of steps on its graph. The symmetries in trigonometry allow for many such relationships to be derived, often providing convenient computational shortcuts.

Example : A Computational Shortcut
Solve $\cos 4\theta = \cos\theta$ for $0 \leq \theta \leq \pi$.

$$\cos 4\theta = \cos\theta \quad \Rightarrow \quad 4\theta = 2k\pi \pm \theta \quad \Rightarrow \quad 3\theta = 2k\pi \quad \text{or} \quad 5\theta = 2k\pi$$

Thus,

$$\theta = \frac{2k\pi}{3} \quad \text{or} \quad \frac{2k\pi}{5}$$

$$\therefore \theta = 0, \ \frac{2\pi}{3}, \frac{2\pi}{5}, \frac{4\pi}{5}$$

which satisfies the given domain of θ.

In this example, we have utilised the properties of cosine:

$$\cos\theta = \cos(-\theta) \quad \text{and} \quad \cos\theta = \cos(2\pi + \theta)$$

to determine the solutions quickly. This means we did not need to expand out $\cos 4\theta$ into a quartic in $\cos\theta$ which saved us considerable time.

If you did not notice, we could also have used the factor formula to solve this quickly. This is the focus of the next section.

Of course, this can also be viewed as a series of graph transformations, for instance the transformation of the sine graph to represent the equation $\sin(\pi + x)$: A horizontal shift of the sine graph by π, followed by reflecting the graph in the x-axis. But in my personal opinion, the idea of taking steps is a more intuitive approach to this concept.

5 Factor Formula

One of the key methods for simplifying trigonometric equations involves using factor formulae. Factor formulae come in two forms: "sum-to-product" and "product-to-sum", and can help simplify certain expressions significantly.

We begin by deriving one of the sum-to-product identities:

$$\cos A - \cos B = -2\sin\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right)$$

It is called "sum-to-product" because it transforms a sum of terms into a product of terms.

First, we rewrite the angles A and B within the brackets:

$$\cos A - \cos B = \cos\left(\frac{A+B}{2} + \frac{A-B}{2}\right) - \cos\left(\frac{A+B}{2} - \frac{A-B}{2}\right)$$

Then we expand them using the addition angle formula:

$$\cos\left(\frac{A+B}{2} + \frac{A-B}{2}\right) - \cos\left(\frac{A+B}{2} - \frac{A-B}{2}\right)$$

$$= \left[\cos\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right) - \sin\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right)\right]$$

$$-\left[\cos\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right)+\sin\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right)\right]$$

The two cos terms cancel out, leaving the sin terms:

$$\cos A-\cos B=-2\sin\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right)$$

A similar method is used to derive the other sum-to-product formulae. Product-to-sum formulae can be derived in a similar manner.

We solve the same example as above, now using the factor formula.

Example : Factor Formula

Solve for the general solutions of $\cos 4x=\cos x$.

$$\cos 4x-\cos x=-2\sin\left(\frac{4x+x}{2}\right)\sin\left(\frac{4x-x}{2}\right)$$

$$=-2\sin\left(\frac{5x}{2}\right)\sin\left(\frac{3x}{2}\right)$$

We equate each factor to zero:

$$\sin\left(\frac{5x}{2}\right)=0 \quad \text{or} \quad \sin\left(\frac{3x}{2}\right)=0$$

From $\sin\left(\frac{3x}{2}\right)=0$, we get:

$$\frac{3x}{2}=k\pi \quad\Rightarrow\quad x=\frac{2k\pi}{3}$$

From $\sin\left(\frac{5x}{2}\right)=0$, we get:

$$\frac{5x}{2}=k\pi \quad\Rightarrow\quad x=\frac{2k\pi}{5}$$

Thus, the general solutions are:

$$x=\frac{2k\pi}{3} \quad \text{or} \quad x=\frac{2k\pi}{5}$$

CHAPTER 21

BINOMIAL THEOREM

1 Overview

The binomial theorem provides a method for expanding powers of binomials (binomial means two terms), such as $(a+b)^n$. It is given by the formula:

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^{n-r} b^r$$

where $\binom{n}{r}$ is the binomial coefficient, also written as ${}^nC_r$, and represents the number of ways to choose r items from a set of n items.

There are multiple ways to arrive at the binomial formula, each with its own intuitive interpretation. In this chapter we will explore two of these approaches, and how the Binomial Theorem is related to similar concepts that we know.

2 Combinatorial Approach

We often see the symbol ${}^nC_r$ in binomial expansions, resembling choose in the topic of Permutations and Combinations. But do we know why it appears? What does this symbol represent in the context of binomials? What are we choosing, and where are we choosing from?

Consider a simpler example of multiplying two binomials:

$$(3x^2+3)(2x-1) = 6x^3 - 3x^2 + 6x - 3$$

Notice that each individual term on the right-hand side is a product of two specific terms from the left-hand side, one chosen from each bracket. For example, the term $-3x^2$ is the product of $3x^2$ from the first bracket and -1 from the second bracket. This idea generalizes to binomial expansion: whenever we multiply brackets, we are selecting one term from each bracket and summing up all such possible combinations.

Now, let's apply this idea to $(a+b)^n$, which is shorthand for multiplying $(a+b)$ with itself n times:

$$(a+b)^n = (a+b)(a+b)\cdots(a+b)$$

To form a term containing b^n, we need to choose b from all n brackets. The number of ways to do this is ${}^nC_n$, which equals 1, since there is only one way to choose all n b's from n brackets. Notice that since we have chosen b from all brackets, we would not be able to choose a from any bracket, since we can only choose exactly one term from each bracket. The corresponding term is ${}^nC_n a^0 b^n = b^n$.

Also notice that here we intentionally did not multiply by ${}^nC_0$ to represent choosing a from 0 brackets. This is because there are only two options from each bracket. The act of not choosing b means that we are definitely choosing a, since each bracket must provide exactly one term: if it is not b, then it must be a. In fact, this is also why ${}^nC_n = {}^nC_0$ and ${}^nC_r = {}^nC_{n-r}$ for that matter, since the number of ways to choose r objects that I want is the same as choosing $n-r$ objects that I do not want. Hence we only need to account for the choosing once.

Now consider the term ab^{n-1}. This term arises by choosing b from $n-1$ brackets and a from exactly one bracket. The number of ways to do this is ${}^nC_{n-1} = {}^nC_1$, as we are choosing 1 a from n brackets. Thus our term is ${}^nC_{n-1}a^1b^{n-1} = {}^nC_1 a^1 b^{n-1}$, where both forms are mathematically equivalent.

This reasoning continues for all terms in the expansion. The binomial coefficient ${}^nC_r$ represents the number of ways to choose r b's (or equivalently $n-r$ a's) from n brackets. This is why each term in the expansion has the form ${}^nC_r a^{n-r} b^r$.

3 Maclaurin Series Approach

We can also use the Maclaurin series to arrive at the Binomial expansion. The Maclaurin series for a function $f(x)$ is given by:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots$$

Applying this to $(1+x)^n$, we obtain:

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

which matches the binomial expansion when $a = 1$ and $b = x$.

In this case, the expansion is finite even though it may appear to go on indefinitely. This is because at some point, the coefficient becomes $\frac{n(n-1)(n-2)\cdots(n-n)}{(n+1)!}$ where the $(n-n)$ implies that the whole term equals zero. From this point onwards, all terms are zero by similar reasoning, thus giving a finite expansion which matches the Binomial Theorem exactly.

A caveat however, is that this series is only valid where $|x| < 1$ to ensure convergence. This is due to the inner mechanisms of the Maclaurin series, which is simply an approximation for small x. This is covered in the chapter of Series Expansions.

4 Binomial Distribution

The binomial distribution is closely related to the concepts of the binomial theorem. In fact, it applies the same idea of choosing from a set, but in the context of probability and random experiments. It models the probability of obtaining a fixed number of successes in a sequence of independent experiments, where each experiment has two possible outcomes (usually called "success" and "failure").

The binomial distribution formula is:

$$P(X = r) = \binom{n}{r} p^r (1-p)^{n-r}$$

where:

- n is the total number of trials (or experiments),
- r is the number of successes you want to observe,
- p is the probability of success on each trial,
- $1 - p$ is the probability of failure on each trial.

The binomial coefficient $\binom{n}{r}$ appears here because we are "choosing" r successes from n trials. The key idea is similar to that of the binomial expansion, where we were choosing terms from different brackets. In the binomial distribution, we are choosing specific outcomes (successes) from multiple independent trials.