A calculus of the absurd

Teymour Aldridge

Contents

Chapter 1

Introduction

These are my mathematics notes. Please direct any queries, comments, suggestions, etc to teymour@reasoning.page (or just say hi - I don't bite, promise!) **Be warned; these are a massive work in progress.**

This document could have been alternatively titled

$$\int \frac{d}{dx}\left(|\sqrt{2}|x\right) dx = |\sqrt{2}|x + c$$

"The central theme of Anna Karenina," he said, "is that a rural life of moral simplicity, despite its monotony, is the preferable personal narrative to a daring life of impulsive passion, which only leads to tragedy." "That is a very long theme," the scout said. "It's a very long book," Klaus replied.

— *Lemony Snicket,* the slippery slope, book eight of a series of unfortunate events

Chapter 2

Solving problems

2.1 Introduction

These are some helpful resources on solving maths problems which I've found on the internet.

2.2 General things to keep in mind

In his book "How to Solve It" the mathematician George Pólya suggests these steps for solving mathematics problems:

> If you cannot solve the proposed problem, look around for an appropriate related problem:
>
> - work backwards
> - work forwards
> - narrow the condition
> - widen the condition
> - seek a counterexample
> - guess and test
> - divide and conquer
> - change the conceptual mode

2.3 Specific heuristics

This list was written by A.H. Schoenfeld:

- **draw a diagram** (if at all possible)

- **examine special cases**
  - choose special values to exemplify the problem and get a "feel" for it
  - examine limiting cases to explore the range of possibilities
  - set any integer parameters equal to $1, 2, 3, \ldots$ in sequences and look for an inductive pattern.
- **try to simplify the problem**
  - exploiting symmetry
  - without loss of generality
- **consider essentially equivalent problems**
- replace the conditions with equivalent ones
- recombine the elements of the problem in different ways
- introduce auxiliary elements (e.g. substitutions)
- reformulate the problem
- change of perspective or notation
- considering argument by contradiction or contrapositive
- assuming you have a solution, and determining its properties (i.e. what *would* a valid answer look like)
- **consider slightly modified problems**
  - choose sub-goals (obtain partial fulfilment of the conditions)
  - relax a condition and then try to re-impose it
  - decompose the domain of the problem and work on it case by case
- **consider broadly modified problems**
  - construct an analogous problem with fewer variables
  - hold all but one variable fixed to determine that variable's impact
  - try to exploit any related problems which have similar
    * form
    * "givens"
    * conclusions
- **verifying solutions**
  - Does it pass these simple tests?

    * Does it use all the pertinent data?
    * Does it conform to reasonable estimates or predictions?
    * Does it withstand tests of symmetry, dimension analysis, or scaling?
  – Does it pass these general tests?
    * Can it be obtained differently?
    * Can it be substantiated by special cases?
    * Can it be reduced to known results?
    * Can it be used to generate something you know?

2.4 Psychological tactics

> The landlady hurried into the backyard, put the mousetrap on the ground (it was an old-fashioned trap, a cage with a trapdoor) and called to her daughter to fetch the cat. The mouse in the trap seemed to understand the gist of these proceedings; he raced frantically in his cage, threw himself violently against the bars, now on this side and then on the other, and in the last moment he succeeded in squeezing himself through and disappeared in the neighbour's field. There must have been on that side one slightly wider opening between the bars of the mousetrap... I silently congratulated the mouse. He solved a great problem, and gave a great example.
>
> That is the way to solve problems. We must try and try again until eventually we recognize the slight difference between the various openings on which everything depends. We must vary our trials so that we may explore all sides of the problem. Indeed, we cannot know in advance on which side is the only practicable opening where we can squeeze through.
>
> The fundamental method of mice and men is the same: to try, try again, and to vary the trials so that we do not miss the few favourable possibilities. It is true that men are usually better in solving problems than mice. A man need not throw himself bodily against the obstacle, he can do so mentally; a man can vary his trials more and learn more from the failure of his trials than a mouse.

—George Pólya, mice and men

From "The Art and Craft of Problem Solving" by Paul Zeitz.

- Get Your Hands Dirty: This is easy and fun to do. Stay loose and experiment. Plug in lots of numbers. Keep playing around until you see a pattern. Then play around some more, and try to figure out why the pattern you see is happening. It is a well-kept secret that much high-level mathematical research is the result of low-tech "plug and chug" methods. The great Carl Gauss, widely regarded as one of the greatest mathematicians in history..., was a big fan of this method. In one investigation, he painstakingly computed the number of integer solutions to $x^2 + y^2 \leqq 90,000$.

- Penultimate Step: Once you know what the desired conclusion is, ask yourself, "What will yield the conclusion in a single step?" Sometimes a penultimate step is "obvious," once you start looking for one. And the more experienced you are, the more obvious the steps are. For example, suppose that A and B are weird, ugly expressions that seem to have no connection, yet you must show that $A = B$. One penultimate step would be to argue separately that $A \geqq B$ AND $B \geqq A$.

  Perhaps you want to show instead that $A \neq B$. A penultimate step would be to show that A is always even, while B is always odd. Always spend some time thinking very explicitly about possible penultimate steps. Of course, sometimes, the search for a penultimate step fails, and sometimes it helps one instead to plan a proof strategy.

- Wishful Thinking and Make It Easier: These strategies combine psychology and mathematics to help break initial impasses in your work. Ask yourself, "What is it about the problem that makes it hard?" Then, make the difficulty disappear! You may not be able to do this legally, but who cares? Temporarily avoiding the hard part of a problem will allow you to make progress and may shed light on the difficulties. For example, if the problem involves big, ugly numbers, make them small and pretty. If a problem involves complicated algebraic fractions or radicals, try looking at a similar problem without such terms. At best, pretending that the difficulty isn't there will lead to a bold solution. At worst, you will be forced to focus on the key difficulty of your problem, and possibly formulate an intermediate question, whose answer will help you with the problem at hand. And eliminating the hard part of a problem, even temporarily, will allow you to have some fun and raise your confidence. If you cannot solve the problem as written, at least you can make progress with its easier cousin!

Chapter 3

Sequences and series

3.1 Sigma notation

One way of writing long sums (commonly used at GCSE), is to use an elipsis (the ... symbol). For example, if we were to write the sum of all the positive integers from 1 to 21 we could write it out in full as

$$\begin{aligned} &1+2+3+4+5+6+7+8+9+10+11+12 \\ &+13+14+15+16+17+18+19+20+21 \end{aligned} \tag{3.1}$$

The other way we could write it is as

$$1+2+3+\ldots+21 \tag{3.2}$$

Here we've used ... to stand for all the terms between 2 and 21. This notation works, but there's another way we could write this sum; using sigma notation! In this case, we could write this as

$$\sum_{i=1}^{21} i \tag{3.3}$$

Sigma notation is not as bad as it looks! All this means (when read aloud) is "the sum of all the values of i where i starts at 1 and ends at 21 (inclusive[1])".

[1] i.e. we include 1 and 21

In general there are three main parts to sigma notation - the place where we start counting from, the place where we finish counting and the "variable of indexation". In a more general case, we would have something of the form

$$\sum_{i=0}^{n} [\text{some expression depending on } i] \tag{3.4}$$

This would mean that we start at $i = 0$ and find the value of whatever the expression depending on i is. Say, for the sake of example, it happened to be[2] $3i^2$.

[2] This is just an example - the expression could be anything!

In this case, we would have $3 \cdot 0^2 = 0$. We would then add this to the value of the expression at 1 ($3 \cdot 1^2$), at 2 ($3 \cdot 2^2$), at 3 ($3 \cdot 3^2$), and so on (all the way to n - $3 \cdot n^2$).

3.2 "Telescoping" series

I think this is a further maths topic.

Sometimes, we have sums where everything cancels out nicely. A specific form of this is called a "telescoping" series, and the term refers to anything in the form

$$\sum_{k=1}^{n} [f(k+a) - f(k)] \tag{3.5}$$

Here, a is a constant natural number (e.g. $1, 2, 3, ...$) and k is the index of summation.

The key idea is that eventually the $-f(k)$ will "catch up" to the $+f(k+a)$. If a were equal to 1, for example, we would have that

$$\sum_{k=1}^{n} [f(k+1) - f(k)] = f(2) - f(1) + f(3) - f(2) + f(4) - f(3) + ... \tag{3.6}$$

A lot of terms are going to cancel here! When we add $f(2)$ and $-f(2)$ we get 0. The same thing happens for $f(3)$ and $-f(3)$, and so on. Every $-f(k)$ term in our series cancels out the $f(k)$ term from the previous term. This means that all the terms, *except* the first and the last terms are going to cancel each other out, leaving us with just

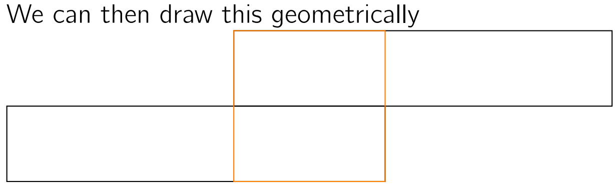
$$-f(1) + f(n) = f(n) - f(1) \tag{3.7}$$

One way to visualise this it to rewrite the sum.

$$\sum_{k=1}^{n} [f(k+a) - f(k)] = \sum_{k=1}^{n} [f(k+a)] - \sum_{k=1}^{n} [f(k)] \tag{3.8}$$

$$= \sum_{k=a}^{n+a} [f(k)] - \sum_{k=1}^{n} [f(k)] \tag{3.9}$$

We can then draw this geometrically



The orange section in the middle represents all the common terms. When we subtract the two series from each other, we are left with the last a terms from the sum of all the $+f(k+a)$ sum and the first a terms from the $-f(k)$ sum.

3.3 Arithmetic series

An arithmetic sequence goes something like

$$a, a+d, a+2d, a+3d, ..., a+(n-1)d \tag{3.10}$$

An arithmetic series is the sum of this sequence. If we want to find the sum of this series, it goes something like

$$S_n = [a] + [a+d] + [a+2d] + ... + [a+(n-1)d] \tag{3.11}$$

If we add all the as and the ds seperately, we get that

$$S_n = [n \cdot a] + [d + 2d + 3d + ... + (n-1)d] \tag{3.12}$$
$$= [n \cdot a] + [(1+2+3+...+(n-1))\, d] \tag{3.13}$$

Here, the question becomes one of the value of

$$1+2+3+...+(n-1) \tag{3.14}$$

To find this, we can get creative in how we group the terms in the sum. We can think about some smaller sequences, [3] for example when $n=5$

[3] A useful heuristic when working with sums of sequences is to first consider the sum of a few terms and then try to generate to n terms

$$1+2+3+4+5 = 15 \tag{3.15}$$

Or when $n=6$

$$1+2+3+4+5+6 = 21 \tag{3.16}$$

In general, the "trick" here is to group the terms quite imaginatively. In each case, we can add up the two terms on opposite sides of the sequence (e.g. $1+6 = 7$, as is $2+3$ and as is $3+4$).

Taking the general case, we have

$$1+2+3+...+(n-3)+(n-2)+(n-1) \tag{3.17}$$

When we add up the terms on opposite ends we get that

$$1+(n-1) = n \tag{3.18}$$
$$2+(n-2) = n \tag{3.19}$$
$$3+(n-3) = n \tag{3.20}$$
$$k+(n-k) = n \tag{3.21}$$

Because every two terms sum to n, for the sum from 1 to $n-1$ we have

$$\frac{1}{2}(n-1)(n) \tag{3.22}$$

This is because we have $n-1$ total items in our sequence, and every *two* terms sum to n, so the total sum is half the number of terms (the number of paired items adding together to give n).

Therefore, returning to our arithmetic series, overall

$$S_n = n \cdot a + \frac{1}{2}(n-1) \cdot n \cdot d \tag{3.23}$$

$$= n\left[a + \frac{1}{2}(n-1)d\right] \tag{3.24}$$

$$= \frac{n}{2}\left[a + \frac{1}{2}(n-1)d\right] \tag{3.25}$$

Equation **??** is also what is given in the formula booklet.

3.4 Geometric series

A geometric series is anything which looks like

$$a + ar + ar^2 + ... + ar^{n-1} \tag{3.26}$$

Here, a is the starting term, r is called the "common ration" (this name might make a bit more sense if you consider how each term is r times smaller than the next term) and n is the total number of terms. We can write the series in $\sum$-notation as (it's worth expanding this and checking that it is indeed the same as Equation **??**)

$$\sum_{j=1}^{n}\left[ar^{j-1}\right] \tag{3.27}$$

3.4.1 Finding the sum of a geometric series

This is one of those things which is much harder to find than it is to verify that one has indeed found the correct thing. **For A Level maths the sum is given in the formula booklet, and it's not necessary to be able to show why it is true.**

If we call the value of Equation **??** S_n (read as "the sum of n terms of the series), then we can multiply S_n by r and get that

$$rS_n = ar + ar^2 + ... + ar^n \tag{3.28}$$

From here we can proceed by finding the value of $S_n - rS_n$

$$S_n - rS_n = (a + ar + ar^2 + ... + ar^{n-1}) - (ar + ar^2 + ... + ar^n) \quad (3.29)$$
$$= a - ar^n \quad (3.30)$$

This can then be rearranged a bit by factoring out the S_n and then dividing through by $1 - r$.

$$S_n(1 - n) = a - ar^n \quad (3.31)$$
$$S_n = \frac{a - ar^n}{1 - r} \quad (3.32)$$
$$S_n = \frac{a(1 - r^n)}{1 - r} \quad (3.33)$$

3.4.2 Sum to infinity

The sum to infinity (which I *think* is in the formula booklet) doesn't always converge to a specific value (because the terms get bigger and bigger). For example, the value of

$$1 + 2 + 4 + 8 + ... \quad (3.34)$$

is infinite, but the value of

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + ... \quad (3.35)$$

is *not* infinite, and can be found. If the size (i.e. it doesn't matter if the value is positive or negative) of the common ratio is less than one, then the terms in the sequence get successively smaller and smaller. This is because multiplying two decimal numbers together makes a smaller number [4].

[4] e.g. $\frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$ and $\frac{1}{16}$ is smaller than $\frac{1}{4}$

If $|r| < 1$ (i.e. if r is greater than -1 and less than 1), then as n tends to infinity (written $n \to \infty$) then r^n tends to 0 (written $r^n \to 0$). We can apply this to Equation **??** and obtain that

$$S_\infty = \frac{a}{1 - r} \quad (3.36)$$

3.4.3 "Hidden" geometric series.

Sometimes geometric series can be hiding in plain sight!

For example, this series is actually a geometric series! [5]

[5] This requires knowledge of logarithms and exponents, which are explored lower down in this document.

$$\log_3 \left(3^1 \cdot 3^{\frac{1}{2}} \cdot 3^{\frac{1}{4}} \cdot ... \right) \quad (3.37)$$

Applying the log rules, we get

$$\log_3\left(3^1\right) + log_3\left(3^{\frac{1}{2}}\right) + \log_3\left(3^{\frac{1}{4}}\right) + ... \tag{3.38}$$

Which simplifies to

$$1 + \frac{1}{2} + \frac{1}{4} + ... \tag{3.39}$$

Which is just an infinite geometric series that converges to 2 (as $|r| < 1$).

3.5 The binomial theorem

3.5.1 Derivation

Note that this definitely isn't on any A Level specification.

A binomial is something in the form

$$(a + b)^n$$

What we're interested in is how to expand this bracket for large values of n. Small values of n are not too bad, e.g. $(a + b)^2$ [6] can be expanded like this

[6] If you've no idea how to expand this, review a GCSE textbook.

$$\begin{aligned}(a + b)^2 &= (a + b)(a + b)\\ &= (a + b)a + (a + b)b\\ &= a^2 + ab + ab + b^2\\ &= a^2 + 2ab + b^2 \end{aligned} \tag{3.40}$$

What about $(a + b)^3$?

$$\begin{aligned}(a + b)^2 &= (a + b)(a + b)^2\\ &= (a + b)(a^2 + 2ab + b^2)\\ &= (a + b)a^2 + (a + b)2ab + (a + b)b^2\\ &= a^3 + a^2b + 2a^2b + 2ab^2 + ab^2 + b^3\\ &= a^3 + 3a^2b + 3ab^2 + b^3 \end{aligned} \tag{3.41}$$

There's a very nice pattern which emerges as we get to higher powers. Let's try to find a nice expression for how to expand any binomial.

We can use the cases of the expansions of $(a + b)^1$, $(a + b)^2$ and $(a + b)^3$ to guess what form all binomials take. In general, the powers of a seem to start at a^n and then go down by one each term. The powers of b seem to do the opposite; they start at b^0 and increase up to b^n. In general, we have something which looks a bit like

$$(a+b)^n = c_1^n a^n + c_2^n a^{n-1} b + c_3^n a^{n-2} b^2 + ... + c_{n-1}^n ab^{n-1} + c_n^n b^n \tag{3.42}$$

Where c_k^n means the coefficient [7] of the kth item in the expansion of the nth power. What are the coefficients, however, and how do we compute them?

Let's think about what happens to the coefficients when we go from an expansion (e.g. $(a+b)^3$ whose value we do know) to an expansion which is one power higher, that we don't know. We can write this in the "general case" by letting $(a+b)^n$ stand for the expansion we do know, and $(a+b)^{n+1}$ for the one we don't. [8]

[7] That is to say, the number by which we multiply the variable. e.g. the coefficient of x^3 in $555x^3 + 14x + 15$ would be 555.

[8] **Protip: do this expansion by hand.**

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&= (a+b)[c_1^n a^n + c_2^n a^{n-1} b + c_3^n a^{n-2} b^2 + ... + c_{n-2}^n a^2 bn - 2 \\
&\quad + c_{n-1}^n ab^{n-1} + c_n^n b^n] \\
&= [c_1^n a^{n+1} + c_2^n a^n b + c_3^n a^{n-1} b^2 + ... + c_{n-2}^n a^3 bn - 2 \\
&\quad + c_{n-1}^n a^2 b^{n-1} + c_n^n ab^n] \\
&\quad + [c_1^n a^n b + c_2^n a^{n-1} b^2 + c_3^n a^{n-2} b^3 + ... + c_{n-2}^n a^2 bn - 1 + c_{n-1}^n ab^n \\
&\quad + c_n^n b^{n+1}] \\
&= c_1^n a^{n+1} + (c_1^n + c_2^n) a^n b + (c_2^n + c_3^n) a^{n-1} b^2 + ... \\
&\quad + (c_{n-2}^n + c_{n-1}^n) a^2 b^{n-1} + (c_{n-1}^n + c_n^n) ab^n + c_n^n b^{n+1}
\end{aligned} \tag{3.43}$$

The new coefficients clearly depend on the coefficients of the previous binomial [9]. But we already know that we can work out the coefficients of the expansion of the binomial of degree $n+1$, if we multiply it by the binomial expansion of n. To avoid having to do all that work, it would be better if we could write down the coefficients of all the terms given only some number n without having to multiply out all the brackets by hand.

[9] i.e. for $(a+b)^n$ this would be $(a+b)^{n-1}$

Using C_k^n to stand for the kth coefficient of the nth power, we can also write $(a+b)^{n+1}$ as follows

$$(a+b)^{n+1} = C_0^{n+1} a^{n+1} + C_1^{n+1} a^{n-1} b + ... + C_n^{n+1} ab^n + C_{n+1}^{n+1} b^{n+1} \tag{3.44}$$

If we compare $(a+b)^{n+1}$ to $(a+b)^n$, the first thing that's clear is that there's an a^{n+1} and b^{n+1} which don't exist in $(a+b)^n$. Other than those terms, all the terms in the expansion exist in both $(a+b)^n$ and $(a+b)^{n+1}$. Looking at Equation **??**, we can see that the kth term of the expansion of $(a+b)^{n+1}$ (apart from the first and last ones) is equal to

$$(c_{k-1}^n + c_k^n) \tag{3.45}$$

For example, the coefficient for $a^{n-1}b^2$ is $c_2^n + c_3^n$ (and as that is the third term in the series, that's what would be expected.)

In general we can write that

$$C_k^{n+1} = C_{k-1}^n + C_k^n \tag{3.46}$$

which doesn't seem very helpful in computing the coefficients.

At this stage, it's probably easiest to proceed with "proof by divine inspiration" [10].

[10] This is code for "the answer is hard to work out, but easy to check". How this can be proved is explored in the "Discrete Mathematics" section.

$$C_k^n = \frac{n!}{k!(n-k)!} \tag{3.47}$$

Overall,

$$(a+b)^n = \sum_{k=0}^{n} C_k^n a^{n-k} b^k \tag{3.48}$$

Note that C_k^n is also often written as $\binom{n}{k}$ (pronounced "n" choose "k").

Example: Find the value of $(x-\sqrt{3})^4 + (x+\sqrt{3})^4$.

These expressions are pretty symmetric (i.e. lots of stuff will cancel when they're expanded). Expanding the first one gives

$$\begin{aligned}(x-\sqrt{3})^4 &= x^4 + 4x^3(-\sqrt{3})^1 + 6x^2(-\sqrt{3})^2 + 4x(-\sqrt{3})^3 + (-\sqrt{3})^4 \\ &= x^4 - 4x^3\sqrt{3} + 6 \cdot 3x^2 - 4x\sqrt{3}^3 + 3^2 \\ &= x^4 - 4x^3\sqrt{3} + 18x^2 - 4x\sqrt{3}^3 + 9\end{aligned}$$

Note that because the other expression has no negative numbers, everything that was negative in the expansion of $(x-\sqrt{3})^4$ will instead be positive, so

$$(x+\sqrt{3})^4 = x^4 + 4x^3\sqrt{3} + 18x^2 + 4x\sqrt{3}^3 + 9$$

and thus that overall,

$$(x-\sqrt{3})^4 + (x+\sqrt{3})^4 = 2x^4 + 36x^2 + 18$$

Example: Find the sum of all the coefficients in the expression $(1+x)^n$.

Expanding this,

$$(1+x)^n = 1 + \binom{n}{1}x + \binom{n}{2}x^2 + ... + \binom{n}{n}x^n$$

Then, set $x = 1$ in this expression [11].

[11] This "trick" is something that also shows up in other places. e.g. partial fractions (see the partial fractions section of the "algebra" topic for details).

$$2^n = 1 + \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{n}$$
$$2^n - 1 = \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{n} \tag{3.49}$$

3.6 The general binomial series

In the case where $|x| < 1$, the expansion of $(1+x)^r$ is given by

$$(1+x)^r = 1 + rx + \frac{r(r-1)}{2!}x^2 + \frac{r(r-1)(r-2)}{3!} + ... \tag{3.50}$$

This can be proved using Maclaurin series (in the "differential calculus" section).

There are some binomial expansions which look like they can't be expanded using this formula, but can. For example

$$\frac{1}{\sqrt{4-x}} \tag{3.51}$$

can be expanded, but only after some rearranging. First, the expression can be rewritten using indices to give

$$(4-x)^{-\frac{1}{2}} \tag{3.52}$$

This is almost, but not quite, in the form $(1+x)$ - the four needs to be taken out first to give

$$\left(4\left(1-\frac{x}{4}\right)\right)^{-\frac{1}{2}} \tag{3.53}$$

and then

$$4^{-\frac{1}{2}}\left(1+\left(-\frac{x}{4}\right)\right)^{-\frac{1}{2}} \tag{3.54}$$

Which can then be expanded.

Chapter 4

Algebra

4.1 Introduction

This is not a chapter about modern (or abstract) algebra, but rather secondary school algebra.

4.2 Fractions

Fractions can be surprisingly confusing. For example, what is the value of the expression directly below (assuming $x \neq 0$, as we can't divide by 0)?

$$\frac{1}{\left(\frac{1}{x}\right)} \tag{4.1}$$

Here's a reasonably good way to find the answer - multiply everything by 1.

$$\frac{1}{\left(\frac{1}{x}\right)} = \frac{1}{\left(\frac{1}{x}\right)} \times \frac{x}{x} \tag{4.2}$$

$$= \frac{x}{\left(\frac{x}{x}\right)} \tag{4.3}$$

$$= x \tag{4.4}$$

We can then apply this principle to more complex fractions.

4.3 Quadratics

The quadratic formula [12] states that for a quadratic $ax^2 + bx + c = 0$,

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{4.5}$$

[12] This is the result of completing the square for $ax^2 + bx + c$ and using it solve the equation $ax^2 + bx + c = 0$ for x

The $b^2 - 4ac$ gives quite a bit of information away. It's called the "discriminant", and sometimes written as Δ.

If $b^2 - 4ac > 0$ then we have exactly two solutions to the quadratic.

If $b^2 - 4ac = 0$ (which also means that $\sqrt{b^2 - 4ac} = 0$), then there will only be one solution (because adding zero does nothing [13] - see the equation directly below for details).

[13] This is often handy for algebraic manipulation.

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-b}{2a} \text{ if } b^2 - 4ac = 0$$

If $b^2 - 4ac < 0$ then there are no real solutions to the quadratic.

4.4 The factor and remainder theorems

The remainder theorem states that for any number a, the remainder when we divide a polynomial (which can be written as $f(x)$) by $(x - a)$ is equal to $f(a)$. [14]

[14] We can prove this without too much trouble. $\frac{f(x)}{(x-a)} = Q(x) + \frac{R(x)}{x-a}$ Where $Q(x)$ is the quotient and $R(x)$ is the remainder. This is the same as $f(x) = Q(x)(x - a) + R(x)$ We can then set $x = a$. $f(a) = R(a)$ Because $Q(x)(a - a) = 0$.

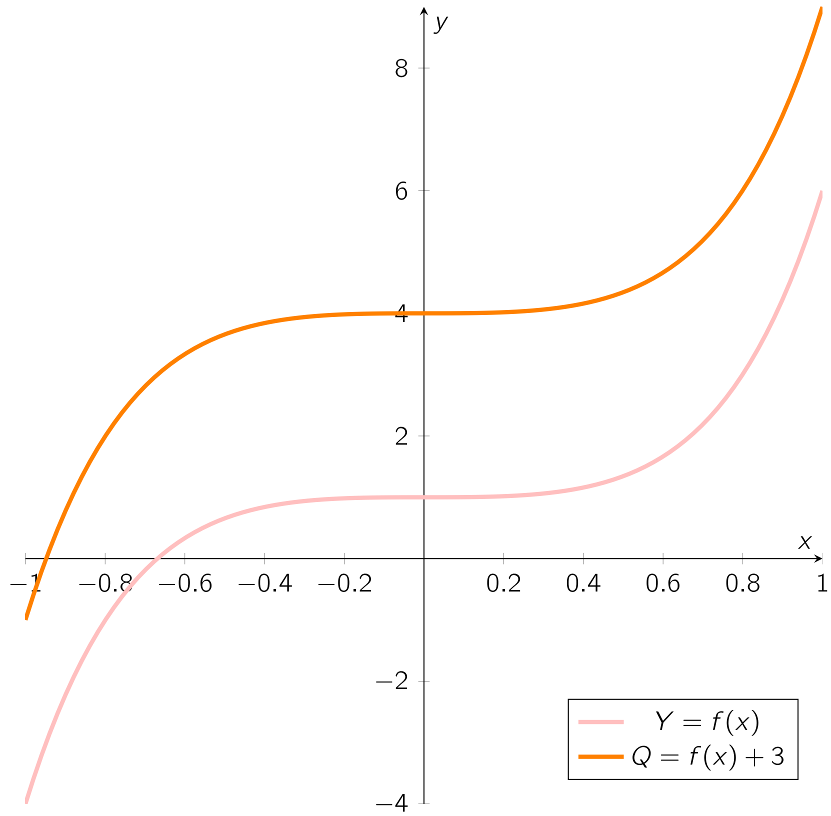
A special case of the remainder theorem is known as the "factor theorem" and it relates the roots of a polynomial to its factors. If $(x - a)$ divides $f(x)$ with no remainder, then by the remainder theorem, this means that $f(a) = 0$ and so a is a root as well as a factor. Roots are factors, and factors are roots.

4.5 Transformations of functions

4.5.1 Y-axis transformations

These are the easier case (at least in my view) to think about. When transforming a function $f(x)$ in the y-axis, there are two key transformations to be aware of - stretching and translating.

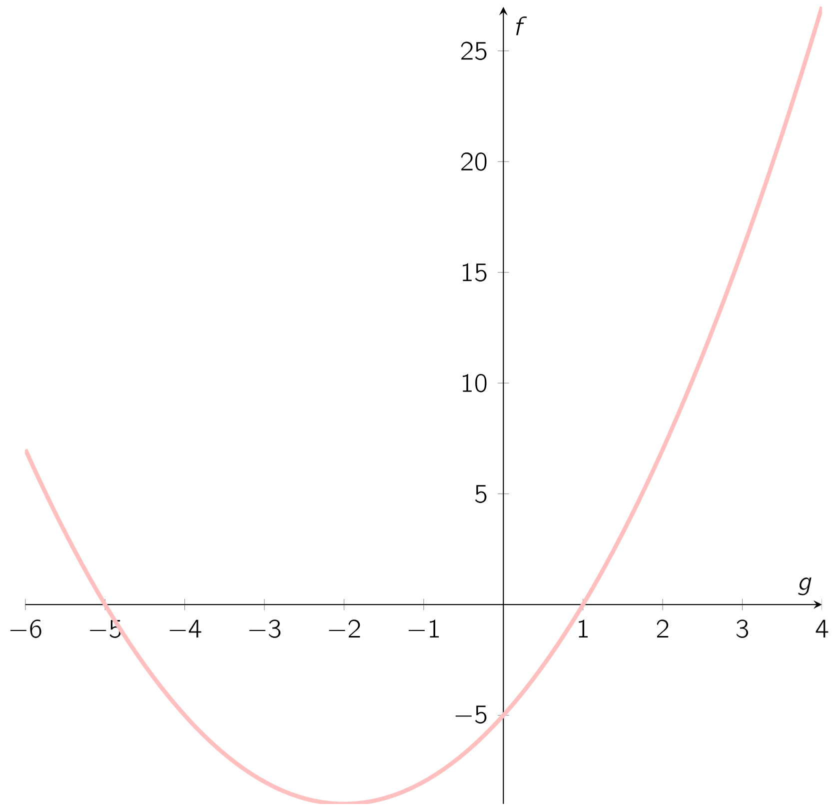
To translate a function in the y-axis we can just add something to it, e.g. to shift the graph of $y = f(x)$ three units up, define a variable, e.g. $Q = f(x) + 3$ - the 2D graph of this function will then be shifted three units above. This is illustrated on the graph below:



4.5.2 X-axis transformations

To transform a function $f(x)$ in the X-axis, we just evaluate $f(g(x))$, where $g(x)$ is a function which maps values from the x-y plane (i.e. the usual set of axis we plot things on) to one in the $g(x)$-y plane (i.e. like the usual set of axis we plot things on, *except* that wherever we had $x = a$ (where a stands for any number) we now want $g(x) = a$).

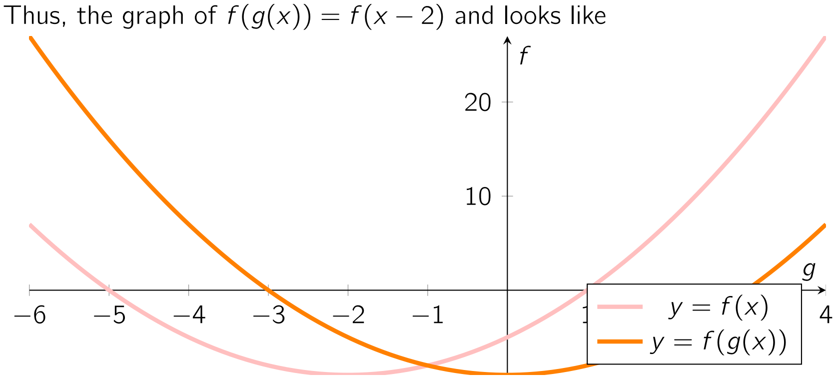
This deserves a bit of explanation. Let's imagine that $g(x) = x - 2$. If we plot $f(g(x))$ against $g(x)$, we might get something like this (for this specific $f(x)$)



We don't want a graph of $f(g(x))$ against $g(x)$, though! We want one of $f(g(x))$ against x. To do this, we need to work out how to write $g(x)$ in terms of x, and then work out where every point on the $g(x)$-axis should be on the x-axis.

As $x - 2 = g(x)$ if we add two to each side, we obtain that $x = g(x) + 2$. This means that if we shift every point on the $g(x)$-axis two to the right then we would have the X-axis!

Thus, the graph of $f(g(x)) = f(x - 2)$ and looks like



We can transform the X-axis in many ways, another one is stretching the graph. For example, if we set $g(x) = \frac{1}{2}x$, then to work out where every point on the $g(x)$-axis should be on the X-axis, we first rearrange $g(x)$, obtaining that

$$x = 2g(x) \tag{4.6}$$

and thus we *stretch* (not, as commonly misconceived, squish) the graph. I try to visualise it as the graph stretching as the infinite number of points on the axis are doubled (moved twice as far away as they once were).

4.6 Partial fractions

When we add fractions (here $a, b, c, d \in \mathbb{R}$[15]), something like this happens

$$\frac{a}{b} + \frac{c}{d} = \frac{ad}{bd} + \frac{cb}{db} \tag{4.7}$$
$$= \frac{ad + cb}{dc} \tag{4.8}$$

[15] Even though they're in the set of real numbers, they don't have to be written down directly as numbers (like 1.2 or $2^3 2$) - they could be algebraic expressions (such as $x - 3$), so long as when the value of those algebraic expressions is computed, the results are real numbers.

Sometimes we have fractions which are the "added together" form, which we'd like to turn into the "not added together" form. For example in the case of this fraction

$$\frac{x+3}{(x+4)(x-8)}$$

We can break it apart. We can use Equation **??** to predict (if we set $b = x+4$ and $d = x-8$, and $ab+cb = x+3$) that "splitting up" (aka partial fraction) will look like

$$\frac{x+3}{(x+4)(x-8)} = \frac{a}{x+4} + \frac{b}{x-8}$$

To solve this, we take advantage of the fact that because the values of a and b are constants[16]. This means that they must be true for *every* value of x, so if we plug in specific values of x, and then find what a and b are given those values of x we know that those values of a and b will be true for all values of x.

[16] i.e. their value is always the same

$$\frac{x+3}{(x+4)(x-8)} = \frac{a}{x+4} + \frac{b}{x-8}$$

times by $(x+4)(x-8)$, then $x+3 = a(x-8) + b(x+4)$

Find find a

set $x = -4$, then $-4+3 = a(-4-8) + b(-4+4)$

$-1 = -12a$

$a = \frac{1}{12}$

Then find b

set $x = 8$, then $8+3 = a \cdot (8-8) + b(8+4)$

$12b = 11$

$b = \frac{11}{12}$

Thus overall

$$\frac{x+3}{(x+4)(x-8)} = \frac{\frac{1}{12}}{x+4} + \frac{\frac{11}{12}}{x-8}$$

$$\frac{x+3}{(x+4)(x-8)} = \frac{1}{12(x+4)} + \frac{11}{12(x-8)}$$

We can also check the answer by adding the fractions back together again:

$$\begin{aligned}\frac{1}{12(x+4)} + \frac{11}{12(x-8)} &= \frac{(x-8)+11(x+4)}{12(x+4)(x-8)} \\ &= \frac{12x-36}{12(x+4)(x-8)} \\ &= \frac{x-3}{(x+4)(x-8)}\end{aligned}$$

Which is what we started out with!

There's another case, however, which we haven't looked at yet; what if there's a repeated root [17].

In this case we have something like

$$\frac{x+3}{(x+4)^2}$$

if we try to apply the previous technique, we do something like:

$$\frac{x+3}{(x+4)^2} = \frac{A}{(x+4)^2} + \frac{B}{(x+4)^2}$$

$$x+3 = A+B$$

[17] Remember, from the factor theorem (see above if not) that roots, i.e. when a polynomial is zero, are also factors (e.g. $x+4$ is a factor of $(x+4)(x-8)$, and 4 is a root.)

This is impossible to solve (because clearly values of A and B which satisfy one value of x are not going to satisfy all values of x) - oops! The trick here is to do some clever algebra. We can rewrite $x+3$ in terms of $x+4$ (because $x+3=(x+4)-1$), and from there it becomes much easier.

$$\begin{aligned}\frac{x+3}{(x+4)^2} &= \frac{(x+4)-1}{(x+4)^2} \\ &= \frac{x+4}{(x+4)^2} + \frac{-1}{(x+4)^2} \\ &= \frac{1}{x+4} - \frac{1}{(x+4)^2}\end{aligned} \tag{4.9}$$

The final thing is what happens when some other terms are mixed in, as in

$$\frac{x+3}{(x+4)^2(x-8)}$$

We could try something like

$$\frac{x+3}{(x+4)^2(x-8)} = \frac{A}{(x+4)^2} + \frac{B}{x-8}$$

however, this fails pretty quickly if we try to add the right-hand side back together

$$\begin{aligned}\frac{A}{(x+4)^2} + \frac{B}{x-8} &= \frac{A(x-8)+B(x+4)^2}{(x+4)^2(x+3)} \\ &= \frac{Ax-8A+Bx^2+8Bx+16B}{(x+4)^2(x+3)} \\ &= \frac{Ax-8A+Bx^2+8Bx+16B}{(x+4)^2(x+3)} \\ &= \frac{Bx^2+(A+8B)x+(16B-8A)}{(x+4)^2(x+3)}\end{aligned}$$

We know that when we add the fractions together, we should get $x+3$ as the numerator. Therefore the numerator of the partial fractions added together should be the same as the numerator of the original fraction.

$$Bx^2+(A+8B)x+(16B-8A) = x+3$$

We can "compare coefficients"[18] and obtain that

$$\begin{cases} Bx^2 = 0x^2 \\ (A+8B)x = x \\ (16B-8A) = 3 \end{cases}$$

[18] Note: this is explored further in the (not yet written) "algebra techniques" section.

Which simplifies to

$$\begin{cases} B = 0 \\ (A + 8B) = 1 \\ (16B - 8A) = 3 \end{cases}$$

If we substitute $B = 0$ into the other equations, we get that

$$\begin{cases} A = 1 \\ -8A = 3 \implies A = -\frac{3}{8} \end{cases}$$

Which means that the equations are inconsistent [19], and that there are no solutions satisfying all the equations at the same time.

Instead, we need to try to express this in a form similar to Equation **??**, i.e. something like $Ax + B$.

$$\frac{x+3}{(x+4)^2(x-8)} = \frac{Ax+B}{(x+4)^2} + \frac{C}{x-8}$$

$$x + 3 = (Ax + B)(x - 8) + C(x + 4)^2$$

Find C first because it's the easiest one to find

$$11 = (8A + B)(8 - 8) + 12^2 C$$

$$C = \frac{11}{144}$$

Now find A and B

$$x + 3 = (Ax + B)(x - 8) + \frac{11}{144}(x + 4)^2$$

$$x + 3 = Ax^2 - 8Ax + Bx - 8B + \frac{11}{144}x^2 + \frac{88}{144}x + \frac{176}{144}$$

$$(\frac{11}{144} + A)x^2 + (B - 1 + \frac{88}{144})x + (\frac{176}{144} - 8B - 3) = 0$$

$$A + \frac{11}{144} = 0 \implies A = -\frac{11}{144}$$

$$B - 1 + \frac{88}{144} = 0 \implies B = 1 - \frac{88}{144} \implies B = \frac{56}{144}$$

There's a caveat to all this, unfortunately! If the degree [20] of the numerator is equal to or greater than the numerator, we first have to divide (using polynomial long division[21]) the numerator by the denominator to get an expression which isn't a fraction, plus a remainder (where the degree of the numerator *is* less than the denominator).

[19] i.e. are no solutions to the equations - more on inconsistent/consistent equations in the "linear algebra" section (though that part of the "linear algebra" section is yet to be written)

[20] i.e. the highest power of the polynomial - e.g. for $x^2 + x + 1$ the degree would be 2 and for $z^3 + 8z + 5$ it would be 3

[21] Also in the yet-to-be-written algebra section

4.7 The modulus function

The modulus function is a very happy function[22]. This is because it's always

[22] Sorry.

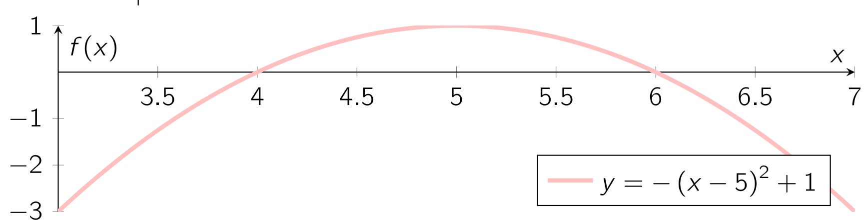
positive, and never takes on any negative values!

We can write it as $|x|$. For example, all of these expressions are true

$$
\begin{aligned}
|12| &= 12 \\
|-12| &= 12 \\
|12-6| &= |6-12| \\
|-x| &= x
\end{aligned}
$$

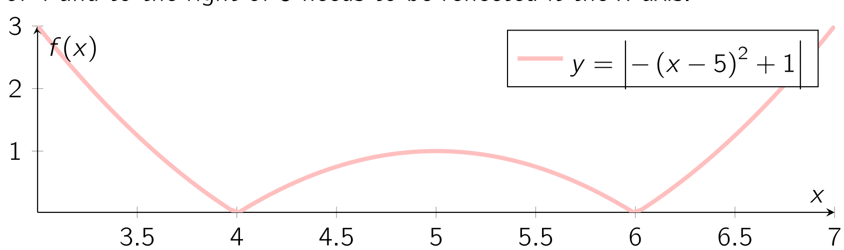
4.7.1 Graphically

When graphing a value inside a modulus function, it is often helpful to first sketch the function without the modulus. For example, when graphing $y = \left|-(x-5)^2+1\right|$, the graph of $y = 2x - 8$ would look like this:



As the modulus function must never be unhappy (i.e. take a negative value), we need to turn that frown upside down![23] Using the quadratic formula to find the roots, we get that they are at $x = 4$ and $x = 6$. Therefore, everything to the left of 4 and to the right of 5 needs to be reflected it the x-axis.

[23] Ok, I'll stop, I promise.



4.7.2 As a function

One way to define the modulus function is, well, as a function. As the function always needs to be positive, we can write that

$$
|x| = \begin{cases} x & x \geqq 0 \\ -1 \cdot x & x < 0 \end{cases} \tag{4.10}
$$

What this means is that in the case where x (the input to the modulus function) is bigger than zero, we just return x. If x is smaller than zero, we multiply it by

−1 in order to make it positive.

4.7.3 Algebraically

To write the modulus function as an algebraic expression, we're interested in functions which are always positive. The one which (probably?) springs to mind is squaring. Using this, we can define the modulus function as [24]

$$|x| = \sqrt{x^2} \tag{4.11}$$

This is really helpful for solving some equations.

[24] Note that this is just the square root of the magnitude. It is also important to **never give into the temptation to say "oh this is just the square root of a square number, it's just the original number." This is wrong**. In the expression $\sqrt{x^2}$ when we square x, both x and $-x$ are mapped to the same value, so stating that the square root of the square is just the original value is not true for negative numbers.

4.7.4 Some examples

Example: Find the complete set of values satisfying the equation $|x-2| \leqq |2x-6|$. [25]

[25] This question came from the OCR A June 2019 Single Maths Pure and Mechanics Paper

Solution: We can use the definition $|x| = \sqrt{x^2}$ to rewrite the inequality above as

$$\sqrt{(x-2)^2} \leqq \sqrt{(2x-6)^2} \tag{4.12}$$

Squaring both sides gives the following result

$$(x-2)^2 = (2x-6)^2 \tag{4.13}$$

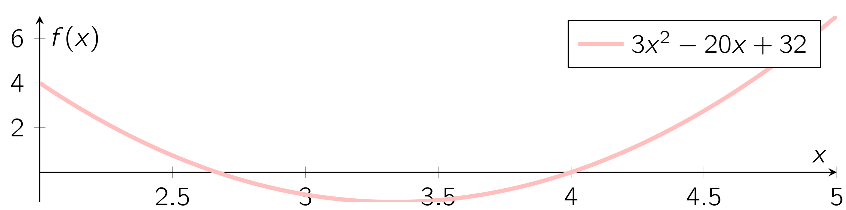
which can then be expanded to give

$$x^2 - 4x + 4 \leqq 4x^2 - 24x + 36 \tag{4.14}$$

and then, being careful not to subtract anything wrong (which I always do) the equation can be reduced to this quadratic

$$0 \leqq 3x^2 - 20x + 32 \tag{4.15}$$

We can then sketch this to work out when it would be greater than zero (or zero would be less than the curve).



From the graph, the inequality is true whenever x is less than the first time it is zero, and whenever the inequality is greater than the second time it is zero. To find when $3x^2 - 20x + 32 = 0$, we can use the quadratic formula, and thus obtain an equation for the values of x.

$$x = \frac{20 \pm \sqrt{20^2 - 4(3)(32)}}{2(3)}$$

Subtracting a positive number from a positive number (note that $\sqrt{20^2 - 4(3)(32)}$ is a positive number) gives a smaller value than adding something to a positive number. Therefore, we can deduce (with aid from a calculator) that the smaller value of x is $\frac{3}{8}$ and the larger value of x is 4, leaving us with two regions in which the inequality is satisfied.

$$x \leqq \frac{3}{8} \text{ and } x \geqq 4$$

A slightly harder example: Find the complete set of values satisfying the inequality

$$||x - 1| - 5| < 3$$

[26]

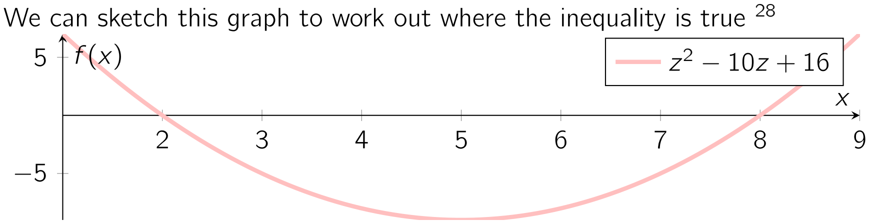
Solution: This question is a bit fiddly. The first step is to square both sides (which is fine, because the values we're taking square roots of are always positive).

$$(|x-1| - 5)^2 < 3^2 \tag{4.16}$$
$$(x-1)^2 - 10\sqrt{(x-1)^2} + 25 < 9 \tag{4.17}$$
$$(x-1)^2 - 10\sqrt{(x-1)^2} + 16 < 0 \tag{4.18}$$

This is just a quadratic [27], and we can substitute $z = \sqrt{(x-1)^2}$ to get that

$$z^2 - 10z + 16 < 0 \tag{4.19}$$
$$(z-8)(z-2) < 0 \tag{4.20}$$

We can sketch this graph to work out where the inequality is true [28]



From the graph, the inequality is true whenever $2 < z < 8$ (i.e. z is between the roots).

Thus we have $2 < \sqrt{(x-1)^2} < 8$. This means that both of these inequalities are true:

$$\sqrt{(x-1)^2} < 8 \tag{4.21}$$
$$\sqrt{(x-1)^2} > 2 \tag{4.22}$$

[26] This question came from `https://madasmaths.com`

[27] Review the section on hidden quadratics if you're not sure what this is. **TODO**: write this section

[28] The other way to do this is to think about where one of the brackets is positive, and the other negative, but in my experience it's a more error-prone method.

If we square both sides, we get that

$$(x-1)^2 < 64 \tag{4.23}$$
$$(x-1)^2 > 4 \tag{4.24}$$

Here's a time to be careful; $(x-1)^2$ is always positive. When we take the square root, however, we can have either the positive or negative square root. Note that the negative square root is just -1 times $\sqrt{\text{whatever}}$, and thus we need to flip the inequality.

Overall then we have these four inequalities.

$$x-1 < 8 \tag{4.25}$$
$$x-1 > -8 \tag{4.26}$$
$$x-1 > 2 \tag{4.27}$$
$$x-1 < -2 \tag{4.28}$$

Which can all be arranged to obtain that the following four inequalities must all be true for the inequality to hold

$$x < 9 \tag{4.29}$$
$$x > -7 \tag{4.30}$$
$$x > 3 \tag{4.31}$$
$$x < -1 \tag{4.32}$$

Examining these, it is clear that there are two regions in which this is true: $-7 < x < -1$ and $3 < x < 9$. This is the complete set of regions which satisfy the inequality. We can write it using set theory notation as

$$\{-7 < x < -1\} \cup \{3 < x < 9\} \tag{4.33}$$

where $\cup$ denotes the "union" operator (which means that overall we have the set of objects in either $\{-7 < x < -1\}$, $\{3 < x < 9\}$, or in both of them - which in this case is nothing, as the two ranges don't overlap).

Example: solve the equation

$$\left|\frac{1}{9}(8t-9)\right| = \left|\frac{1}{3}(2t-11)\right| \tag{4.34}$$

Solution: this can be solved by using the fact that $|x| = \sqrt{x^2}$, but the numbers are not nice, and I made so many arithmetic errors (it was painful). The other way is to think about the definition of the modulus function.

Instead, we can think about the definition of the modulus function. For each side of the equation we have

$$\left|\frac{1}{9}\left(8t-9\right)\right| = \begin{cases} \frac{1}{9}\left(8t-9\right) & \frac{1}{9}\left(8t-9\right) \geqq 0 \\ -\frac{1}{9}\left(8t-9\right) & \frac{1}{9}\left(8t-9\right) < 0 \end{cases} \tag{4.35}$$

$$\left|\frac{1}{3}(2t-11)\right| = \begin{cases} \frac{1}{3}(2t-11) & \frac{1}{3}(2t-11) \geqq 0 \\ -\frac{1}{3}(2t-11) & \frac{1}{3}(2t-11) < 0 \end{cases} \tag{4.36}$$

Thinking about the cases, the graphs can intersect either

- Where the argument to the modulus function is less than zero, and thus has been multiplied by negative one, for both functions.
- Where the argument of one of the functions is less than zero, and the other is greater than zero.
- Where the argument of both functions is greater than zero.

In the first case, the values of t where the two curves intersect would be when

$$(-1)\cdot\frac{1}{9}\left(8t-9\right) = (-1)\cdot\frac{1}{3}(2t-11) \tag{4.37}$$

this is just the same as

$$\frac{1}{9}\left(8t-9\right) = \frac{1}{3}(2t-11) \tag{4.38}$$

Which is the same equation as for case three.
The only other case is when

$$-\frac{1}{9}\left(8t-9\right) = \frac{1}{3}(2t-11) \tag{4.39}$$

Note that it doesn't matter which one is less than zero (and has thus been multiplied by -1), as we can just multiply both sides by -1 to get from one to the other.

Thus solving Equation **??**, we have

$$\frac{1}{9}\left(8t-9\right) = \frac{1}{3}(2t-11) \tag{4.40}$$

$$8t-9 = 3(2t-11) \tag{4.41}$$

$$8t-9 = 6t-33 \tag{4.42}$$

$$2t = -24 \tag{4.43}$$

$$t = -12 \tag{4.44}$$

And solving Equation **??** we have

$$-\frac{1}{9}(8t-9) = \frac{1}{3}(2t-11) \quad (4.45)$$
$$-(8t-9) = 3(2t-11) \quad (4.46)$$
$$-8t+9 = 6t-33 \quad (4.47)$$
$$-14t = -42 \quad (4.48)$$
$$t = 3 \quad (4.49)$$

4.8 Parametric equations

If you've ever been to a science museum, then you may have seen a kind of device where if you turn a handle connected to a cog, that cog spins a bunch of other cogs. Although all the cogs spin at different rates, they're all ultimately driven by the cog which you're spinning.

This is a bit like how parametric equations work - we create a "parameter" (the cog which you spin, and often named t) and then x and y (or whatever the axes are called) are a bunch of other cogs connected to the initial cog.

For example, we can write the equation of the unit circle in terms of x and y.

$$x^2 + y^2 = 1 \quad (4.50)$$

But we could also write it as two separate equations - one for x in terms of a new variable we'll introduce, t, and one for y in terms of t.

$$x = \cos(t) \quad (4.51)$$
$$y = \sin(t) \quad (4.52)$$

We can get from the parametric equations (the ones in terms of t) to the Cartesian equations using a little algebra. Adding together x^2 and y^2 gives this equation.

$$x^2 + y^2 = \cos^2(t) + \sin^2(t) \quad (4.53)$$

As $\cos^2(t) + \sin^2(t) = 1$, the overall result is that

$$x^2 + y^2 = 1 \quad (4.54)$$

which is the Cartesian equation of a circle with modulus one!

4.9 Roots of polynomials

This is a further maths topic.

Any quadratic (which we can write as $ax^2 + bx + c$) that has roots α and β can be equivalently written as $(x - \alpha)(x - \beta)$.

When we expand $(a - \alpha)(x - \beta)$ we get a quadratic.

$$x^2 - (\alpha + \beta)x + \alpha\beta \tag{4.55}$$

We can then compare the coefficients of Equation **??** with those of ax^2+bx+c.

$$x^2 - (\alpha + \beta)x + \alpha\beta = x^2 + \frac{b}{a}x + \frac{c}{a} \tag{4.56}$$

This gives a sete of equations relating the roots and the coefficients of a polynomial.

$$-(\alpha + \beta) = \frac{b}{a} \implies \alpha + \beta = -\frac{b}{a} \tag{4.57}$$

$$\alpha\beta = \frac{c}{a} \tag{4.58}$$

Something similar is also the case for higher-degree polynomials [29].

[29] TODO: add a proof for this

Example: A quadratic equation $x^2 - 8x + 12 = 0$ has roots α and β. Find a quadratic with roots α^2 and β^2.

Solution 1: To find the coefficients of our new quadratic, we need to find the value of $\alpha^2\beta^2$ and $\alpha^2 + \beta^2$.

$$\begin{aligned}\alpha^2\beta^2 &= (\alpha\beta)^2 \\ &= 12^2 \\ &= 144\end{aligned}$$

$$\begin{aligned}\alpha^2 + \beta^2 &= (\alpha + \beta)^2 - 2(\alpha\beta) \\ &= 8^2 - 2(12) \\ &= 40\end{aligned}$$

Therefore, a quadratic with roots α^2 and β^2 is $x^2 - 40x + 144 = 0$.

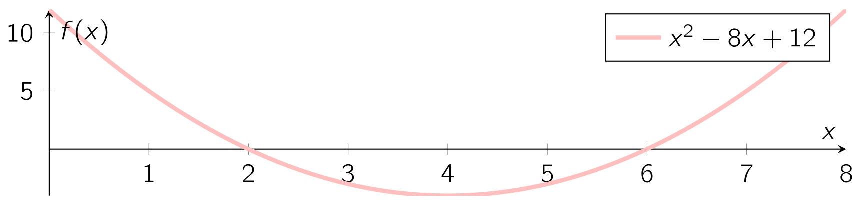
Solution 2: We can also do this using a substitution. First, note that for our original quadratic, we know that $x = \alpha$ is a root. We want a new polynomial, however, where it is not $x = \alpha$ which is a root, but rather $x = \alpha^2$ that is a root. Consider the graph of our function (below)

What we want to do is transform the position of the roots. [30] Let's look at one of the roots (the one at $x = 2$) - it should clearly end up at $2^2 = 4$ (as this is what the question is asking for). If we call our quadratic $P(x)$, and think about the point x, at the point x, we'd like to have a y-value of $P(\sqrt{x})$, rather than the current $P(x)$.[31]

Therefore, our new quadratic should be

[30] If this is leading to thoughts about transformations of graphs (an earlier topic in the algebra section) then, I mean, yes!

[31] Because that way $\sqrt{x} \mapsto x$, which also means that $x \mapsto x^2$



$$P(\sqrt{x}) = (\sqrt{x})^2 - 8\sqrt{x} + 12$$
$$= x - 8\sqrt{x} + 12$$

We're interested in when this function happens to be zero, so we want [32]

$$P(\sqrt{x}) = 0$$
$$x - 8\sqrt{x} + 12 = 0$$
$$x + 12 = 8\sqrt{x}$$
$$(x+12)^2 = 64x$$
$$x^2 + 24x + 144 = 64x$$
$$x^2 - 40x + 144 = 0$$

This is the same as the other method which relied upon manipulating the roots directly! In general: use whichever method is nicer.

[32] To manipulate this equation we use a trick where something is of the form $x_{\text{nasty}} + y_{\text{nice}} + z_{\text{nice}} + \ldots = 0$, so we move all the "nice" stuff over to one side and then apply a function to both sides in order to eliminate the "nasty" stuff.

4.9.1 A harder example

Example [33]: the cubic C, with roots α, β, and γ is given by

$$8x^3 + 12x^2 + 2x - 3 = 0 \tag{4.59}$$

The integer function S_n is defined as

$$S_n = (2\alpha + 1)^n + (2\beta + 1)^n + (2\gamma + 1)^n \tag{4.60}$$

Find the values of S_3 and S_{-2}.

Solution: The easier (in my opinion), way to solve this is by using a substitution. In the case of S_3 we have our old root $x = x$ and we want to transform it (on the same axis) to the position $(2x+1)^3$, which defines a seperate axis, X. Therefore, to write the x-axis in terms of the X-axis, we rearrange

$$X = (2x+1)^3 \tag{4.61}$$

[33] From `https://madasmaths.com/archive/maths_booklets/further_topics/various/roots_of_polynomial_equations.pdf`

to be in terms of X, after which we can then substitute X for x in the polynomial.

$$(2x+1)^3 = X \tag{4.62}$$
$$2x+1 = X^{\frac{1}{3}} \tag{4.63}$$
$$2x = X^{\frac{1}{3}} - 1 \tag{4.64}$$
$$x = \frac{X^{\frac{1}{3}} - 1}{2} \tag{4.65}$$

Using this we can then substitute into the original polynomial

$$8x^3 + 12x^2 + 2x - 3 = 0 \tag{4.66}$$
$$\implies 8\left(\frac{X^{\frac{1}{3}}-1}{2}\right)^3 + 12\left(\frac{X^{\frac{1}{3}}-1}{2}\right)^2 + 2\left(\frac{X^{\frac{1}{3}}-1}{2}\right) - 3 = 0 \tag{4.67}$$

This can then be simplified, a lot.

$$8\left(\frac{X^{\frac{1}{3}}-1}{2}\right)^3 + 12\left(\frac{X^{\frac{1}{3}}-1}{2}\right)^2 + 2\left(\frac{X^{\frac{1}{3}}-1}{2}\right) - 3 = 0 \tag{4.68}$$
$$\implies \left(X^{\frac{1}{3}}-1\right)^3 + 3\left(X^{\frac{1}{3}}-1\right)^2 + \left(X^{\frac{1}{3}}-1\right) - 3 = 0 \tag{4.69}$$
$$\implies \left(X - 3X^{\frac{2}{3}} + 3X^{\frac{1}{3}} - 1\right) + 3\left(X^{\frac{2}{3}} - 2X^{\frac{1}{3}} + 1\right) + \left(X^{\frac{1}{3}} - 1\right) - 3 = 0 \tag{4.70}$$
$$\implies \left(X - 3X^{\frac{2}{3}} + 3X^{\frac{1}{3}} - 1\right) + \left(3X^{\frac{2}{3}} - 6X^{\frac{1}{3}} + 3\right) + \left(X^{\frac{1}{3}} - 1\right) - 3 = 0 \tag{4.71}$$
$$\implies X + \left(-3X^{\frac{2}{3}} + 3X^{\frac{2}{3}}\right) + \left(+3X^{\frac{1}{3}} - 6X^{\frac{1}{3}} + X^{\frac{1}{3}}\right) + (-1 + 3 - 1 - 3) = 0 \tag{4.72}$$
$$\implies X - 2X^{\frac{1}{3}} - 2 = 0 \tag{4.73}$$

Here, we pull the familiar trick[34] and rearrange

$$X - 2 = 2X^{\frac{1}{3}} \tag{4.74}$$

From here, we cube both sides and march onwards

[34] Where familiar means "did it once in the previous example".

$$
\begin{aligned}
(X-2)^3 &= \left(2X^{\frac{1}{3}}\right)^3 && (4.75)\\
\implies X^3 + 3\left(X^2\right)(-2) + (3)(X)\left((-2)^2\right) + (-2)^3 &= 8X && (4.76)\\
\implies X^3 - 6X^2 + 12X - 8 &= 8X && (4.77)\\
\implies X^3 - 6X^2 + 4X - 9 &= 0 && (4.78)
\end{aligned}
$$

Because we know this polynomial has the desired roots, and the sum of the roots is equal to $-\frac{b}{a}$, the value of S_3 is

$$-\frac{-6}{1} = 6 \tag{4.79}$$

Chapter 5

Proof

We must not believe those, who today, with philosophical bearing and deliberative tone, prophesy the fall of culture and accept the ignorabimus [that we cannot know whether something is true or false]. For us there is no ignorabimus, and in my opinion none whatever in natural science. In opposition to the foolish ignorabimus our slogan shall be "Wir müssen wissen - wir werden wissen" [we must know — we will know]

— *David Hilbert*, radio address in 1930

The human mind is incapable of formulating (or mechanizing) all its mathematical intuitions, i.e., if it has succeeded in formulating some of them, this very fact yields new intuitive knowledge, e.g., the consistency of this formalism. This fact may be called the "incompletability" of mathematics. On the other hand, on the basis of what has been proved so far, it remains possible that there may exist (and even be empirically discoverable) a theorem proving machine which in fact is equivalent to mathematical intuition, but cannot be proved to be so, nor even be proved to yield only correct theorems of finitary number theory

Kurt Goedel, remarks on his incompleteness theorems

5.1 Proving identities

To prove an identity, pick one of the sides of the identity, and then apply the rules of algebra to work to the other side. Once you've picked a side, you have to stick with it; you can't use both sides (because you'd be accepting what we want to prove as being true in order to prove that it's true which would then not be a proof).

5.2 Proof by induction

There is a rather formal mathematical definition of this, which reads (perhaps not too pleasantly, though)

> If a proposition $P(n)$ is defined for all positive integral values of n in such a way that $P(1)$ is valid, and the validity of $P(k)$ implies that of $P(k+1)$ for an arbitrary positive integer k, then $P(n)$ is valid for all n.

(this definition comes from Meserve's book on algebra[**?**]). One way of thinking about induction is by considering an analogy to dominoes. If we have a row of dominoes, and we knock one domino over, and knocking any given domino over means that the next one over, then all the dominoes will fall over.

A general formula for solving any induction problem involving property $P(n)$ reads something like:

- The basis step - show that $P(a)$ is true.
- The inductive step - show that if $P(k)$ is true, then this implies that $P(k+1)$ is also true.
- The completion step - write some words stating that as $P(a)$ is true and $P(k)$ being true implies $P(k+1)$ is true, $P(n)$ must be true for all integer n with $n \geqq a$.

5.2.1 Divisibility

A not uncommon type of induction problem might read something like: prove by induction that for all $n \in \mathbb{N}$

$$2^{n+1} + 5 \cdot 9^n \tag{5.1}$$

is divisible by 7.

The first step is the basis case, which is not too hard here. If we define

$$P(n) = 2^{n+1} + 5 \cdot 9^n \tag{5.2}$$

Then $P(1) = 2^2 + 5 \cdot 9$ which means that $P(1) = 49$ (which is divisible by 7, as $7 \cdot 7 = 49$).

The next step is the inductive step. Here we assume that $P(k)$ is true, and then consider $P(k+1)$.
[35]

$$P(k+1) = 2^{k+1+1} + 5 \cdot 9^{k+1} \tag{5.3}$$

The key thing here is to apply the inductive hypothesis (aka the fact that $P(k)$ is divisible by 7). This is pretty much always the first thing to do!

We know that $P(k) = 2^{k+1} + 5 \cdot 9^k$. Currently, this doesn't help very much as we have no way to apply it to the equation above. What we can do, however, is rearrange it to express either 2^{k+1} or $5 \cdot 9^k$ in terms of $P(k)$ and the other term. In this case, we'll choose to eliminate 2^{k+1}, but we could do it the other way, and it would be fine.

$$P(k) = 2^{k+1} + 5 \cdot 9^k \tag{5.4}$$
$$2^{k+1} = P(k) - 5 \cdot 9^k \tag{5.5}$$

We can now apply this to $P(k+1)$, by taking out a factor of 2 from 2^{k+1+1} and then substituting.

$$\begin{aligned} P(k+1) &= 2^{k+1+1} + 5 \cdot 9^{k+1} && (5.6)\\ &= 2 \cdot 2^{k+1} + 5 \cdot 9^{k+1} && (5.7)\\ &= 2\left(P(k) - 5 \cdot 9^k\right) + 5 \cdot 9^{k+1} && (5.8) \end{aligned}$$

Now we can factorise the expression, and hopefully find a factor of 7 lurking somewhere!

$$\begin{aligned} P(k+1) &= 2\left(P(k) - 5 \cdot 9^k\right) + 5 \cdot 9^{k+1} && (5.9)\\ &= 2P(k) - 10 \cdot 9^k + 5 \cdot 9^{k+1} && (5.10)\\ &= 2P(k) - 10 \cdot 9^k + 5 \cdot 9 \cdot 9^k && (5.11)\\ &= 2P(k) + 35 \cdot 9^k && (5.12) \end{aligned}$$

This means that $P(k)$ being divisible by 7 implies that $P(k+1)$ is also divisible by 7.

Now onto the completion step! This is pretty much the same every time: " Thus $P(n)$ implies $P(n+1)$ and as $P(1)$ is true, $P(n)$ holds for all $n \in \mathbb{N}$ where $n \geqq 1$. "

[35] People wiser than me have suggested that a clever "trick" we can pull here is to show that the *difference* between $P(k+1)$ and $P(k)$ is divisible by 7 (which implies that $P(k)$ being divisible implies that $P(k+1)$ is also divisible). I've never found it *particularly* helpful.

5.2.2 Practice questions

Some practice questions are available at `https://1drv.ms/b/s!AqQu_3R15Myn6i-xTfV9BmFUZuaF?e=hWB6Zd`.

5.3 Proof by contradiction

The key idea behind proof by contradiction (also known in Latin as *reductio ad absurdum*) is that we assume the opposite of what we're trying to prove is true, and then show that this causes a contradiction (e.g. it implies that $1 + 1 = 3$ or some other absurdity).

5.3.1 Proof that $\sqrt{2}$ is irrational

A classic example here is the proof that $\sqrt{2}$ is irrational. [36]

[36] Or, in fact, any value of k where k is not a square number.

To prove this, we start by writing down the definition of the rational numbers, which is that for any rational number we can write it in the form

$$\frac{p}{q} \tag{5.13}$$

where both p and q are integers and the greatest common divisor of p and q is 1. To prove this we assume that $\sqrt{2}$ is rational, and thus we can write that

$$\sqrt{2} = \frac{p}{q} \tag{5.14}$$

Which implies that

$$2p^2 = q^2 \tag{5.15}$$

As if the square of any integer is even, the integer itself must be even [37] we can write $q = 2a$ (for some suitable value of a). Therefore, we can write

[37] This is because if the integer was odd - i.e. we could write it in the form $2k + 1$, then the square would also be odd because $(2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$ which is one more than an even number and thus also odd. As the square is even, the integer must therefore also be even.

$$\sqrt{2} = \frac{2a}{q} \tag{5.16}$$

Then we can square both sides and manipulate them a bit.

$$\sqrt{2} = \frac{2a}{q} \tag{5.17}$$

$$2 = \frac{4a^2}{q^2} \tag{5.18}$$

$$q^2 = 2a^2 \tag{5.19}$$

Therefore, q is also even! However, for a number to be rational, the greatest common divisor must be 1 - here the greatest common divisor is 2: we have derived a contradiction and $\sqrt{2}$ must be irrational.

5.3.2 Proof that there are infinitely many primes

Let us assume that there are *not* infinitely many primes. In that case, we can state that there is an ordered set (i.e. every element in the set is less than the previous element in the set) with a finite number of elements, $\{p_1, p_2, p_3, ..., p_n\}$ (where p_k is the *kth* prime). If we multiply all the prime numbers together and add one to them, then we end up with a new number $P := \{p_1\} \cdot \{p_2\} \cdot \{p_3\} \cdot ... \cdot \{p_n\} + 1$. This number must have a prime factor (as every number does - this is an assumption we make, which can be proved separately), which must be one of $\{p_1, p_2, p_3, ..., p_n\}$. This means that this number must divide both $\{p_1\} \cdot \{p_2\} \cdot \{p_3\} \cdot ... \cdot \{p_n\}$ and P, so it must also divide the difference (i.e. it must divide 1), however, the smallest prime is 2 which does not divide 1. Therefore, we have found a number which does not have prime factors (which is impossible!) This means we have found an absurdity, and there must be infinitely many primes.

Chapter 6

Discrete mathematics

TODO

6.1 Combinatorics

6.1.1 Listing all permutations of a string

Sometimes, we have a string (e.g. "MATHS"), of which we'd like to list out all the possible permutations. To list these, we have to do this systematically. A nice way to do this is using Heap's algorithm. To generate all the permutations of a string, we first pick one letter. When we generate all the permutations of this string, we will have different strings, each of which has that letter at a different position. For example, in the case of "MATHS", we will have all the following (where "_" means we don't know what letter will be at that position.)

- "M____"
- "_M___"
- "__M__"
- "___M_"
- "____M"

The key idea of Heap's algorithm is that in each of these cases, the valid choices for the four unknown items ("_") are any of the permutations of the four letters in question. This means that we've found a way to write the permutations of five letters (which we don't know) in terms of the permutations of four letters arranged with a known letter in-between them. From here, we can apply the same algorithm many times - for the four-letter permutation we can reduce it to a three letter permutation, and so on until we have only one letter left. In this case, there is only one possible permutation.

There are too many possibilities here to list out all the permutations, but for the first one ("M_ _ _ _"), the process would go something like this:

- "MA_ _ _"
- "M_A_ _"
- "M_ _A_"
- "M_ _ _A"

If we then pick the first one here (the process for the rest is the same) we have

- "MAT_ _"
- "MA_T_"
- "MA_ _T"

proceeding with the first one again

- "MATH_"
- "MAT_H"

and then finally

- "MATHS"

There are a *lot* more possibilities than that, but I've only listed out some of them; to obtain the full list, just continue the process.

6.1.2 Finding the number of permutations

How many ways can we arrange distinct people in a straight line? For n people in a straight line, we can put n people in the first position, $n-1$ people in the second position, $n-2$ in the third, and so on. Overall, then the number of ways of arranging distinct people in a line is equal to

$$(n)(n-1)(n-2)...1 = n! \tag{6.1}$$

TODO

6.1.3 From English to maths

One thing that can be tricky in combinatorics is working out what words in English mean in terms of combinatorial operations. Here's a handy dictionary.

One strategy I find very useful in solving combinatorics problems is to write out a description of what I'm after in English, and then translate this into combinatorial operations (e.g. permutations, combinations, etc.). [38]

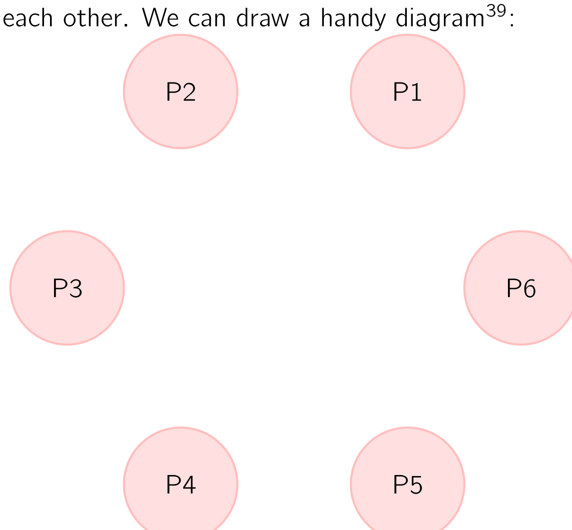
[38] This strategy works really well throughout mathematics, but it's especially helpful here.

English	Combinatorics
This can happen in way A or in way B	number of ways of A + number of ways B
To have "whatever" I need both A and then B	number of ways of A · number of ways B
For every item in this set of n objects there are k ways of obtaining it.	k^n
I have a group of n things, and I want to pick k of them, but I don't care about the order in which I get them	$\binom{n}{k}$
I have a group of n different things, and I want to pick k of them, and I do care about the order in which I get them	$\frac{n!}{(n-k)!}$

6.1.4 Arranging people in a circle

Sometimes we want to arrange people in a circle. If you ever feel the need to procrastinate over the organising of a dinner party for n people, you can consider the number of unique ways in which you can seat your guests.

The key thing to note, is that a circle is "just" a straight line, where we have joined the ends. This means that the people at either end of the line now sit next to each other. We can draw a handy diagram[39]:



Note that there is no way to have a "start" of a circle (whereas we can clearly have a front and back of a line). Because rotating people doesn't actually change their arrangements *relative to each other*, we can say that (for example, all of these arrangements would be the same)

- 123456
- 234561
- 345612

[39] This usually helps with almost any maths problem

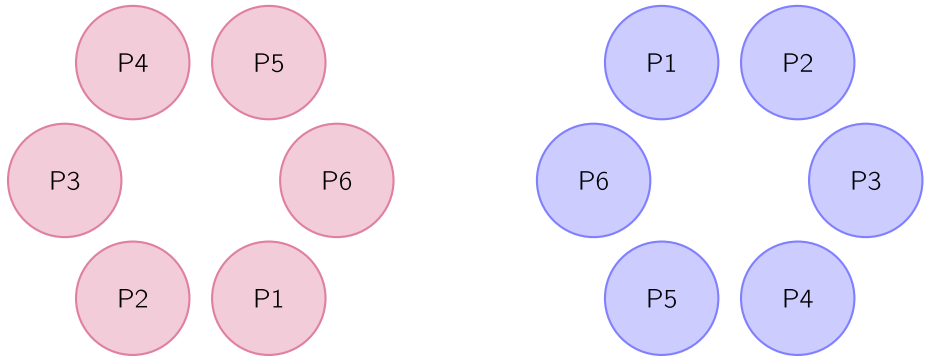


Figure 6.1: Different possible arrangements of people around a table

- 456123
- 561234
- 612345

If we draw out some of these arrangements, it's clear that the position of the people (again *relative to each other*, not the "start" because a circle doesn't have a start) hasn't changed. Person 1 is still next to Person 2 and 6, Person 2 is still next to Person 1 and 3 (and so on).

How many ways *can* we arrange people around a table, though? Well, when we were looking at the number of ways we could arrange people in a straight line, we were thinking about this *relative to* the start and end of the line. We don't have any handy features of geography [40] to define the arrangement of the people at the table in relation to.

[40] Fun fact: the Austrian foreign minister once described Italy as a "mere feature of geography"

The way we can get out of this pickle is by defining one! Let's pick a person, and then think about how we can seat everyone else at the table *relative to that person*. For n people, we can seat the first person at an arbitrary seat (which there is one way of doing). We can then work our way round the circle. If we pick a direction (left or right) to mean "next", we can then establish that there are $n-1$ options for the person immediately next to the first person. There are then $n-2$ options for the person two spaces away, $n-3$ for the person three spaces away, and so on. By the time we're back to the person on the other side of the first person, we have only one choice. Overall, this means that:

$$\text{There are } (n-1)! \text{ ways to seat } n \text{ people in a circle.} \tag{6.2}$$

6.2 Graph theory

TODO

6.3 Algorithmic complexity

TODO

Chapter 7

Sets and numbers

7.1 Sets

Sometimes it is helpful to talk about collections of things. Some things have common attributes, which means that any reasoning we apply to one object with this attribute can be applied to any other object with these attributes. Sets are about *abstraction* - we can focus less on the individual particularities of different objects, and instead focus more on their commonalities.

7.1.1 Russell's paradox

Note: not on any A Level specifications

7.2 Numbers

The "natural numbers" are what would happen if someone were to spontaneously break out counting: $0, 1, 2, 3, 4, 5, ...$ (and so on). These are called the "natural numbers" and in mathematics notation denoted as $\mathbb{N}$. To say a number is a natural number, we can say that it is in the set of natural numbers, and can write this as $x \in \mathbb{N}$ (if we let x be our number).

From the natural numbers, a logical next step is the set of "integers". These are just the numbers $..., -3, -2, -1, 0, 1, 2, 3, ...$ (going infinitely far in either direction).

After this, there are the "rational numbers". A "rational" number is a number which can be written as $n = p/q$. So 1 or 55 or 0.5 or 0.75 or $\frac{55}{155}$ or 0.5553234242352353522534234342 are all rational numbers. This set can be denoted using the symbol $\mathbb{Q}$.

Some numbers aren't rational, however. $\sqrt{2}$ or π are "irrational" numbers [41]. The "real" numbers, $\mathbb{R}$, include the "irrational" numbers.

Note that $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

[41] There's a proof of this in the proof section.

Chapter 8

Trigonometry

8.1 Trigonometric functions

You've probably come across the following formulae:[42] [43]

$$\cos(\theta) = \frac{\text{adjacent}}{\text{hypotenuse}} \tag{8.1}$$

$$\sin(\theta) = \frac{\text{opposite}}{\text{hypotenuse}} \tag{8.2}$$

$$\tan(\theta) = \frac{\text{opposite}}{\text{adjacent}} \tag{8.3}$$

The way we work out the actual values of $\cos(\theta)$, $\sin(\theta)$ and $\tan(\theta)$ is by making things as easy as possible for ourselves; we draw a triangle inside a circle with radius one. From here, we know that

$$\sin(\theta) = \frac{y}{1}$$

$$\cos(\theta) = \frac{x}{1}$$

Note that because this is the unit circle, we have

$$x^2 + y^2 = 1$$

And if we substitute $\cos(\theta)$ and $\sin(\theta)$ we get that

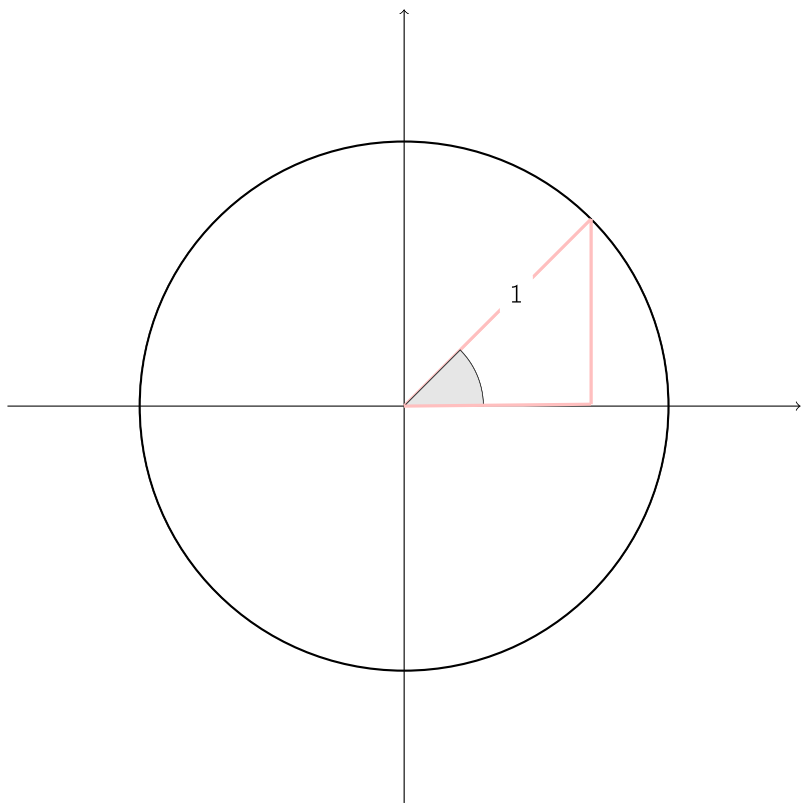
$$\cos^2(\theta) + \sin^2(\theta) = 1$$

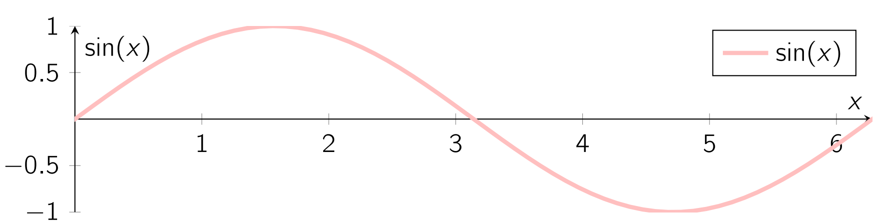
Below you can find high-precision, to-scale plots of the graphs [44] of both $\sin(x)$ and $\cos(x)$ as well as a diagram of the unit circle.

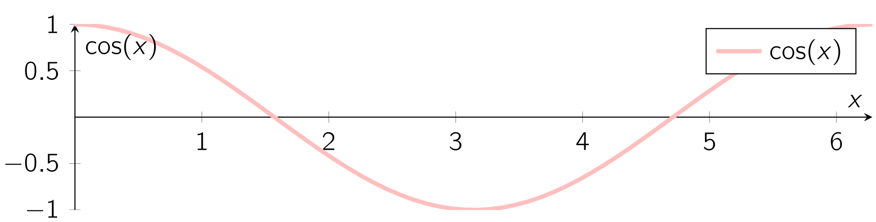
[42] Often remembered using the sort-of mnemonic "SOH-CAH-TOA" (i.e. "COS=OPPOSITE/ADJACENT, COS=ADJACENT/HYPOTENUSE, TAN=OPPOSITE/ADJACENT).

[43] The Greek letter θ is often used for angles in the same way as the variable x is used to denote unknowns.

[44] **Protip**: learn how to draw the graphs without having to thtink about it!


1


1
0.5
−0.5
−1
sin(x)
sin(x)
x
1
2
3
4
5
6


1
0.5
−0.5
−1
cos(x)
cos(x)
x
1
2
3
4
5
6

8.2 Trigonometric identities

This section goes through a bunch of trigonometric identities. The first one is the "Pythagorean identity" which is that

$$sin^2(\theta) + cos^2(\theta) \equiv 1$$

It looks a bit like Pythagoras' theorem! [45]

The next identities are the "addition formulae" which state that

$$\cos(\alpha \mp \beta) \equiv \cos(\alpha)\cos(\beta) \mp \sin(\alpha)\sin(\beta) \quad (8.4)$$

$$\sin(\alpha \pm \beta) \equiv \cos(\alpha)\sin(\beta) \pm \sin(\alpha)\cos(\beta) \quad (8.5)$$

These are in the formula booklet. [46]

A special case of these are the "double angle formulae" which are what we get if we set $\alpha, \beta = x$ in Equations **??** and **??**. [47]

$$cos(2x) \equiv \cos^2(x) - \sin^2(x) \quad (8.6)$$

This is derived from Equation **??** by replacing α and β with x, and then simplifying a bit:

$$\cos(x + x) \equiv \cos(x)\cos(x) - \sin(x)\sin(x) \quad (8.7)$$

$$\cos(2x) \equiv \cos^2(x) - \sin^(x) \quad (8.8)$$

$$sin(2x) \equiv 2\sin(x)\cos(x) \quad (8.9)$$

This is derived for **??** in a similar way to how the double-angle formula for cos is derived: replace α and β with x, and simplify.

$$\sin(x + x) \equiv \cos(x)\sin(x) + \sin(x)\cos(x) \quad (8.10)$$

$$\sin(2x) \equiv 2\cos(x)\sin(x) \quad (8.11)$$

What about $tan(\theta)$? Don't memorise identities for $tan(\theta)$ because it's equal to $\frac{sin(\theta)}{cos(\theta)}$. Just use the identities for $sin(\theta)$ and $cos(\theta)$!

8.3 Spangles

Spangles are "special" angles. They're special because they show up a lot. Their values are given in this table [48].

[45] Why this is true was explored in the previous section.

[46] The "easy" way to prove this is using complex numbers. I was going to point out that they can be proven by using triangles, however, as previously mentioned geometry is not my thing. For a geometric proof see `https://www.youtube.com/watch?v=2SlvKnlVx7U`. In the "complex numbers" section of this document there's a proof of this identity which uses the properties of complex numbers.

[47] These are useful for the integration of $\cos^2(x)$ and $\sin^2(x)$.

[48] There are numerous problems with the formatting of this table which I will one day get around to fixing.

	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\sin(x)$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\cos(x)$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
$\tan(x)$	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	undefined

Don't memorise the table! All you need to remember is that

$$\sin\left(\frac{\pi}{6}\right) = \frac{1}{2} \tag{8.12}$$

From there, you can work out the rest of the values for $\sin(x)$, as the number being rooted just goes up by one (from $\frac{\sqrt{1}}{2}$ to $\frac{\sqrt{2}}{2}$ to $\frac{\sqrt{3}}{2}$). The values of $\cos(x)$ do the same thing, but the other way round. For $\tan(x)$, as

$$\tan(x) = \frac{\sin(x)}{\cos(x)} \tag{8.13}$$

the values of $\tan(x)$ can be computed from the values of $\sin(x)$ $\cos(x)$.

8.4 General solutions to trigonometric equations

When is this equation true?

$$\sin(x) = \frac{\sqrt{3}}{2} \quad 0 < x < 2\pi$$

Using the spangles (in the previous section), we know that it's true when $x = \frac{\pi}{3}$. In the interval in question, however, this isn't the only point where it's true! If you look at the graph of $\sin(x)$, it's clear that there's another solution to this in the interval $\left(\frac{\pi}{2}, \pi\right)$. Because of sin's symmetry, we know that this solution will be at $\pi - \frac{\pi}{3}$.

We can generalise this to any point (and for different trig functions). For any point a which is in the range of $\sin(x)$, we know

$$\sin(\theta) = a \implies \theta = \arcsin(a) \tag{8.14}$$

We can call the value returned by any inverse trig function the "principle value". It is usually the angle closest to 0. However, this value is not necessarily the only possible value. For one, we know that $\sin(x)$ and $\cos(x)$ repeat every 2π (or 360°) so for any value of θ which solves $\sin(\theta) = a$ or $\cos(\theta) = a$, then that value plus or minus any multiple of 2π (or 360°) will also solve the equation.

The other thing to bear in mind is that $\sin(\theta)$ has an axis of symmetry in the lines $x = \frac{\pi}{2}$ and $x = \frac{3\pi}{2}$. This means that if θ is the principal value solving $\sin(\theta) = a$, then $\sin(\pi - \theta) = a$ is also a solution.

For $\cos(\theta)$, something very similar is the case, except that the axis of symmetry is in the line $x = \pi$, and thus if θ is the principal value solving $\cos(\theta) = a$, then $\cos(2\pi - \theta) = a$ is also a solution.

Overall, we can write that for $\sin(\theta) = a$

$$\theta = \begin{cases} \arcsin(a) \pm 2\pi \cdot k \\ \pi - \arcsin(a) \pm 2\pi \cdot k \end{cases} \quad k \in \mathbb{N} \tag{8.15}$$

And that for $\cos(\theta)$

$$\theta = \begin{cases} \arccos(a) \pm 2\pi \cdot k \\ 2\pi - \arccos(a) \pm 2\pi \cdot k \end{cases} \quad k \in \mathbb{N} \tag{8.16}$$

This is also equivalent to

$$\theta = \pm \arccos(a) \pm 2\pi \cdot k \quad k \in \mathbb{N} \tag{8.17}$$

8.5 Sums of trig functions as a single trig function

Example: Express $3\cos(\theta) + 4\sin(\theta)$ in the form $R\sin(\theta + \alpha)$.
Solution: start by applying the angle addition formula for $\sin(\theta)$ (Equation **??**).

$$\begin{aligned} R\sin(\theta+\alpha) &= R\cos(\alpha)\sin(\theta) + R\sin(\alpha)\cos(\theta) \\ &= \quad 1\sin(\theta) + \quad 3\cos(\theta) \end{aligned}$$

From here, comparing coefficients gives

$$\begin{cases} R\cos(\alpha) = 1 \\ R\sin(\alpha) = 3 \end{cases}$$

This means

$$\begin{aligned} &R^2\cos^2(\alpha) + R^2\sin^2(\alpha) = 1 + 3^2 \\ &R^2(\cos^2(\alpha) + sin^2(\alpha) = 10 \\ &R^2 = 10 \\ &R = \sqrt{10} \end{aligned}$$

as well as that

$$\begin{aligned} &\frac{R\sin(\alpha)}{R\cos(\alpha)} = 3 \\ &\tan(\alpha) = 3 \\ &\alpha = \arctan(3) \end{aligned}$$

So the solution is

$$sin(\theta) + 3cos(\theta) = \sqrt{10}\sin\left(x + \arctan(3)\right)$$

Note that this technique is very useful for solving equations of the form $A\cos(\theta) + B\sin(\theta) = c$, as we just rewrite the left hand side as a single trigonometric function, and then use the method for solving such trig functions[49].

[49] Explored in the section above.

Chapter 9

Exponentials and logarithms

9.1 Exponentials

Hopefully you're vaguely aware that a^b means "a multiplied by itself b times" (for $b \in \mathbb{N}$)[50]. From this definition, there are a bunch of useful facts we can derive.

[50] If you're not sure about this notation, review the section on "Sets and Numbers"

$$a^{b+c} = a^b a^c \tag{9.1}$$

A somewhat non-rigorous argument for this being true is as follows: $a^b = a * a * a * ... * a$ (a times itself b times). When we multiply a^b by a^c, which is equal to $a^c = a * a * a * ... * a$ (a times itself c times), we are then multiplying a times itself b times by a times itself c times. Overall, therefore we are multiplying a by itself $b + c$ times.

$$\left(a^b\right)^c = a^{bc} \tag{9.2}$$

To see that this is true, first note that we start by multiplying a by itself b times $(a * a * a * ... * a)$. We then raise this to the power of c, so $(a * a * a * ... * a)^c$, which means we have $(a * a * a * ... * a) * (a * a * a * ... * a) * ... * (a * a * a * ... * a)$. In total, there are $c * b$ lots of a (every bracket is b lots of a, and there are c of the brackets, so overall there are $c * b$ lots of a).

9.2 Logarithms

These seem scary at first, but they're not actually too bad.

A logarithm has a "base", and a "power". When $\log_a(b)$ is written, it means "what needs to be raised to the power of a to get b?" For example, $log_2(8) = 3$, as $2^3 = 8$.

The definition of a logarithm is that $z = \log_b(w)$ if and only if $w = b^z$. From here, we can prove a bunch of facts about the logarithm function.

For example, if we let $z = \log_b(w)$ and $p = \log_b(q)$ then we can then express $\log(wq)$ in terms of z and p.

$$\log(wq) = \log(b^z b^p)$$

We can then use one of the law of powers, that $b^x b^y = b^{x+y}$ [51] to write that

[51] This is explored above.

$$\log(b^z b^p) = \log(b^{z+p})$$

After this, we can use the definition of the log function to simplify the right-hand side of the previous equation.

$$\log(b^{z+p}) = z + p$$

And from our earlier definitions of $z = \log_b(w)$ and $p = \log_b(q)$ we can say that [52]

[52] This is particularly powerful because it means that we can write any multiplication as a sum (and there's a *lot* more algebra that can be applied to sums than products).

$$\begin{aligned} \log(wq) &= z + p \\ &= \log_b(w) + \log_b(q) \end{aligned} \tag{9.3}$$

9.3 Euler's number

9.3.1 Definition of e

Note: the A Level doesn't actually require any knowledge of how e is defined.

Euler's number is defined in a number of different ways. One way which is quite nice, is to think about compound interest. When you deposit money with a bank, it lends that money to other people, with interest (they borrow money from the bank and then pay back the money, plus a percentage fee). The bank then pays back some of this money to you (or they used to).

We can write a mathematical formula to represent the amount of money that we have after a certain amount of time. Every year, the amount of money in the bank account in question increases by $1 + p$ (where r is the annual rate of interest, e.g. $5\% = 0.05$ or $0.05\% = 0.00005$). Therefore, after t years the amount of money we have, assuming that we started with I units would be

$$A = I(1 + r)^t \tag{9.4}$$

Most banks, however, don't apply interest once a year. Instead, they apply it monthly. If we introduce a new variable, n, then we can write the amount of money we have after t years as

$$A = I(1 + \frac{r}{n})^{nt} \tag{9.5}$$

We can now consider an absurd scenario that only a mathematician can pretend is likely to have any relevance to real life[53] and think about what happens when we apply our interest rate an infinite number of times over one year.

mehow, the results of thought experiment do a remarkable number of vorld consequences

We can use a limit to represent this:

$$\lim_{n\to\infty} I\left(1+\frac{r}{n}\right)^{n} \tag{9.6}$$

If we try to simplify things a bit, and set all the constants (I and r) equal to 1, we can then write that

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n} \tag{9.7}$$

There's nothing special about the letter e, it's just what this limit is called in maths (in the same way that there's nothing special about "gravity" - it's just a word that is commonly understood to mean that all objects are attracted to each other because they have mass).

Chapter 10

Vectors (the geometric interpretation)

"This is some kind of a plot, right?" Slothrop sucking saliva from velvet pile. "Everything is some kind of a plot, man," Bodine laughing. "And yes but, the arrows are pointing all different ways"

— Thomas Pynchon, gravity's rainbow*. Usually somewhere near p. 600 (depending on edition).*

I make no secret of the fact that I think geometry should be wrapped up, and placed in a compost heap. From there it might helpfully decompose and from the limited nutrients of its remains give birth to something better.

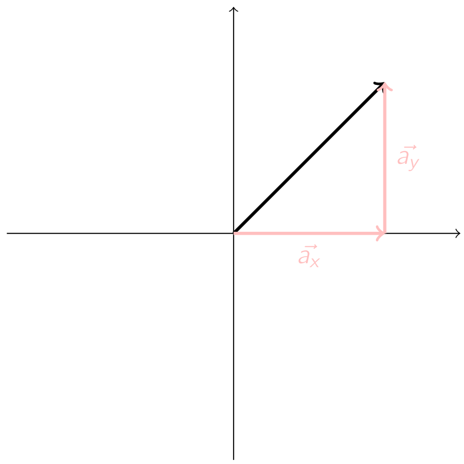
Unfortunately, today is not the day. In any 2D co-ordinate system, we may reference any point in that 2D space through two values. Usually, these are referred to as x and y.

10.1 Distances

Consider the vector $\vec{a}$. Let's say that $a = (1, 1)$. How far is a from the origin?

[54]

[54] The distance between a vector and the origin is the same thing as the magnitude of the vector.

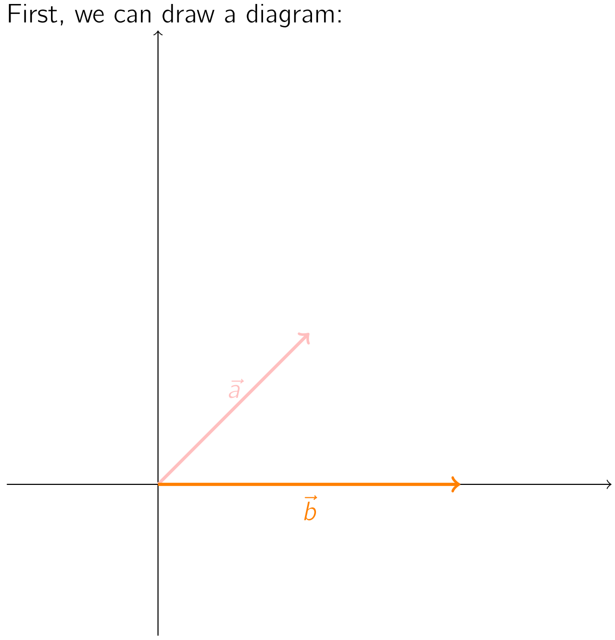


To find the distance between the origin (the vector $(0, 0)$) and $\vec{a}$ we can use Pythagoras' theorem.

$$\|a\| = \sqrt{(a_x)^2 + (a_y)^2} \tag{10.1}$$

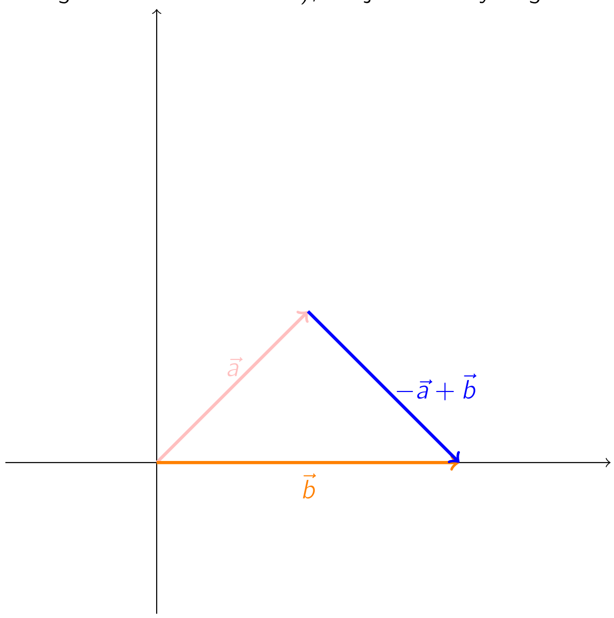
How about if we want to find the distance between the vector $\vec{a} = (1, 1)$ and another vector $\vec{b} = (2, 0)$? What we can do is find the *vector* between the two points (this is written as $\vec{ab}$), and then use Pythagoras' theorem in the same way we did above?

First, we can draw a diagram:



We have no clue how to find the vector between $\vec{a}$ and $\vec{b}$ ($\vec{ab}$). To find the vector between $\vec{a}$ and $\vec{b}$ can draw a diagram and think about what we do know.

Remember that we can read $\vec{a}$ as "move from the origin to $(1, 1)$" and $\vec{b}$ as "move from the origin to $(2, 0)$". Then, to move from $\vec{a}$ to $\vec{b}$ we want to move "from $\vec{a}$ to the origin" and "from the origin to $\vec{b}$". This is $-\vec{a} + \vec{b}$. To find the distance (aka magnitude of this vector), we just use Pythagoras's theorem.



Therefore the distance between $\vec{a}$ and $\vec{b}$ is

$$\begin{aligned} \left\| -\vec{a} + \vec{b} \right\| &= \left\| \begin{bmatrix} -1 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix} \right\| \\ &= \sqrt{(-1+2)^2 + (-1)^2} \\ &= \sqrt{2} \end{aligned}$$

10.2 Circles

All the points which are a common distance from a single point collectively form a circle. How far is every point on a circle from the centre? The radius! If we have two points q and p in space, we can draw a right-angled triangle between them, and obtain that the distance between these two points is given by the formula

$$\|-c + q\| = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2} \quad (10.2)$$

which is true because of Pythagoras' theorem [55].

A circle can be thought of as all the points which are the same distance (i.e. the radius) from a common point (i.e. the centre).

[55] This was explained further in the section above.

Thus in general we can write the equation of a circle which has a centre at (c_x, c_y) and radius r as

$$\sqrt{(x - c_x)^2 + (y - c_y)^2} = r \tag{10.3}$$

and square both sides to obtain that

$$(x - c_x)^2 + (y - c_y)^2 = r^2 \tag{10.4}$$

Example: For which values of b does the minimum point of $y = x^2 - 2x + b$ lie inside the circle $(x - 2)^2 + (y + 3)^2 = 20$?

Solution: The first step here is to find the minimum point of $y = x^2 - 2x + b$, which can be done by completing the square.

$$x^2 - 2x + b = (x - 1)^2 + b - 1 \tag{10.5}$$

Looking at the above equation, we know that the minimum point is when y is at its smallest possible value. This is when $(x-1)^2$ is at its smallest possible value (which is zero, as $x^2 \geqq 0$, and thus zero is the smallest value it can be), which is when $x = 1$ and thus $(x - 1)^2 = 0$. Any other value of x would give a bigger value. When $x = 1$, we have $y = b - 1$ and thus we have found the minimum point $(1, b - 1)$.

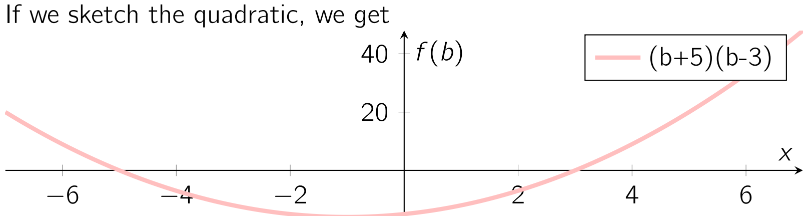
If you recall that the equation of a circle essentially specifies a *distance* from the centre, then our circle is just saying that any point *on* the circle is exactly $\sqrt{20}$ units from the centre. We're not interested in points *on* the circle, though, we're interested in any points *within* the circle. These are any points where the distance between $(1, b - 1)$ and the centre of the circle $(2, -3)$ is less than $\sqrt{20}$. Thus we can write down an inequality.

$$\sqrt{(1 - 2)^2 + ((b - 1) - (-3))^2} < \sqrt{20} \tag{10.6}$$

The left-hand side is just Pythagoras' theorem (sketch the two points in space, and then draw a right angled triangle between them). With this inequality, we can then start to manipulate, starting by squaring both sides. Usually it's dangerous to square both sides of an inequality, as negative numbers become positive, and thus break the whole thing. However, in this equation all the values must be positive (as adding two squares is just adding two positive numbers, which will always be positive, and $\sqrt{20} > 0$) so it's fine. Thus,

$$\begin{aligned}
&(1 - 2)^2 + ((b - 1) - (-3))^2 < 20 \\
&1 + (b + 2)^2 < 20 \\
&b^2 + 4b + 4 + 1 < 20 \\
&b^2 + 4b - 15 < 0
\end{aligned}$$

If we sketch the quadratic, we get



After using the quadratic formula to find the roots, we can tell that the minimum point of the curve is within the circle for all values of b which are in the range $-6.359 < b < 2.359$.

10.3 Vector lines

This is a further maths topic

10.4 Planes

This is a Further Maths topic.

TODO

In the meantime, `http://www.mit.edu/~hlb/1802/pdf/MIT18_02SC_notes_6.pdf` was pretty helpful.

Chapter 11

Differential calculus

She [Leni Poekler] even tried, from what little calculus she'd picked up, to explain it to Franz as Δt approaching zero, eternally approaching, the slices of time growing thinner and thinner, a succession of rooms each with walls more silver, transparent, as the pure light of the zero comes nearer.
— *Thomas Pynchon*, Gravity's Rainbow. *Usually somewhere around p. 160 (depending on edition)*

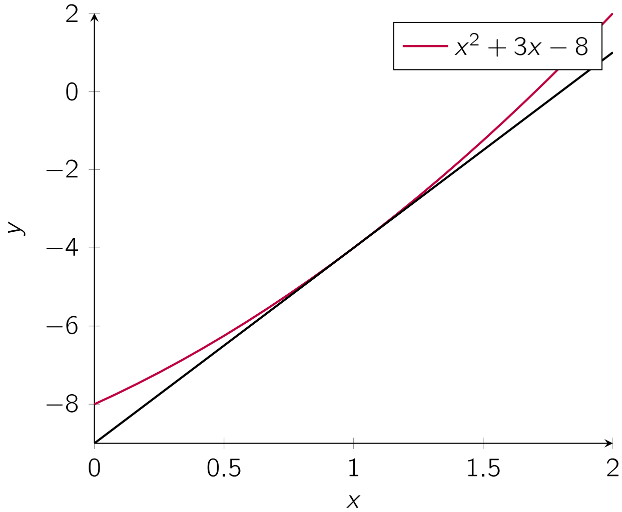
11.1 Definition of the derivative

For a single-variable function, e.g. $f(x)$, the derivative tells us the "gradient" at any given point on a function. It can also be thought of as a ratio which specifies how much the value of $f(x)$ changes when we change x.

Another way to think about the gradient is to think about the slope that the tangent [56] to the curve at a given point would have. This is the same as the gradient at that point.

[56] A tangent to a curve is a straight line which touches the curve at only a single point.

For example, in the curve below, the straight line (in grey) is a tangent to the curve at the point $x = 1$.



The derivative gives us a formula for the gradient of a tangent to the curve at any point on the curve [57]. It initially looks difficult to compute this. A good approach is to attempt to approximate this.

First, let's think about what we're trying to find. The gradient of a straight line which goes through the points (x_1, y_1) and (x_2, y_2) is given by the formula below. [58]

$$m = \frac{\Delta y}{\Delta x} \tag{11.1}$$

$$= \frac{y_2 - y_1}{x_2 - x_1} \tag{11.2}$$

Because we don't know how to find the gradient at *one* point on the curve, a good way to approximate this is to find the gradient of the curve using points which are *close* to each other.

Let's pick some points and work out the corresponding gradients.

x_1	$f(x_1)$	x_2	$f(x_2)$	h	gradient
1	-4	1.1	-3.49	0.1	5.1
1	-4	1.01	-3.9499	0.01	5.01
1	-4	1.001	-3.994999	0.001	5.001
1	-4	1.0001	-3.9995	0.0001	5.0001
1	-4	1.00001	-3.99995	1E-05	5.00001

As we pick values of x_2 which are closer and closer to the point $x = 1$ at which we are trying to find the gradient, it is clear that the gradient gets closer and closer to 5. The lines we draw (if you sketch them) [59] also become closer and closer to the tangent line.

This is how we compute derivatives, but before we can do that, we need to introduce some new mathematics, called the "limit." The idea of a limit is that it

[57] Technically, the derivative of some functions are not defined for the whole function, but that's not relevant here.

[58] This is explained further up in the document. TODO: actually explain this

[59] Or, if you look at these diagrams.

gives us the value of a function, as the independent variable approaches a given value.

For example,

$$\lim_{x \to 3} 3x = 9 \tag{11.3}$$

Which means that as we get close to $x = 3$ (but not necessarily, *at* $x = 3$) the value of the function tends to 9. Limits are most useful for functions where we don't know the value at a given point, but we do know the values around that point. For example, suppose we had a function $f(x) = 3x$, *except* at $x = 3$ where it was undefined. In that case $f(3)$ is undefined, but $\lim_{x \to 3} f(x) = 3$.

The derivative of a function $f(x)$, is defined as a limit.

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{11.4}$$

Let's work out the derivative of $f(x) = x^2 + 3x - 8$ at any point on the curve.

$$\begin{align}
\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} &= \lim_{h \to 0} \frac{((x+h)^2 + 3(x+h) - 8) - (x^2 + 3x - 8)}{h} \tag{11.5}\\
&= \lim_{h \to 0} \frac{x^2 + 2hx + h^2 + 3x + 3h - 8 - x^2 - 3x + 8}{h} \tag{11.6}\\
&= \lim_{h \to 0} \frac{2hx + h^2 + 3h}{h} \tag{11.7}\\
&= \lim_{h \to 0} 2x + h + 3 \tag{11.8}\\
&= 2x + 3 \tag{11.9}
\end{align}$$

This gives us a formula for the gradient anywhere on this polynomial!

We can write the derivative in a number of ways. The best[60] is

[60] Fight me.

$$\frac{dy}{dx}$$

Where "dy" means a really small change in y (or an infinitesimal of y), and "dx" means (you guessed it) a really small change in x (or an infinitesimal of x).

Note that the variables don't have to be y and x! They could be any function and its dependent variable, for example if we had a function $q(o) = o^2 + 24o - 100$ we could write its derivative in either of the forms below. Note that the second form is a handy way for denoting the derivative of an expression.

$$\frac{dq(o)}{do} = \frac{d}{do}[o^2 + 24o - 100] \tag{11.10}$$

There are also a bunch of other ways which are usually worse, but are still used (usually for brevity, because they're much shorter than writing out $\frac{dy}{dx}$ every time).

$$\dot{y} \text{ or } y' \text{ or } y'(x)$$

Again, the function could be called something other than $y(x)$. For example in the case of $q(o)$ we'd write

$$\dot{(q)} \text{ or } q' \text{ or } q'(x)$$

11.2 Derivatives of sums

Let's suppose we have a function $f(x) = q(x) + r(x)$, then the derivative of $f(x)$ is [61]

[61] Note that this relies on the property that the limit of two things added together is the same as the sum of the limits of the two things

$$\lim_{x\to a}(z(x)+q(x)) = \lim_{x\to a} z(x)+\lim_{x\to a} q(x)$$

Where $z(x)$ and $q(x)$ are any functions of x whose limit is defined as $x \to a$.

$$\frac{df}{dx} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \tag{11.11}$$

$$= \lim_{h\to 0} \frac{q(x+h) + r(x+h) - q(x) - r(x)}{h} \tag{11.12}$$

$$= \lim_{h\to 0} \frac{q(x+h) - q(x) + r(x+h) - r(x)}{h} \tag{11.13}$$

$$= \lim_{h\to 0} \frac{q(x+h) - q(x)}{h} + \lim_{h\to 0} \frac{r(x+h) - r(x)}{h} \tag{11.14}$$

$$= \frac{dq}{dx} + \frac{dr}{dx} \tag{11.15}$$

That is to say that

$$\frac{d}{dx}(a(x) + b(x)) = \frac{d}{dx}(a) + \frac{d}{dx}(b) \tag{11.16}$$

11.3 Differentiating polynomials

The first thing we'll look at is how to differentiate $f(x) = x^n$.

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{(x+h)^n - x^n}{h} \tag{11.17}$$

$$= \lim_{h\to 0} \frac{x^n + \binom{n}{1}x^{n-1}h + \binom{n}{2}x^{n-2}h^2 + ... + h^n - x^n}{h} \tag{11.18}$$

$$= \lim_{h\to 0} \frac{\binom{n}{1}x^{n-1}h + \binom{n}{2}x^{n-2}h^2 + ... + h^n}{h} \tag{11.19}$$

$$= \lim_{h\to 0} \binom{n}{1}x^{n-1} + \binom{n}{2}x^{n-2}h + ... + h^{n-1} \tag{11.20}$$

$$= nx^{n-1} \tag{11.21}$$

We can then combine this with the rule for the derivatives of sums from above to find the derivatives of any polynomial.

For example, we can find the derivative of x^2+3x-8 (which was the example used above).

$$\frac{d}{dx}(x^2+3x-8) = \frac{d}{dx}[x^2] + \frac{d}{dx}[3x] + \frac{d}{dx}[-8] \tag{11.22}$$
$$= 2x+3 \tag{11.23}$$

Why is $\frac{d}{dx}(-8) = 0$?

Let's suppose we have a function $f(x) = c$, then the derivative of $f(x)$ is just

$$\lim_{h\to 0} \frac{-8-(-8)}{h} = \lim_{h\to 0} \frac{0}{h} \tag{11.24}$$
$$= 0 \tag{11.25}$$

11.4 The product rule

Proving this is a little tricky, and needs some ingenuity [62]. The product rule gives us a way to find the derivative of a function which is the product of two functions $f(x) = a(x)b(x)$.

[62] Sometimes nicknamed "proof by divine inspiration".

The trick here is to "add zero":

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \frac{a(x+h)b(x+h)-a(x)b(x)}{h} \tag{11.26}$$
$$= \lim_{h\to 0} \frac{a(x+h)b(x+h)-a(x+h)b(x)+a(x+h)b(x)-a(x)b(x)}{h} \tag{11.27}$$
$$= \lim_{h\to 0} \frac{a(x+h)(b(x+h)-b(x))+b(x)(a(x+h)-a(x)}{h} \tag{11.28}$$
$$= \lim_{h\to 0} \frac{a(x+h)(b(x+h)-b(x))}{h} + \lim_{h\to 0} \frac{b(x)(a(x+h)-a(x)}{h} \tag{11.29}$$
$$= \lim_{h\to 0} a(x+h)\frac{(b(x+h)-b(x))}{h} + \lim_{h\to 0} b(x)\frac{(a(x+h)-a(x)}{h} \tag{11.30}$$

If (as it does) $h \to 0$ then $a(x+h) \to a(x)$, we can rewrite the limit as

$$a(x)\lim_{h\to 0} \frac{b(x+h)-b(x)}{h} + b(x)\lim_{h\to 0} \frac{a(x+h)-a(x)}{h} = a(x)\frac{d}{dx}[b(x)] + b(x)\frac{d}{dx}[a(x)] + b(x) \tag{11.31}$$

Overall, we therefore can say that the derivative of a function $f(x) = a(x)b(x)$ is

$$\frac{df}{dx} = \frac{d}{dx}[a(x)]b(x) + a(x)\frac{d}{dx}[b(x] \tag{11.32}$$

11.5 The chain rule

The chain rule is used to find the derivatives of "functions of a function". Mathematically, these are written as

$$\frac{d}{dx}[y(u(x))]$$

and it's possible to find this just by known the derivatives of $y(u)$ and $u(x)$.

The proof is a little involved, so for the moment you can find it at http://kruel.co/math/chainrule.pdf.

The result we're after is that

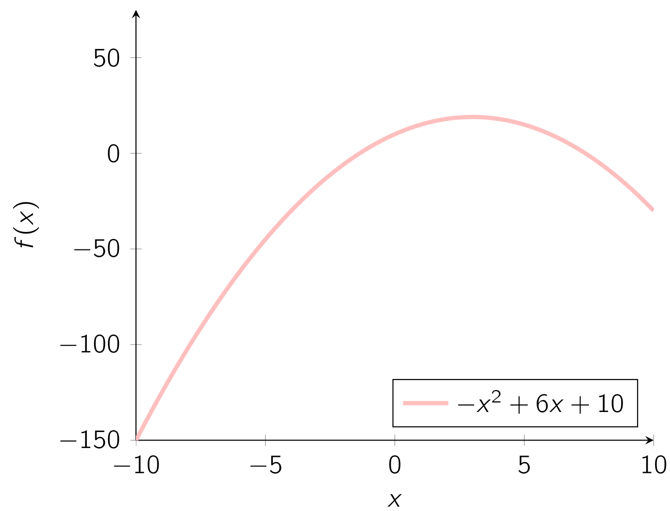
$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} \tag{11.33}$$

11.6 Maxima and minima of functions

Most people's lives have ups and downs that are relatively gradual, a sinuous curve with first derivatives at every point. They're the ones who never get struck by lightning. No real idea of cataclysm at all. But the ones who do get hit experience a singular point, a discontinuity in the curve of life—do you know what the time rate of change is at a cusp? Infinity, that's what! A-and right across the point, it's minus infinity! How's that for sudden change, eh? Infinite miles per hour changing to the same speed in reverse, all in... the Δt across the point. That's getting hit by lightning, folks.

—Thomas Pynchon, Gravity's Rainbow*. Usually somewhere around p. 660 (depending on edition). Note: some profanity has been removed.*

Consider this curve



What is the gradient at the turning point? Well in the instant where the curve's gradient is turning from having a positive gradient to a negative one, there will be a point where the gradient is zero. Before this point, the gradient is increasing. After this point, the gradient is negative [63]. It's the point in the middle where there's no change, and where the "turning point" of the curve is.

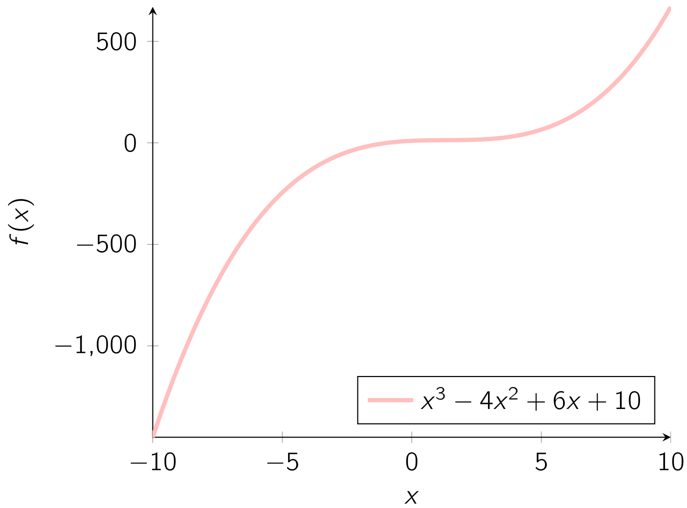
[63] Note that this is only true **for some curves** - if the curve was the other way up, then the gradient would be negative before the point where the gradient is zero, and positive after that point.

We can therefore write that for a function $f(x)$, the turning points of the curve are some of the points where (turning points are points where the sign of the derivative either side of the stationary point - i.e. when the derivative is zero - changes, so from positive to negative or negative to positive; more on how to spot the difference is given further down)

$$\frac{df(x)}{x} = 0 \tag{11.34}$$

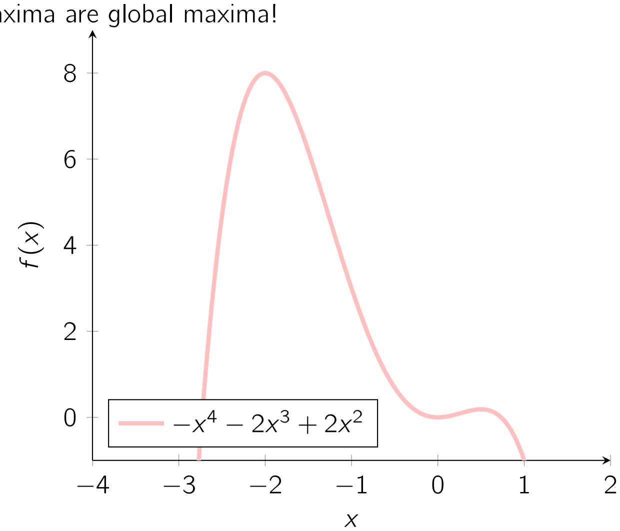
How do the turning points relate to the minima and maxima of the curve? A turning point (which we can write in co-ordinate form as $(x, f(x))$) has a value of $f(x)$ that is either *smaller* or *larger* than the points around it. This means that all turning points are either *local* minima or maxima. What this means is that they're bigger/smaller than the other points immediately around them, but not necessarily bigger/smaller than all the points on the whole curve.

Sometimes, for example, the curve doesn't have a maximum or minimum because *it keeps growing*! This curve, for example, rides off towards the sunset of infinity:



Other times, we do have a global maxima/minima [64] , however, not all local maxima are global maxima!

[64] a *global* maxima is a point on the curve such that there is no other point that is bigger than it



Another thing to note is that a local maxima requires

$$\frac{d^2f(x)}{dx^2} < 0 \text{ and } \frac{dy}{dx} = 0 \tag{11.35}$$

and a local minima requires that

$$\frac{d^2f(x)}{dx^2} > 0 \text{ and } \frac{dy}{dx} = 0 \tag{11.36}$$

If we have instead that

$$\frac{d^2f(x)}{dx^2} = 0 \text{ and } \frac{dy}{dx} = 0 \tag{11.37}$$

we don't have a turning point - we have a stationary point!

11.7 Implicit differentiation

This is "just" an application of the chain rule!

What is

$$\frac{d}{dx}y^2 \tag{11.38}$$

This depends very much on what y is. Usually, however, when y is written, what is really meant is $y(x)$. Therefore, here, we really have

$$\frac{d}{dx}\left[(y(x))^2\right] \tag{11.39}$$

If we define a new function $f(y)$ equal to y^2, we then have

$$\frac{d}{dx}\left[f(y(x))\right] \tag{11.40}$$

We can then use the chain rule [65]

$$\frac{d}{dx}\left[f(y(x))\right] = \frac{df}{dy}\cdot\frac{dy}{dx} \tag{11.41}$$

As we know that $\frac{d}{dy}\left[f(y)\right] = 2y$, we can then write the derivative of y^2 with respect to x as

$$2y\frac{dy}{dx} \tag{11.42}$$

11.7.1 Optimisation using implicit differentiation

Example: The point $P = (1, 1)$ exists in the x-y plane. The circle C is defined as

$$(x-4)^2 + (y-4)^2 = 81 \tag{11.43}$$

Which point on C is the furthest from $(1, 1)$?

Solution: The main thing here is to clearly define the problem in terms of algebraic relations (i.e. equations) which make it easy to solve the problem using differentiation.

What we're trying to maximise is the distance of some unknown point - which we'll call (x, y) - from the point $(1, 1)$. We're not after any old point, though! For our points x and x we also require that $(x-4)^2 + (y-4)^2 = 81$ (as they must lie on the circle).

We can define a function which outputs the distance between $(1, 1)$ and any two points (x, y) as

[65] Explained in the section above.

$$f(x) = \sqrt{(x-1)^2 + (y-1)^2} \tag{11.44}$$

[66]

[66] Note that even though y appears in this function, the function is still just a function of x as y is a function of x - we just write y instead of $y(x)$ as it saves space.

We then know that this function's turning points (and thus the minima and maxima) will be when

$$\frac{d}{dx}\left[f(x)\right] = 0 \tag{11.45}$$

We can find the derivative using implicit differentiation

$$\frac{d}{dx}\left[\sqrt{(x-1)^2 + (y-1)^2}\right] = \frac{1}{2}\frac{\frac{d}{dx}\left[(x-1)^2 + (y-1)^2\right]}{\sqrt{(x-1)^2 + (y-1)^2}} \tag{11.46}$$

$$= \frac{1}{2}\frac{\left[2(x-1) + 2(y-1)\frac{dy}{dx}\right]}{\sqrt{(x-1)^2 + (y-1)^2}} \tag{11.47}$$

and we want to know when this is equal to zero, which means that we can write that

$$\frac{1}{2}\frac{\left[2(x-1) + 2(y-1)\frac{dy}{dx}\right]}{\sqrt{(x-1)^2 + (y-1)^2}} = 0 \tag{11.48}$$

$$\left[2(x-1) + 2(y-1)\frac{dy}{dx}\right] = 0 \tag{11.49}$$

What we'd usually try to do here is substitute either x for y or y for x into the function which gives us the relationship between x and y (in this case, C, as defined in Equation **??**). Currently, though, this isn't possible as there's a $\frac{dy}{dx}$ throwing a spanner in the works. If we differentiate Equation **??**, however, we can then express $\frac{dy}{dx}$ in terms of x and y, and thus eliminate it from the equation.

$$\frac{d}{dx}\left[(x-4)^2 + (y-4)^2\right] = \frac{d}{dx}\left[81\right] \tag{11.50}$$

$$2(x-4) + 2(y-4)\frac{dy}{dx} = 0 \tag{11.51}$$

$$(x-4) + (y-4)\frac{dy}{dx} = 0 \tag{11.52}$$

$$\frac{dy}{dx} = \frac{-(x-4)}{(y-4)} \tag{11.53}$$

$$\frac{dy}{dx} = \frac{4-x}{y-4} \tag{11.54}$$

Returning to Equation **??**, we can now elimate $\frac{dy}{dx}$.

$$2(x-1)+2(y-1)\frac{dy}{dx}=0 \tag{11.55}$$
$$(x-1)+(y-1)\frac{dy}{dx}=0 \tag{11.56}$$
$$\frac{dy}{dx}=\frac{1-x}{y-1} \tag{11.57}$$
$$\frac{4-x}{y-4}=\frac{1-x}{y-1} \tag{11.58}$$
$$(4-x)(y-1)=(1-x)(y-4) \tag{11.59}$$
$$4y-4-xy+x=y-4-xy+4x \tag{11.60}$$
$$3y=3x \tag{11.61}$$
$$y=x \tag{11.62}$$

We can now subsitute this equation into C (aka Equation **??**), and find the values we've been after all this time.

$$(x-4)^2+(y-4)^2=81 \tag{11.63}$$
$$(x-4)^2+(x-4)^2=81 \tag{11.64}$$
$$2(x-4)^2=81 \tag{11.65}$$
$$(x-4)^2=\frac{81}{2} \tag{11.66}$$
$$x-4=\pm\sqrt{\frac{81}{2}} \tag{11.67}$$
$$x=4\pm\sqrt{\frac{81}{2}} \tag{11.68}$$

Because $x=y$ there are two cases: in the first

$$x,y=4+\sqrt{\frac{81}{2}} \tag{11.69}$$

In the second case instead

$$x,y=4-\sqrt{\frac{81}{2}} \tag{11.70}$$

Plugging the two possible values of x and y into $f(x)$ (hello again distance function - last seen in Equation **??**), we get (using a handy pocket calculator) that $x,y=4+\sqrt{\frac{81}{2}}$ is the further point from $(1,1)$, and thus the furthest point is

$$x,y=4+\sqrt{\frac{81}{2}} \tag{11.71}$$

11.8 A bunch of trigonometric limits

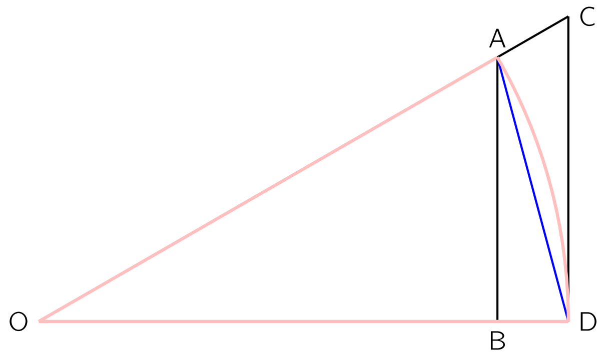
There are some useful properties about what happens to the values of trig functions when the angles we input into them tend to zero[67] These are that

[67] They're also in the formula sheet, where they are called the "small angle approximations".

$$\lim_{x \to 0} \frac{\sin(x)}{x} = 1 \text{ and } \lim_{x \to 0} \frac{1 - \cos(x)}{x} = 0$$

and it isn't immediately clear why this is true. I hope you like geometry, because if (like me) you don't, then proving this is just a bit painful (note: the proof is definitely not in the A Level, so you can just skip to the next section where the actual differentiation of trig functions - which is in the A Level - is).

First, we can draw a diagram of the unit circle, with a bunch of additional triangles



from which we can work out the areas of some of the shapes. The area of the small triangle (ADO) is just half the base times the height. The height is AB, which (by applying trigonometry to ABO) is just $\sin(\theta)$ and thus the area of ADO is $\frac{1}{2} 1 \sin(\theta) = \frac{\sin(\theta)}{2}$. The area of the section of the circle, with angle θ is $\pi 1^2 \frac{\theta}{2\pi}$, i.e. $\frac{\theta}{2}$. The area of OCD is $\frac{\tan(\theta)}{2}$.

From here, we can write down an inequality, and use a principle called various things (including the "squeeze principle" and the "sandwich principle") but meaning one thing; if for every value of x it is true that $a(x) < b(x) < c(x)$, then if as $x \to p$ both $a(x)$ and $c(x)$ both tend towards L, then $b(x)$ will also tend towards l.

Our inequality states that the area of ABO is smaller than the section of the circle, which in turn is smaller than the area of triangle OCD. Thus we have that

$$\frac{\sin(\theta)}{2} < \frac{\theta}{2} < \frac{\tan(\theta)}{2} \tag{11.72}$$

which we can then start to rearrange. Firstly, all the 2s can go

$$\sin(\theta) < \theta < \tan(\theta) \tag{11.73}$$

Then, we can divide through by $\sin(\theta)$, which gives that [68]

[68] Remember that

$$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)}$$

$$1 < \frac{\theta}{\sin(\theta)} < \frac{1}{\cos(\theta)} \tag{11.74}$$

The middle bit looks pretty close to what we're trying to prove! If we apply $f(x) = \frac{1}{x}$ to each part of the inequality, we get that

$$1 > \frac{\sin(\theta)}{\theta} > \cos(\theta) \tag{11.75}$$

69

From here, we can now take a limit as $\theta \to 0$.

$$\lim_{\theta\to 0} \cos(\theta) < \lim_{\theta\to 0} \frac{\sin(\theta)}{\theta} < \lim_{\theta\to 0} 1 \tag{11.76}$$

From the graph of $\cos(\theta)$, as $\theta \to 0$ we can see that $\cos(\theta) \to 1$, and so by the squeeze principle (mentioned above) because limits on either side of the middle are equal to 1, then

$$\lim_{\theta\to 0} \frac{\sin(\theta)}{\theta} = 1 \tag{11.77}$$

To then prove that

$$\lim_{\theta\to 0} \frac{1-\cos(\theta)}{\theta} = 0$$

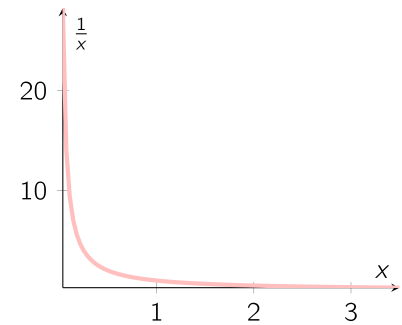
we can just rewrite in terms of Equation **??**, by multiplying by $1 + \cos(\theta)$.

$$\begin{aligned}
\lim_{\theta\to 0} \frac{1-\cos(\theta)}{\theta} &= \lim_{\theta\to 0} \frac{1-\cos(\theta)}{\theta}\frac{1+\cos(\theta)}{1+\cos(\theta)} \\
&= \lim_{\theta\to 0} \frac{1-\cos^2(\theta)}{\theta}\frac{1}{1+\cos(\theta)} \\
&= \lim_{\theta\to 0} \frac{\sin^2(\theta)}{\theta}\frac{1}{1+\cos(\theta)}
\end{aligned}$$

But we want to apply $\lim_{\theta\to 0} \frac{\sin(\theta)}{\theta} = 1$ somewhere! If we split $\sin^2(\theta)$ into $\sin(\theta)\sin(\theta)$ we then have that

$$\lim_{\theta\to 0} \frac{\sin^2(\theta)}{\theta}\frac{1}{1+\cos(\theta)} = \lim_{\theta\to 0} \sin(\theta)\frac{\sin(\theta)}{\theta}\frac{1}{1+\cos(\theta)}$$

Which is equal to 0, because as θ approaches 0, so does $\sin(\theta)$.



69 Why did we flip the inequality? If you look at the graph of $\frac{1}{x}$ above, and pick any two values of x you like (e.g. 2 and 3), and then take the reciprocal bigger values become *smaller*! This means that while we previously had $3 > 2$, we now have $\frac{1}{3} < \frac{1}{2}$. This is because $f(x) = \frac{1}{x}$ is a *decreasing* function.

11.9 Derivatives of trigonometric functions

11.9.1 Derivative of sin(x)

What is the derivative of sin(x)? First, we can use the definition of the limit and a little algebra.

$$\frac{d}{dx}[\sin(x)] = \lim_{h\to 0} \frac{\sin(x+h) - \sin(x)}{h} \tag{11.78}$$

$$= \lim_{h\to 0} \frac{\sin(x)\cos(h) + \sin(h)\cos(x) - \sin(x)}{h} \tag{11.79}$$

We want to rewrite this in terms of the two limits we found in the previous section!

$$\begin{aligned}\lim_{h\to 0} \frac{\sin(x)\cos(h) + \sin(h)\cos(x) - \sin(x)}{h} &= \lim_{h\to 0}\left[\frac{\sin(x)(\cos(h)-1)}{h} + \frac{\sin(h)}{h}\cos(x)\right] \\ &= \lim_{h\to 0}\left[-\sin(x)\frac{(1-\cos(h))}{h} + \frac{\sin(h)}{h}\cos(x)\right] \\ &= \lim_{h\to 0}[0 + 1\cos(x)] \\ &= \cos(x)\end{aligned}$$

11.9.2 Derivative of inverse trig functions

Use the chain rule!

11.9.3 Derivative of arctan(x)

From the definition of arctan(x) we know that

$$\arctan(\tan(x)) = x \tag{11.80}$$

Using the chain rule, we then obtain that

$$\begin{aligned}\frac{d}{d(\tan(x))}(\arctan(\tan(x)))\frac{d}{dx}(\tan(x)) &= 1 \\ \frac{d}{d(\tan(x))}(\arctan(\tan(x)))\left(\sec^2(x)\right) &= 1\end{aligned}$$

Therefore,

$$\frac{d}{d(\tan(x))}(\arctan(\tan(x))) = \frac{1}{\sec^2(x)} \tag{11.81}$$

If we substitute $u = \tan(x)$, we get that

$$\begin{aligned}\frac{d}{du}(\arctan(u)) &= \frac{1}{1+\tan^2(x)} \text{ by the Pythagorean identity}\\ &= \frac{1}{1+u^2}\end{aligned}$$

Or, equivalently (as renaming the variable u has no effect on the inequality),

$$\frac{d}{dx}(\arctan(x)) = \frac{1}{1+x^2} \tag{11.82}$$

11.10 Maclaurin series

Some functions can be written as an "infinite power series", which is a sum in the form

$$a + bx + cx^2 + dx^3 + \ldots$$

How would we find the values of $a, b, c, d, \ldots$ for a specific function?

Differentiation! Let's take $\sin(x)$:

$$\begin{aligned}\sin(x) &= c_1 + c_2x + c_3x^2 + c_4x^3 + \ldots\\ \cos(x) &= c_2 + 2c_3x + 3c_4x^2 + \ldots\\ -\sin(x) &= 2c_3 + 3\cdot 2c_4x + 4\cdot 3c_5x^2\ldots\\ -\cos(x) &= 3\cdot 2\cdot 1c_4 + 4\cdot 3\cdot 2c_5x\\ \sin(x) &= 4\cdot 3\cdot 2\cdot 1c_5x\\ &\ldots\end{aligned}$$

How do we find the value of a coefficient? Another algebra trick - plug in a specific value for x, in this case 0. This is because every term in the power series, except the constant one, depends on x and thus is zero when x is zero. In order to find the value of the next constant, we can just differentiate, which brings all the powers down by one.

Therefore

$$\begin{aligned}c_1 &= \sin(0)\\ c_2 &= \cos(0)\\ 2c_3 &= -\sin(0)\\ 3\cdot 2\cdot 1c_4 &= -\cos(x)\\ 4\cdot 3\cdot 2\cdot 1c_5 &= \sin(x)\\ &\ldots\end{aligned}$$

This leads directly to the general case (i.e. the value of c_n)

$$c_n = \frac{\sin^{(n)}(0)}{n!} \tag{11.83}$$

Where $sin^{(n)}(0)$ means the value of the *nth* derivative at the point $x = 0$. Therefore, overall, we can write the power series of any function [70] as

$$f(x) = \sum_{s=0}^{\infty} \left[\frac{f^{(s)}(0)}{s!} x^s \right] \tag{11.84}$$

Maclaurin series of common functions

This is just a list, which is also given in the formula booklet.

TODO

[70] There are some additional conditions - the function must be infinitely differentiable (we can keep differentiating forever) and each derivative must be defined at the point $x = 0$.

Chapter 12

Integral calculus

in the dynamic space of the living Rocket, the double integral has a different meaning. To integrate here is to operate on a rate of change so that time falls away: change is stilled .... "Meters per second" will integrate to "meters." The moving vehicle is frozen, in space, to be- come architecture, and timeless. It was never launched. It will never fall

—*Thomas Pynchon,* Gravity's Rainbow*. Usually somewhere around p. 301 (depending on edition.)*

12.1 Introduction

12.2 Integration by parts

12.3 Integration by substitution

12.4 Integral arithmetic

This technique goes by different names, but integral arithmetic captures the basic idea pretty well; sometimes it is very helpful to treat integrals as algebraic objects in order to find their value.

A very common example of this is where, by integrating $f(x)$ (or any other integrable function) with respect to x, we can arrive with an equation of the form (here we define k to stand for "an integral we know to directly find the value of")

$$\int f(x)dx = k_0 + k_1 + k_2 + ... + k_n + a \int f(x)dx \tag{12.1}$$

It is important that $a \neq 1$ (because if a is equal to one then we cannot solve for $\int f(x)dx$), in which case we can just subtract $a \int f(x)dx$ from both sides, to solve for $\int f(x)dx$.
Example: find the value of

$$\int e^{2x} \cos(x) dx$$

Solution: Start by integrating by parts

$$\int \underbrace{\cos(x)}_{v} \underbrace{e^{2x}}_{du} dx = \underbrace{\frac{e^{2x}}{2}}_{v} \underbrace{\cos(x)}_{v} - \int \underbrace{\frac{e^{2x}}{2}}_{u} \underbrace{[-\sin(x)]}_{dv} dx$$

Then integrate $\int \frac{e^{2x}}{2}[-\sin(x)]$ by parts.

$$\int \underbrace{[-\sin(x)]}_{u} \underbrace{\frac{e^{2x}}{2}}_{dv} = \underbrace{[-\sin(x)]}_{u} \underbrace{\frac{e^{2x}}{4}}_{v} - \int \underbrace{\frac{e^{2x}}{4}}_{v} \underbrace{[-\cos(x)]}_{du} dx$$

Overall then, we have

$$\int \cos(x) e^{2x} dx = \frac{e^{2x}}{2} \cos(x) - \frac{e^{2x}}{4}[-\sin(x)] - \frac{1}{2} \int \frac{e^{2x}}{2} \cos(x) dx$$

And we can add $\int \frac{e^{2x}}{2} \cos(x) dx$ to both sides, giving that

$$\frac{5}{4} \int \cos(x) e^{2x} dx = \frac{e^{2x}}{2} \cos(x) + \frac{e^{2x}}{4} \sin(x)$$

and then after multiplying both sides by $\frac{4}{5}$, we get that

$$\int \cos(x) e^{2x} dx = \frac{2e^{2x} \cos(x) + e^{2x} \sin(x)}{5}$$

Integrating by parts can get really messy - good presentation is key.

12.5 Integration of trigonometric functions

12.5.1 Integral of $\cos^2(x)$

Example: Find the value of

$$\int \cos^2(x) dx$$

Solution: We can't integrate $\cos^2(x)$ directly, so we need to rewrite it first. Of the trigonometric identities, the one which looks most useful is the double-angle formula [71]. After rearranging this (see the footnote for details), we can the write that

[71] i.e. $\cos(2x) = \cos^2(x) - \sin^2(x) \implies \cos(2x) = 2\cos^2(x) - 1 \implies \cos^2(x) = \frac{\cos(2x)+1}{2}$)

$$\begin{aligned}\int \cos^2(x)dx &= \int \frac{\cos(2x)+1}{2}dx \\ &= \frac{1}{2}\int \cos(2x)+1dx \\ &= \frac{1}{2}\left[\frac{1}{2}\sin(2x)+x\right]+c \\ &= \frac{1}{4}\sin(x)+\frac{1}{2}x\end{aligned}$$

12.5.2 Integral of $\frac{\sin(x)}{\cos(x)+\cos^3(x)}$

Example: Find the value of[72]

$$\int \frac{\sin(x)}{\cos(x)+\cos^3(x)}dx$$

Solution: We substitute $u = \sin(x)$. If this seems like a wacky thing to do, there are a bunch of good reasons:

- We know that the derivative of $\cos(x)$ is $-\sin(x)$, so we can replace $\sin(x)$ by $-\frac{du}{dx}$, which we then integrate w.r.t x, so the overall integral is integrated w.r.t u.
- As we can replace $\cos(x)$ with u and $\cos(x)$ with u^3, this simplifies the expression considerably.

Proceeding,

$$\begin{aligned}\int \frac{\sin(x)}{\cos(x)+\cos^3(x)}dx &= \int \frac{-\frac{du}{dx}}{u+u^3}dx \\ &= -\int \frac{1}{u+u^3}du \\ &= -\int \frac{1}{u(1+u^2)}du\end{aligned}$$

Which we can split into partial fractions:

$$\begin{aligned}\frac{1}{u(1+u^2)} &= \frac{A}{u}+\frac{Bu+C}{1+u^2} \\ 1 &= A(1+u^2)+(Bu+C)u \\ 0 &= A+Au^2+Bu^2+Cu-1 \\ 0 &= (A+B)u^2+Cu+(A-1)\end{aligned}$$

[72] This question came from `https://madasmaths.com`

From where we can come up with some simultaneous equations

$$\begin{cases} A + B = 0 \\ C = 0 \\ A - 1 = 0 \end{cases}$$

From this we can deduce that $A = 1$, $B = -1$ and $C = 0$. Overall, then, we have

$$\begin{aligned} -\int \frac{1}{u} - \frac{u}{1+u^2} du &= \int \frac{u}{1+u^2} - \frac{1}{u} du \\ &= \frac{1}{2}\ln(1+u^2) - \ln(u) + c \\ &= \ln\left(\frac{\sqrt{1+u^2}}{u}\right) + c \\ &= \ln\left(\sqrt{\frac{1+u^2}{u^2}}\right) + c \\ &= \frac{1}{2}\ln\left(\frac{1+u^2}{u^2}\right) + c \end{aligned}$$

Note that another way to do the last four steps is

$$\begin{aligned} \frac{1}{2}\ln(1+u^2) - \ln(u) &= \frac{1}{2}\left[\ln(1+u^2) - 2\ln(u)\right] \\ &= \frac{1}{2}\left[\ln\left(\frac{1+u^2}{u^2}\right)\right] \end{aligned}$$

which is possibly nicer.

12.5.3 Integral of $\frac{1-\cos(x)}{1+\cos(x)}$

Example: find the value of

$$\int \frac{1-\cos(x)}{1+\cos(x)} dx$$

Solution: Anything which looks like $1 + \cos(x)$ can be turned into $1 - \cos^2(x)$ which is also known[73] as $\sin^2(x)$. This is by multiplying by $1 - \cos(x)$ (because $(x+y)(x-y) = x^2 - y^2$, also known as the difference of two squares). We can proceed by multiplying by one.

[73] By the Pythagorean identity.

$$\begin{aligned}\int \frac{1-\cos(x)}{1+\cos(x)}dx &= \int \frac{1-\cos(x)}{1+\cos(x)} \cdot \frac{1-\cos(x)}{1-\cos(x)}dx \\ &= \int \frac{1-2\cos(x)+\cos^2(x)}{\sin^2(x)}dx \\ &= \int \csc^2(x) - 2\cot(x)\csc(x) + \cot^2(x)dx\end{aligned}$$

Note that while a lot of these integrals are in the formula sheet [74] some of them are not (i.e. we have no idea what the integral of $\cot^2(x)$ is at present). To get around this, we can use the Pythagorean identity to reduce this problem (the answer to which we don't know) to a problem which we do know the answer to! Take $\cos^2(x)+\sin^2(x) = 1$, divide through by $\sin^2(x)$ and obtain that $\cot^2(x)+1 = \csc^2(x) \implies \cot^2(x) = \csc^2(x) - 1$.

Then,

$$\begin{aligned}\int \csc^2(x) - 2\cot(x)\csc(x) + \cot^2(x)dx &= \int 2\csc^2(x) - 2\cot(x)\csc(x) - 1 \\ &= -2\cot^2(x) + 2\csc(x) - x + c\end{aligned}$$

[74] At least, in the OCR A formula booklet, it is given that $\frac{d}{dx}[\cot(x)] = -\csc^2(x)$ and $\frac{d}{dx}[\csc(x)] = -\csc(x)\cot(x)$

12.5.4 Integral of $\sqrt{1-\cos(x)}$

Example: find the value of

$$\int \sqrt{1-\cos(x)}dx \tag{12.2}$$

Solution: As it is, we can't integrate this, therefore the only possible option is to rewrite it somehow, into a form we can integrate.

A way to think about this which isn't particularly elegant, but is usually quite effective, is to think about all the possible identities which could work, and try each one.

e.g. thinking about identities which involve $\cos(x)$, we have [75]

- $\cos^2(x) + \sin^2(x) \equiv 1$, which doesn't help here because there's no $\cos^2(x)$ anywhere.
- $\cos(a+b) \equiv \cos(a)\cos(b) - \sin(a)\sin(b)$ - also no help
- $\cos(2x) \equiv [\cos(x)]^2 - [\sin(x)]^2$, and its two other forms (rearranging using the first bullet point), $\cos(2x) \equiv 1-2[\sin(x)]^2$ and $\cos(2x) \equiv 2[\cos(x)]^2-1$.

[75] These are discussed in the trigonometry section of this document.

The last identity looks quite useful, because we have a $\cos(2x)$, and a 1, and we can rewrite it in the form

$$1 - \cos(2x) \equiv 2\left[sin(x)\right]^2 \tag{12.3}$$

The $1 - \cos(2x)$ looks quite a lot, but not exactly, like our integral, but we can fix that by simply replacing [76] x with $\frac{x}{2}$, giving

[76] This is fine, because when we derived the double-angle formula from the addition formula we set $a = x$ and $b = x$ in the equation $\cos(a+b) \equiv \cos(a)\cos(b) - \sin(a)\sin(b)$, but we could have just as well have set $a = \frac{x}{2}$ and $b = \frac{x}{2}$, which would give $\cos(x) \equiv \left[\cos\left(\frac{x}{2}\right)\right]^2 - \left[\sin\left(\frac{x}{2}\right)\right]^2$ which we can then rearrange to give the result we need.

$$1 - \cos(x) \equiv 2\left[\sin\left(\frac{x}{2}\right)\right]^2 \tag{12.4}$$

Applying this to our integral, we have

$$\int \sqrt{1-\cos(x)}dx = \int \sqrt{2\left[\sin\left(\frac{x}{2}\right)\right]^2}dx \tag{12.5}$$

$$= \int \sqrt{2}\sin\left(\frac{x}{2}\right)dx \tag{12.6}$$

The last part we can do by inspection (as the derivative of $-\cos(x)$ is equal to $\sin(x)$, by the chain rule, the derivative of $-\cos\left(\frac{x}{2}\right)$ is equal to $-\frac{1}{2}\cos(\frac{x}{2})$, and therefore the integral of $\sin\left(\frac{x}{2}\right)$ with respect to x is just $-2\cos\left(\frac{x}{2}\right)$).

Therefore, we have

$$\int \sqrt{1-\cos(x)}dx = \sqrt{2}\int \sin\left(\frac{x}{2}\right)dx \tag{12.7}$$

$$= -2\sqrt{2}\cos\left(\frac{x}{2}\right) + c \tag{12.8}$$

$$\tag{12.9}$$

Chapter 13

Polar coordinates

TODO

Chapter 14

Differential equations

14.1 Separation of variables

14.2 Integrating Factors

e^x shows up a lot in differential equations, because it has properties that are helpful when we differentiate it. One way in which it helps us is in solving first-order linear differential equations, which are equations of the form

$$\frac{dy}{dx} + p(x)y = q(x)$$

This can be solved using the product rule. If we define a function $f(x)$, we can write by the product rule that the derivative of $ye^{f(x)}$ is

$$\frac{dy}{dx}e^f + e^f\frac{df}{dx}y \tag{14.1}$$

This doesn't immediately look like our equation, but if we multiply through by e^f, we get that

$$\frac{dy}{dx}e^{f(x)} + p(x)e^{f(x)}y = q(x)e^{f(x)} \tag{14.2}$$

What we can do here is write that the left hand side is equal to the derivative of $ye^{f(x)}$. This only works, however, if the derivative of $f(x)$ is equal to $p(x)$. [77] If it is, we can write that

$$\frac{d}{dx}\left[ye^{f(x)}\right] = q(x)e^{f(x)} \tag{14.3}$$

And thus we can solve the equation by integrating.

[77] This is because $\frac{d}{dx}\left[ye^{f(x)}\right] = \frac{d}{dx}ye^{f(x)} + y\frac{d}{dx}\left[e^{f(\ldots}\right.$
$= \frac{dy}{dx}e^{f(x)} + y\frac{d}{dx}[f(x)\ldots$
And if $f(x) = \int p(x)dx$
then $\frac{d}{dx}[f(x)] = p(x)$
And thus $\frac{d}{dx}\left[ye^{f(x)}\right] = \frac{dy}{dx}e^{f(x)} + p(x)e^{f(x)}y$ which is just the left-hand side of the equation.

14.3 Substitution

Linear differential equations [78] are easy to solve, which is nice. Some equations are less easy to solve, but in the same way that there are "hidden quadratics" [79] we can also have "hidden first-order linear differential equations" (which is much more of a mouthful). By making a substitution we can rewrite many equations which don't look linear at first sight as linear differential equations.

[78] The ones in the form where they can be solved using an integrating factor

[79] Have another look at the algebra section if you're not sure about what hidden quadratics are.

14.4 Second-Order Differential Equations

14.4.1 Introduction

The easiest second-order differential equations to solve are those which we can integrate directly, for example

$$\frac{d^2y}{dx^2} = \cos(x) \tag{14.4}$$

When we integrate this once, we get that

$$\int \frac{d^2y}{dx^2} dx = \int \cos(x) dx \tag{14.5}$$

$$\frac{dy}{dx} = \sin(x) + c \tag{14.6}$$

and then integrating again, we get that

$$\int \frac{dy}{dx} dx = \int \sin(x) + c dx \tag{14.7}$$

$$y(x) = -\cos(x) + cx + d \tag{14.8}$$

Which is the general solution to this particular second-order differential equation. With first-order differential equations we only have one constant, and we can determine the value of the constant given a single point on the curve. [80]

For a second-order differential equation, however, we have two constants, so we definitely can't solve the equation given only point. We need either two points on the curve, or one point on the original curve and one point on its first derivative.

[80] In more formal notation, for the differential equation

$$\frac{dy}{dx} = f(x)$$

with solution $y(x) = F(x)+c$ we can determine the value of c given a value of x, x_0 and the corresponding value of y, y_0 at that point.

14.4.2 Homogenous second-order ODEs

These are a bit icky, because they're not solvable in general. Fortunately a lot of them are solvable.

Homogenous (i.e. everything is a function of only one variable) second-order, linear differential equations can be solved without *too* much difficulty. We can reduce an equation in the form

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0 \tag{14.9}$$

to a quadratic by setting $y = e^{\lambda x} \implies \frac{dy}{dx} = \lambda e^{\lambda x} \implies \frac{d^2y}{dx^2} = \lambda^2 e^{\lambda x}$.
From here,

$$a\lambda^2 e^{\lambda x} + b\lambda e^{\lambda x} + ce^{\lambda x} = 0$$
$$a\lambda^2 + b\lambda + c = 0 \text{ Which is fine as } e^x > 0$$
$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Here there are a number of possibilities for the value of the discriminant [81].

- $b^2 - 4ac > 0$, which is the "straightforward" case
- $b^2 - 4ac = 0$, in which case there's only one value of λ
- $b^2 - 4ac < 0$, in which case we can use complex numbers and trigonometry.

In the case where $\Delta > 0$ [82] we have two possible solutions to the differential equation,

$$e^{\frac{-b \pm \sqrt{b^2-4ac}}{2a}x} \tag{14.10}$$

In the case where $\Delta = 0$ we have only

$$y(x) = e^{-\frac{b}{2a}x}$$

And in the case where $\Delta < 0$ we have the same case as in Equation **??**, except that there's a complex part to the root[83]

[81] If you've no idea what this is, review the earlier section on quadratics.

[82] note that Δ means "the discriminant"

[83] And, as Euler might point out, $e^{i\theta} = \cos(\theta) + i\sin(\theta)$

14.4.3 Why two linearly independent solutions?

[84]

Let's suppose we have a differential equation involving $y(x)$ and its first and second derivative. We may want to solve this subject to the initial conditions

$$y(x_0) = y_0 \text{ and } \left.\frac{dy}{dx}\right|_{x=x_0} = \dot{y}_0 \tag{14.11}$$

In order to be able to solve for *every* possible initial condition, we need a linear combination of two "linearly independent" solutions, using which we can satisfy any possible initial condition. This means that if we have two solutions $y_1(x)$ and $y_2(x)$ and neither can be written as a multiple of the other, then the "general solution" (i.e. the one with the constants in it, like $+c$ for first-order differential equations when we integrate them) is of the form

[84] If two vectors (which we can call v_1 and v_2) are linearly independent, then the only values of α_1 and α_2 which solve the equation $\alpha_1 v_1 + \alpha_2 v_2 = 0$ are $\alpha_1 = \alpha_2 = 0$. If there were a different combination, then we could write that $v_1 = -\frac{\alpha_2}{\alpha_1} v_2$ And thus they wouldn't be linearly independent, as one of the vectors is a multiple of the other. This can be generalised to n vectors.

$$y(x) = \alpha y_1(x) + \beta y_2(x) \tag{14.12}$$

The good news is that for homogenous second-order differential equations we do have two linearly independent solutions![85] For example if our roots of the auxiliary are $\alpha \pm \beta$, then we would have that the two solutions are [86]

$$y(x) = e^{(\alpha+\beta)x} \text{ and } y(x) = e^{(\alpha-\beta)x} \tag{14.13}$$

Proof that we need two linearly independent solutions, and that two linearly independent solutions are sufficient to solve any initial condition (note: no A Level exam board examines this). The reason we *need* two linearly independent solutions is that if we have two sets of linearly independent initial conditions we certainly can't write both of them the as linear combinations of a single solution. [87] For example, suppose we have two possible initial conditions, one where

$$y(x_0) = 1 \text{ and } \frac{dy}{dx}\Big|_{x=x_0} = 0 \tag{14.14}$$

and another where

$$y(x_0) = 0 \text{ and } \frac{dy}{dx}\Big|_{x=x_0} = 1 \tag{14.15}$$

These are linearly independent solutions, and we can't write them both as a linear combination of a single solution. Thus, there is at least one case where we need at least two linearly independent solutions.

The next thing to prove is that if we have two solutions ($y_1(x)$ and $y_2(x)$) which are linearly independent, then for suitable values of α and β we can satisfy any initial condition using a solution of the form

$$y(x) = \alpha y_1(x) + \beta y_2(x) \tag{14.16}$$

Does this even solve the differential equation though? Yes! We can prove this with a bunch of algebra (there's an easier way to do this by introducing some new notation, but that's for later[88])

$$y = \alpha y_1(x) + \beta y_2(x) \implies \frac{dy}{dx} = \alpha\frac{dy_1}{dx} + \beta\frac{dy_2}{dx} \tag{14.17}$$
$$\implies \frac{d^2y}{dx^2} = \alpha\frac{d^2y_1}{dx^2} + \beta\frac{d^2y_2}{dx^2} \tag{14.18}$$

and thus that [89].

And thus $y = \alpha y_1(x) + \beta y_2(x)$ is a solution to Equation **??**

If that was messy, it gets worse. [90]

[85] Except in the case where $\Delta = 0$, but there's a way to get around that (more on that later).

[86] Which are linearly independent, as one of them cannot be written as a multiple (of something linear) of the other.

[87] This is a *lot* like how we can write any vector in a 2D space in terms of two vectors, so long as those vectors are not parallel (same for 3D space, except with three vectors).

[88] Note: I have yet to write about this new notation

[89] Note that both $\alpha[a\frac{d^2y_1}{dx^2} + b\frac{dy_1}{dx} + cy_1]$ and $\beta[a\frac{d^2y_2}{dx^2} + b\frac{dy_2}{dx} + cy_2]$ are zero, because we know that $y_1(x)$ and $y_2(x)$ solve the equation $a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$

[90] I've had nightmares about drowning in a differential equation algebra soup. Literal soup made of algebra - it was a *very* strange dream.

$$a[\alpha \frac{d^2y_1}{dx^2} + \beta \frac{d^2y_2}{dx^2}] \tag{14.19}$$

$$+b[\alpha \frac{dy_1}{dx} + \beta \frac{dy_2}{dx}] \tag{14.20}$$

$$+c[\alpha y_1(x) + \beta y_2(x)] = p[a\frac{d^2y_1}{dx^2} + b\frac{dy_1}{dx} + cy_1] \tag{14.21}$$

$$+ q[a\frac{d^2y_2}{dx^2} + b\frac{dy_2}{dx} + cy_2]$$

$$= 0 \tag{14.22}$$

From Equation **??** we know that the derivative of our solution will be

$$\frac{dy}{dx} = \alpha \frac{dy_1}{dx} + \alpha \frac{dy_2}{dx} \tag{14.23}$$

as we want to show that this can satisfy any set of initial conditions, where $y(x_0) = y_0$ and $\dot{y}(x_0) = \dot{y}_0$, we can start by writing a set of simultaneous equations

$$\begin{cases} \alpha y_1(x_0) + \beta y_2(x_0) = y_0 \\ \alpha \dot{y}_1(x_0) + \beta \dot{y}_2(x_0) = \dot{y}_0 \end{cases} \tag{14.24}$$

we can write these in matrix form and obtain that

$$\begin{pmatrix} y_1(x_0) & y_2(x_0) \\ \dot{y}_1(x_0) & \dot{y}_2(x_0) \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} y_0 \\ \dot{y}_1 \end{pmatrix} \tag{14.25}$$

These equations can be solved whenever the determinant of the 2x2 matrix above is not equal to zero, i.e. whenever

$$y_1(x_0)\dot{y}_2(x_0) - y_2(x_0)\dot{y}_1(x_0) \neq 0 \tag{14.26}$$

Then the two equations have a unique solution. The easiest way to go from here is to prove this by contradiction (as we have lots of techniques for dealing with equalities (=) and not very many for dealing with inequalities involving $\neq$). We can proceed by assuming that two linearly independent solutions are *not* sufficient to determine the general solution of any second-order differential equation, and write that

$$y_1(x_0)\dot{y}_2(x_0) - y_2(x_0)\dot{y}_1(x_0) = 0 \tag{14.27}$$

We can now manipulate this a little

$$y_1(x_0)\dot{y}_2(x_0) = y_2(x_0)\dot{y}_1(x_0) \tag{14.28}$$

Dividing through leads to the formula

$$\frac{y_1(x_0)}{y_2(x_0)} = \frac{\dot{y}_1(x_0)}{\dot{y}_2(x_0)} \tag{14.29}$$

Here, though it looks like we have a contraction. Why? Let's set

$$c = \frac{y_1(x_0)}{y_2(x_0)} \tag{14.30}$$

and

$$d = \frac{\dot{y}_1(x_0)}{\dot{y}_2(x_0)} \tag{14.31}$$

Then we can write that

$$y_1(x_0) = cy_2(x_0) \tag{14.32}$$

and that

$$\dot{y}_1(x_0) = d\dot{y}_2(x_0) \tag{14.33}$$

But we specified earlier that $y_1(x)$ and $y_2(x)$ are linearly independent! And now we've found that in Equation **??** that they're *not* linearly independent, and thus we've found that assuming that two linearly independent solutions *doesn't* solve any set of initial conditions for a second-order differential equation leads to a contradiction. Hence, two linearly independent solutions *are* sufficient.

14.5 Systems of differential equations

Sometimes we have multiple variables which change at the same time. For example, we might have variables x and y, which both depend on variable t. We might then have a system of equations involving these:

TODO

Chapter 15

Complex numbers

This is a further maths topic.

15.1 Introduction

What happens when you square root a negative number? Well, according to a bunch of people in history, nothing. They claimed it was like dividing by zero, *non possibilis*[91]!

Modern mathematicians, however, think that things do happen when you square root negative numbers (and this can easily be seen after looking at real-world applications).

To denote the square root of a negative number, we declare a new symbol, i [92], which is defined as

$$i^2 = -1 \tag{15.1}$$

Using this definition, we can (for example) write

$$\begin{aligned}
&\sqrt{-1} = i \\
&\sqrt{-15} = \sqrt{15}\sqrt{-1} = \sqrt{15}i \\
&\sqrt{-49} = \sqrt{49}\sqrt{-1} = 7i \\
&15 + \sqrt{-49} = 15 + \sqrt{49}\sqrt{-1} = 7i \\
&x + y\sqrt{-1} = x + yi
\end{aligned}$$

In general, we can write a complex number as $x + yi$. Here x denotes the "real" part of the complex number (as there is no i anywhere to be seen) and the y denotes the imaginary part (as there is an i).

For a complex number, z, we can write the real part of z as $\Re(z)$ and the imaginary part of z as $\Im(z)$. Note that many (most?) people don't use the weird

[91] Not possible!

[92] Note that some people also write this using the letter j.

symbols that this document uses, but instead Re(z) and Im(z) to denote the real and imaginary components of a complex number respectively.

Using this definition, we can work out what happens when we do some basic operations.

We can add complex numbers:

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

We can also multiply them

$$\begin{aligned}\frac{a + bi}{c + di} &= \frac{a + bi}{c + di} \cdot \frac{c - di}{c - di} \\ &= \frac{ac - adi + bci - bdi^2}{c^2 - d^2 i^2} \\ &= \frac{ac + bd + (bc - ad)i}{c^2 + d^2}\end{aligned}$$

15.1.1 The complex conjugate

The *complex conjugate* of $x + yi$ is just $x - yi$ - that is, when we take the complex conjugate of a number, the real part remains unchanged, but the sign of the complex part is flipped (i.e. positive numbers become negative and vice versa).

The complex conjugate can be denoted in a number of ways. For a complex number z, we can write the conjugate as z^* (it can also be equivalently written as $\bar{z}$).

One interesting property of the complex conjugate is that for $x = a + bi$

$$\begin{aligned}zz^* &= (a + bi)(a - bi) && (15.2) \\ &= a^2 + b^2 && (15.3) \\ &\geqq 0 && (15.4)\end{aligned}$$

In **??** we used the difference of two squares (combined with the fact that $i \cdot i = -1$).

The result of this multiplication is a real number!

15.1.2 Dividing complex numbers

Another operation we can perform is division. To divide two complex numbers (which we can call w and z), we do something similar to rationalising the denominator (except for complex numbers).

For example, in the specific case where $z = 1 + i$ and $w = 3 + 4i$, we have can make the denominator "real" through multiplication by the complex conjugate.

$$\frac{z}{w} = \frac{1+i}{3+4i} \tag{15.5}$$
$$= \frac{1+i}{3+4i} \cdot \frac{1+i}{3-4i} \tag{15.6}$$
$$= \frac{(1+i)(3-4i)}{25} \tag{15.7}$$

15.1.3 The absolute value of a complex number

The absolute value of a complex number $z = x + yi$ is equal to $\sqrt{x^2 + y^2}$ (which is equivalent to zz^*). The modulus of a complex number is equal to its Pythagorean length (from the origin).

15.1.4 Properties of the complex conjugate and absolute value

- $(zw)^* = z^*w^*$
- $\left(\frac{z}{w}\right)^*$
- $(z^*)^*$
- $|z| = |z^*|$
- $|zw| = |z|\,|w|$
- $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$
- $|z + w| \leqq |z| + |w|$

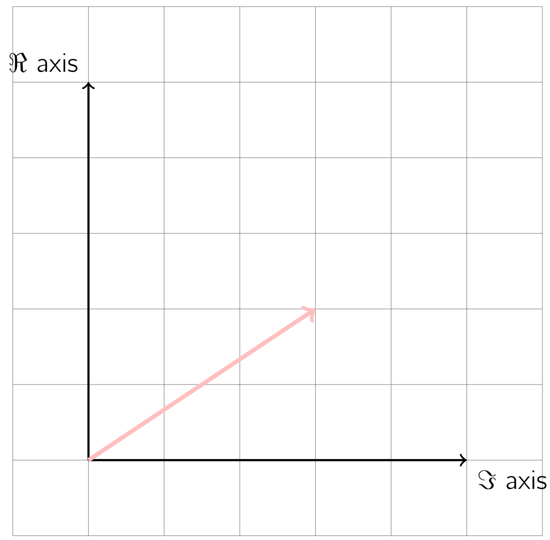
15.2 The Argand diagram

15.2.1 Plotting complex numbers

Complex numbers can be interpreted in a Cartesian fashion[93]. Instead of an x and a y co-ordinate, we can say that we have a real ($\Re$) and an imaginary (Im) co-ordinate.

For example, if we have the expression $3 + 2\sqrt{-1} = 3 + 2i$, we could say that the co-ordinates are $(3, 2)$, and plot these on a graph.

[93] Where every point in the plain is uniquely specified by numerical co-ordinates



We label the vertical axis as ℑ (or Im) - for "imaginary" - and the horizontal axis as ℜ (or Re) - for "real".

15.2.2 Loci on the Argand diagram

Consider the expression $|z - i| = 1$. One way to think about this is using vectors. Here z is a variable, standing for any vector in the complex plane. Read aloud, this equation means something along the lines of "the distance between i and z is equal to 1". Why the distance? Because z and i can be treated as vectors (and as explored earlier in the vectors section) $z - i$ is the vector from i to z (which means that the length of the vector $z - i$ is the same as the distance between i and z) so the absolute value of this is the distance between i and z.

15.2.3 Polar co-ordinates and complex numbers

We can also define complex numbers differently (i.e. not using Cartesian co-ordinate system). Instead of identifying each complex number using an x (real) and y (imaginary) co-ordinate, we can instead define it using an angle (θ) and a distance from the origin (r).

15.2.4 Graphical interpretation of operations on complex numbers

15.3 The trigonometric form of a complex number

When we say $z := x + iy$ (also known as "z is defined as $x + iy$"), x and y can be anything! For example, x and y could equal $\cos(\theta)$ and $\sin(\theta)$ respectively. By setting different values of theta, we can obtain any point on a circle with radius 1, and centre $(0, 0)$. For example, if we have $z := \cos(\theta) + i\sin(\theta)$, then to obtain the number 1, we can just set $\theta = 0$. To be able to obtain every complex number, however, we need to introduce a second variable (which we can call r). This "scales" the circle, so that for each value of r, the values of θ between 0 and 2π (2π is not inclusive) we can obtain a circle with that radius. Overall, we can write any complex number in the form

$$z := r\left(\cos(\theta) + i\sin(\theta)\right) \tag{15.8}$$

To find the trigonometric form of a given complex number there are two ways.

To convert a complex number (e.g. $1 + 3i$) into trigonometric form, the first way algebra is to use algebra (in particular "comparing coefficients"). By comparing $r\cos(\theta) + ri\sin(\theta)$ with $1 + 3i$, we obtain that $r\cos(\theta) = 1$ and that $r\sin(\theta) = 3$. Thus

$$\begin{aligned} r^2\cos^2(\theta) + r^2\sin^2(\theta) &= r^2(\cos^2(\theta) + \sin^2(\theta) \\ &= r^2 \\ &= 2 \text{ (from } 1 + 3i) \end{aligned}$$

from which we can deduce that $r = \pm\sqrt{2}$. The other thing we can do is divide through, thus obtaining that $\frac{r\sin(\theta)}{r\cos(\theta)} = \frac{1}{1}$, and thus that $\theta = \arctan(1)$ (which is $\frac{\sqrt{2}}{2}$). Overall, we can then write that

$$1 + i = \sqrt{2}e^{\frac{\sqrt{2}}{2}i} \tag{15.9}$$

which we can also do for *any* complex number.

The second way involves geometry (I still need to find where I originally wrote my notes on this, but the method is to draw the complex number - not necessary, but usually helpful - and to then find the modulus and argument of the complex number).

15.4 The exponential form of a complex number

Consider the Maclaurin series for $\cos(x)$, $i\sin(x)$ and e^{ix}. What happens when we add $\cos(x) + i\sin(x)$? Well a fair amount of algebra to start with! After that, however, we do get an interesting result, though.

This means that [94]

[94] This is sometimes referred to as "Euler's formula"

$$e^{ix} = \cos(x) + i\sin(x) \tag{15.10}$$

Which provides a link between trigonometry and complex numbers! This turns out to be very useful in proving trig identities.

$$\begin{aligned}\cos(x) + i\sin(x) &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{ix}{1!} - \frac{ix^3}{3!} + \frac{ix^5}{5!} - \frac{ix^7}{7!} \\ &= 1 + \frac{ix}{1!} - \frac{x^2}{2!} - \frac{ix^3}{3!} + \frac{x^4}{4!} + \frac{ix^5}{5!} - \frac{x^6}{6!} - \frac{ix^7}{7!} \\ &= 1 + \frac{ix}{1!} + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \frac{(ix)^4}{4!} + \frac{(ix)^5}{5!} + \frac{(ix)^6}{6!} + \frac{(ix)^7}{7!} \\ &= e^{ix}\end{aligned}$$

For any complex number, where r is the modulus and θ is the argument, we have

$$re^{\theta i} = \cos(\theta) + i\sin(\theta) \tag{15.11}$$

We can use this to write other complex numbers, such as $1 + i$ in the form $re^{i\theta}$.

15.5 Properties of the exponential form of a complex number

It is substantially easier to multiply together complex numbers in exponential form (than their Cartesian counterparts). For example, if we want to multiply $W := r_0e^{\theta_0}$ and $Z := r_1e^{\theta_1}$ then we can perform the operation using the laws of indices

$$WZ = \left(r_0e^{\theta_0}\right)\left(r_1e^{\theta_1}\right) \tag{15.12}$$

$$= r_0 \cdot r_1 \cdot e^{\theta_0+\theta_1} \tag{15.13}$$

This means that when we multiply two complex numbers, we obtain a new complex number whose argument is the sum of the arguments of the two (complex) numbers which we multiplied together, and whose modulus is just the product of the two (complex) numbers we multiplied together.

This allows us to perform some transformations. For example, if we multiply a complex number by $i = e^{\frac{\pi}{2}}$, then we are effectively just rotating that complex number by $\frac{\pi}{2}$ rad (aka 90°).

15.6 Helpful things to know about the exponential form

Complex numbers sometimes form geometric series (particularly when trigonometric functions are translated into exponential form). For example,

$$2+2e^{\frac{\pi}{10}i}+2e^{\frac{2\pi}{10}i}+2e^{\frac{3\pi}{10}i}+2e^{\frac{4\pi}{10}i} \tag{15.14}$$

forms a geometric series with $a=2$ and $r=2e^{\frac{\pi}{10}i}$. This means that the sum (as shown in the "sequences and series" section of these notes) is equal to

$$\frac{2\left(1-\left(e^{\frac{\pi}{10}}\right)^5\right)}{1-e^{\frac{\pi}{10}}} \tag{15.15}$$

15.7 Cool stuff with trigonometry

Example: Show that

$$\cos(a+b)=\cos(a)\cos(b)-\sin(a)\sin(b)$$

Solution:

We can write $\cos(a+b)$ as the real part of $e^{(a+b)i}$. This because $\cos(a+b)$ is equal to the real part of $\cos(a+b)+i\sin(a+b)$, which is equal to $e^{(a+b)i}$.

$$\begin{aligned}\cos(a+b)&=\Re(e^{(a+b)i})\\&=\Re(e^{ai}e^{bi})\\&=\Re((\cos(a)+i\sin(a))(\cos(b)+i\sin(b)))\\&=\Re(\cos(a)\cos(b)+i\cos(a)\cos(b)+i\sin(a)\cos(b)+i^2\sin(a)\sin(b))\\&=\cos(a)\cos(b)-\sin(a)\sin(b)\end{aligned}$$

Example: Express $\sin(3x)$ in terms of $\sin(x)$.
Solution: Firstly, we can write $\sin(3x)$ as the equation

$$\sin(3x)=\Im(\cos(3x)+i\sin(3x)) \tag{15.16}$$

We can then apply De Moivre's theorem[95] to rewrite the expression in terms of $\sin(x)$ and $\cos(x)$

[95] $\cos(nx)+i\sin(nx)=(\cos(x)+i\sin(x))^n$

$$\Im(\cos(3x)+i\sin(3x))=\Im((\cos(x)+i\sin(x))^3) \tag{15.17}$$

We can now expand the binomial obtained, which leads to the result that

$$\Im((\cos(x)+i\sin(x))^3)=\Im((\cos^3(x)+3\cos^2(x)i\sin(x)+3\cos(x)i^2\sin^2(x)+i^3\sin^3(x))) \tag{15.18}$$

Then, we can tidy this up a bit, leading to the expression

$$\Im((\cos^3(x)+3\cos^2(x)\sin(x)i-3\cos(x)\sin^2(x)-i\sin^3(x))) \tag{15.19}$$

We are only interested in the imaginary parts of the expansion, so it is therefore equal to just

$$3\cos^2(x)\sin(x) - \sin^3(x) \tag{15.20}$$

We want $\sin(3x)$ in terms of $\sin(x)$, however! There's a rogue gatecrasher [96] in the previous expression - the $\cos(x)$! Fortunately we can remove the $\cos^2(x)$ without too much difficulty using the Pythagorean identity.

[96] A handy way to remember whether $\cos(x)$ or $\sin(x)$ shows a certain property is (as previously mentioned, TODO: mention) that $\sin(x)$ generally behaves "better" than $\cos(x)$.

$$3(1-\sin^2(x))\sin(x) - \sin^3(x) = 3\sin(x) - 3\sin^3(x) - \sin^3(x) \tag{15.21}$$
$$= 3\sin(x) - 4\sin^3(x) \tag{15.22}$$

Thus, we can say that

$$\sin(3x) = 3\sin(x) - 4\sin^3(x) \tag{15.23}$$

15.8 The roots of unity

The nth roots of unity are the nth roots of one. How can there be more than one (1) nth root of one? Well, some of them are complex, of course!

Note that [97]

[97] In this document the natural numbers include zero!

$$e^{\pm 2n\pi i} = 1 \text{ where } n \in \mathbb{N}$$

This is because of Euler's formula.

Therefore, if we take the nth roots of both sides, we get that

$$e^{\pm \frac{2\pi}{n} i} = 1^{\frac{1}{n}} \text{ where } n \in \mathbb{N}$$

Which we can use to compute the nth roots of unity. Note that by the fundamental theorem of algebra for the nth roots of unity, there are n different values.

Example: by considering the ninth roots of unity, show that [98]

[98] I believe this question comes from the textbook "Further Pure Mathematics" [**?**]

$$\cos\left(\frac{2\pi}{9}\right) + \cos\left(\frac{4\pi}{9}\right) + \cos\left(\frac{6\pi}{9}\right) + \cos\left(\frac{8\pi}{9}\right) = -\frac{1}{2}$$

Solution

Using Euler's formula [99] we can write the sum of $\cos(\frac{2\pi}{9})+\cos(\frac{4\pi}{9})+\cos(\frac{6\pi}{9})+\cos(\frac{8\pi}{9})$ as

[99] $e^{\theta i} = \cos(\theta) + i\sin(\theta)$, see above for more

$$\frac{1}{2}\left[e^{\frac{2\pi}{9}i} + e^{-\frac{2\pi}{9}i} + e^{\frac{4\pi}{9}i} + e^{-\frac{4\pi}{9}i} + e^{\frac{6\pi}{9}i} + e^{-\frac{6\pi}{9}i} + e^{\frac{8\pi}{9}i} + e^{-\frac{8\pi}{9}i}\right]$$

This is actually the sum of eight of the nine roots of unity in disguise! Note that if we add 2π to anything in the form e^{ai}, this has no effect (again, Euler's formula and the fact that 2π radians is a full rotation). Therefore, the previous expression is the same as

$$\frac{1}{2}\left[e^{\frac{2\pi}{9}i}+e^{\frac{16\pi}{9}i}+e^{\frac{4\pi}{9}i}+e^{\frac{14\pi}{9}i}+e^{\frac{6\pi}{9}i}+e^{\frac{12\pi}{9}i}+e^{\frac{8\pi}{9}i}+e^{\frac{10\pi}{9}i}\right]$$

which can be re-ordered as

$$\frac{1}{2}\left[e^{\frac{2\pi}{9}i}+e^{\frac{4\pi}{9}i}+e^{\frac{6\pi}{9}i}+e^{\frac{8\pi}{9}i}+e^{\frac{10\pi}{9}i}+e^{\frac{12\pi}{9}i}+e^{\frac{14\pi}{9}i}+e^{\frac{16\pi}{9}i}\right]$$

which are all the ninth roots of unity (except 1).

Thus we can write that

$$\begin{aligned}\left[\cos\left(\frac{2\pi}{9}\right)+\cos\left(\frac{4\pi}{9}\right)+\cos\left(\frac{6\pi}{9}\right)+\cos\left(\frac{8\pi}{9}\right)\right] &= -1+1+\omega+\omega^2+...+\omega^8\\ &= -1+\frac{1-\omega^9}{\omega-1}\\ &= -1+\frac{1-1}{\omega-1}\ (\omega^9=1 \text{ as } \omega \text{ is a 9th root of 1})\\ &= -1\end{aligned}$$

which means that

$$\cos\left(\frac{2\pi}{9}\right)+\cos\left(\frac{4\pi}{9}\right)+\cos\left(\frac{6\pi}{9}\right)+\cos\left(\frac{8\pi}{9}\right)=-\frac{1}{2}$$

Chapter 16

Hyperbolic functions

16.1 Definitions

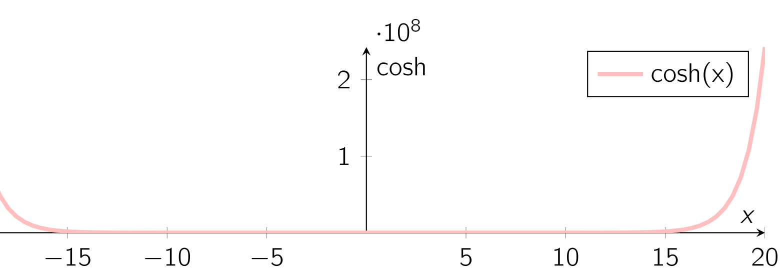
There are definitely nice ways to think about these (including their relation to hyperbolic geometry, etc.).

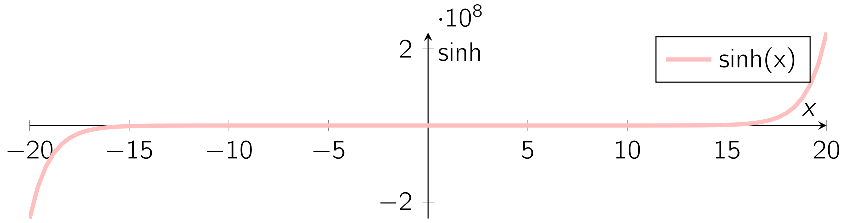
The most efficient method is to cut to the chase and define the hyperbolic functions:

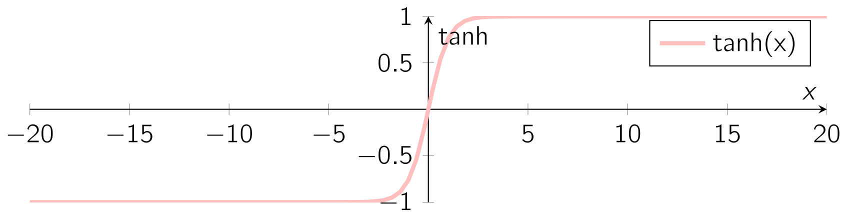
$$\cosh(x) = \frac{e^x + e^{-x}}{2} \tag{16.1}$$
$$= \frac{e^{2x} + 1}{2e^x} \text{ multiplying by } \frac{e^x}{e^x} \tag{16.2}$$

$$\sinh(x) = \frac{e^x - e^{-x}}{2} \tag{16.3}$$
$$= \frac{e^{2x} - 1}{2e^x} \text{ multiplying by } \frac{e^x}{e^x} \tag{16.4}$$

When plotted, they look like this:







16.2 Properties

16.2.1 Odd/even nature

Note that as with their analogous trigonometric functions [100], $\sinh(x)$ is an odd function, and $\cosh(x)$ is an even function.

[Proof that sinh(x) is an odd function]Proof that $\sinh(x)$ is an odd function

$$\sinh(-x) = \frac{e^{(-x)} - e^{-(-x)}}{2} \tag{16.5}$$
$$= \frac{e^{(-x)} - e^{-(-x)}}{2} \tag{16.6}$$
$$= \frac{e^{-x} - e^{x}}{2} \tag{16.7}$$
$$= -\sinh(x) \tag{16.8}$$

Proof that $\cosh(x)$ is an odd function

$$\cosh(-x) = \frac{e^{(-x)} + e^{-(-x)}}{2} \tag{16.9}$$
$$= \frac{e^{-x} + ex}{2} \tag{16.10}$$
$$= \cosh(x) \tag{16.11}$$

Proof that $\tanh(x)$ is an odd function

[101]

[100] $\sin(x)$ is analogous to $\sinh(x)$ and $\cos(x)$ is analogous to $\cosh(x)$

[101] This can be proved by rewriting $\tanh(x)$ in terms of e^x, but as we've already proved the even/odd nature of $\cos(x)$ and $\sin(x)$ it makes sense to use that!

$$\tanh(-x) = \frac{\sinh(-x)}{\cosh(-x)} \quad (16.12)$$
$$= \frac{-\sinh(x)}{\cosh(x)} \quad (16.13)$$
$$= -\frac{\sinh(x)}{\cosh(x)} \quad (16.14)$$
$$= -\tanh(x) \quad (16.15)$$

16.2.2 Inverse functions

All the hyperbolic functions have inverses, although for some of them, it is necessary to restrict the domain. [102]

The inverse functions are in the formula booklet of most maths exam boards. Otherwise, they can be derived.

16.2.3 Inverse function of sinh(x)

Taking $\sinh(x)$, as an example, we know that

$$\sinh(x) = \frac{e^x - e^{-x}}{2} \quad (16.16)$$

We can write x in terms of $\sinh(x)$, by multiplying the fraction through by $\frac{e^x}{e^x}$.

$$\sinh(x) = \frac{e^{2x} - 1}{2e^x} \quad (16.17)$$

Here we can substitute $u = e^x$

$$\sinh(x) = \frac{u^2 - 1}{2u} \quad (16.18)$$

And then we can rearrange this to give a quadratic in terms of u.

$$u^2 - 2u\sinh(x) - 1 = 0 \quad (16.19)$$

By applying the quadratic formula, we can solve this for u

$$u = \frac{-(-2\sinh(x)) \pm \sqrt{(-2\sinh(x))^2 - 4(1)(-1)}}{2} \quad (16.20)$$

We can then simplify this a bit

$$u = \frac{2\sinh(x) \pm \sqrt{4}\sqrt{\sinh(x)^2 + 1}}{2} \quad (16.21)$$

If we then reverse the substitution[103], replacing u with e^x.

[102] As noted in the "functions" section (although I might not have uploaded that section yet?), functions can only have inverses if they are one-to-one (i.e. for every value the function outputs, there can only be one possible input which would have resulted in that output.) The graph of $y = \cosh(x)$ is not one-to-one for the whole domain (some outputs correspond to multiple inputs), so we have to restrict its value.

[103] Arguably, the substitution didn't help here, but substitutions do often make it easier to spot the structure of a problem by simplifying problems in a way which makes them look more like a previously seen problem.

$$e^x = \sinh(x) \pm \sqrt{\sinh(x)^2 + 1} \tag{16.22}$$

The next step is to take the natural logarithm of both sides

$$x = \ln\left(\sinh(x) \pm \sqrt{\sinh(x)^2 + 1}\right) \tag{16.23}$$

here, because $a < \sqrt{a^2 + 1}$ [104] we cannot have the negative case (for the $\pm$), as the domain of the logarithm function requires that the input is greater than zero.

[104] This can be proved by considering the cases when $a > 0$, $a = 0$ and $a < 0$

And we have found the inverse function!

$$\text{arsinh}(x) = \ln\left(x + \sqrt{x^2 + 1}\right) \tag{16.24}$$

16.3 Relationship to trig functions

Using Euler's formula, it is possible to write both $\cos(\theta)$ and $\sin(\theta)$ in terms of e^x. As $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, and $e^{-ix} = \cos(-\theta) + i\sin(-\theta) = \cos(\theta) - i\sin(\theta)$ we can either add or subtract these two quantities in order to write both trigonometric functions in terms of e.

For cos, we can add e^{ix} and e^{-ix}.

$$e^{ix} + e^{-ix} = \cos(\theta) - i\sin(\theta) + \cos(\theta) + i\sin(\theta) \tag{16.25}$$
$$= 2\cos(\theta) \tag{16.26}$$

Thus we can say that

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2} \tag{16.27}$$

for all values of x. [105].

[105] Which looks remarkably like a hyperbolic function!

We can do a similar thing for $\sin(x)$.

$$e^{ix} - e^{-ix} = \cos(\theta) - i\sin(\theta) - (\cos(\theta) + i\sin(\theta)) \tag{16.28}$$
$$= -2i\sin(\theta) \tag{16.29}$$

Which means that

$$\sin(x) = \frac{e^{ix} - e^{-ix}}{-2i} \tag{16.30}$$

If, in either of these equations we set $x = i\theta$ we can write $\sin(i\theta)$ and $\cos(i\theta)$ in terms of $\sinh(x)$ and $\cosh(x)$, respectively.

$$\sin(i\theta) = \frac{e^{i(i\theta)} - e^{-i(i\theta)}}{-2i} \tag{16.31}$$
$$= \frac{e^{-\theta} - e^{\theta}}{-2i} \tag{16.32}$$
$$= i\frac{e^{\theta} - e^{-\theta}}{2} \text{ multiplying by } \frac{i}{i} = 1 \tag{16.33}$$
$$= i\sinh(\theta) \tag{16.34}$$

$$\cos(i\theta) = \frac{e^{i(i\theta)} + e^{-i(i\theta)}}{2} \tag{16.35}$$
$$= \frac{e^{-\theta} + e^{\theta}}{2} \tag{16.36}$$
$$= \cosh(\theta) \tag{16.37}$$

16.4 Identities

In general, all the trigonometric identities also hold for hyperbolic functions, albeit with minor modifications.

The basic rule of thumb is that wherever you see $\pm\sin^2(x)$ in "normal" trigonometry, replace it with $\mp\sinh^2(x)$ (so $\sin^2(x)$ would be replaced with $-\sinh^2(x)$, and $-\sin^2(x)$ with $\sinh^2(x)$). This comes from the relationship $\sin(ix) = i\sinh(x)$ which means that $\sin^2(ix) = i^2\sinh^2(x) = -\sinh^2(x)$. Why this is true was explored in the previous section.

To prove identities, the usual thing to try is to write everything in terms of the exponential function and go from there.

Example: Show that

$$\cosh(2x) = \cosh^2(x) + \sinh^2(x)$$

Solution: Usually when proving identities it's easiest to start with the more "complicated" [106] side.

[106] This is a purely qualitative distinction, but it's usually the side that makes you think either "yuck" or "what a fun challenge" depending on your view of mathematics.

$$\cosh^2(x) + \sinh^2(x) = \left(\frac{e^{2x}+1}{2e^x}\right)^2 + \left(\frac{e^{2x}-1}{2e^x}\right)^2 \tag{16.38}$$
$$= \frac{2e^{4x}+2}{4e^{2x}} \tag{16.39}$$
$$= \frac{e^{4x}+1}{2e^{2x}} \tag{16.40}$$
$$= \cosh(2x) \tag{16.41}$$

Note that for the binomials in Equation **??** we know that the middle terms will cancel because the brackets are the same, except for a -1 in one and a 1 in the other.

Chapter 17

Linear algebra

This is a further maths topic.

17.1 Linearity

Linear algebra is about stuff that's "linear" (yes, that definition isn't very helpful, I know). More concretely, something is linear if it is in the form $y = ax + b$, whereas something like $y = ax^2 + bx + c$ would be quadratic, something like $y = e^x$ would be exponential, and so on.

Linear things have nice properties. For example, we would call an operator "linear" [107] if:

$$f(a + b) = f(a) + f(b) \tag{17.1}$$

[107] For example, in probability theory (which is explored further down in the document), the expected value of a discrete random variable is linear because $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

Linear algebra gives us a bunch of tools for working with these kinds of equations. You might think these don't look very interesting (compared to possibly more exciting-looking other equations) but don't let the apparent simplicity of these equations fool you. We can represent all sorts of interesting things using linear algebra (for example, statistics, computer graphics, machine learning/deep learning, and more!)

A fairly common use of linear algebra is to solve a "system" of equations. A "system" means that all the equations refer to the same variables.

17.2 Matrices

17.2.1 Matrix multiplication

Example: given that

$$A = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \tag{17.2}$$

consider A, A^2, A^3 and A^4 to make a conjecture about A^n, and then prove this conjecture by induction.

Solution

Although it's tempting to reach for one's calculator to compute A, A^2, A^3 and A^4, it makes things harder.

When we multiply A by A, we get that

$$A^2 = \begin{pmatrix} 2^2 & 2 \cdot 3 + 3 \cdot 2 \\ 0 & 2^2 \end{pmatrix} \tag{17.3}$$

If we then multiply this by A again, we get that

$$A^3 = \begin{pmatrix} 2^3 & 2(2 \cdot 3 + 3 \cdot 2) + 3 \cdot 2^2 \\ 0 & 2^3 \end{pmatrix} \tag{17.4}$$

We can multiply out the bracket and get that

$$\begin{aligned} 2(2 \cdot 3 + 3 \cdot 2) + 3 \cdot 2^2 &= 2^2 \cdot 3 + 2^2 \cdot 3 + 3 \cdot 2^2 && (17.5) \\ &= 2^2 \cdot 3 + 2^2 \cdot 3 + 2^2 \cdot 3 && (17.6) \\ &= 3 \cdot 2^2 \cdot 3 && (17.7) \end{aligned}$$

Simplifying here is a bad idea, as we're looking for patterns, and these are easier to spot when we have more information to work with (which we do when we think about the factors of our numbers.)

Multiplying to find A^4 we get that

$$A^4 = \begin{pmatrix} 2^4 & 2(3 \cdot 2^2 \cdot 3) + 3 \cdot 2^3 \\ 0 & 2^4 \end{pmatrix} \tag{17.8}$$

We can simplify this a bit, because

$$2(3 \cdot 2^2 \cdot 3) + 3 \cdot 2^3 = 3 \cdot 2^3 \cdot 3 + 3 \cdot 2^3 \qquad = 4 \cdot 2^3 \cdot 3 \tag{17.9}$$

This implies that

$$A^4 = \begin{pmatrix} 2^4 & 4 \cdot 2^3 \cdot 3 \\ 0 & 2^4 \end{pmatrix} \tag{17.10}$$

Looking at these four matrices, the general shape is something like

$$A^n = \begin{pmatrix} 2^n & n \cdot 2^{n-1} \cdot 3 \\ 0 & 2^n \end{pmatrix} \tag{17.11}$$

There's another way of doing this, but the maths behind it isn't on the A Level specification. It's the way I originally did it.

From the initial multiplication, we can see that the general shape of A^n is

$$\begin{pmatrix} 2^n & ? \\ 0 & 2^n \end{pmatrix} \tag{17.12}$$

Here the ? denotes the confusion about what $A_{1,2}$ is. In terms of notation, it's slightly more handy to call ? something more like $f(n)$.

We can also write the matrix A^n as AA^{n-1}.

$$A^n = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2^{n-1} & f(n-1) \\ 0 & 2^{n-1} \end{pmatrix} \tag{17.13}$$

We can apply the standard rules of matrix multiplication to the above.

$$A^n = \begin{pmatrix} 2^n & 2f(n-1) + 3 \cdot 2^{n-1} \\ 0 & 2^n \end{pmatrix} \tag{17.14}$$

We can compare the value of A^n we have computed here (in terms of $f(n-1)$) to the one we'd like to know (that of $f(n)$). This gives

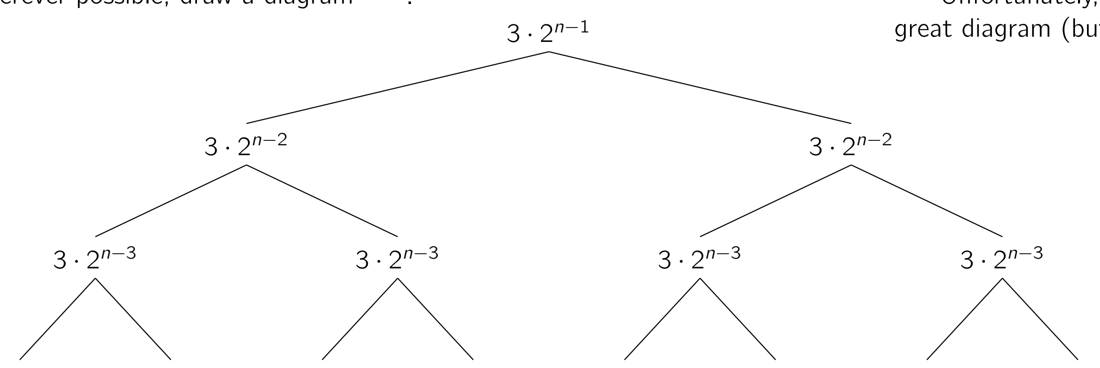
$$f(n) = 2f(n-1) + 3 \cdot 2^{n-1} \tag{17.15}$$

We also know that $f(1) = 3$. To find a closed-form expression for $f(n)$, we can just keep substituting:

$$\begin{aligned} f(n) &= 2(2f(n-2) + 3 \cdot 2^{n-2}) + 3 \cdot 2^{n-1} && (17.16) \\ &= 2^2 f(n-2) + 3 \cdot 2^{n-1} + 3 \cdot 2^{n-1} && (17.17) \\ &= 2^2(2f(n-3) + 3 \cdot 2^{n-3}) + 3 \cdot 2^{n-1} + 3 \cdot 2^{n-1} && (17.18) \\ &= 2^3 f(n-3) + 3 \cdot 2^{n-1} + 3 \cdot 2^{n-1} + 3 \cdot 2^{n-1} && (17.19) \\ &= n \cdot 3 \cdot 2^{n-1} && (17.20) \end{aligned}$$

The last step isn't necessarily immediately obvious (doing the expansion by hand might make it easier to understand). The other way to think about this is by using one of the helpful heuristics looked at in the heuristics section, in this case "wherever possible, draw a diagram" [108].

[108] Unfortunately, it's not a great diagram (but I tried).



The sum of each node $f(n)$ in the tree is $2f(n-1)+3\cdot 2^{n-1}$ - the $2f(n-1)$s are drawn as children of the parent node. We want to find the sum of all the nodes in the tree. To do this, note that we have n levels of the tree and at each level the nodes sum to $3\cdot 2^{n-1}$, and thus overall we have that $f(n)$ is equal to the sum of all the nodes, which is equal to $n\cdot 3\cdot 2^{n-1}$.

It's all downhill (in difficulty) from here!

For the proof by induction, if we set

$$P(n) = \begin{pmatrix} 2^n & n\cdot 2^{n-1}\cdot 3 \\ 0 & 2^n \end{pmatrix} \tag{17.21}$$

then we want to show that $A^n = P(n)$ for every natural number.

We start with the basis case: as $A^1 = A$, it is clear that $A^1 = P(1)$.

For the inductive step, we assume that $P(k) = A^k$, and then we consider $P(k+1)$, which is equal to A^1A^k. To this expression, we can now apply the inductive hypothesis (always be looking at $P(k+1)$ to see where $P(k)$ has been hidden!) and thus this expression is equal to $AP(k)$. We can then carry out the multiplication

$$\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 2^k & k\cdot 2^{k-1}\cdot 3 \\ 0 & 2^k \end{pmatrix} = \begin{pmatrix} 2\cdot 2^k & 2(k\cdot 2^{k-1}\cdot 3)+3\cdot 2^k \\ 0 & 2\cdot 2^k \end{pmatrix} \tag{17.22}$$

This can then be rewritten as

$$\begin{pmatrix} 2^{k+1} & k\cdot 2^k\cdot 3+3\cdot 2^k \\ 0 & 2^{k+1} \end{pmatrix} \tag{17.23}$$

and as $k\cdot 2^k\cdot 3+3\cdot 2^k$ is equal to $2^k\cdot 3\cdot(k+1)$ we have that

$$\begin{pmatrix} 2^{k+1} & 2^k\cdot 3\cdot(k+1) \\ 0 & 2^{k+1} \end{pmatrix} \tag{17.24}$$

The final step is to rewrite the 2^k as $2^{(k+1)-1}$, and then we have

$$\begin{pmatrix} 2^{k+1} & (k+1)\cdot 2^{(k+1)-1}\cdot 3 \\ 0 & 2^{k+1} \end{pmatrix} \tag{17.25}$$

Thus as $P(n)$ is true for $n=1$ and $P(n)$ implies that $P(n+1)$ is true, $P(n)$ must hold for all n.

17.3 Transformations

17.3.1 Invariance

Chapter 18

Real analysis

18.1 Sequences

A sequence is a lot of numbers, one after the other. Usually the terms are related in some way to the previous terms. For example, in this sequence[109], each term is the sum of the two previous terms[110]

$$1, 2, 3, 5, 8, 13, \ldots \tag{18.1}$$

We can also write the nth term of a sequence x as $x(n)$, and denote the entire sequence as $\{x_n\}$

18.1.1 Convergence of sequences

Sometimes a sequence can *converge*, which informally means that the value of the *n*th term in the sequence gets closer and closer to a specific value - infinitely close in fact!

Can we say more than "this sequence *looks* like it's getting closer and closer to this value" though? Yes! Another way of saying that something is getting "closer" to a value, is to say that the distance between the two is decreasing. Using the modulus function [111] we can write the distance between any two points $x(n)$ and a as

$$|a(n) - a| \tag{18.2}$$

What about "really close"? The way to do this is to say that for every value of $\epsilon > 0$ where ϵ is a real number and as $n \to \infty$ we have that

$$|x(n) - a| < \epsilon \tag{18.3}$$

then $\{x_n\}$ converges to a. As we can pick any $\epsilon > 0$ (but not $\epsilon = 0$) what we are in effect saying is that the distance is infinitely close to zero, but not equal to zero. That is, it *converges* to 0 but is not equal to zero.

[109] Known as the Fibonacci sequence

[110] If we set that the first term is 1 and the second 2 then to get the third term, we add the previous two, so $1 + 2 = 3$. This process continues for the rest of the terms in the sequence.

[111] There is a section on the modulus function in the "Algebra" chapter of this document, and another section on how we can write distances in terms of the modulus function in the vectors chapter.

18.1.2 A sequence can only converge to one value

Before proving this, it's helpful to formulate precisely what we want to prove.

What [112]we want to show is that *for every* converging sequence, there exists a unique value to which it converges. The only way this could not be true is if *there exists* at least one sequence which converges to more than one value.

To prove this by contradiction, let us assume that there is indeed a sequence $\{x_n\}$ which converges to both a and b, where $a \neq b$. In this case we have that for all $\epsilon > 0$ that as $n \to \infty$,

$$|x_n - a| < \epsilon \text{ and } |x_n - b| < \epsilon \tag{18.4}$$

We would like to show that $a = b$, or that $a - b = 0$ (as this would be a contradiction). We know that

$$|a - b| = |a - x_n + x_n - b| \tag{18.5}$$

Using the triangle inequality we can write that
[113]

$$|a - x_n + x_n - b| \leqq |a - x_n| + |x_n - b| \tag{18.6}$$
$$= \epsilon + \epsilon \tag{18.7}$$
$$= 2\epsilon \tag{18.8}$$

Therefore, overall we have that

$$|a - b| \leqq 2\epsilon \tag{18.9}$$

However, we can pick any value which is greater than 0 for ϵ here. If as we are assuming, $a \neq b$ then for any constant $c > 0$ we also have that $c\,|a - b|$ is greater than 0. If we set c to something less than $\frac{1}{2}$ (e.g. $\frac{1}{3}$) then we would have that

$$|a - b| \leqq \frac{2}{3}\,|a - b| \tag{18.10}$$

This is definitely not true, and hence we have derived a contradiction; thus there can only be one.

[112] Note that *there exists* is the negation of *for every*. It helps me to imagine a haystack. If you tell me that "there exists a needle in this haystack", then the only way I can disprove you is to search through all of (for every) the haystack and show that there is no needle. If I say "for every piece of hay in this haystack, none of them is hiding a needle", then you can disprove me by showing that at least one piece of hay is hiding a needle.

[113] Unfortunately, this equation is a bit ambiguous - what I'm trying to say is that everything on the left-hand side is less than or equal to everything on the right-hand side and that everything on the right-hand side is equal.

Chapter 19

Probability

19.1 Some facts about variance

The "variance" of a discrete random variable[114] is a measure of "spread" (how far apart values in a distribution are). It gives the expected value of the square of the distance of the observed values (in the outcome space) from the mean (expected value of the distribution). That's a mouthful to say, so it can be easier to write this as a formula.
[115]

$$\mathrm{Var}[X] = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] \tag{19.1}$$

There is an equivalent way in which the variance can be expressed which is a bit easier to use when trying to calculate the variance of a discrete random variable by hand:

$$\begin{aligned}
\mathrm{Var}[X] &= \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] && && (19.2)\\
&= \mathrm{E}\left[X^2 - 2X\,\mathrm{E}[X] + \mathrm{E}[X]^2\right] && \text{(step 1)} && (19.3)\\
&= \mathrm{E}[X^2] - 2\,\mathrm{E}[X]\,\mathrm{E}[X] + \mathrm{E}[X^2] && \text{(step 2)} && (19.4)\\
&= \mathrm{E}\left[X^2\right] - (\mathrm{E}[X])^2 && && (19.5)
\end{aligned}$$

When we went from step 1 to step 2, we took advantage of the fact that $\mathrm{E}[X]$ is constant; in effect, we grouped our expression as$\mathrm{E}\left[(2\,\mathrm{E}[X])\,X\right]$,[116] and then used the linearity of expectation [117] to rewrite it as $(2\,\mathrm{E}[X])\,\mathrm{E}[X]$.

[114] Note: you're not imagining things, I still need to add the section I have written defining these.

[115] If it's not clear why $X - \mathrm{E}[X]$ gives the signed distance between X and $\mathrm{E}[X]$, take a look at the "vectors" chapter.

[116] Bear in mind that $\mathrm{E}[X]$ is a constant

[117] If this means nothing to you, please be aware that I have yet to write this section.

19.2 The binomial distribution

19.3 The geometric distribution

I think this is only in Further Maths.

19.4 The normal distribution

19.5 Continuous probability distributions

19.6 Transformations of continuous random variables

Imagine we have a random variable X, with probability density function

$$f_X(x) = \begin{cases} 4x^3 & 0 < x \leqq 1 \\ 0 & \text{otherwise} \end{cases} \tag{19.6}$$

and we want to find the probability density function for the random variable

$$Y = \frac{1}{X^4} \tag{19.7}$$

The first thing to do is to find the cumulative probability function for X, (by integrating)

$$F_X(x) = \begin{cases} 0 & x \leqq 0 \\ \int_0^x 4x^3 dx & 0 < x \leqq 1 \\ 1 & x > 1 \end{cases} \tag{19.8}$$

$$= \begin{cases} 0 & x \leqq 0 \\ x^4 & 0 < x \leqq 1 \\ 1 & x > 1 \end{cases} \tag{19.9}$$

Then, we can find the cumulative probability function for Y in terms of the cumulative probability function for X. From the definition of the cumulative probability function we know that

$$F_Y(y) = P(Y \leqq y) \tag{19.10}$$

Then as $Y = \frac{1}{X^4}$ we can substitute X for Y

$$P(Y \leqq y) = P(\frac{1}{X^4} < y) \tag{19.11}$$

We can then manipulate this into some function of $P(X < g(y))$ (where $g(y)$ is a function we need to determine).

$$P(Y \leqq y) = P\left(\frac{1}{X^4} < y\right) \tag{19.12}$$
$$= P\left(X^4 > \frac{1}{y}\right) \tag{19.13}$$
$$= P\left(X > \sqrt[4]{\frac{1}{y}}\right) \tag{19.14}$$
$$= 1 - P\left(X < \sqrt[4]{\frac{1}{y}}\right) \tag{19.15}$$

Note that in the second step we flipped the inequality because we took the reciprocal of both functions, and the reciprocal function makes bigger values smaller (and vice-versa) so to keep the inequality true, we had to flip the signs. From here, we plug into $F_X(x)$.

$$P(Y \leqq y) = 1 - P\left(X < \sqrt[4]{\frac{1}{y}}\right) \tag{19.16}$$
$$= 1 - F_X\left(\sqrt[4]{\frac{1}{y}}\right) \tag{19.17}$$
$$= \begin{cases} 1 - 0 & x \leqq 0 \\ 1 - \left(\sqrt[4]{\frac{1}{y}}\right)^4 & 0 < x \leqq 1 \\ 1 - 1 & x > 1 \end{cases} \tag{19.18}$$
$$= \begin{cases} 1 & x \leqq 0 \\ 1 - \frac{1}{y} & 0 < x \leqq 1 \\ 0 & x > 1 \end{cases} \tag{19.19}$$

We also need to rewrite bounds in terms of y, rather than x. As $Y = \frac{1}{X^4}$ we can write

$$X = Y^{-\frac{1}{4}} \tag{19.20}$$

And thus that

$$P(Y \leqq y) = \begin{cases} 1 & y^{-\frac{1}{4}} \leqq 0 \\ 1 - \frac{1}{y} & 0 < y^{-\frac{1}{4}} \leqq 1 \\ 0 & y^{-\frac{1}{4}} > 1 \end{cases} \tag{19.21}$$

$$= \begin{cases} 1 & y \geqq \infty \\ 1 - \frac{1}{y} & y < \infty \text{ and } y \geqq 1 \\ 0 & y < 1 \end{cases} \tag{19.22}$$

$$= \begin{cases} 1 - \frac{1}{y} & y \geqq 1 \\ 0 & y < 1 \end{cases} \tag{19.23}$$

Note that here it is assumed that $\frac{1}{0} = \infty$ (it makes it nice and easy to work with the bounds).

As we now have the cumulative probability function for Y, the final step is to differentiate to get the probability density function.

$$\frac{d}{dx}\left[P(Y \leqq y)\right] = \begin{cases} \frac{d}{dx}\left[1 - \frac{1}{y}\right] & y \geqq 1 \\ 0 & y < 1 \end{cases} \tag{19.24}$$

$$= \begin{cases} \frac{d}{dx}\left[-y^{-1}\right] & y \geqq 1 \\ 0 & y < 1 \end{cases} \tag{19.25}$$

$$= \begin{cases} y^{-2} & y \geqq 1 \\ 0 & y < 1 \end{cases} \tag{19.26}$$

$$= f_Y(y) \tag{19.27}$$

19.7 Hypothesis testing

Here's one vision for how science should work: I theorise that A has some effect on B. Of course, A might not have any effect on B. To do this, we can construct two hypotheses: the "null hypothesis" which is the possibility that there is no effect and the "alternate hypothesis" which is mutually exclusive[118] to the null hypothesis. [**?**]

[118] i.e. it is *not* possible for both the null and alternate hypothesis to be true

For example, let's suppose that we have a jar filled with objects of type "A" and "B". Last time I checked, 37% of the objects were of type A. I suspect, however, that since last time I checked the proportion of objects that of type "A" has decreased.

From this, we can come up with some hypotheses

Null hypothesis: there is no change (i.e. the proportion is still 37%).

Alternate hypothesis: the proportion of items which are of type "A" has decreased (i.e. the proportion is $< 37\%$).

If we have some data, we can try to work out how likely or unlikely the data is under the null hypothesis - if getting that data is really unlikely under the null hypothesis then we can reject the null hypothesis in favour of the alternate hypothesis[119].

[119] You **mustn't** ever *disprove* the null hypothesis, you can only uphold/reject it. We're not saying that we've *disproved* the hypothesis, we're saying that it's either *likely* or *unlikely* to be true.

Chapter 20

Statistics

Sir Humphrey Appleby: *[Looking amused] Statistics? You can prove anything with statistics.*
Jim Hacker: *Even the truth?*
Sir Humphrey Appleby: *Ye-es*
—*From* Yes Prime Minister.

20.1 p-values

TODO

20.2 Testing for a median using the binomial distribution

In any set of data, we'd expect half the values to lie either above or below the median[120]. The probability distribution of the number of values above (or below) the median (X) can be defined as

[120] This is from how the median is defined.

$$X \sim B\left(n, \frac{1}{2}\right) \tag{20.1}$$

We can use this to carry out a hypothesis test to see if a set of data has a given median; we count up the total number of observations, and how many are above or below the median. We can then work out the probability of how likely it would be to see that value (or a more extreme value) under the null hypothesis.

20.3 Wilcoxon Matched-Pairs

20.4 Wilcoxon Signed-Rank

Appendices

Appendix A

Appendices

These contain a list of resources, as well as solutions to some miscellaneous problems.

Appendix B

Resources

These are all resources which I find useful. They are not in any particular order.

- This document! https://notes.reasoning.page/html/main or https://notes.reasoning.page/main.pdf
- Madasmaths (available at https://madasmaths.com) - a veritable goldmine of mathematics questions.
- George Pólya's "How To Solve It" (which is discussed, briefly, in the "solving problems" section of this document)
- Paul Zeitz's "The Art and Craft of Problem Solving"
- Math Centre https://mathcentre.ac.uk (in particular their "teach yourself" section, available at https://mathcentre.ac.uk/types/#h4)
- HELM (Helping Engineers Learn Mathematics): features notes and practice questions. Available at https://www.lboro.ac.uk/departments/mlsc/student-resources/helm-workbooks/
- Isaac Physics (https://isaacphysics.org), for both the practice questions as well as their notes (https://isaacphysics.org/concepts?stage=all).
- Colin Beveridge's blog, "Flying Colours Maths" which I originally discovered because of a post explaining how to find all the invariant lines under a matrix tranformation (https://www.flyingcoloursmaths.co.uk/ask-uncle-colin-invariant-lines/)
- BlackPenRedPen, a YouTube channel (https://www.youtube.com/channel/UC_SvYP0k05UKiJ_2ndB02IA)
- Stephen Miller's "The Probability Lifesaver" - really good book on probability

- Adrian Banner's "The Calculus Lifesaver"
- Georgi Shilov's "Linear Algebra"
- Jiří Lebl's "Notes on Diffy Qs" (https://www.jirka.org/diffyqs/diffyqs.pdf)
- The American Institute of Mathematics maintain a list of open access textbooks of which they approve (https://aimath.org/textbooks/approved-textbooks/).
- Ian Anderson's "A First Course in Discrete Mathematics"
- Philip J. Davis and Reuben Hersh's "The Mathematical Experience"
- "Thinking Mathematically" by John Mason et al. (https://www.goodreads.com/book/show/1241663.Thinking_Mathematically)
- I found Alan Schoenfeld's "Mathematical Thinking and Problem Solving" interesting, but I can see why many other people might not.
- Thomas Cormen et al. "Introduction to Algorithms"
- Advanced Problems in Mathematics: Preparing for University https://www.openbookpublishers.com/product/342

Appendix C

Assorted problems

Solutions to problems which have yet to be categorised.

C.0.1 Cones problem

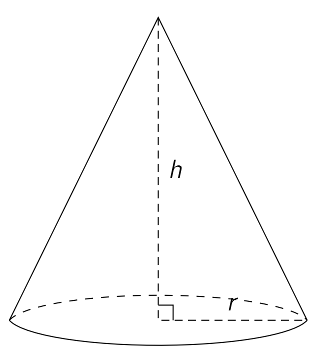
Problem: Two similar cones, X and Y, have surface areas 270cm^2 and 120cm^2 respectively.

The volume of cone X is 1215cm^3.

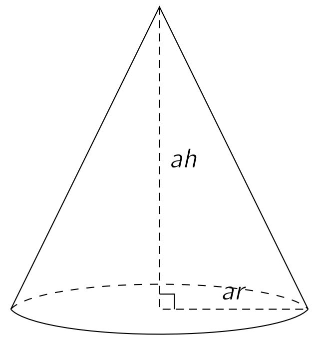
Show that the volume of cone Y is 360cm^3.

Solution:

We have two cones, let's let X be this one, with radius r and height h



Then we can draw Y (which is smaller than X, but not drawn to scale here). As they are similar cones, the radius of Y is equal to ar (where a is the scale factor between the two cones) and the height of Y is equal to ah



First, the volume of a cone is $\frac{\pi r^2 h}{3}$ and the surface area of a cone is $\pi r(r + \sqrt{h^2 + r^2})$. These can be derived using integration (but not in Single Maths A Level at least).

Using this, we can write this equation relating the radius and height of X to the value given in the question

$$\pi r(r + \sqrt{h^2 + r^2}) = 270 \tag{C.1}$$

We can also do the same for Y

$$\pi ar(ar + \sqrt{(ah)^2 + (ar)^2}) = 120 \tag{C.2}$$

This can be simplified a bit by factoring the a^2 from the square rooted part of the equation, giving

$$\pi ar(ar + \sqrt{a^2}\sqrt{(h)^2 + (r)^2}) = 120 \tag{C.3}$$

and thus that

$$\pi a^2 r(r + \sqrt{(h)^2 + (r)^2}) = 120 \tag{C.4}$$

We can solve this for a by dividing the two equations, giving that

$$\frac{\pi a^2 r(r + \sqrt{(h)^2 + (r)^2})}{\pi r(r + \sqrt{h^2 + r^2})} = \frac{120}{270} \tag{C.5}$$

and thus, after cancelling, that

$$a^2 = \frac{9}{4} \tag{C.6}$$

which means that

$$a = \frac{3}{2} \tag{C.7}$$

Given this, we can then turn to the volume formula.

We know that the volume of X is

$$\pi r^2 h = 1215 \tag{C.8}$$

and the formula for the volume of Y is equal to

$$\pi(ar)^2(ah) = a^3\pi r^2 h \tag{C.9}$$

Thus, by multiplying Equation **??** through by a^3, we have

$$a^3\pi r^2 h = a^3 1215 \tag{C.10}$$

The left-hand side is just the volume of Y, and the right-hand side is just

$$a^3 1215 = \left(\frac{2}{3}\right)^3 \cdot 1215 \tag{C.11}$$

$$= 360 \tag{C.12}$$

Which is what we needed to show.

C.0.2 Divisibility of abc_{10}

Question: "When we represent 789 in decimal, we mean 7 hundreds plus 8 tens plus 9 units. If abc_{10} is taken to represent a decimal number (not a times b times c as in algebra), we mean the value $100a + 10b + c$. Show that if $a + b + c$ is divisible by 9, then abc_{10} is divisible by 9."

Solution: Anything that is divisible by 9 can be written as $9 \cdot$ something. As this is algebra, we can use a letter (e.g. m) to represent "something". As we are assuming that $a + b + c$ is divisible by 9, we can write that

$$a + b + c = 9m \tag{C.13}$$

Where m is some number (which depends on the value of a, b and c). Then we can consider abc_{10}. First, we can write this (as given in the question) as

$$abc_{10} = 100a + 10b + c \tag{C.14}$$

To show that this is divisible by 9 we need to write it as $9 \cdot$ something. We need to apply the fact that $a + b + c = 9m$ here (because we are trying to show that, given this assumption, abc_{10} is divisible by 9). We want to somehow extract an $a+b+c$, here. Currently, we have a c, which is great (because it's what we're after), and all we need is a b and a c, which we can get by breaking $100a$ into $99a + a$ and $10b$ into $9b + b$. When we apply this to the expression as a whole, our desired result pops out fairly quickly.

$$abc_{10} = 100a + 10b + c \tag{C.15}$$
$$= 99a + a + 9b + b + c \tag{C.16}$$
$$= 9(11a + b) + a + b + c \tag{C.17}$$
$$= 9(11a + b) + 9m \tag{C.18}$$
$$= 9(11a + b + m) \tag{C.19}$$

Therefore, if $a + b + c = 9m$ (i.e. $a + b + c$ is divisible by 9) abc_{10} is also divisible by 9.