Contents

1 Proofs

1.1 Induction

We saw 2 examples of proof by induction in level 4. Proof by induction goes as follows:

1. Establish that a statement is true if $n = 1$

2. Establish that IF a statement is true for $n = k$ THEN it is true for $n = k + 1$

3. Conclude that the statement is true for every positive integer n.

Example. We will prove by induction that $\sum_{r=1}^{n} r^2 = \frac{n(n+1)(2n+1)}{6}$.

Step 1. We know that if $n = 1$ then the left hand side is 1, and the right hand side is $\frac{1(1+1)(2*1+1)}{6} = 1$ so we have verified the "base case".

Step 2. Suppose that

$$\sum_{r=1}^{k} r^2 = \frac{k\left(k+1\right)\left(2k+1\right)}{6}$$

Then we note that

$$\sum_{r=1}^{k+1} r^2 = \frac{k\left(k+1\right)\left(2k+1\right)}{6} + \left(k+1\right)^2 = \frac{k\left(k+1\right)\left(2k+1\right) + 6\left(k+1\right)^2}{6} = \frac{\left(k+1\right)\left(k\left(2k+1\right) + 6\left(k+1\right)\right)}{6}$$

$$= \frac{\left(k+1\right)\left(2k^2+7k+6\right)}{6} = \frac{\left(k+1\right)\left(k+2\right)\left(2k+3\right)}{6} = \frac{\left(k+1\right)\left(\left(k+1\right)+1\right)\left(2\left(k+1\right)+1\right)}{6}$$

This establishes that if the formula is true for k, then it is true for k+1. Since it is true at 1, it is true for all n, hence it is proved.

Example. We will prove by induction that for every positive integer n, $f\left(n\right) := 5^n + 4n + 7$ is divisible by 8.

To do this, we will verify it for $n = 1$. We see that

$$f\left(1\right) = 5^1 + 4 + 7 = 16$$

So now suppose it is true for k, then

$$f\left(k+1\right) = 5^{k+1} + 4\left(k+1\right) + 7 = 5^{k+1} + 4k + 11$$

The trick is that if we consider $5f\left(k\right)$ this will look a lot like $f\left(k+1\right)$. We see that

$$5f\left(k\right) = 5^{k+1} + 20k + 35$$

Now we see that

$$f(k+1) = 5f(k) - 16k - 24 = 5f(k) - 8(2k+3)$$

Therefore it is now clear that if $f(k)$ is divisible by 8 then $f(k+1)$ must also be divisible by 8, so the result is proven by induction.

Example. Suppose we have a sequence u_n satisfying $u_n = 7u_{n-1} - 12u_{n-2}$ and we are given that $u_1 = 7$ and $u_2 = 25$. We will prove here by induction that it follows that the general formula for u_n is $u_n = 3^n + 4^n$. However, we will see in the last section how we can solve recurrence relations like these directly, and it is useful to know that one can check the solution we get by induction.

Since u_n depends on the previous 2 values instead of the previous 1 value, we need to verify 2 base cases instead of 1 in order to proceed with the induction step. We first verify that if $n = 1$ then $3^1 + 4^1 = 7$ and $3^2 + 4^2 = 25$. Now we suppose it is true for $n = k-1$ and $n = k-2$, then

$$u_k = 7\left(3^{k-1} + 4^{k-1}\right) - 12\left(3^{k-2} + 4^{k-2}\right) = \frac{7}{3}3^k + \frac{7}{4}4^k - \frac{12}{3^2}3^k - \frac{12}{4^2}4^k$$
$$= \left(\frac{7}{3} - \frac{12}{9}\right)3^k + \left(\frac{7}{4} - \frac{12}{16}\right)4^k = 3^k + 4^k$$

So the result is proved for all n. You get the idea.

Example. We can prove $n^3 - n$ is always divisible by 6 similarly. It is 0 when $n = 1$ trivially so just take the difference between 2 terms. We get

$$(n+1)^3 - n^3 - (n+1) + n = ... = 3n^2 + 3n = 3n(n+1)$$

And clearly one of n and $n+1$ is even so the difference between consecutive terms, and thus the whole thing, is divisible by 6.

1.2 Examples of irrationality proofs

We saw in level 3 how to prove $\sqrt{2}$ is irrational. We can use the same method to prove that $\sqrt{3}$ is irrational. It is not automatic that if two numbers are irrational then so is their sum (Consider $\sqrt{2} + \left(2 - \sqrt{2}\right)$) and in fact nobody in the world has proved that $e + \pi$ is irrational as of writing this (although I am completely certain that it is true). What we will prove here as an example is that $\sqrt{2} + \sqrt{3}$ is irrational.

To do this, note that if we suppose $\sqrt{2} + \sqrt{3} = \frac{a}{b}$ for a contradiction then $\left(\sqrt{2} + \sqrt{3}\right)^2 = \frac{a^2}{b^2}$. Expanding the left hand side we get

$$\sqrt{2}^2 + \sqrt{3}^2 + 2\sqrt{2}\sqrt{3} = 5 + 2\sqrt{6}$$

But since we can prove $\sqrt{6}$ is irrational using the same trick as from level 3, it follows that no such integers a^2 and b^2 can exist, as if $5 + 2\sqrt{6} = \frac{p}{q}$ then $\sqrt{6} = \frac{1}{2}\left(\frac{p}{q} - 5\right)$ which can in fact be written as a fraction.

We will now prove that $\log_2(3)$ is irrational. It's quite a fun trick.

Suppose $\log_2(3) = \frac{a}{b}$, then multiplying both sides by b gives $b\log_2(3) = a$ so $\log_2\left(3^b\right) = a$. Taking 2 to the power of both sides we have the claim that there are integers a and b such that $3^b = 2^a$, which is nonsense as if $a \geq 1$ (which is true since $\log_2(3)$ is not 0) then we have that an odd number equals an even number which is not possible.

2 Complex numbers

2.1 Properties

We will briefly introduce some notation here.

$\mathbb{N}$ is the set of natural numbers (ie, positive integers), $\mathbb{Z}$ is the set of (positive or negative or zero) integers, $\mathbb{Q}$ is the set of rational numbers, $\mathbb{R}$ is the set of real numbers, and $\mathbb{C}$ is the set of complex numbers (I will explain what these are now).

Recall from level 4 that complex numbers are numbers of the form $a + bi$ where a, b are real and $i = \sqrt{-1}$. We can draw these on an "Argand diagram" where a is the x-coordinate and b is the y-coordinate. We discussed in level 4 that multiplying by i is like rotating by 90 degrees anticlockwise, and we more generally discussed what multiplying complex numbers does geometrically.

Definition. The modulus of a complex number z, written $|z|$ is the distance the point representing z is from the origin on an Argand diagram. By Pythagoras, it follows that if $z = a + bi$ then $|z| = \sqrt{a^2 + b^2}$.

Definition. The argument of a complex number z, written $\arg(z)$, is as follows:

- If $z = 0$ it is undefined. If z is a positive real number $\arg(z) = 0$. If z is a negative real number $\arg(z) = \pi$

- If z has positive real part then $\arg(z)$ is the angle, in radians, between the line joining z and the origin and the line joining the points 0 and 1.

- If z has negative real part then $\arg(z)$ is minus the angle, in radians, between the line joining z and the origin and the line joining the points 0 and 1.

From this definition we see that $\arg(z)$ takes values from $-\pi$ to π.

This definition is convenient because in level 4 we saw that walking an angle θ along the unit circle lands us at the point $\cos(\theta) + i\sin(\theta)$. Well actually, if you think about it, this θ is exactly the argument of the complex number (it just may be different by multiples of a full rotation). As I also mentioned in level 4, this is equal to $e^{i\theta}$.

Now given a complex number z, it is equal to $|z|w$ where w is on the unit circle, and $w = \cos(\arg(z)) + i\sin(\arg(z))$.

Example. We will consider $z = 1 + \sqrt{3}i$. Note that $|z| = \sqrt{1^2 + \sqrt{3}^2} = \sqrt{4} = 2$. And $\arg(z)$ involves solving for the angle. We can do this by calling the angle θ and considering figure 1.

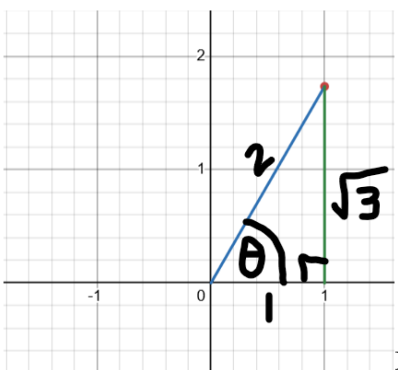

Figure 1

We see that we must have $\tan(\theta) = \sqrt{3}$ and thus $\theta = \frac{\pi}{3}$. Therefore we can write $z = 2\left(\cos\left(\frac{\pi}{3}\right) + i\sin\left(\frac{\pi}{3}\right)\right)$, or $z = 2e^{i\frac{\pi}{3}}$.

Definition. The complex conjugate of a complex number $a + bi$ is $a - bi$. Geometrically we are taking the complex number on an Argand diagram and reflecting it about the real axis.

Note that if we have a complex number on the unit circle of the form $e^{i\theta}$ then its conjugate will be $e^{-i\theta}$ since it will be as though we traversed minus the angle.

If z is a complex number then its conjugate is conventionally written as either z^* or $\bar{z}$.

One thing to know about the complex conjugate is:

Proposition. Let $f(x)$ be a polynomial with real coefficients and suppose that z is a root, then the conjugate of z is also a root.

Proof. The fact is not too surprising and the proof is not difficult, but as far as I can tell it's not proven in A level textbooks so see level 6.1.

$\square$

Using this fact, we will do an example.

Example. Find all roots of $x^4 + x^3 - 5x^2 - 7x + 10$ given that $i - 2$ is a root.

By the proposition above, $-i - 2$ is also a root. Therefore by the factor theorem, $(x - i + 2)(x + i + 2)$ is a quadratic factor of the polynomial. By the difference of squares rule this expands to $(x + 2)^2 - i^2 = x^2 + 4x + 5$. Therefore we can do polynomial long division (level 3) to try to find the other 2 roots. We want to find

$$\frac{x^4 + x^3 - 5x^2 - 7x + 10}{x^2 + 4x + 5}$$
$$= x^2 + \frac{-3x^3 - 10x^2 - 7x + 10}{x^2 + 4x + 5}$$
$$= x^2 - 3x + \frac{2x^2 + 8x + 10}{x^2 + 4x + 5}$$
$$= x^2 - 3x + 2 = (x - 1)(x - 2)$$

Therefore all 4 roots of the polynomial above are 1, 2, $i - 2$ and $-i - 2$.

2.2 Proving trig identities

We basically did this in level 4.

De Moivre's theorem says that for any integer n,

$$(\cos(\theta) + i\sin(\theta))^n = \cos(n\theta) + i\sin(n\theta)$$

This is easy to see from our geometric ideas about complex numbers from level 4. It is pretty ironic, though, that most texts will prove this by induction using the addition formulae that we need the ideas from the proof of this to even prove in the first place, which is so circular.

We can use this to find, say, a formula for $\sin(5\theta)$ as follows.

We note that

$$(\cos(\theta) + i\sin(\theta))^5 = \cos(5\theta) + i\sin(5\theta)$$

We will carefully expand the left hand side, using $i^2 = -1$, to get

$$\cos^5(\theta) + 5i\sin(\theta)\cos^4(\theta) - 10\sin^2(\theta)\cos^3(\theta) - 10i\sin^3(\theta)\cos^2(\theta) + 5\sin^4(\theta)\cos(\theta) + i\sin^5(\theta) = \cos(5\theta) + i\sin(5\theta)$$

As in level 4 we take the imaginary part of both sides. We conclude that

$$5\sin(\theta)\cos^4(\theta) - 10\sin^3(\theta)\cos^2(\theta) + \sin^5(\theta) = \sin(5\theta)$$

We can also find a formula for (eg) $\cos^5(\theta)$. It's a bit trickier but I'll walk through the steps.

First, note that for any z, simply expanding gives

$$\left(z + \frac{1}{z}\right)^5 = z^5 + 5z^3 + 10z + \frac{10}{z} + \frac{5}{z^3} + \frac{1}{z^5}$$

We will work with $e^{i\theta}$ for now. As we saw in level 4, we have to be a bit careful saying that De Moivre's theorem is true just because $\left(e^{i\theta}\right)^n = e^{i(n\theta)}$. However, in this case we're fine because n is an integer. By definition $a^b = e^{b(\log(a)+2ni\pi)}$ for any integer n, but if b is an integer then adding multiples of $2nbi\pi$ to the exponent will multiply the result by 1 so we will not have multi-valuedness issues. Really these proofs should be left for the proofs level but we did all the work in level 4 and level 6 is hard enough already.

We now have, from the above identity,

$$\begin{aligned}\left(e^{i\theta}+e^{-i\theta}\right)^5 &= e^{5i\theta}+5e^{3i\theta}+10e^{i\theta}+10e^{-i\theta}+5e^{-3i\theta}+e^{-5i\theta} \\ &= \left(e^{5i\theta}+e^{-5i\theta}\right)+5\left(e^{3i\theta}+e^{-3i\theta}\right)+10\left(e^{i\theta}+e^{-i\theta}\right)\end{aligned}$$

Recall that $\cos(\theta)=\cos(-\theta)$ and that $\sin(\theta)=-\sin(-\theta)$. Note that this implies that

$$e^{i\theta}+e^{-i\theta}=(\cos(\theta)+i\sin(\theta))+(\cos(-\theta)+i\sin(-\theta))=2\cos(\theta)$$

We get a similar result for the terms $e^{3i\theta}+e^{-3i\theta}$ and $e^{5i\theta}+e^{-5i\theta}$, as well as the left hand side which was $\left(e^{i\theta}+e^{-i\theta}\right)^5$. Thus the equation above becomes

$$(2\cos(\theta))^5=2\cos(5\theta)+10\cos(3\theta)+20\cos(\theta)$$

Thus,

$$\begin{aligned}\cos^5(\theta) &= \frac{1}{2^5}(2\cos(5\theta)+10\cos(3\theta)+20\cos(\theta)) \\ &= \frac{1}{16}(\cos(5\theta)+5\cos(3\theta)+10\cos(\theta))\end{aligned}$$

To find a formula for $\sin^n(\theta)$ we can consider $z-\frac{1}{z}$ instead of $z+\frac{1}{z}$ and then do the exact same procedure.

So yeah this is basically how that works.

2.3 Roots of unity

We know that in the real numbers, -1 and 1 are the 2 numbers such that if you square them you get 1, but 1 is the only cube root of 1.

However, in the complex numbers, there are actually 2 other cube roots of 1. These are the points shown in Figure 2.

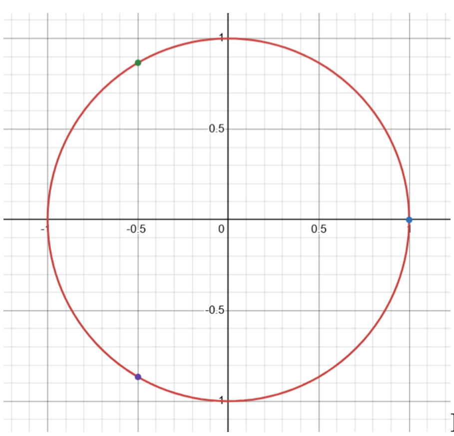

Figure 2

These points are located at 120 and 240 degrees along the unit circle. Therefore you should be able to see why they cube to 1 and no other points on the unit circle do. Note that points off the unit circle will certainly not cube to 1 since their cube will have the wrong modulus.

A similar pattern happens for 4th and 5th and higher roots of 1.

Now we note that every complex number can be written as $z=re^{i\theta}$ where $r=|z|$ and $\theta=\arg(z)$. Note that if $z=0$ then θ will just be 0 or whatever you want it to be since it doesn't matter - assume for what follows that $z\neq 0$.

Now what happens is that if n is a positive integer, $z^n = r^n e^{in\theta}$. If we want to find the n'th roots of z, then we take any n'th root and multiply it by all the n'th roots of 1 and we will have n total roots. It's easy to see that this works and that it gives all of the roots, it's just that if you then take it to the power of n you need the modulus and argument to match the original number.

If $z = re^{i\theta}$ then its roots are given by the formula

$$\sqrt[n]{r}e^{i\frac{\theta+2\pi k}{n}}$$

over integers k ranging from 0 to n-1.

Example. Let's find all cube roots of $1+i$. I stepped through a similar example more carefully before so I will just write that $1+i = \sqrt{2}e^{i\frac{\pi}{4}} = 2^{\frac{1}{2}}e^{i\frac{\pi}{4}}$. Therefore, one cube root will be $2^{\frac{1}{6}}e^{i\frac{\pi}{12}}$. But we need to multiply by $e^{i\frac{2\pi}{3}}$ and $e^{-i\frac{2\pi}{3}}$ to get 2 more roots. Therefore our 3 cube roots of $1+i$ (with the fractions in the powers simplified) are as follows:

$$2^{\frac{1}{6}}e^{i\frac{\pi}{12}}$$

$$2^{\frac{1}{6}}e^{i\frac{3\pi}{4}}$$

$$2^{\frac{1}{6}}e^{i\frac{-7\pi}{12}}$$

Note that the n'th roots of any non-zero complex number make a regular n-gon if you join them by edges, which is pretty cute, and A level further maths exam boards like to set problems around this fact so it's good to know it.

We know that the sum of all the roots of unity will be 0. This is not unexpected, and we will later see that since the roots of the unity are all roots of the polynomial $z^n - 1$, and the sum of the roots is given by the second leading coefficient which is 0, that proves the result.

2.4 Loci in the complex plane

Definition. A locus is a set of points.

There are several types of loci that come up in A level that you ought to know so I will go through them quickly. Here z is the non-fixed variable.

Example. The locus in the complex plane of $|z-a| = |z-b|$ is the set of points that are the same distance from a and b. Visually it is easy to see that this is just the line perpendicular to the line joining a and b passing through the point $\frac{a+b}{2}$. So for example, the locus of $|z-1| = |z+1|$ is just the imaginary axis.

Example. The locus of $\arg(z) = \frac{\pi}{3}$ is just the half-line containing all complex numbers of the form $re^{i\frac{\pi}{3}}$ where $r > 0$. Something like $\arg(z-a) =$ constant would be a similar half-line but starting at the point a instead of at the origin.

Example. The locus of $|z-a| = b$ is a circle of radius b centered at a.

Example. Let a and b be fixed complex numbers, then the locus of $\arg(z-b) = \arg(z-a)$ is essentially as follows:

It is the line through a and b minus the line segment joining a and b. This is because we need to be on the line in order for the angles to match, as for the angles to match we need $z-a$ and $z-b$ to be parallel, however if we are on the line segment *between* a and b we will have $\arg(z-b) = \arg(z-a) \pm \pi$, I hope this is sufficiently clear.

Example. Figure 3 shows the region of $\frac{\pi}{4} < \arg(z-(1+2i)) < \frac{\pi}{2}$

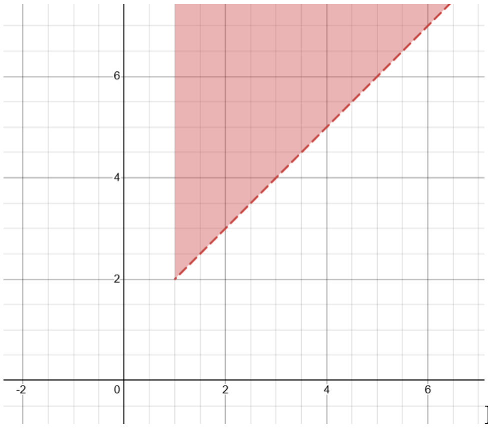

Figure 3

Example. Figure 4 shows the locus from figure 3 intersected with the region $2 < |z - (1+2i)| < 3$

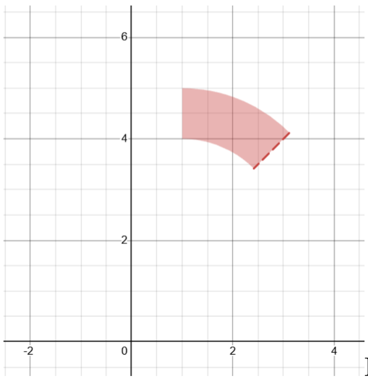

Figure 4

Those are the "simple" examples. Now I will show the 2 more tricky examples.

Example. I will find, step by step, the locus of points such that $|z-2| = 2|z+1|$. This means we need the point to be twice as far away from 2 as it is from -1 and we want to describe the set of all such points. What if I told you it will turn out to be a perfect circle!

Write $z = x + iy$. Then we have $|x + iy - 2| = 2|x + iy + 1|$. A general principle is that absolute values are nasty but squared absolute values are nice. We have, by Pythagoras:

$$|x - 2 + iy| = 2|x + 1 + iy|$$

$$(x-2)^2 + y^2 = 4\left((x+1)^2 + y^2\right)$$

$$x^2 - 4x + 4 + y^2 = 4\left(x^2 + 2x + 1 + y^2\right)$$

After a bit of algebra clean up,

$$3x^2 + 12x + 3y^2 = 0$$

$$x^2 + 4x + y^2 = 0$$

We know from level 3 how to deal with equations for a circle, and it is now clear that that is exactly what this is!

$$x^2 + 4x + 4 + y^2 = 4$$

$$(x+2)^2 + y^2 = 2^2$$

Thus we have a circle of radius 2 centered about the point $(-2, 0)$.

Example. Consider the locus of points such that $\arg\left(\frac{z-1}{z+1}\right) = \frac{\pi}{4}$.

Suppose for now that both have non-negative argument i.e. z has non-negative imaginary part. Then we can rewrite this as

$$\arg(z-1) - \arg(z+1) = \frac{\pi}{4}$$

by the identity that $\arg(ab) = \arg(a) + \arg(b)$ whenever a, b and ab all have non-negative imaginary part unless ab is positive real and a, b are both negative real (This is easy to see since the angle of the product of complex numbers is the sum of the angles, we impose the non-negative condition to not have issues where things differ by multiples of 2π), applied to $a = \frac{z-1}{z+1}$ and $b = z+1$. We will never run into this problem because the right hand side is $\frac{\pi}{4}$ - as long as it stays strictly between 0 and π we're fine.

Visually, we are saying that we need the lines in figure 5 to meet at an angle of $\frac{\pi}{4}$.

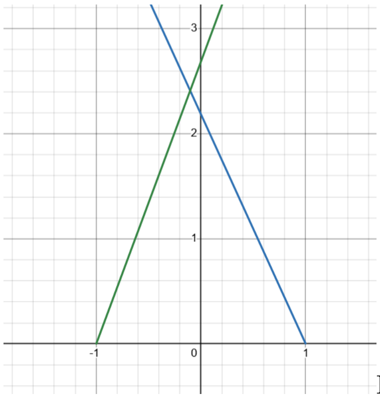

Figure 5

Now recall the circle theorems from level 2. We know that by the property shown in Figure 37 of level 2, that what we want is very similar to if the locus was a circle. What happens in this case is that we note that if $z = i$ then $\arg(z-1) - \arg(z+1) = \frac{\pi}{2}$, so we have double the angle we want. Now I know this is not satisfying, but we will guess that the locus is indeed a circle centered there, since certainly for all such points on the circle, by the "double angle" circle theorem from level 2, we are on the locus, as long as we are indeed on the same segment of the circle (in this case, above the real line). If the intersection is inside the circle we note that visually, the angle between the lines shown in figure 5 will be too large, and outside the circle they will be too small.

This is a very long winded way of saying the locus will be the part of the circle above $z = 0$. See figure 6:

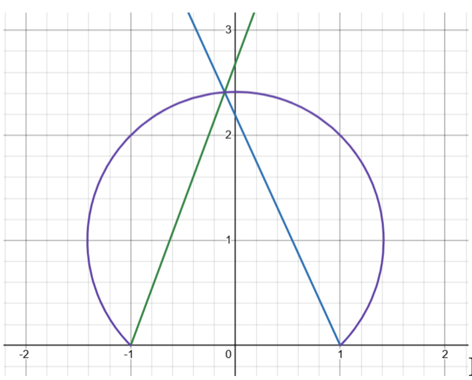

Figure 6

If we are below the real line, then the angle in question will be $3\frac{\pi}{4}$ since everything will effectively be flipped, as the theorems from level 2 only applied on the SAME segment of the circle, and now by the theorem the angle should be half of the angle at the center which is now 270 degrees instead of 90 degrees, see figure 7 for a visual of this:

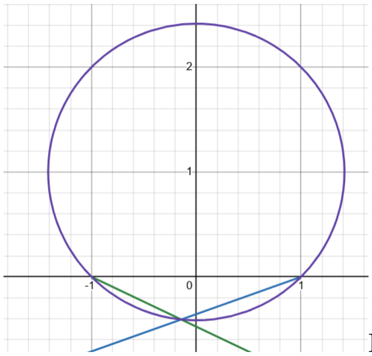

Figure 7

So indeed the locus is as in figure 6.

In general, a locus like $\arg\left(\frac{z-a}{z-b}\right) = c$ will be a circle through a and b which we can find as above. If $c < \frac{\pi}{2}$ it will be the segment with the majority of the circle. If $c > \frac{\pi}{2}$ it will be the minority. If $c = \frac{\pi}{2}$ it will be the half circle.

Note: The reason it is only a half circle even in that last case is that, for example, if we have $\arg\left(\frac{z-1}{z+1}\right) = \frac{\pi}{2}$, then the argument ratio is actually $-\frac{\pi}{2}$ at a point below the real line like $-i$.

2.5 Transformations in the complex plane

We can do simple transformations like scaling, rotating and moving the complex plane. I think I've shown enough similar things in levels 2 and 3 that you would get the idea, you know what it means for a circle to be scaled by 2x and moved 1 unit to the right or whatever. Lets do some less trivial transformations.

Example. We will prove that if we consider the circle $|z-1| = 1$, then the place it is sent to under the map $z \mapsto \frac{1}{z}$ will be a straight line.

We write $w = \frac{1}{z}$ so $z = \frac{1}{w}$. Therefore $z - 1 = \frac{1}{w} - 1 = \frac{1-w}{w}$. Thus we have $\frac{|1-w|}{|w|} = 1$ so $|1-w| = |w-1| = |w|$. From section 1 we know that the locus of this will be the line $Re(w) = \frac{1}{2}$.

We can do similar things in general.

Example. Consider the circle $|z| = 2$. We will do the transformation $w = \frac{2+z}{1-z}$ and find where the circle lands under this transformation.

In level 3 we gave an example of how to find the inverse of functions like this. With care, we find that $z = \frac{w-2}{w+1}$. Therefore we want to solve $\left|\frac{w-2}{w+1}\right| = 2$, which is equivalent to $|w-2| = 2|w+1|$. Conveniently, we found the locus of this exact thing in a previous example, it is the circle $|w+2| = 2$.

3 Algebra

3.1 Roots of polynomials

We can extend what we did with quadratics in level 3. To be honest this is probably just in the A level syllabus because it is easy to set questions on it. Basically, if we have a cubic polynomial with roots α, β, γ then it will be

$$(x-\alpha)(x-\beta)(x-\gamma) = x^3 - (\alpha+\beta+\gamma)x^2 + (\alpha\beta+\alpha\gamma+\beta\gamma)x - \alpha\beta\gamma$$

In general, what will happen is that the n'th coefficient after the leading one will have a - if n is odd and a + if n is even followed by the sum of all distinct products of n roots, this is just the result of expanding the brackets.

I'll do a quick example of how we might use this.

Example. The polynomial $x^3 + 5x^2 - 6x + 1 = 0$ has roots α, β, γ.

Note that by the expansion above we know

$$\alpha + \beta + \gamma = -5$$
$$\alpha\beta + \alpha\gamma + \beta\gamma = -6$$
$$\alpha\beta\gamma = -1$$

We will find a cubic polynomial with roots $A = \alpha + \beta$, $B = \beta + \gamma$ and $C = \gamma + \alpha$.

First, note that $A + B + C = 2(\alpha + \beta + \gamma) = -10$.

Now note that

$$AB + BC + CA = (\alpha + \beta)(\beta + \gamma) + (\beta + \gamma)(\gamma + \alpha) + (\gamma + \alpha)(\alpha + \beta)$$
$$= 3(\alpha\beta + \alpha\gamma + \beta\gamma) + \left(\alpha^2 + \beta^2 + \gamma^2\right) = 3(\alpha\beta + \alpha\gamma + \beta\gamma) + (\alpha + \beta + \gamma)^2 - 2(\alpha\beta + \alpha\gamma + \beta\gamma)$$
$$= (\alpha\beta + \alpha\gamma + \beta\gamma) + (\alpha + \beta + \gamma)^2 = (-6) + (-5)^2 = 19$$

Now note that

$$ABC = (\alpha + \beta)(\beta + \gamma)(\gamma + \alpha) = 2\alpha\beta\gamma + \alpha^2\beta + \alpha^2\gamma + \beta^2\alpha + \beta^2\gamma + \gamma^2\alpha + \gamma^2\beta$$

Because of the fact that

$$(\alpha + \beta + \gamma)(\alpha\beta + \alpha\gamma + \beta\gamma) = 3\alpha\beta\gamma + \alpha^2\beta + \alpha^2\gamma + \beta^2\alpha + \beta^2\gamma + \gamma^2\alpha + \gamma^2\beta$$

We have

$$ABC = (\alpha + \beta + \gamma)(\alpha\beta + \alpha\gamma + \beta\gamma) - \alpha\beta\gamma = 31$$

Therefore our polynomial is

$$x^3 - (A + B + C)x^2 + (AB + BC + CA)x - (ABC)$$
$$x^3 + 10x^2 + 19x - 31$$

3.2 Sums of polynomials

In section 1 we established by induction a formula for $\sum_{i=1}^{n} i^2$. There is also a formula for $\sum_{i=1}^{n} i^3$. Now here's a fun story.

When I was in year 3, I was going home from school on the bus and I tried adding up the cube numbers in my head. I got 1, 9, 36, 100, 225,... and I noticed something. They were all square numbers, and they were exactly the squares of triangle numbers. This made me very excited. It turns out that this is actually the formula. I mean the triangle numbers (1, 1+2, 1+2+3, 1+2+3+4, ...) or (1,3,6,10,15,...) which we saw in level 3 are given by $\frac{n(n+1)}{2}$.

The square of this is $\frac{n^2(n+1)^2}{4}$. We will check by induction that this works. We already checked the first few cases when I gave the numerical values, so I will just add $(n+1)^3$ to the expression to verify.

$$\frac{n^2(n+1)^2}{4} + (n+1)^3 = \frac{\left(n^2 + 4(n+1)\right)(n+1)^2}{4} = \frac{(n+2)^2(n+1)^2}{4}$$

This verifies it.

Now we can find the general term of sums like $\sum_{i=1}^{n} \left(i^3 + 7i^2 - 10i + 5\right)$ in the "obvious" way. It won't be nice, but we can just kind of add things together.

Example.

$$\sum_{i=1}^{n} i^3 - 2i^2 + i = \frac{n^2(n+1)^2}{4} - \frac{n(n+1)(2n+1)}{3} + \frac{n(n+1)}{2} = n(n+1)\left[\frac{n(n+1)}{4} - \frac{2n+1}{3} + \frac{1}{2}\right]$$

$$= n(n+1)\left[\frac{n^2}{4} - \frac{5n}{12} + \frac{1}{6}\right] = \frac{1}{12}n(n+1)\left[3n^2 - 5n + 2\right] = \frac{1}{12}n(n+1)(3n-2)(n-1)$$

Example. We can solve for unknowns like given that $\sum_{i=1}^{n} 3i^2 - 5i = 90$ find n. We find that

$$\sum_{i=1}^{n} 3i^2 - 5i = \frac{n(n+1)(2n+1)}{2} - \frac{5n(n+1)}{2} = n^3 - n^2 - 2n$$

Therefore we want to solve $n^3 - n^2 - 2n - 90 = 0$. This may be unsatisfying since we don't really have a general method to solve cubic equations like this but we spot that 5 is an integer solution, and when we divide by $n-5$ we get $n^2 + 4n + 18$ so there are no other roots. You get the idea that we can solve for unknowns.

3.3 The method of differences

The method of differences, or telescoping sums, works when we want to find a sum like $\sum_{i=1}^{n} f(i)$ where $f(i)$ can be written as $g(i) - g(i+1)$ or $g(i+1) - g(i)$. It helps to do an example.

Consider the sum $\sum_{i=1}^{n} \frac{1}{i(i+1)}$. We can decompose $\frac{1}{i(i+1)}$ into partial fractions and we get $\frac{1}{i} - \frac{1}{i+1}$. We therefore have that

$$\sum_{i=1}^{n} \frac{1}{i(i+1)} = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + ... + \left(\frac{1}{n} - \frac{1}{n+1}\right)$$

Now we notice something: We have $-\frac{1}{2}$ and $+\frac{1}{2}$. These cancel. We have $-\frac{1}{3}$ and $+\frac{1}{3}$. These also cancel. In the end we get, after cancellation,

$$\sum_{i=1}^{n} \frac{1}{i(i+1)} = 1 - \frac{1}{n+1}$$

This is the essence of the method. We'll do one more example

$$\sum_{i=1}^{n} \frac{1}{4i^2 - 1} = \sum_{i=1}^{n} \frac{1}{2}\left(\frac{1}{2i-1} - \frac{1}{2i+1}\right) = \frac{1}{2}\sum_{i=1}^{n}\left(\frac{1}{2i-1} - \frac{1}{2(i+1)-1}\right)$$

In general as we saw above if the thing in the sum is $g(i) - g(i+1)$ then as above the result is $g(1) - g(n+1)$. In this case it works if $g(i) = \frac{1}{2i-1}$ so the sum is

$$\frac{1}{2}\left(1 - \frac{1}{2n+1}\right) = \frac{n}{2n+1}$$

3.4 Inequalities

It is possible to solve certain inequalities. I'll show by example.

Example. Suppose we want to solve the inequality $\frac{1}{x-3} < \frac{x}{x+5}$.

One way to do it is to get rid of the fractions. Normally we would just do this by multiplying both sides by the denominators. However, we have to be a bit careful as inequalities are only preserved if we multiply by something positive. We are always fine if we multiply by something squared, so we will just multiply in this case by the denominators squared. We now want to solve

$$\frac{1}{x-3}(x-3)^2(x+5)^2 < \frac{x}{x+5}(x-3)^2(x+5)^2$$

This is almost equivalent to the original inequality except at the points 3 and -5 where we would have multiplied both sides by 0 so a strict inequality would turn into an equality. However note that our original functions were undefined at those points anyway. Tidying up, the thing we want to solve is

$$(x-3)(x+5)^2 < x(x-3)^2(x+5)$$

This is easier if we move everything to one side:

$$(x-3)(x+5)^2 - x(x-3)^2(x+5) < 0$$

We see some common factors. We want to solve

$$(x-3)(x+5)(x+5-x(x-3)) < 0$$

$$(x-3)(x+5)\left(-x^2+4x+5\right) < 0$$

$$(x-3)(x+5)(1-x)(x-5) < 0$$

This we can do. We know this will be a graph that crosses the origin exactly at the points 3, -5, 1, 5. To determine if the inequality holds between these critical points, we just check any of them. For example between 1 and 5 we try plugging in $x=2$ and we see that the left hand side of the inequality is <0 so the whole interval from 1 to 3 must also be a solution since the graph is continuous and does not pass the origin anywhere there. Similarly we see that $x>5$ and $x<-5$ is also a solution. and in the other intervals we do not have a solution. As for at the critical points, we already noted that at two of them the functions aren't even defined, and at the other 2 we see that in this case we will have an exact equality. All this means the solution set is exactly

$$\{1<x<3\} \cup \{x>5\} \cup \{x<-5\}$$

Turning it into a polynomial is a reliable way to solve this that often works. Another way to solve it is to check a point between critical points where the critical points are where the graphs either cross eachother or are not continuous/defined. We could do this in the above example: We get -5 and 3 as critical points for free as the functions we started with are not defined there, and then to find the other critical points we could solve $\frac{1}{x-3} = \frac{x}{x+5}$. Since this is an equality now, we can just multiply by the common denominators - no squaring needed. We get $x+5 = x(x-3)$ which simplifies to $x^2-4x-5=0$. The roots of this equation are -1 and 5 which are therefore our other 2 critical points.

Example. We can solve inequalities involving absolute values such as $|x^2-1| < 2(x+1)$. The process is roughly as follows:

- Find where x^2-1 is positive or 0 and where it is negative.

- We solve, only considering points in the places where $x^2-1 \geq 0$, the inequality $x^2-1 < 2(x+1)$.

- In the places where $x^2-1<0$ when we take the absolute value we get $1-x^2$ so we would need to solve the inequality $1-x^2 < 2(x+1)$ only considering solutions in places where $x^2-1<0$ (In this case it is when $-1<x<1$).

I don't feel like doing it but we need it for A level further maths so I put it here. It's probably just in the syllabus since it's easy to set questions on it, I don't think mathematicians solve absolute value inequalities like ever as far as I know.

4 Calculus

4.1 Limits and l'hopital's rule

Suppose we want to calculate some limit, such as

$$\lim_{x\to 0} \frac{e^x-1}{x\cos(x)}$$

L'hopital's rule works under the following condition:

If we have $\lim_{x\to a} \frac{f(x)}{g(x)}$ but $f(a)=g(a)=0$, and in an interval around a we have $g(x)\neq 0$ and $f(x)$ and $g(x)$ have continuous derivatives, then the limit is equal to $\frac{f'(a)}{g'(a)}$, whenever that is actually defined.

We will not prove this here, but in level 6.1 I give a nice visual argument and show that this theorem implies the other cases we will show shortly, and in level 6.2 I give a formal proof.

Thus, in the example above, the limit would be

$$\frac{e^x}{\cos(x) - x\sin(x)} \text{ evaluated at } x = 0$$

which is just 1.

When $\frac{f'(a)}{g'(a)}$ is not defined we instead consider $\frac{f'(x)}{g'(x)}$ approaching a limit as $x \to a$ and what that limit is. We will prove this version in level 6.2. It is then possible to apply the rule again if necessary and then we get $\frac{f''(a)}{g''(a)}$.

Some limits are easier, for example if we have polynomial divided by polynomial and they share a root you can effectively simplify it: $\lim_{x\to 3} \frac{x^2-9}{x-3}$ is just 6 since $\frac{x^2-9}{x-3}$ is exactly $x+3$ away from 3.

Now we will do a more interesting limit. We will find $\lim_{x\to 0^+} x\ln(x)$ which is $\lim_{x\to 0^+} \frac{\ln(x)}{1/x}$. Note that here we only consider going to 0 from the right because ln is not defined for negative values for our purposes. This is not a $0/0$ form like I stated, this is an ∞/∞ form. However, in the level 6.1 video on the topic I prove that if we can prove the rule in the $0/0$ case we have it in all the cases. Note that the limit may also be as $x \to \infty$ but I deal with these cases in level 6 as well. The analogous situation to "an interval around ∞" is the unbounded interval of all numbers greater than something. In the ∞/∞ case, in addition not wanting the denominator to be non-zero near the point, we want the numerator to be non-infinity in some interval around the point.

Therefore, we can apply the rule to get

$$\lim_{x\to 0^+} x\ln(x) = \lim_{x\to 0^+} \frac{\ln(x)}{1/x} = \lim_{x\to 0^+} \frac{1/x}{-1/x^2} = \lim_{x\to 0^+} -x = 0$$

Now, since e^x is continuous and we have $x^x = e^{x\ln(x)}$, it follows that $x^x \to 1$ as $x \to 0$.

Example. Let's find $\lim_{x\to\infty} \frac{x^4}{e^x}$. Since we have an $\frac{\infty}{\infty}$ form and the conditions for the theorem hold, we can instead find $\lim_{x\to\infty} \frac{4x^3}{e^x}$ by differentiating the top and bottom. We repeat - we can differentiate 3 more times and then we have $\lim_{x\to\infty} \frac{24}{e^x}$ in the end, which is clearly 0.

Since this proof works with "4" replaced with any finite number, the slogan to remember is "exponentials beat polynomials".

4.2 Hyperbolic trigonometric functions

Definition. $\sinh(x) = \frac{e^x - e^{-x}}{2}$ and $\cosh(x) = \frac{e^x + e^{-x}}{2}$.

Figure 7 shows the graph of sinh.

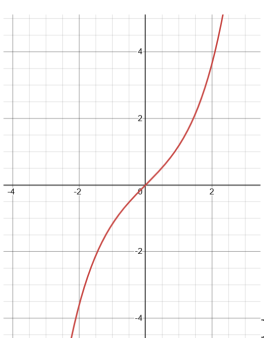

Figure 7

Figure 8 shows the graph of cosh.

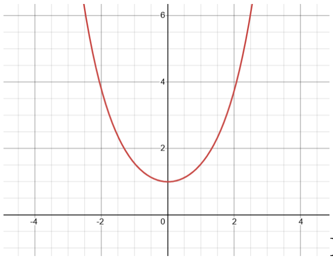

Figure 8

It is easy to see that these will be what the graphs look like if we take the average of e^x and $\pm e^{-x}$, which is literally what we are doing. Note that cosh is ≥ 1 everywhere since e^x curves upwards and is increasing so by visualizing it you can see that it will deviate from 1 by being above at $+x$ more than it will deviate from 1 by being below at $-x$.

By standard differentiation rules we see that $\frac{d}{dx}\sinh(x) = \cosh(x)$ and $\frac{d}{dx}\cosh(x) = \sinh(x)$.

We define $\tanh(x) = \frac{\sinh(x)}{\cosh(x)}$, $\operatorname{sech}(x) = \frac{1}{\cosh(x)}$, $\operatorname{csch}(x) = \frac{1}{\sinh(x)}$, $\coth(x) = \frac{\cosh(x)}{\sinh(x)}$. We can also differentiate these using standard differentiation rules.

Ok you may be thinking what are we doing and why. We will see soon how these are connected to the familiar trigonometric functions.

There are several identities analogous to trigonometric identities that are easy to prove directly from the definition. I give a few examples of such identities, the rest are proved similarly. You can even prove them from the "connection to trig" that we will see soon if you wish.

Proposition. $\cosh^2(x) - \sinh^2(x) = 1$.

Proof.

$$\cosh^2(x) - \sinh^2(x) = \left(\frac{e^x + e^{-x}}{2}\right)^2 - \left(\frac{e^x - e^{-x}}{2}\right)^2 = \frac{e^{2x} + e^{-2x} + 2}{4} - \frac{e^{2x} + e^{-2x} - 2}{4} = \frac{2}{4} - \frac{-2}{4} = 1$$

□

Proposition. $\sinh(a+b) = \sinh(a)\cosh(b) + \cosh(a)\sinh(b)$.

Proof.

$$\sinh(a)\cosh(b) + \cosh(a)\sinh(b)$$
$$= \frac{1}{4}\left(e^a - e^{-a}\right)\left(e^b + e^{-b}\right) + \frac{1}{4}\left(e^a + e^{-a}\right)\left(e^b - e^{-b}\right)$$
$$= \frac{1}{4}\left(e^{a+b} - e^{-(a+b)} - 2e^{b-a} + 2e^{a-b}\right) + \frac{1}{4}\left(e^{a+b} - e^{-(a+b)} + 2e^{b-a} - 2e^{a-b}\right)$$
$$= \frac{1}{2}\left(e^{a+b} - e^{-(a+b)}\right) = \sinh(a+b)$$

□

Proposition. $\tanh^2(x) + \operatorname{sech}^2(x) = 1$.

Proof. $\cosh^2(x) - \sinh^2(x) = 1$ by 2 propositions ago. Therefore

$$\cosh^2(x) = 1 + \sinh^2(x)$$

Dividing both sides by $\cosh^2(x)$ gives

$$1 = \frac{1}{\cosh^2(x)} + \frac{\sinh^2(x)}{\cosh^2(x)}$$

$$= \operatorname{sech}^2(x) + \tanh^2(x)$$

□

We can have inverses for these. The inverse for sinh is arsinh. The inverse for cosh (restricted to positive reals only since cosh is not one-to-one) is arcosh and the inverse for tanh is artanh.

Proposition. $\operatorname{arsinh}(x) = \ln\left(x + \sqrt{x^2+1}\right)$

Proof. Let $y = \operatorname{arsinh}(x)$ so $x = \sinh(y)$. Then the good news is that $\operatorname{arsinh}(x)$ has a proper inverse since it is a one-to-one function.

$$x = \frac{e^y - e^{-y}}{2}$$

By multiplying by $2e^y$ and rearranging

$$e^{2y} - 2xe^y - 1 = 0$$

By the quadratic formula

$$e^y = \frac{2x \pm 2\sqrt{x^2+1}}{2} = x \pm \sqrt{x^2+1}$$

Note that $\sqrt{x^2+1} > x$ so to have a real valued thing we will always pick the $+$ sign. Thus, $e^y = x + \sqrt{x^2+1}$, so done by taking logs.

□

Since arcosh is not a one-to-one function, we need to be more careful about its inverse. We will pick the positive value of a such that $\cosh(a) = x$ to be $\operatorname{arcosh}(x)$.

Proposition. $\operatorname{arcosh}(x) = \ln\left(x + \sqrt{x^2-1}\right)$

Proof. Let $y = \operatorname{arcosh}(x)$ so $x = \cosh(y)$.

$$x = \frac{e^y + e^{-y}}{2}$$

By multiplying by $2e^y$ and rearranging

$$e^{2y} - 2xe^y + 1 = 0$$

By the quadratic formula

$$e^y = \frac{2x \pm 2\sqrt{x^2-1}}{2} = x \pm \sqrt{x^2-1}$$

Here is where we have to be careful. Note that if $x > 1$ then

$$\left(x + \sqrt{x^2-1}\right)\left(x - \sqrt{x^2-1}\right) = 1$$

and

$$x + \sqrt{x^2-1} > x - \sqrt{x^2-1}$$

Thus it follows that $x + \sqrt{x^2-1}$ is the unique one of these roots that is > 1, and thus the only one with a positive logarithm. Thus, taking logs of both sides we get

$$\operatorname{arcosh}(x) = \ln\left(x + \sqrt{x^2-1}\right)$$

□

Now we will investigate tanh. Note that $\tanh = \frac{\sinh}{\cosh} = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x}-1}{e^{2x}+1}$. Lets note a few properties:

- tanh has derivative $\frac{\cosh^2 - \sinh^2}{\cosh^2} = \operatorname{sech}^2$ which is always positive, thus it is increasing.

- $\tanh(0) = 0$

- $\tanh(-x) = \frac{e^{-2x}-1}{e^{-2x}+1} = -\frac{e^{2x}}{e^{2x}}\frac{1-e^{-2x}}{e^{-2x}+1} = -\frac{e^{2x}-1}{e^{2x}+1} = -\tanh(x)$

- As $x \to \infty$ the numerator and denominator of tanh is dominated by the exponential term, thus tanh is approximately 1.

Given all this, figure 9 shows the graph of tanh.

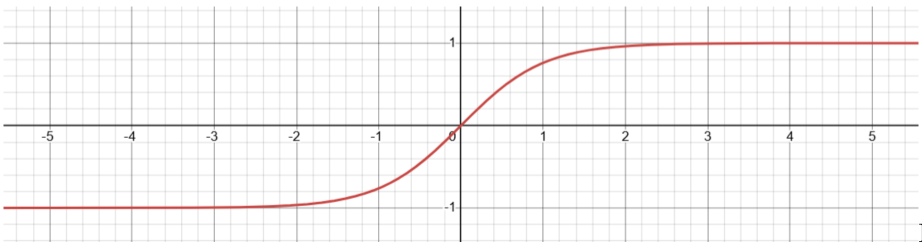

Figure 9

We will now investigate the inverse of tanh which is artanh. We have

$$y = \text{artanh}(x)$$

$$x = \frac{e^{2y}-1}{e^{2y}+1}$$

$$xe^{2y} + x = e^{2y} - 1$$

$$(x-1)e^{2y} = -x - 1$$

$$e^{2y} = \frac{1+x}{1-x}$$

$$y = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right)$$

So that is a formula for artanh.

Note that using the log formulas and differentiating we get

$$\frac{d}{dx}\text{arsinh}(x) = \frac{1}{\sqrt{1+x^2}}$$

I'll show this one: If y is arsinh then $\frac{dx}{dy} = \frac{d(\sinh(y))}{dy} = \cosh(y) = \sqrt{1+x^2}$ so $\frac{dy}{dx} = \frac{1}{\sqrt{1+x^2}}$

We can similarly get

$$\frac{d}{dx}\text{arcosh}(x) = \frac{1}{\sqrt{x^2-1}}$$

$$\frac{d}{dx}\text{artanh}(x) = \frac{1}{1-x^2}$$

This is useful for integration.

Now it is finally time I explain how these are connected to the trigonometric functions.

Note that

$$\frac{e^{ix}+e^{-ix}}{2} = \frac{\cos(x) + i\sin(x) + \cos(-x) + i\sin(-x)}{2}$$

Using familiar symmetry properties of sin and cos we see this is just $\cos(x)$.

So effectively $\cos(x) = \cosh(ix)$ and $\cosh(x) = \cos(ix)$. Although we have not explicitly defined cos for complex numbers, the $\frac{e^{ix}+e^{-ix}}{2}$ formula is exactly how this is done.

We also can define $\sin(x) = \frac{\cos(x)-\cos(-x)+i\sin(x)-i\sin(-x)}{2i} = \frac{e^{ix}-e^{-ix}}{2i}$. Since $\frac{1}{i} = -i$ this means that $\sin(x) = -i\sinh(ix)$ and $-i\sin(ix) = \sinh(x)$. The way to remember how these identities go is to make sure it is consistent with sinh and sin both being approximately equal to x near 0.

4.3 Trigonometric and hyperbolic substitutions

The trigonometric substitution is as follows. If we have something like $\sqrt{a^2-x^2}$ then $x = a\sin(u)$ is often a convenient substitution to make because (as long as $\cos(u) > 0$), $\sqrt{a^2-x^2}$ becomes $\sqrt{a^2\left(1-\sin^2(u)\right)} = a\cos(u)$.

Example. We will calculate $\int \sqrt{1-x^2}dx$ in 2 ways (for x ranging from 0 to 1).

The first way is by a geometric picture. Note that $y = \sqrt{1-x^2}$ is the equation for a semicircle, so the area under the circle from 0 to T is the sum of the area of the triangle and the arc shown in figure 10.

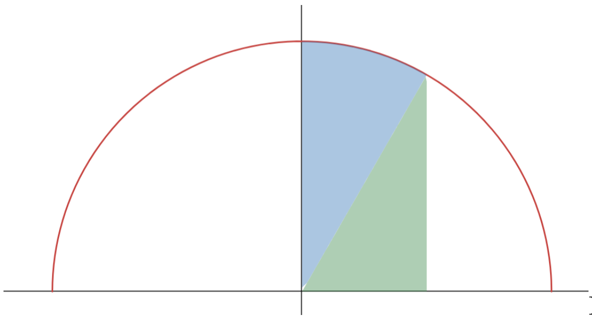

Figure 10

The area of the arc is $\frac{1}{2}$ times the angle in radians, and this angle is exactly $\arcsin(T)$ by considering the fact that it is the same as the angle at the top of the green triangle, the sin of which is opposite over hypotenuse, which is T. Thus the area of the arc is $\frac{1}{2}\arcsin(T)$.

The area of the green triangle is half the base times the height, which is $\frac{1}{2}T\sqrt{1-T^2}$. Therefore the total area is

$$\frac{1}{2}\left(\arcsin(T) + T\sqrt{1-T^2}\right)$$

So this is an antiderivative of $y = \sqrt{1-x^2}$.

Now I will calculate it using the substitution $x = \sin(u)$ where u can range from 0 to $\frac{\pi}{2}$ and x ranges from 0 to 1.

We get

$$\int \sqrt{1-x^2}dx = \int \cos(u)\,dx = \int \cos(u)\frac{dx}{du}du = \int \cos^2(u)\,du = \int \frac{1+\cos(2u)}{2}du$$
$$= \frac{1}{2}u + \frac{\sin(2u)}{4} + c = \frac{1}{2}\arcsin(x) + \frac{\sin(u)\cos(u)}{2} = \frac{1}{2}\arcsin(x) + \frac{1}{2}x\sqrt{1-x^2}$$

exactly as we expected.

Note that we can solve something like $\frac{1}{\sqrt{a^2-x^2}}$ using a similar substitution but this is unnecessary since we already know that it looks like the antiderivative of arcsin: A substitution $x = au$ would turn $\frac{1}{\sqrt{a^2-x^2}}dx$ into $\frac{a}{\sqrt{a^2-a^2u^2}}du = \frac{1}{\sqrt{1-u^2}}du$ which integrates to arcsin(u) since arcsin is a known antiderivative of this function (level 3).

Here is an example of a hyperbolic substitution:

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

Put $x = \frac{1}{2}\sinh(u)$ so we have

$$= 2\int_{x=0}^{1} \frac{1}{2}\sqrt{\sinh^2(u)+1}dx = \int_{u=0}^{\text{arsinh}(2)} \cosh(u)\,dx$$

Now $\frac{dx}{du} = \frac{1}{2}\cosh(u)$ so

$$= \frac{1}{2}\int_0^{\text{arsinh}(2)} \cosh^2(u)\,du = \int_0^{\text{arsinh}(2)} \frac{e^{2u}}{8} + \frac{1}{4} + \frac{e^{-2u}}{8}du$$

Now $\text{arsinh}(2) = \ln\left(2+\sqrt{1+2^2}\right) = \ln\left(2+\sqrt{5}\right)$. Thus we have

$$\int_0^{\ln\left(2+\sqrt{5}\right)} \frac{e^{2u}}{8} + \frac{1}{4} + \frac{e^{-2u}}{8}du = \left[\frac{e^{2u}}{16} + \frac{1}{4}u - \frac{e^{-2u}}{16}\right]_0^{\ln\left(2+\sqrt{5}\right)}$$

$$= \frac{e^{2\ln\left(2+\sqrt{5}\right)}}{16} + \frac{1}{4}\ln\left(2+\sqrt{5}\right) - \frac{e^{-2\ln\left(2+\sqrt{5}\right)}}{16} - \left[\frac{e^0}{16} + \frac{1}{4}0 - \frac{e^{-0}}{16}\right]$$

$$= \frac{1}{4}\ln\left(2+\sqrt{5}\right) + \frac{1}{8}\sinh\left(2\ln\left(2+\sqrt{5}\right)\right)$$

By the sinh addition formula this is just

$$= \frac{1}{4}\ln\left(2+\sqrt{5}\right) + \frac{1}{4}\sinh\left(\ln\left(2+\sqrt{5}\right)\right)\cosh\left(\ln\left(2+\sqrt{5}\right)\right)$$

Since $\ln\left(2+\sqrt{5}\right)$ is arsinh(2) it follows that $\sinh\left(\ln\left(2+\sqrt{5}\right)\right) = 2$ and $\cosh\left(\ln\left(2+\sqrt{5}\right)\right) = \sqrt{1+2^2} = \sqrt{5}$. Thus the final answer is

$$= \frac{1}{2}\left(\sqrt{5} + \frac{1}{2}\ln\left(2+\sqrt{5}\right)\right)$$

4.4 Polar coordinates

Definition. Polar coordinates are a system where θ is the angle like the argument of a complex number and r is the distance from the origin, and you can have curves defined by $r = f(\theta)$. However, note that we can sometimes have θ outside the typical $-\pi$ to π range and allow for multivalued functions. To show what I mean, figure 11 shows the graph of $r = \theta$.

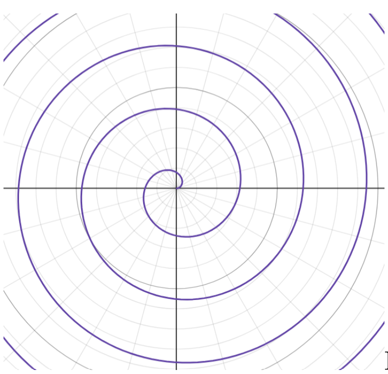

Figure 11

Note that $x = r\cos(\theta)$ and $y = r\sin(\theta)$. In addition, we have $\tan(\theta) = \frac{y}{x}$ and $r = \sqrt{x^2+y^2}$.

This way, we can turn a polar curve into a curve defined parametrically by θ. We already know for parametric curves that they will be vertical when $\frac{dx}{d\theta} = 0$, $\frac{dy}{d\theta} \neq 0$ and horizontal when $\frac{dy}{d\theta} = 0$, $\frac{dx}{d\theta} \neq 0$. I won't go through an example specifically for polar coordinates because hopefully you get the idea by now.

We can find the area enclosed by a polar curve as follows. Figure 12 shows a polar curve with the region it encloses when $a < \theta < b$ shaded in red.

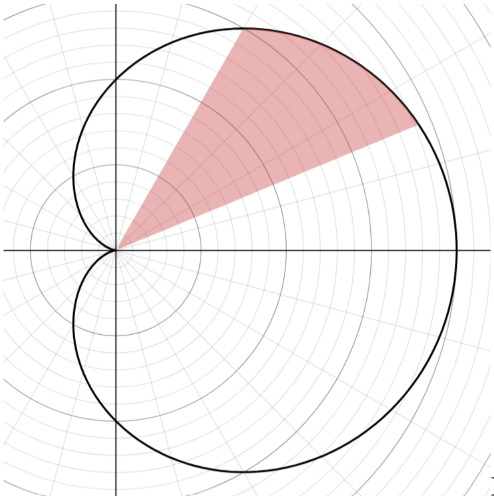

Figure 12

Now a formula which is not hard to prove visually but that we will prove in level 6.1 says that this area is equal to $\int_{\theta=a}^{b} \frac{1}{2} r^2 d\theta$.

The shape in Figure 12 is called a cardioid and has equation $r = 1 + \cos(\theta)$ and I will find the area of the whole thing, which is equal to

$$\int_{\theta=0}^{2\pi} \frac{1}{2} (1 + \cos(\theta))^2 \, d\theta$$

$$= \int_{\theta=0}^{2\pi} \frac{1}{2} + \cos(\theta) + \frac{1}{2} \cos^2(\theta) \, d\theta$$

$$= \int_{\theta=0}^{2\pi} \frac{1}{2} + \cos(\theta) + \frac{1}{2} \frac{1 + \cos(2\theta)}{2} d\theta$$

Note that $\cos(\theta)$ and $\cos(2\theta)$ both integrate to 0 between 0 and 2π since the up and down areas of their graphs cancel out. Therefore we just have to find

$$= \int_{\theta=0}^{2\pi} \frac{1}{2} + \frac{1}{2} \frac{1}{2} d\theta = \int_{\theta=0}^{2\pi} \frac{3}{4} d\theta = \frac{3}{2}\pi$$

So yeah that's how you do that.

4.5 Finding arc lengths

Let $f(x)$ be a function defined between 2 endpoints a and b such that its derivative is continuous. Then this function is some curve, and its length is given by the following formula, which I can't prove since we are actually *defining* length this way. Before I give the formula I will motivate it with an INTUITION.

Let δx be a small displacement. Then as in Figure 13, we want the length of the curve from x to $x+\delta x$ to be approximately the length of the following tangent line:

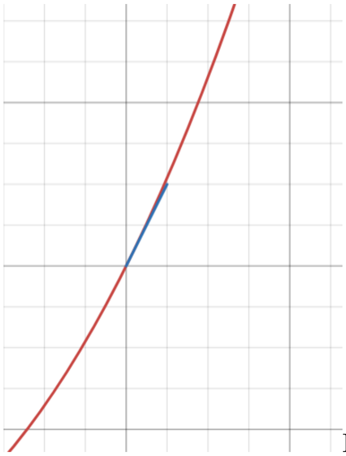
Figure 13

By Pythagoras, the length of this tangent line is exactly $\delta x\sqrt{1+\left(\frac{dy}{dx}\right)^2}$. Therefore the total length of f from a to b is defined to be

$$\int_a^b \sqrt{1+\left(\frac{dy}{dx}\right)^2}dx$$

Example. Let's find the arc length of the parabola $y = x^2$ from $x = 0$ to $x = 1$. This will therefore be

$$\int_0^1 \sqrt{1+(2x)^2}dx = \int_0^1 \sqrt{1+4x^2}dx$$

Very conveniently and totally not intentionally to reduce the work for me, this is exactly an example we did in section 4.3, and the answer came out to be

$$= \frac{1}{2}\left(\sqrt{5}+\frac{1}{2}\ln\left(2+\sqrt{5}\right)\right)$$

which is about 1.4789 units. So that is how long that piece of parabola would be if it was like a rope and you pulled it.

If the curve is defined parametrically with $x(t)$ and $y(t)$ then provided they have continuous derivatives with respect to t everywhere from a to b including at the end points we define the length of this curve to be $\int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2}dt$.

Note that it is not immediately obvious that this is well defined: We would have to prove that if we define the same curve parametrizing it in 2 different ways we will get the same result. This only works if both parametrizations as function of t are a one-to-one continuously differentiable map including at the endpoints: See level 6.1 for a proof.

Example. Consider a circle defined by $x(t) = \cos(t)$ and $y(t) = \sin(t)$, then its length will be

$$\int_0^{2\pi} \sqrt{\left(\frac{d\cos(t)}{dt}\right)^2+\left(\frac{d\sin(t)}{dt}\right)^2}dt = \int_0^{2\pi} \sqrt{(-\sin(t))^2+(\cos(t))^2}dt$$

$$= \int_0^{2\pi} 1dt = 2\pi$$

It is nice to see evidence that this definition is sensible.

4.6 Finding surface areas of revolution

I will give a formula that right now I'm saying is the definition, and I'll give intuition for why it is sensible. In level 8.2 when we define surface area more generally I will show that it agrees with this definition.

Let $f:[a,b]\to\mathbb{R}$ be such that $f\geq 0$ everywhere and has a continuous derivative, then the surface area of revolution of $f:[a,b]\to\mathbb{R}$ is given by the formula $\int_a^b 2\pi f(x)\sqrt{1+(f'(x))^2}dx$

The intuition for this is shown in figure 14:

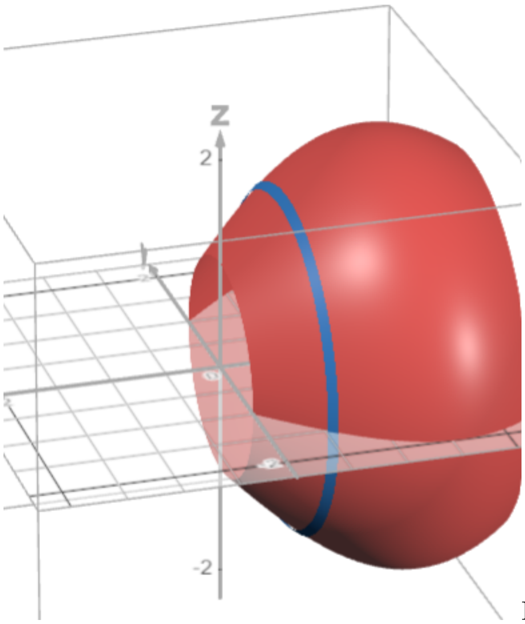

Figure 14

If we have a function at x then we can approximate the surface area of revolution between x and $x+\delta x$ by a part of a cone which has surface area approximately equal to $2\pi f(x)\sqrt{1+(f'(x))^2}\delta x$ - Ie, equal to the small width times the circumference times what the “arc length” element would be.

Example. the surface area of a sphere would correspond to f being $\sqrt{1-x^2}$ on $[-1,1]$ and here f' is $\frac{-x}{\sqrt{1-x^2}}$. Thus we have

$$2\pi\int_{-1}^{1}\sqrt{1-x^2}\sqrt{1+\frac{x^2}{1-x^2}}dx=2\pi\int_{-1}^{1}\sqrt{1-x^2+x^2}dx=4\pi$$

so that is equal to the surface area of a sphere.

4.7 Weierstrass substitution

Let $t=\tan\left(\frac{x}{2}\right)$. Then by the double angle formula we have $\tan(x)=\frac{2t}{1-t^2}$ for $t\neq\pm 1$.

Using t we can get a formula for $\cos(x)$ in terms of t: By the double angle formula,

$$\cos(x)=2\cos^2\left(\frac{x}{2}\right)-1=\frac{2}{\sec^2\left(\frac{x}{2}\right)}-1=\frac{2}{1+t^2}-1=\frac{1-t^2}{1+t^2}$$

Also,

$$\sin(x)=\pm\sqrt{1-\left(\frac{1-t^2}{1+t^2}\right)^2}=\pm\sqrt{\frac{4t^2}{(1+t^2)^2}}=\pm\frac{2t}{1+t^2}$$

Luckily by considering the graphs we see that the signs of t and $\sin(x)$ always coincide. Thus, everywhere t is defined the formula for $\sin(x)$ is just $\frac{2t}{1+t^2}$ which has the same sign as t.

We also get by taking the reciprocal that $\sec(x)=\frac{1+t^2}{1-t^2}$, $\csc(x)=\frac{1+t^2}{2t}$, $\cot(x)=\frac{1-t^2}{2t}$.

Example. It follows from the csc and cot formulae that $\csc(x)-\cot(x)=\tan\left(\frac{x}{2}\right)$.

Definition. The Weierstrass substitution, which can often simplify integrals, is to set $t = \tan\left(\frac{x}{2}\right)$. Note that $\frac{dt}{dx} = \frac{1}{2}\sec^2\left(\frac{x}{2}\right) = \frac{1}{2}\left(1+t^2\right)$, so we can replace dx by $\frac{2}{1+t^2}dt$.

Example. Lets use it to solve $\int \sec(x)\,dx$ (between 0 and $\frac{\pi}{2}$). Then this is

$$\int \frac{1+t^2}{1-t^2}dx = \int \frac{2}{1-t^2}dt = \int \frac{1}{1-t} + \frac{1}{1+t}dt = -\ln(1-t) + \ln(1+t) + c = \ln\left(\frac{(1+t)^2}{1-t^2}\right) + c = \ln(\sec(x) + \tan(x)) + c$$

4.8 Improper integrals

Under standard definitions of Riemann integrals we want integrals to be of bounded functions on bounded domains. We say an integral is improper if we drop one of these conditions. In this case, we have to take a limit.

Example.

$$\int_0^\infty e^{-x}dx = \lim_{M\to\infty}\int_0^M e^{-x}dx = \lim_{M\to\infty} 1 - e^{-M} = 1$$

The limit might not exist, for example

$$\lim_{M\to\infty}\int_0^M \cos(x)\,dx = \lim_{M\to\infty}\sin(M) = \text{does not exist}$$

Example.

$$\int_0^\infty \frac{1}{1+x^2}dx = \lim_{M\to\infty}\int_0^M \frac{1}{1+x^2}dx = \lim_{M\to\infty}\arctan(M) - \arctan(0) = \frac{\pi}{2}$$

Example.

$$\int_0^1 \ln(x)\,dx = \lim_{h\to 0}\left((1\ln(1) - 1) - (h\ln(h) - h)\right)$$

Since we showed earlier that $h\ln(h) \to 0$ the result is -1, which makes sense since it is - the area from $\int_0^\infty e^{-x}dx$, just rotated.

Note that we CANNOT say that $\int_{-\infty}^{\infty} f(x)\,dx = \lim_{M\to\infty}\int_{-M}^{M} f(x)\,dx$, we need to do one limit at a time. The correct way would be

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{M\to\infty}\int_0^M f(x)\,dx + \lim_{M_2\to\infty}\int_{-M_2}^0 f(x)\,dx$$

The "wrong" definition can fail because, for example, it need not be the case that $\lim_{M\to\infty}\int_{-M}^{M} f(x)\,dx = \lim_{M\to\infty}\int_{-M}^{2M} f(x)\,dx$.

In general when doing improper integrals we split the domain so we are only doing one limit at a time.

4.9 Use of Taylor series

See level 4 for the definition and derivation of a Taylor series about a point, and a proof that they are differentiable and therefore continuous.

Here are the standard ones:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + ...$$

The last one is only valid when $|x| < 1$ but the rest are valid everywhere - in levels 4 and 6 there is plenty of discussion of why. Here, though, I will derive the one for ln. Note that the derivative of ln is $\frac{1}{x}$, the second derivative is $-\frac{1}{x^2}$, the third is $\frac{2}{x^3}$, the fourth is $-\frac{2*3}{x^4}$. In general the n'th derivative is $-(-1)^n \frac{(n-1)!}{x^n}$ (note that $(-1)^n$ is -1 when n is odd and 1 when n is even), which at $x = 1$ is just $-(-1)^n (n-1)!$. We then divide by $n!$ to get the n'th Taylor coefficient, and $\frac{-(-1)^n(n-1)!}{n!} = \frac{-(-1)^n}{n}$ for $n \geq 1$, and for $n = 0$ we have $\ln(1+0) = 0$.

By using this, it is possible to find certain limits, and in fact it is clear why it is equivalent to L'hopital's rule.

Example. Consider $\frac{\ln(1+x)+\sin(x)}{e^x-1}$. Then the Taylor series of the numerator and the denominator is $\frac{x-\frac{x^2}{2}+...+x-\frac{x^3}{3}+...}{1+x+\frac{x^2}{2}+...-1}$. Considering only the terms in lowest powers of x will give $\frac{2x+\text{Terms in x^2 and higher}}{x+\text{Terms in x^2 and higher}}$ which is equal to $\frac{2+\text{Terms in x and higher}}{1+\text{Terms in x and higher}}$ whenever is not 0, which approaches 2 as $x \to 0$ as power series are continuous, so this is how we find the limit using Taylor series.

Example. We can cleverly use Taylor series to find $\sum_{n=1}^{\infty} \frac{1}{n*2^n}$. To do this note that

$$\sum_{n=1}^{\infty} \frac{1}{n*2^n} = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n \frac{1}{n} = \sum_{n=1}^{\infty} (-1)^n \left(-\frac{1}{2}\right)^n \frac{1}{n} = -\ln\left(1-\frac{1}{2}\right) = \ln(2)$$

Definition. A Taylor series is called a Maclaurin series if we take it about $x = 0$.

Note that we can find the Taylor series of composite functions as follows (see level 6.1 for proof of why this works). Let f and g have valid Taylor series such that the Taylor series of f about $g(x_0)$ is

$$f = a_0 + a_1(x - g(x_0)) + a_2(x - g(x_0))^2 + ...$$

Then the Taylor series of $f(g(x))$ about x_0 is as follows:

$$f = a_0 + a_1(g(x) - g(x_0)) + a_2(g(x) - g(x_0))^2 + ...$$

But we need to find Taylor series for each of these.

This is generally a bad strategy since we would have to find the Taylor series of a squared function, which is possible (we will see a trick to find derivatives of products of functions in section 4.11), but the following examples are easier

Example. Let's find a Taylor series for $e^{\sin(x)}$ about $x = 0$.

$$e^x * e^{\frac{-x^3}{6}} * e^{\frac{x^5}{120}} * ...$$
$$= \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + ...\right) * \left(1 - \frac{x^3}{6} + ...\right) * (1 + ...) * ...$$

We only want, say the term up to x^4, so we will ignore all higher terms. We get

$$1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + ... - \frac{x^3}{6} - \frac{x^4}{6} - ... + ...$$
$$= 1 + x + \frac{x^2}{2} - \frac{x^4}{8} + ...$$

Example. Consider $\ln\left(\frac{\sqrt{1+x}}{1+x^2}\right)$. Then this is just $\frac{1}{2}\ln(1+x) - \ln(1+x^2)$ which is

$$\frac{x}{2} - \frac{x^2}{4} + \frac{x^3}{6} - \frac{x^4}{8} + ... - x^2 + \frac{x^4}{2} - ... = \frac{x}{2} - \frac{5x^2}{4} + \frac{x^3}{6} + \frac{3x^4}{8} + ...$$

4.10 Reduction formulae

A reduction formula is basically a formula that relates integrals that we can find by integrating by parts. It's best if I do an example.

Example. Let $I_n = \int_0^\infty x^n e^{-x} dx$. Then integrating by parts (taking the limit implicitly and noting that exponentials beat polynomials) gives:

Set $u = x^n$, $dv = e^{-x}$, $v = -e^{-x}$, $du = nx^{n-1}$

$$I_n = \left[-x^n e^{-x}\right]_0^\infty + n \int_0^\infty x^{n-1} e^{-x} dx = 0 + nI_{n-1}$$

Therefore, for integers n, and the fact that $I_0 = 1$, we deduce a formula $I_n = n!$. This gives a generalization of the factorial to non-integers.

Example. This one is less trivial. Let $I_n = \int_0^{\frac{\pi}{2}} \sin^n(x)\, dx$. As it turns out, we will use this one in level 6.3.

To do integration by parts set $u = \sin^{\mathrm{n}-1}(\theta)$ and $dv = \sin(\theta)$ so that $v = -\cos(\theta)$and $du = (n-1)\sin^{n-2}(\theta)\cos(\theta)$

$$I_n = \left[-\cos(\theta)\sin^{\mathrm{n}-1}(\theta)\ \right]_0^{\frac{\pi}{2}} + (n-1)\int_0^{\frac{\pi}{2}} \sin^{n-2}(\theta)\cos^2(\theta) d\theta$$

Where we had to be careful to not flip any minus signs the wrong way even when we had double negatives to deal with.

Now

$$I_n = (n-1)\int_0^{\frac{\pi}{2}} \sin^{n-2}(\theta) - \sin^n(\theta) d\theta$$

because the bracketed part vanishes to 0 and we have the pythagorean identity $\cos^2(\theta) = 1 - \sin^2(\theta)$.

Thus $I_n = (n-1)(I_{n-2} - I_n)$ so $nI_n = (n-1) I_{n-2}$ so we have our recurrence relation $I_n = \frac{n-1}{n} I_{n-2}$ for n at least 2.

We can therefore find

$$\int_0^{\frac{\pi}{2}} \sin^7(x)\, dx$$

It is equal to

$$I_7 = \frac{6}{7} I_5 = \frac{4}{5}\frac{6}{7} I_3 = \frac{2}{3}\frac{4}{5}\frac{6}{7} I_1 = \frac{16}{35}\int_0^{\frac{\pi}{2}} \sin(x)\, dx = \frac{16}{35}$$

No trig identities needed - that would have been horrible.

4.11 The Leibniz rule

Let $f(x)$ and $g(x)$ be differentiable functions. Then there is a formula, called the Leibniz rule, for $\frac{d^n}{dx^n}(f(x)\, g(x))$ which gives the n'th derivative of $f(x)g(x)$.

The formula is as follows:

$$\frac{d^n}{dx^n}(f(x)\, g(x)) = \sum_{r=0}^{n} \binom{n}{r} \frac{d^r}{dx^r} f(x) \frac{d^{(n-r)}}{dx^{(n-r)}} g(x)$$

Theorem. The formula above is true

Proof. The case $n = 1$ is just the product rule. Now proceed by induction. Suppose it is true when $n = k$.

$$\frac{d^{k+1}}{dx^{k+1}}(f(x)\, g(x)) = \frac{d}{dx}\frac{d^k}{dx^k}(f(x)\, g(x)) = \frac{d}{dx}\sum_{r=0}^{k} \binom{k}{r} \frac{d^r}{dx^r} f(x) \frac{d^{(k-r)}}{dx^{(k-r)}} g(x)$$

$$= \sum_{r=0}^{k} \binom{k}{r} \frac{d}{dx} \left(\frac{d^r}{dx^r} f(x) \frac{d^{(k-r)}}{dx^{(k-r)}} g(x) \right) = \sum_{r=0}^{k} \binom{k}{r} \left(\frac{d^{(r+1)}}{dx^{(r+1)}} f(x) \frac{d^{(k-r)}}{dx^{(k-r)}} g(x) + \frac{d^r}{dx^r} f(x) \frac{d^{(k+1-r)}}{dx^{(k+1-r)}} g(x) \right)$$

To simplify this, let $h(r) = \frac{d^r}{dx^r} f(x) \frac{d^{(k+1-r)}}{dx^{(k+1-r)}} g(x)$. Then we have

$$= \sum_{r=0}^{k} \binom{k}{r} (h(r+1) + h(r)) = \sum_{r=0}^{k+1} \left(\binom{k}{r} + \binom{k}{r-1} \right) (h(r))$$

$$= \sum_{r=0}^{k+1} \binom{k+1}{r} (h(r))$$

Since this is an identity we proved about binomial coefficients in level 3. Thus, the result follows.

□

This is useful for finding derivatives, and thus Taylor series, of products. Here is an example of using the rule:

$$\frac{d^3}{dx^3} x^5 e^x = x^5 e^x + 3\left(5x^4 e^x\right) + 3\left(20x^3 e^x\right) + 60x^2 e^x = \left(x^5 + 15x^4 + 60x^3 + 60x^2\right) e^x$$

However, in the case of $x^5 e^x$ the Taylor series is easy to find: Just take the one for e^x and add 5 to every power of x.

5 Differential equations

5.1 Integrating factor method

This method can be used to solve differential equations of the form $\frac{dy}{dx} + f(x) y = g(x)$.

We find an antiderivative of $f(x)$, and we don't worry about $+c$ we just need any antiderivative. So we have $\int f(x)\, dx$ and we multiply the entire equation by $e^{\int f(x) dx}$, which we call our integrating factor. Since we are multiplying by something that is never 0 we don't change anything. We get

$$e^{\int f(x) dx} \left(\frac{dy}{dx} + f(x) y \right) = g(x) e^{\int f(x) dx}$$

We then note that by the product and chain rules this is equivalent to saying

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

It helps to see an example.

We will solve the differential equation $\frac{dy}{dx} = y + x$. The steps are as follows:

$$\frac{dy}{dx} = y + x$$

$$\frac{dy}{dx} - y = x$$

The coefficient multiplying y is -1 so the integrating factor is $e^{\int (-1) dx} = e^{-x}$. We get

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

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

$$ye^{-x} = \int xe^{-x}dx$$

We can do this integral by parts. We have $u = x$, $dv = e^{-x}$, $du = 1$, $v = -e^{-x}$. We have

$$ye^{-x} = -xe^{-x} - \int -e^{-x}dx = -(x+1)e^{-x} + c$$

Therefore,

$$y = ce^{x} - x - 1$$

That is the general solution.

We could put an initial condition like $y = 0$ when $x = 0$ and we will get $y = e^{x} - x - 1$

Example. Now we will solve

$$\frac{dy}{dx} + \tan(x)\, y = \cos(x)$$

on the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$

Here our integrating factor is $e^{\int \tan(x)dx}$.

We calculated the integral of $\tan(x)$ in an example somewhere in level 3, we found that it was $\ln(\sec(x))$. Therefore our integrating factor is $e^{\ln(\sec(x))} = \sec(x)$.

We now get

$$\sec(x)\left(\frac{dy}{dx} + \tan(x)\, y\right) = 1$$

$$\sec(x)\frac{dy}{dx} + \sec(x)\tan(x)\, y = 1$$

I don't think I've shown the sec derivative on this website yet. We note that

$$\frac{d}{dx}\sec(x) = \frac{d}{dx}(\cos(x))^{-1} = (-\sin(x))\left(-\cos(x)^{-2}\right) = \frac{\sin(x)}{\cos(x)}\frac{1}{\cos(x)} = \sec(x)\tan(x)$$

Therefore our differential equation can be written as

$$\frac{d}{dx}(\sec(x)\, y) = 1$$

Integrating both sides gives

$$\sec(x)\, y = x + c$$

Therefore

$$y = (x+c)(\cos(x))$$

5.2 Second order differential equations

Ok here we go.

We can solve differential equations like $a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$. There is an easy way and a hard way. To keep in the spirit of the A level textbooks, I will show you the hard way where you memorize a million rules of magic, and defer the easy way, where there is one rule and one rule only, to level 6.1.

Definition. A differential equation is said to be n'th order if the n'th derivative is the highest one that appears in the equation. So the above equation is second order.

There is one thing we do in both the easy way and the hard way, which is to solve the equation $ax^2 + bx + c = 0$ which we call the auxiliary equation and consider the roots. In the hard way we proceed as follows.

We memorize several magic rules that you have to rote memorize for the hard way, rather than the 1 rule from the easy way that makes all of these rules drop out nicely.

Rule 1. If the roots are real and not repeated and they are a and b our solution is $C_1e^{ax} + C_2e^{bx}$. We assert that this is unique because the equation is second order so we have two arbitrary constants. We ignore the fact that you can map the real numbers in a one-to-one fashion to sets of two real numbers so the notion of "number of constants" is not well defined, or even simpler just say $C_1e^{ax} + C_2e^{bx} + 0C_3 + 0C_4^2$. Yeah I will be making fun of A level textbooks and general texts on the topic as I do this like WHY DON'T THEY JUST SHOW US THE EASY WAY.

In the case $a = -b$ you can also write the solution as $(C_1 + C_2)\cosh(ax) + (C_1 - C_2)\sinh(ax)$ which is the same as $A_1\cosh(ax) + A_2\sinh(ax)$ where A_1 and A_2 are also arbitrary constants.

Rule 2. If the roots are real and repeated and both equal to a then we assert that the solution is $(C_1x + C_2)e^{ax}$. Why the x? Because by magic and because I said so.

Rule 3. This one is less surprising if you accept rule 1. If the roots are $a \pm bi$ we assert that the solution is $(C_1\sin(bx) + C_2\cos(bx))e^{ax}$.

If you're not convinced that the easy way will compress these THREE MAGIC rules into ONE NON-MAGIC rule then don't worry, there will be plenty more magic rules. Very soon we will start guessing random solutions and assert that they are the general solution.

We will now consider equations like $a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x)$.

Rule 4. To solve this we find the complementary function (ie the general solution if $f(x)$ is replaced with 0), and find a particular integral (I will explain what this means shortly and it will come with its own set of magic rules), and just add them together!

Rule 5. To find the particular integral we guess, and then try to solve for the constants involved in our guess by trying to make the particular integral satisfy the differential equation. If $f(x)$ is a constant we guess a constant. If $f(x)$ is a polynomial of degree n then we guess a polynomial of degree n.

Rule 6. If $f(x)$ is an exponential or a sin or cos term, then we guess that the particular integral is a multiple of it.

Rule 7. If f is an exponential term that matches a term of the complementary function we guess that it is that multiplied by x.

Rule 8. If f is a $\sin(Ax)e^{Bx}$ or $\cos(Ax)e^{Bx}$ term that appears in our complementary function or a combination of 2 such terms, try multiplying by x because like I said that works because I said so and by magic. We guess

$$x\left(C_1\sin(Ax)e^{Bx} + C_2\cos(Ax)e^{Bx}\right)$$

as a particular integral where we will later solve for C_1 and C_2.

I love memorizing three forms of the complementary function and a suite of guessing rules instead of just making one substitution. I will give it to them that these rules make the solution short, it just annoys me that 8 rules can be compressed into 1.

Ok lets do an example and cover each of the cases.

Example. Lets solve $y'' - 5y' + 6y = 0$. In this case we just apply rule 1. The roots of the auxiliary equation $x^2 - 5x + 6 = 0$ are 2 and 3 so we get $Ae^{2x} + Be^{3x}$ as our general solution where A and B are arbitrary constants.

Example. Lets solve $y'' - 4y' + 4y = x^2$. In this case we apply rule 4 (roots of auxiliary equation $x^2 - 4x + 4 = 0$ are 2 and repeated), so we need to find our complementary function which falls under rule 2 and is therefore $(A + Bx)\, e^{2x}$.

To find the particular solution we apply rule 5, we guess that it is of the form $g(x) = px^2 + qx + r$. To solve for the coefficients we need to make $g'' - 4g' + 4g = x^2$. Note that $g' = 2px + q$ and $g'' = 2p$ so we have

$$2p - 4(2px + q) + 4\left(px^2 + qx + r\right) = x^2$$

Therefore

$$(4p)\, x^2 + (4q - 8p)\, x + (2p + 4r - 4q) = x^2$$

Equating coefficients gives the equations $4p = 1$ (so $p = 1/4$), then $4q - 8p = 0$ (so $q = 1/2$), then $2p + 4r - 4q$ (so $r = 3/8$). Therefore our particular solution will be $\frac{1}{4}x^2 + \frac{1}{2}x + \frac{3}{8}$.

Applying rule 4 again it follows that the general solution is

$$y = (A + Bx)\, e^{2x} + \left(\frac{1}{4}x^2 + \frac{1}{2}x + \frac{3}{8}\right)$$

Oh yeah and I should probably set some initial conditions. Lets say that when $x = 0$, $y = \frac{67}{8}$ and $\frac{dy}{dx} = \frac{17}{2}$. These numbers were picked because I messed this up before and wanted to make it so I had to change as little as possible.

We note that if $x = 0$ then y reduces to $A + \frac{3}{8}$ so we need $A = 8$. Now finding the derivative of y I will use the product rule. We get

$$\frac{dy}{dx} = (B + 2A + 2Bx)\, e^{2x} + \left(\frac{1}{2}x + \frac{1}{2}\right)$$

If x is 0 and A is 8 this reduces to

$$\frac{dy}{dx} = (B + 16) + \frac{1}{2}$$

Since $A = 8$ we have

$$\frac{dy}{dx} = B + \frac{33}{2}$$

Therefore to make this $\frac{17}{2}$ we need $B = -8$. Therefore the solution that satisfies these initial conditions is

$$y = (8 - 8x)\, e^{2x} + \left(\frac{1}{4}x^2 + \frac{1}{2}x + \frac{3}{8}\right)$$

Although I will not claim that this is a universal theorem in a precise sense, it is true that often in an n'th order differential equation we want n separate initial conditions to make the solution unique. For an A level exam you can take this to be always true.

Example. Lets solve $y'' + 4y = \sin(2x)$.

By rule 3 since the roots of the auxiliary equation $x^2 + 4 = 0$ are $\pm 2i$ our complementary function is $A\sin(2x) + B\cos(2x)$. Normally this would also be the form of our particular solution but since there is a collision the form will actually be

$$x\,(p\sin(2x) + q\cos(2x))$$

To find p and q note that if the particular solution is y we want $y'' + 4y = \sin(2x)$. By the Leibniz rule from earlier, or directly, we find that

$$y'' = x\,(-4p\sin(2x) - 4q\cos(2x)) + 4p\cos(2x) - 4q\sin(2x)$$

So

$$y'' + 4y = 4p\cos(2x) - 4q\sin(2x)$$

Therefore $q = -\frac{1}{4}$ and $p = 0$. So our particular solution is $-\frac{1}{4}x\cos(2x)$. Therefore the general solution is

$$y = -\frac{1}{4}x\cos(2x) + A\sin(2x) + B\cos(2x)$$

We've seen how to plug in initial conditions so I won't do it again, instead I will do 1 more example.

Example. Lets solve $y'' + y' + y = e^x$.

First, we find the complementary function. The auxiliary equation is $x^2 + x + 1 = 0$. Its roots are $-\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$. Therefore the complementary function is as follows:

$$y = e^{-\frac{1}{2}x}\left(A\sin\left(\frac{\sqrt{3}}{2}x\right) + B\cos\left(\frac{\sqrt{3}}{2}x\right)\right)$$

Since there is no "collision" (whatever that is even formally supposed to mean), our guess for the particular solution is just Ce^x. Since the differential equation is so nice, putting this into the equation is easy. If $y = Ce^x$ then y' and y'' are the same we get $3Ce^x = e^x$ so $C = \frac{1}{3}$. Therefore the general solution is

$$y = e^{-\frac{1}{2}x}\left(A\sin\left(\frac{\sqrt{3}}{2}x\right) + B\cos\left(\frac{\sqrt{3}}{2}x\right)\right) + \frac{1}{3}e^x$$

This is horrible to differentiate to get initial conditions, so just pray you don't get a question in an exam where you have to because I guarantee you, you are either a complete genius or you will mess up the algebra somewhere when dealing with differentiating all the terms and stuff.

Those examples basically cover all the cases to know for this subsection.

If a differential equation is not one we immediately know how to solve, we either use a substitution (see below), or one thing that is possible to do is given y at some x to try to approximate y at $x + \delta x$ for some small δx using the information we know. There are ways to do this with varying levels of success but that is beyond the scope of this level. The method of series solutions that we will see soon is also useful for getting approximate solutions, as long as you are careful with convergence issues and stuff - we will address a special case in level 6.1 and even that won't be easy.

5.3 Use of substitutions

It is possible to use substitutions to solve differential equations. I will give a brief example.

Example. Suppose we want to solve $x^2\frac{d^2y}{dx^2} - x\frac{dy}{dx} + y = 0$ for $x > 0$. We will use the substitution $z = \ln(x)$ to do this. We find that

$$\frac{dy}{dz} = \frac{dx}{dz}\frac{dy}{dx} = \frac{dy/dx}{dz/dx} = \frac{dy/dx}{1/x} = x\frac{dy}{dx}$$

We find also that

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

Therefore we have

$$x^2\frac{d^2y}{dx^2} = \frac{d^2y}{dz^2} - \frac{dy}{dz}$$

Thus we want to solve

$$x^2\frac{d^2y}{dx^2} - x\frac{dy}{dx} + y = \frac{d^2y}{dz^2} - 2\frac{dy}{dz} + y = 0$$

From the last section, the roots of the Auxillary equation $x^2 - 2x + 1 = 0$ are both 1. Therefore the general solution is $y = (Az + B) e^z$, so the general solution to the original equation is $y = x (A \ln(x) + B)$.

Ok, I can't resist doing this:

Example. We will use the substitution $u = \frac{dy}{dx} - y$ to solve the differential equation $\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + y = 0$.

To do this, note that $\frac{du}{dx} = \frac{d^2y}{dx^2} - \frac{dy}{dx}$ so our original equation is equivalent to $\frac{du}{dx} - u = 0$. Using an integrating factor of e^{-x} we have

$$\begin{aligned} e^{-x}\frac{du}{dx} - ue^{-x} &= 0 \\ \frac{d}{dx}\left(ue^{-x}\right) &= 0 \\ ue^{-x} &= C \\ u &= Ce^x \\ \frac{dy}{dx} - y &= Ce^x \end{aligned}$$

We can use an integrating factor again.

$$\begin{aligned} \frac{dy}{dx}e^{-x} - ye^{-x} &= C \\ \frac{d}{dx}\left(ye^{-x}\right) &= C \\ ye^{-x} &= C_1x + C_2 \\ y &= (C_1x + C_2)\, e^x \end{aligned}$$

This is an idea of where the rule comes from! If you want to really feel like what I felt when learning this pretend like you didn't see this example and that it's all a bunch of magic rule nonsense.

5.4 Series solutions

It is possible to find solutions to differential equations in the form of a power series.

Whether the power series actually converges is a difficult question. We will see in Level 6.1 that for any differential equation of the following forms

- $\frac{dy}{dx} + f(x)\, y = g(x)$

- $\frac{d^2y}{dx^2} + f(x)\frac{dy}{dx} + g(x)\, y = h(x)$

If f, g, h have convergent power series in an interval about the point of interest, so does our series for y.

Example. As an example, we will find a series solution about $x = 0$ to the equation $\frac{d^2y}{dx^2} = x^2 + y^2$. We will assume the series solution should start $1 + x + ...$ since we need to specify 2 terms. I will showcase the **method of successive differentiation**. Other than convergence issues it is clear why this works.

To find the second derivative, we note that at 0, $y = 1, x = 0$ so $\frac{d^2y}{dx^2} = 1$.

To find the third derivative, we differentiate the entire equation. Doing this gives $\frac{d^3y}{dx^3} = 2x + 2y\frac{dy}{dx}$. Given our information about what y and $\frac{dy}{dx}$ are when $x = 0$, we plug this in and the result is that at $x = 0$, $\frac{d^3y}{dx^3} = 2$.

We can find the fourth derivative. Differentiating again, this time using the product rule, gives $\frac{d^4y}{dx^4} = 2 + 2\left(\frac{dy}{dx}\right)^2 + 2y\frac{d^2y}{dx^2}$. Using the fact that $y = y' = y'' = 1$ we get that the fourth derivative is 6.

I feel like finding the fifth derivative to see if the apparent pattern continues to hold. We get

$$\frac{d^5y}{dx^5} = 2y'y'' + 2yy''' + 2y'y'' = 4y'y'' + 2yy'''$$

Since $y = y' = y'' = 1$ and $y''' = 2$ this derivative is 8.

Therefore, since the coefficient of x^n is $\frac{1}{n!}$ times the n'th derivative in a Taylor series, we get that our series solution is

$$1 + x + \frac{1}{2!}x^2 + \frac{2}{3!}x^3 + \frac{6}{4!}x^4 + \frac{8}{5!}x^5 + ...$$

Simplifying, we have

$$1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \frac{x^5}{15} + ...$$

So, the pattern did not in fact continue.

Example. I will find a power series solution about $x = \pi$ to the differential equation $\frac{dy}{dx} = \sin(x)\, y$. Ok we can solve this but we will pretend we can't. We will say that $y(\pi) = 3$.

Now note that if $x = \pi$ and $y = 3$ then $\frac{dy}{dx} = 3(\sin(\pi)) = 0$.

Differentiating and using above values we get $y'' = \cos(x)\, y + \sin(x)\, y' = -3$.

Differentiating again using the Leibniz rule we get

$$y''' = -\sin(x)\, y + 2\cos(x)\, y' + \sin(x)\, y'' = 0 + 0 + 0 = 0$$

$$y'''' = -\cos(x)\, y - 3\sin(x)\, y' + 3\cos(x)\, y'' + \sin(x)\, y''' = -(-3) + 0 + (-3)(-3) + 0 = 12$$

That is enough terms for me. The solution about $x = \pi$ will be

$$3 - \frac{3}{2!}(x-\pi)^2 + \frac{12}{4!}(x-\pi)^4 + ... = 3 - \frac{3}{2}(x-\pi)^2 + \frac{1}{2}(x-\pi)^4 + ...$$

5.5 Coupled differential equations

There is a way to solve systems of 2 differential equations of a form similar to the following example:

Example. Lets solve the system

$$\frac{dy}{dx} = 2y - z$$

$$\frac{dz}{dx} = 2y + 5z$$

The general approach is to convert this into one 2nd order differential equation. Differentiating the first equation gives

$$\frac{d^2y}{dx^2} = 2\frac{dy}{dx} - \frac{dz}{dx}$$

Substituting the $\frac{dz}{dx}$ from the second equation gives

$$\frac{d^2y}{dx^2} = 2\frac{dy}{dx} - 2y - 5z$$

The first equation rearranged says $z = 2y - \frac{dy}{dx}$. Therefore, substituting this into what we have above we get

$$\frac{d^2y}{dx^2} = 7\frac{dy}{dx} - 12y$$

This is something we can solve. The roots of the auxiliary equation $x^2-7x+12=0$ are 3 and 4 therefore $y = Ae^{3x}+Be^{4x}$ and $\frac{dy}{dx} = 3Ae^{3x}+4Be^{4x}$

Since we know that $z = 2y - \frac{dy}{dx}$ it follows that $z = -Ae^{3x} - 2Be^{4x}$. We could find A and B if given initial values for y and z.

Note that this is a solution because the only step that is not completely reversible is differentiating the first equation, but we explicitly found z so that this y and z satisfy the first equation.

Example. Suppose we have a situation where we start with 1 of A, and A gets converted into B at a rate of 1 unit of A every 5 seconds, and B gets converted into C at a rate of 1 unit of B every 3 seconds. We want to find how much A,B,C there is as a function of time.

Then differential equations that model this are as follows.

$$\frac{dB}{dt} = \frac{1}{5}A - \frac{1}{3}B$$

Since we need to take into account B being both created and destroyed.

$$\frac{dC}{dt} = \frac{1}{3}B$$

But $C = 1 - A - B$ always so rewrite

$$-\frac{dA}{dt} - \frac{dB}{dt} = \frac{1}{3}B$$

Using our equation for $\frac{dB}{dt}$ gives

$$\begin{aligned}
-\frac{dA}{dt} - \frac{1}{5}A + \frac{1}{3}B &= \frac{1}{3}B \\
\frac{dA}{dt} &= -\frac{1}{5}A
\end{aligned}$$

Oh hang on we can just solve for A without even differentiating. We have $A = ke^{-\frac{t}{5}}$. By the initial condition, $k = 1$. Therefore $A = e^{-\frac{t}{5}}$.

So we have the equation

$$\begin{aligned}
\frac{dB}{dt} + \frac{1}{3}B &= \frac{1}{5}e^{-\frac{t}{5}} \\
\frac{dB}{dt}e^{\frac{t}{3}} + \frac{1}{3}Be^{\frac{t}{3}} &= \frac{1}{5}e^{\frac{2t}{15}} \\
\frac{d}{dt}\left(Be^{\frac{t}{3}}\right) &= \frac{1}{5}e^{\frac{2t}{15}} \\
Be^{\frac{t}{3}} &= k_2 + \frac{3}{2}e^{\frac{2t}{15}} \\
B &= k_2e^{-\frac{t}{3}} + \frac{3}{2}e^{-\frac{t}{5}}
\end{aligned}$$

Since $B = 0$ initially we have $k_2 = -\frac{3}{2}$. Thus,

$$B = \frac{3}{2}\left(-e^{-\frac{t}{3}} + e^{-\frac{t}{5}}\right)$$

Using $C = 1 - A - B$ we get

$$C = 1 - \frac{5}{2}e^{-\frac{t}{5}} + \frac{3}{2}e^{-\frac{t}{3}}$$

It makes sense since this starts at 0 and approaches 1 as t gets large.

Figure 15 shows A in red, B in blue and C in green as a graph.

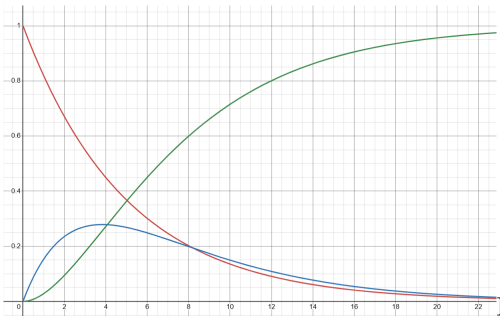

Figure 15

6 Matrices

6.1 Matrix addition and multiplication

Let me just say: I will teach the basics first and then explain the geometric intuition. But textbooks/exam boards that put this in the syllabus WITHOUT explaining intuitively what is going on and seemingly puts it in JUST because it is easy to set questions on it, leaving students to memorize and follow rules but not really understand what a matrix is (which is not all of them, but I've seen this) ARE SO BAD.

Definition. A matrix is a grid of numbers. For example, $\begin{pmatrix} 67 & 0.2 & 3 & 19 \\ 8 & 4 & 5.1 & 0 \\ 3 & 7 & 6.1 & 0 \end{pmatrix}$ is a matrix.

We would say that the above matrix is a 3*4 matrix since it has 3 rows and 4 columns. If the matrix is A we would say that A_{ij} is the thing in the i'th row and j'th column. For example, $A_{21} = 8$.

If two matrices have the same dimensions, we can add them. To do this we just add the components. For example,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 7 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 9 \\ 9 & 12 \end{pmatrix}$$

We can also multiply a matrix by a number by simply multiplying the entry in every position by that number.

Matrix multiplication is a bit different. For it to be possible to multiply 2 matrices together, say we want to multiply a matrix A by a matrix B in the order AB, and A is n*m and B is p*q, then multiplication is possible if and only if $m = p$. The resulting matrix AB will be an $n * q$ matrix.

If two matrices are compatible, then what happens is that to find the element in the i'th row and j'th column of AB, we need to take the i'th row of A and the j'th column of B and effectively take a dot product.

Example. Lets say we want to multiply

$$\begin{pmatrix} 3 & 5 & 2 \\ 8 & -4 & 1 \end{pmatrix} \begin{pmatrix} 6 & 5 \\ 0 & 7 \\ 8 & -9 \end{pmatrix}$$

This is a 2*3 matrix times a 3*2 matrix. 3 and 3 match so multiplication is possible, and the result will be a 2*2 matrix. Lets say this 2*2 matrix is $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then to find a, we need to take the first row of the left matrix and the first column of the right matrix and effectively take their dot product. Specifically,

$$a = (3, 5, 2).(6, 0, 8) = 3 * 6 + 5 * 0 + 2 * 8 = 34$$

We can similarly find $b = 32$, $c = 56$ and $d = 3$.

Note that it is not generally true that $AB = BA$. However, it is true that matrix multiplication is associative, i.e. $A(BC) = (AB)C$. Although we only prove this in level 6.1, it will become surprisingly intuitive in our subsequent work.

Definition. For each n, the identity matrix is the square matrix, which we call I such that it has 1's on the diagonal and 0 everywhere else. Specifically it looks like

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

This is special because if we multiply it on the right or left by anything, we will get the same thing. Specifically, if we have IA then the ij element of this will be the i'th row of I dotted with the j'th column of A, which just picks out the element in the i'th row and j'th column of A. The same thing happens if we have AI. So basically multiplying by I does nothing.

Definition. The transpose of a matrix A, denoted A^T, is A if you swap rows and columns. For example, the transpose of $\begin{pmatrix} 3 & 5 & 2 \\ 8 & -4 & 1 \end{pmatrix}$ is $\begin{pmatrix} 3 & 8 \\ 5 & -4 \\ 2 & 1 \end{pmatrix}$

Although this is not difficult to prove, we defer it to level 6.1 because I'm trying to replicate the stupid textbooks. It is always the case that $A^T B^T = (BA)^T$.

Definition. Given a matrix A, we can define powers of the matrix in the obvious way: $A^2 = AA$, $A^3 = AAA$, etc.

Example. Consider the matrix $\begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}$. I claim that $\begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}^n = \begin{pmatrix} 1-2n & 4n \\ -n & 2n+1 \end{pmatrix}$ and will prove it by induction. By plugging in n=1 we see that that case works.

Thus, suppose that $\begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}^k = \begin{pmatrix} 1-2k & 4k \\ -k & 2k+1 \end{pmatrix}$. Then

$$\begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}^{k+1} = \begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}^k = \begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 1-2k & 4k \\ -k & 2k+1 \end{pmatrix}$$

$$= \begin{pmatrix} -1\,(1-2k) + 4\,(-k) & -1\,(4k) + 4\,(2k+1) \\ -1\,(1-2k) + 3\,(-k) & -1\,(4k) + 3\,(2k+1) \end{pmatrix}$$

$$= \begin{pmatrix} -2k-1 & 4k+4 \\ -k-1 & 2k+3 \end{pmatrix} = \begin{pmatrix} 1-2\,(k+1) & 4\,(k+1) \\ -\,(k+1) & 2\,(k+1)+1 \end{pmatrix}$$

exactly as we set out to prove. Thus we are done by induction.

6.2 Interpretation as linear maps on vectors

Definition. A function f that takes in a vector and outputs a vector is said to be linear if $f\,(aX + bY) = af\,(X) + bf\,(Y)$.

It turns out that a matrix can represent a linear map from vectors to vectors. In the most general case if we want a map from M-dimensional vectors to N-dimensional vectors this will be represented by an N*M matrix.

For example, a linear map from $\mathbb{R}^3$ (i.e. 3 dimensional real vectors) to $\mathbb{R}^2$ can be represented by a 2*3 matrix in the sense that if M is a 2*3 matrix and v is a 3-dimensional column vector, then Mv will be a 2 dimensional column vector and $v \mapsto Mv$ defines a function from $\mathbb{R}^3$ to $\mathbb{R}^2$.

Example. Consider the matrix $M := \begin{pmatrix} 6 & 1 & -5 \\ 2 & -3 & 7 \end{pmatrix}$. Then $M\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ will be $\begin{pmatrix} 6*1+0+0 \\ 2*1+0+0 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \end{pmatrix}$.

If you notice what just happened there, multiplying M by the first of the 3 "basis vectors" (ie the vectors that are 1 unit in the x, y, z directions) basically gives you the first column of M. Multiplying by each basis vector gives the columns of M.

Note that since every vector (in this level we only work in finite dimensional real space) is a sum of multiples of basis vectors, it follows that a linear map is entirely determined by where it sends basis vectors.

Thus, every Matrix is a linear map (since $(aM + bN)v = a(Mv) + b(Nv)$ is true, if you don't see this I give a brief proof in Level 6.1), and every linear map has an associated matrix.

From now on, we will work with mostly with square matrices and column vectors.

Example. Consider the matrix $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Then this sends $(1, 0)$ to $(1, 2)$ and $(0, 1)$ to $(2, 1)$. Visually, it transforms the grid as shown in figure 16.

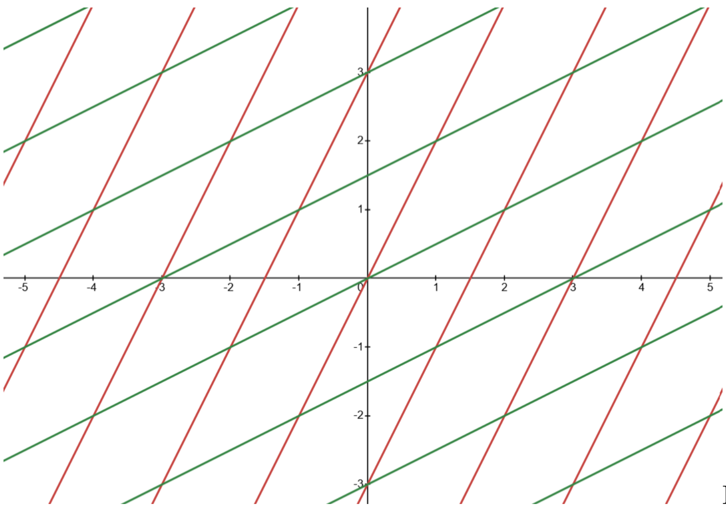

Figure 16

It is now easy to see why $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is an "identity" - it will do nothing and not transform the grid at all.

6.3 Geometric examples

Example. Suppose we are in 2D and we want to consider the transformation that stretches the whole grid by a factor of λ, then this will be

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$

Example. Suppose that we want to stretch by a factor of λ in the x-direction but leave the y-axis unchanged, then this will be

$$\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}$$

Example. Suppose that we want to rotate everything by an angle of θ anticlockwise. Then we will send the x-axis to the point $(\cos(\theta), \sin(\theta))$ and the y-axis to the point $(-\sin(\theta), \cos(\theta))$. Thus the matrix will be given by

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

Example. Suppose we want to reflect everything about the x-axis. Then the x-axis will remain unchanged and $(0, 1)$ will go to $(0, -1)$. Therefore, the matrix will be

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Interestingly, the transformation above is like taking the complex conjugate of every complex number. The transformation where we multiply the whole complex plane by $re^{i\theta}$ can be given by

$$\begin{pmatrix} r\cos(\theta) & -r\sin(\theta) \\ r\sin(\theta) & r\cos(\theta) \end{pmatrix}$$

Example. Suppose we want to find the matrix that represents the transformation where we stretch by a factor of 2 in the x-direction *then* rotate by 30 degrees anticlockwise.

Let M be the matrix representing the 2 stretch and R be the matrix representing the rotation and v be any vector. Then Mv will be v under the stretch transformation, and $R(Mv)$ will be Mv under the R transformation. Since matrix multiplication is associative, this is the same as $(RM)v$ so RM is the matrix we want.

In general, if we want to do A then B we want the product BA. In the above case, RM can be found by:

$$\begin{pmatrix} \cos(30) & -\sin(30) \\ \sin(30) & \cos(30) \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{3} & -\frac{1}{2} \\ 1 & \frac{\sqrt{3}}{2} \end{pmatrix}$$

On the other hand, if we did the transformations in the other order, we would have

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} \sqrt{3} & -1 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

Example. Suppose we want the 3D transformation that keeps the z-axis fixed and rotates the x and y axes by an angle of θ in the direction that would appear anticlockwise from someone in the positive z direction. Then by considering where it must send each of the basis vectors, it must be equal to

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

6.4 The determinant

For a 2*2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ the determinant is defined as $ad - bc$. For a matrix A we denote its determinant as det(A).

For a 3*3 matrix we can find its determinant as follows:

For each element in the first row, consider the 2*2 matrix that you would get if you deleted that row and column and compute its determinant and then multiply it by the element in that row. Add these results together for each element in the first row except for the second element subtract it instead of adding it.

Concretely, given a 3*3 matrix A, its determinant is

$$A_{11}(A_{22}A_{33} - A_{23}A_{32}) - A_{12}(A_{21}A_{33} - A_{23}A_{31}) + A_{13}(A_{21}A_{32} - A_{22}A_{31})$$

Now here is a fact that feels like utter magic that if you are anything like me you will punch the wall over being expected to use this without proof because it just feels like I'm making up a formula and telling you that it works and you don't understand it. See level 6.1 for the proof - it's in a video. Consider a shape in 2D or 3D shape and let it be transformed by a square matrix of the corresponding size. Then the (absolute value of the) determinant will give you the factor by which its area/volume is multiplied. If the determinant is negative the orientation of space is reversed, and if it is 0 then the matrix as a transformation collapses everything into lower dimensions, and in this case we say it is singular.

By this property, it is true for any matrices that $\det(AB) = \det(A)\det(B)$. This property is possible to verify by direct computation.

Example. A 2D hexagon of area 5 is transformed by the matrix $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Find its new area.

Since the determinant is $(2)(2) - (1)(1)$, its area is scaled by 3 and will thus become 15.

Example. Lets compute the determinant of

$$\begin{pmatrix} 1 & 4 & 2 \\ 6 & 3 & -2 \\ -9 & 7 & 5 \end{pmatrix}$$

It will be

$$1 * \det \begin{pmatrix} 3 & -2 \\ 7 & 5 \end{pmatrix} - 4 * \det \begin{pmatrix} 6 & -2 \\ -9 & 5 \end{pmatrix} + 2 * \det \begin{pmatrix} 6 & 3 \\ -9 & 7 \end{pmatrix}$$
$$= 1 * 29 - 4 * 12 + 2 * 69 = 119$$

6.5 Finding inverse matrices

Definition. The inverse of a matrix is exactly what it sounds like - Given a matrix A, the inverse, denoted A^{-1} is the matrix such that $AA^{-1} = A^{-1}A = I$.

I'll give a procedure for finding the inverse. Of course, it will feel like magic because I will *state* the procedure but in the spirit of the *tExTbOoKs* I won't prove it.

Note that if a matrix is singular, it is not invertible - this makes sense because for a function to be invertible it should in principle be one-to-one, which a matrix that collapses everything into lower dimensions certainly isn't. We will see from the formula exactly how we run into trouble if the determinant is 0.

For a 2*2 matrix the formula for the inverse of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $\frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. It is actually possible to check by directly multiplying that this is indeed an inverse.

For a 3*3 matrix the procedure for finding an inverse is utter magic if you don't know the proof.

I'll do an example to demonstrate. First, you calculate the matrix of minors. This is the determinants of the matrix from removing the row and column. Then you apply signs in the pattern $\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$ to get the cofactor matrix, then you take the transpose to get the adjugate matrix, then divide by the determinant to get the inverse matrix.

To do the example consider the matrix from above:

$$\begin{pmatrix} 1 & 4 & 2 \\ 6 & 3 & -2 \\ -9 & 7 & 5 \end{pmatrix}$$

First we need to find the minors:

$$\begin{pmatrix} \det \begin{pmatrix} 3 & -2 \\ 7 & 5 \end{pmatrix} & \det \begin{pmatrix} 6 & -2 \\ -9 & 5 \end{pmatrix} & \det \begin{pmatrix} 6 & 3 \\ -9 & 7 \end{pmatrix} \\ \det \begin{pmatrix} 4 & 2 \\ 7 & 5 \end{pmatrix} & \det \begin{pmatrix} 1 & 2 \\ -9 & 5 \end{pmatrix} & \det \begin{pmatrix} 1 & 4 \\ -9 & 7 \end{pmatrix} \\ \det \begin{pmatrix} 4 & 2 \\ 3 & -2 \end{pmatrix} & \det \begin{pmatrix} 1 & 2 \\ 6 & -2 \end{pmatrix} & \det \begin{pmatrix} 1 & 4 \\ 6 & 3 \end{pmatrix} \end{pmatrix}$$
$$= \begin{pmatrix} 29 & 12 & 69 \\ 6 & 23 & 43 \\ -14 & -14 & -21 \end{pmatrix}$$

Then we apply the signs:

$$\begin{pmatrix} 29 & -12 & 69 \\ -6 & 23 & -43 \\ -14 & 14 & -21 \end{pmatrix}$$

Then we take the transpose:

$$\begin{pmatrix} 29 & -6 & -14 \\ -12 & 23 & 14 \\ 69 & -43 & -21 \end{pmatrix}$$

Then we divide by the determinant:

$$\frac{1}{119}\begin{pmatrix} 29 & -6 & -14 \\ -12 & 23 & 14 \\ 69 & -43 & -21 \end{pmatrix}$$

So yeah thats how you find the inverse.

6.6 Solving equations using matrices

Example. Consider the system of 3 equations

$$x + 4y + 2z = 3$$
$$6x + 3y - 2z = -7$$
$$-9x + 7y + 5z = 2$$

We could solve this by hand but there is a much nicer way.

We can write the 3 equations as 1 matrix equation by considering how matrix multiplication works. The equation can be written as

$$\begin{pmatrix} 1 & 4 & 2 \\ 6 & 3 & -2 \\ -9 & 7 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ -7 \\ 2 \end{pmatrix}$$

Conveiently this matrix is the exact matrix we found the inverse of in the last section. What we can do is to *multiply both sides of the equation on the left by this inverse.* Then what happens is that since it is, well an inverse, we get

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{119}\begin{pmatrix} 29 & -6 & -14 \\ -12 & 23 & 14 \\ 69 & -43 & -21 \end{pmatrix}\begin{pmatrix} 3 \\ -7 \\ 2 \end{pmatrix} = \frac{1}{119}\begin{pmatrix} 101 \\ -169 \\ 466 \end{pmatrix}$$

Ok these are big numbers usually in an exam the numbers are nicer.

Reading off the coefficients, we get $x = \frac{101}{119}$, $y = -\frac{169}{119}$, $z = \frac{466}{119}$ as the solution to the original system of equations.

This always works and gives a unique solution when the matrix is invertible. When the matrix is not invertible, there is a way this can fail.

An equation like $ax + by + cz = d$ (with a, b, c not all 0) always gives a plane. I can explain best what a plane is by saying see figure 17 for an example. It is a flat two dimensional surface that extends infinitely in all dimensions.

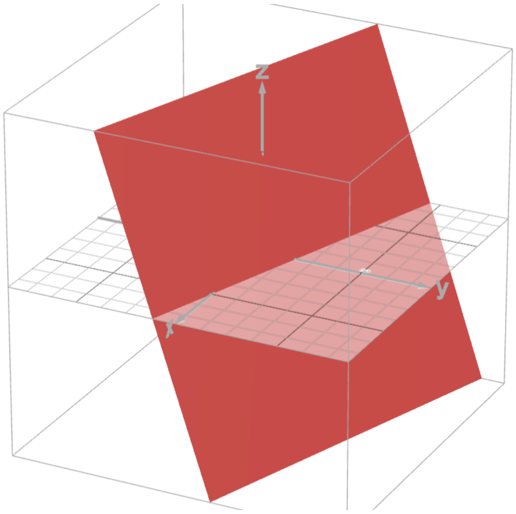

Figure 17

To see why the equation gives a plane: If c is not 0 it is saying $z = Ax + By + C$ for some A, B, C, i.e. z is a function of x, y that will look like a plane if you plot it in 3D. If $c = 0$ then we have $ax + by = d$ which is just a line that extends infinitely in the z direction, again clearly a plane.

Note that it is easy to see that if we change d but keep a,b,c unchanged we will get another plane that cannot intersect the original one, i.e. a parallel plane.

Using this, we can analyze given the equations, exactly how the failure happens.

Example. Suppose the equations are

$$x + 2y + z = 7$$
$$2x + 4y + 2z = 10$$
$$\text{something else}$$

Then the first two planes are parallel regardless of the third one so there will be no places where they intersect, hence no solutions. This is because the second plane can be written as $x + 2y + z = 5$ which is parallel to the first plane.

If the matrix is not invertible, two planes may be parallel and there will be no solutions. Sometimes there may be infinitely many solutions, in which case, depending on if the solution space you get has dimension 1 or 2, the planes either all meet in a line or they all coincide.

There is one more case, however. If there are no solutions but no planes are parallel, they will form a prism like in figure 18.

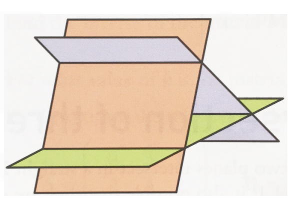
Figure 18

In the normal case where the matrix is invertible, they will intersect at a point like in figure 19.

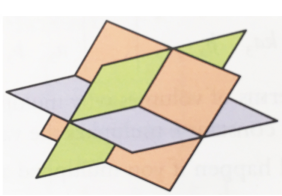
Figure 19

6.7 Invariant lines

Definition. Given a matrix, an invariant line is a line such that when you do a transformation the line stays in place. This is different from a line of invariant points, where *every point* on the line stays in place.

Example. We will prove that the matrix $\begin{pmatrix} 4 & -2 \\ 5 & 3 \end{pmatrix}$ has no invariant lines as follows.

To do this, suppose first that it has an invariant line which is not vertical. Then this will be of the form $y = mx + c$. Ie, we will have

$$\begin{pmatrix} 4 & -2 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} x \\ mx + c \end{pmatrix} = \begin{pmatrix} 4x - 2mx - 2c \\ 5x + 3mx + 3c \end{pmatrix}$$

In order to have an invariant line, we need that for every x, we have $y = mx + c$ in the transformed world, i.e. that

$$5x + 3mx + 3c = m\,(4x - 2mx - 2c) + c$$

ie, we need that

$$5x + 3mx + 3c = 4mx - 2m^2x - 2mc + c$$

For this to be true for every x, we need the coefficients to match: i.e. $5 + 3m = 4m - 2m^2$, i.e. $2m^2 - m + 5 = 0$. This is a quadratic equation with no real roots so we have a contradiction.

Next we need to check for vertical lines of the form $x = k$. But then

$$\begin{pmatrix} 4 & -2 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} k \\ y \end{pmatrix} = \begin{pmatrix} 4k - 2y \\ 5k + 3y \end{pmatrix}$$

We need this point to lie on the line, meaning that $4k - 2y = k$ for all values of y. This clearly is not the case though so we are done.

To foreshadow what we will do in a bit, a line of invariant points of a matrix M will occur if for every v on the line we have $Mv = v$. Ie, $Mv - v = 0$, i.e. $(M - I)\,v = 0$. As demonstrated, a non-zero vector will collapse to 0 by $M - I$ exactly when it is a singular matrix. This is because if the matrix is not singular we can get a unique solution for v as before, and if it is singular there will be a solution (since $v = 0$ works) and it will *not* be unique - I'll prove in level 6.1 that a non-invertible matrix cannot correspond to a system of equations with a unique solution.

6.8 Eigenvalues and eigenvectors

Suppose that we have a matrix A. Then we want to consider which vectors v are scaled by A, i.e. vectors with the property that

- They are non-zero

- There exists a constant λ such that $Av = \lambda v$

Definition. When this happens, we say that v is an Eigenvector of A with Eigenvalue λ

Theorem. λ is an eigenvalue of the matrix A if and only if $\det(A - \lambda I) = 0$

Proof. First, suppose λ is an eigenvalue. Then this means there is a non-zero vector v such that $Av = \lambda v$ which implies that $Av = \lambda I v$ which implies that $(A - \lambda I)\,v = 0$. Since v is a non-zero vector, this means that the system of equations corresponding to the matrix $(A - \lambda I)\,v = 0$ does not have a unique solution, hence $A - \lambda I$ is not invertible so its determinant is 0.

On the other hand, if $A - \lambda I$ is not invertible then as we will see in level 6.1, the system of equations corresponding to the matrix $(A - \lambda I)\,v = 0$ will not have a unique solution, hence there is a non-zero vector v that satisfies it and the reverse chain of logic works.

$\square$

Example. Let's find the eigenvalues and corresponding eigenvectors of the matrix $A = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$.

We need to find

$$\det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & -1 \\ 2 & 4-\lambda \end{pmatrix} = (1-\lambda)(4-\lambda) - (-2) = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3)$$

Thus the eigenvalues are 2 and 3.

This equation, the quadratic equation corresponding to $\det(A - \lambda I) = 0$, is called the *characteristic polynomial* of the matrix A.

To find an eigenvector corresponding to the eigenvalue 2, we need to solve

$$\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

This corresponds to the system of equations $x - y = 2x$ and $2x + 4y = 2y$. We see that both of these equations are true exactly when $y = -x$. Therefore, any multiple of the vector $(1, -1)$ is an eigenvector.

Definition. A normalized eigenvector is one with length 1.

In the example above, to find a normalized eigenvector, we would just divide by the length, i.e. we would have $\frac{1}{\sqrt{2}}(1, -1)$.

To find an eigenvector corresponding to the eigenvalue 3, we need to solve

$$\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \end{pmatrix}$$

This corresponds to the system of equations $x - y = 3x$ and $2x + 4y = 3y$. We see that both of these equations are true exactly when $y = -2x$. Therefore, any multiple of the vector $(1, -2)$ is an eigenvector. A noramlized eigenvector would be $\frac{1}{\sqrt{5}}(1, -2)$.

What we can now do is form the matrix of eigenvectors, where each eigenvector corresponds to a column. In this case the matrix would be $\begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}$.

If P is an eigenvector matrix for A, then

Theorem. $P^{-1}AP = D$ where D is a diagonal matrix like $\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$, where these are the eigenvalues, and they appear in the same order as the order in which we built the columns. Equivalently, $PDP^{-1} = A$.

The reason $A = PDP^{-1}$ is essentially that $\left(PDP^{-1}\right)v$ takes v as input, and if we take v to be separated into eigenvector components, it first moves v such that those are its normal components (by the P^{-1}), scales everything by the eigenvalue factors (by the D), then moves it back (by the P).

Therefore, to "Diagonalize" the matrix A above, a matrix P and a diagonal matrix D such that $P^{-1}AP = D$ is $P = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}$ and $D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$.

In this level/A levels we do not consider the case where eigenvalues are repeated or complex. In level 6.1 we will see why P is actually invertible (provided we have as many distinct eigenvalues as dimension, which is always the case in A level problems, see the vectors and matrices course in level 7.3 for further discussion on exactly when these things work and don't work).

6.9 The Cayley-Hamilton theorem

The Cayley-Hamilton theorem basically says that every matrix satisfies its own characteristic equation. See level 6.1 for the proof.

As an example, consider the A above, $\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$ with characteristic equation $\lambda^2 - 5\lambda + 6$. Then the theorem claims that $A^2 - 5A + 6I = 0$.

Indeed we can verify this for this case: $\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}^2 = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} -1 & -5 \\ 10 & 14 \end{pmatrix}$.

Then $A^2 - 5A + 6I = \begin{pmatrix} -1 & -5 \\ 10 & 14 \end{pmatrix} + \begin{pmatrix} -5 & 5 \\ -10 & -20 \end{pmatrix} + \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix}$. Adding these indeed gives the zero matrix.

Example. Consider the matrix $A = \begin{pmatrix} 3 & 2 & -3 \\ 2 & 3 & -3 \\ 4 & 2 & -4 \end{pmatrix}$.

We want to find its characteristic polynomial. This is

$$\det\begin{pmatrix} 3-\lambda & 2 & -3 \\ 2 & 3-\lambda & -3 \\ 4 & 2 & -4-\lambda \end{pmatrix}$$

$$= (3-\lambda)\left((3-\lambda)(-4-\lambda) - (-3)(2)\right) - 2\left(2(-4-\lambda) - (-3)(4)\right) - 3\left(2(2) - 4(3-\lambda)\right)$$
$$= -\lambda^3 + 2\lambda^2 + \lambda - 2 = \left(\lambda^2 - 1\right)(2-\lambda) = (\lambda-1)(\lambda+1)(2-\lambda)$$

This gives roots -1, 1, 2. We could find eigenvectors by solving equations like before, but that is not the focus.

Note that if we plug in $\lambda = 0$ we will get 2 from the characteristic polynomial which means $\det(A) = \det(A - 0I) = 2$ so A has an inverse.

There is a neat way we can use the Cayley Hamilton theorem to find the inverse. We note that $-A^3 + 2A^2 + A - 2I = 0$, so $I = \frac{-A^3+2A^2+A}{2}$. Multiplying both sides by A^{-1} gives that $A^{-1} = \frac{-A^2+2A+I}{2}$. Thus the only difficult thing to do is finding A^2.

6.10 Symmetric and orthogonal matrices

In this section, see level 6.1 to understand why anything I am saying is true.

Definition. A matrix M is orthogonal if $M^{-1} = M^T$. Equivalently, if when we write out the columns as vectors, they each have length 1 but any 2 distinct ones are perpendicular.

Definition. A matrix M is symmetric if $M = M^T$.

Now here is a not so obvious fact. When diagonalizing a symmetric matrix A, if there are no repeated Eigenvalues, and when building the matrix P we always use normalized Eigenvectors, it is always the case that P is orthogonal and so $P^{-1}AP = D$ and $P^TAP = D$ are equivalent expressions. It's mad that my A level syllabus SPECIFICALLY focuses on diagonalizing SYMMETRIC matrices when the procedure is THE SAME as for general matrices only WE ADD AN ADDITIONAL MAGICAL NOT OBVIOUS FACT ABOUT THE TRANSPOSE UNNECESSARILY WHEN THE THING I HATE THE MOST AND THE WHOLE REASON I EVEN MADE THIS WEBSITE IN THE FIRST PLACE IS HAVING TO USE MAGICAL NOT OBVIOUS FACTS.

With that rant out of the way, lets do an example.

Example. We will diagonalize the symmetric matrix $A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 2 \end{pmatrix}$.

The first step is to determine $\det(A - \lambda I)$. This is

$$\det \begin{pmatrix} 2-\lambda & 0 & 1 \\ 0 & 4-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{pmatrix} = (2-\lambda)\left((4-\lambda)(2-\lambda)\right) + \left(-1(4-\lambda)\right)$$

$$= (4-\lambda)\left((2-\lambda)^2 - 1\right) = (4-\lambda)\left(\lambda^2 - 4\lambda + 3\right) = (4-\lambda)(1-\lambda)(3-\lambda)$$

Thus, the eigenvalues are $1, 3, 4$. We need to find corresponding Eigenvectors. For the eigenvalue 1 we need to solve

$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

This is equivalent to $2x + z = x$ (implies $z = -x$), $4y = y$ (implies $y = 0$) and $x + 2z = z$ (implies $z = -x$ again). Therefore, $(1, 0, -1)$ is an eigenvector that works. Thus, a normalized eigenvector would be $\frac{1}{\sqrt{2}}(1, 0, -1)$.

For the eigenvalue 3 we need to solve

$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \\ 3z \end{pmatrix}$$

This is equivalent to $2x + z = 3x$ (implies $z = x$), $4y = 3y$ (implies $y = 0$) and $x + 2z$ (implies $z = x$ again). Therefore, $(1, 0, 1)$ is an eigenvector that works, thus a normalized eigenvector would be $\frac{1}{\sqrt{2}}(1, 0, 1)$.

For the eigenvalue 4 we need to solve

$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4x \\ 4y \\ 4z \end{pmatrix}$$

This is equivalent to $2x + z = 4x$ (implies $z = 2x$), $4y = 4y$ (implies nothing), and $x + 2z = 4z$ (implies $x = 2z$). Since $x = 2z$ and $z = 2x$ it follows that $x = 4x$ so $x = z = 0$. Therefore, a possible eigenvector, which is already normalized, is $(0, 1, 0)$.

Thus the matrix of normalized eigenvectors is

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}$$

This was in the order 1,3,4. Thus, P is a matrix such that $P^T A P$ is a diagonal matrix D equal to

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

7 Vectors

7.1 Cartesian and vector equations for a line

As you know, a line through a vector v in the direction of w is the set of points of the form $v + \lambda w$ where λ varies. This is the vector equation.

There is also a cartesian equation for a line: For example consider the line

$$\begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -5 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 + 2\lambda \\ 3 - 5\lambda \\ 4 + 3\lambda \end{pmatrix}$$

Definition. Given this line, the direction cosines are the components of the normalized component of the thing multiplying λ which is the direction the lines go in. In the above case to normalize the vector, by Pythagoras we have to divide it by $\sqrt{38}$, so the direction cosines will be $\frac{2}{\sqrt{38}}$ for the x direction, $-\frac{5}{\sqrt{38}}$ for the y direction and $\frac{3}{\sqrt{38}}$ for the z direction.

Thus we have $x = 1 + 2\lambda$ so $\lambda = \frac{x-1}{2}$. By doing the same thing for y and z, we get that an alternative equation for the line is

$$\lambda = \frac{x-1}{2} = \frac{y-3}{-5} = \frac{z-4}{3}$$

This is a "cartesian" equation because it involves x, y, z rather than being a vector equation (which involves vectors directly)

well, this works only when none of the components of the direction vector are 0.

7.2 Cartesian and vector equations for a plane

We saw in the chapter on matrices that $ax + by + cz = d$ is the cartesian equation for a plane. Note, however, that if $n = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ and $r = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, and n is a fixed vector and r is free to vary, then an equivalent vector formulation of a plane is the equation $n.r = d$.

Note that if v and w are both on the plane, then $n.v = n.w$ so $n.(v - w) = 0$ so $v - w$ is perpendicular to n. This means n is the vector perpendicular to all lines that lie on the plane, i.e. the "normal" to the plane. So this is the vector way of determining a plane.

7.3 The cross product

The cross product is an operation that takes 2 3D vectors and outputs another 3D vector. You can find it by the following formula:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}$$

If i, j, k are the basis vectors in the x, y, z directions respectively then a notational trick you can do is to write this compactly as

$$\det \begin{pmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

since finding the determinant of this actually gives the same result as above.

The cross product has the following properties which we will prove in level 6.1, but for now it is yet another magic rule that we're expected to take on trust, it's getting ridiculous I didn't even realize how many there were.

- $a \times b$ is perpendicular to both a and b

- The length of $a \times b$ is the area of the parallelogram with sides a and b that passes through the origin.

These above properties give 2 possibilities for $a \times b$. However, $a \times b$ always follows the "right hand rule": if your index finger points in the a direction and your middle finger points in the b direction then your thumb will point in the $a \times b$ direction.

We can use the cross product and both definitions to solve problems such as

Example. Let's find the equation of the plane passing through $a := (4, 1, 0)$, $b := (-1, 3, -1)$ and $c := (2, 2, 2)$

We want to find the normal vector to the plane. We note that $b - a$ and $c - a$ will both be perpendicular to this normal vector, and a vector perpendicular to both of these is given by $(b-a)\times(c-a)$ which is $\begin{pmatrix}-5\\2\\-1\end{pmatrix}\times\begin{pmatrix}-2\\1\\2\end{pmatrix}=\begin{pmatrix}5\\12\\-1\end{pmatrix}$ using the algebraic formula for the cross product. We can even verify using the dot product that the result is indeed perpendicular to both.

Thus, we know the normal vector so the equation for the plane will be $5x+12y-z=d$ for some unknown d. To find d, just pick any point, say a and note that it is on the plane so we must have $5*4+12*1-0=d$, so $d=32$ and the plane is given by $5x+12y-z=32$.

Example. We want to find the area of the parallelogram spanned by the vectors $(2,0,1)$ and $(0,1,1)$. We can compute their cross product:

$$\begin{pmatrix}2\\0\\1\end{pmatrix}\times\begin{pmatrix}0\\1\\1\end{pmatrix}=\begin{pmatrix}-1\\-2\\2\end{pmatrix}$$

The absolute value of the cross product is $\sqrt{1^2+2^2+2^2}=3$ by Pythagoras so the area is 3.

The parallelogram in question is shown in figure 20.

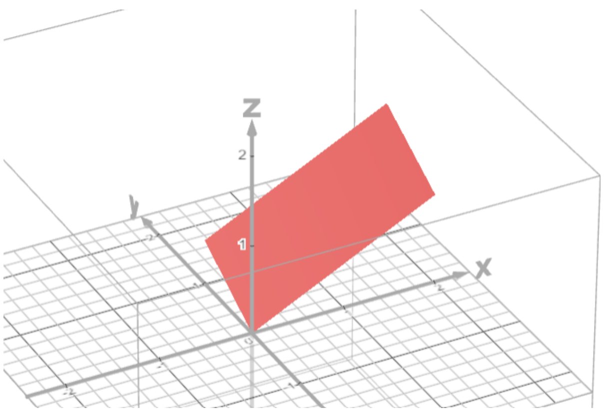

Figure 20

Note that the cross product of 2 parallel vectors, or 2 vectors one of which is 0, is always 0. Because of this, if we have a line of the form $r=a+\lambda b$ where r is an arbitrary point on the line, a,b are fixed and λ is any real number, we can write this as $r-a=\lambda b$ which is equivalent to $(r-a)\times b=0$. This is another way to write a vector equation for a line.

7.4 Finding the shortest distance between 2 lines

To find the shortest distance between 2 lines, we will prove and then use a formula. We will first consider lines that are not parallel.

Visually, from figure 21 it is easy to see that whatever the shortest distance is will in fact be perpendicular to both lines.

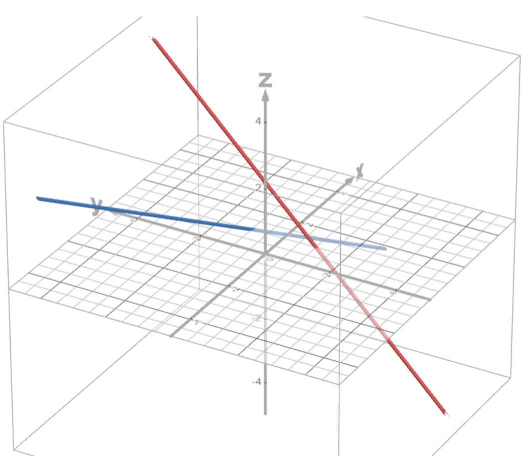

Figure 21

The idea is we want to take the line segment joining *any* point on one line with a point on the other and project it onto the direction perpendicular to both lines.

Recall from level 3 that the projection of a onto the direction of b has length $\frac{|a.b|}{|b|}$, so the projection will be this quantity times the normalized b.

Suppose the lines are given by $a+\lambda c$ and $b+\mu d$. Then we want to project $b-a$ onto the direction perpendicular to both lines which is $c \times d$. Thus, by the formula above, the distance will be $\frac{|(b-a).(c\times d)|}{|c\times d|}$.

Example. The lines shown in figure 21 are the lines

$$\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

and

$$\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}$$

Thus, using the formula, we want

$$\frac{\left| \left(\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right) . \left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \right) \right|}{\left| \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \right|}$$

This is

$$\frac{\left| \begin{pmatrix} -2 \\ 3 \\ -1 \end{pmatrix} . \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} \right|}{\left| \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} \right|} = \frac{|8|}{|2\sqrt{2}|} = 2\sqrt{2}$$

Hence the distance is $2\sqrt{2}$.

If the lines are parallel, then we again want the perpendicular distance, but the problem now is that we cannot take the cross product as we will just get 0. A formula that works is

$$\frac{|(b-a)\times d|}{|d|}$$

where d is the direction that the lines go in.

To prove this, note that if b' is the closest point to a on the line through b, it is clear that $\left|(b'-a)\times \frac{d}{|d|}\right|$ gives the distance, as its length will be the area of a parallelogram with 2 perpendicular sides one of which has the length we want. Thus the distance is given by

$$\frac{|(b'-a)\times d|}{|d|} = \frac{|(b'-b)\times d + (b-a)\times d|}{|d|} = \frac{|(b-a)\times d|}{|d|}$$

since $b'-b$ and d are parallel.

7.5 Finding the distance between a point and a plane

Suppose we have a point x_0, which we consider as a vector, and a plane with equation that r satisfies $r.n = p$ for some vector n and number p, then the closest distance from x_0 to the plane is given by the equation

$$\frac{|n.x_0 - p|}{|n|}$$

Less compactly, if the plane has equation $Ax + By + Cz = d$ then the distance from this plane to the point (x_0, y_0, z_0) is given by

$$\frac{|Ax_0 + By_0 + Cz_0 - D|}{\sqrt{A^2 + B^2 + C^2}}$$

We will prove this formula in level 6.1.

Example. Let's find the closest distance between the plane $3x + 12y - 4z = 7$ and the point $(3, 2, 5)$

Then the formula gives

$$\frac{|3 * 3 + 12 * 2 - 4 * 5 - 7|}{\sqrt{3^2 + 12^2 + 4^2}} = \frac{6}{\sqrt{169}} = \frac{6}{13}$$

So the distance is $6/13$.

7.6 Finding the angle between 2 planes

To explain the procedure to find the angle between 2 planes, we start with figure 22 which shows 2 planes and their normal vectors.

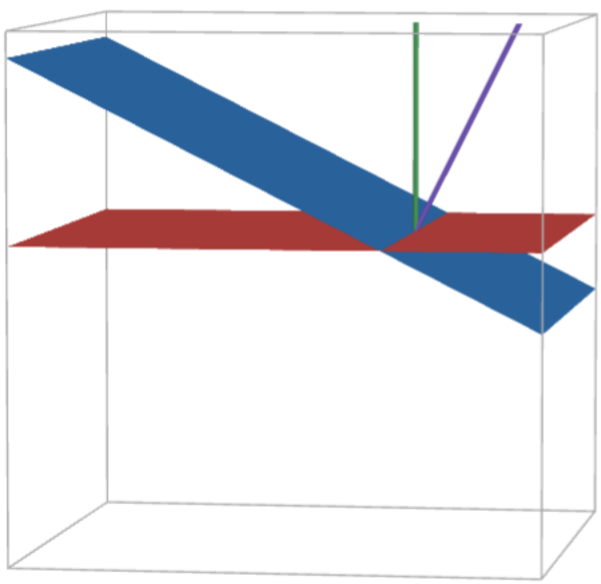

Figure 22

Now hopefully by looking at this picture you are convinced that we can reduce the problem to finding the angle between 2 normal vectors. If their dot product is negative take minus one of them so we can find the *acute* angle between the planes. This is now easy, as given the equation for a plane we can read off the normal vectors from the coefficients of the plane equation. We know how to find the angle between vectors from level 3. Moving on.

7.7 Finding the angle between a line and a plane

This one is simple. Just find the acute angle between the line in question and the plane's normal using the dot product way we've seen in level 3 and then take 90 degrees minus that.

7.8 Finding the line of intersection of 2 planes

We want a line in a direction perpendicular to the normal of both planes, assuming they intersect but do not coincide so the normals are non-parallel. This will be parallel to the line of intersection of the planes. To find a line perpendicular to both normals we take their cross product.

This gives the direction of the line. However, to find the line we need to find any point on the 2 planes. This is a bit fiddly.

We'll have 2 equations in 3 variables with a whole line of solutions and we want to find any solution.

As an example, suppose we want a point on the planes

$$x + y + 2z = 4$$

$$3x - 2y - 4z = 7$$

We can first try assuming $x = 0$ and solving the resulting system of equations, which will be $y + 2z = 4$ and $-2y - 4z = 7$. However, this system of equations does not have compatible solutions. Therefore there are no solutions with $x = 0$ which geometrically means the line of intersection is perpendicular to the x-axis. Note that this failure is atypical and I just wanted to show it. If we assume y is 0 we can find a point just fine. It will be the solution to $x + 2z = 4$ and $3x - 4z = 7$ which in this case turns out to be $z = \frac{1}{2}$ and $x = 3$ as a point on the plane.

Now that we have a point, I'll quickly find the line. A vector perpendicular to both normals is $\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 3 \\ -2 \\ -4 \end{pmatrix} =$ $\begin{pmatrix} 0 \\ 10 \\ -5 \end{pmatrix}$

As a hint: If you find the direction first, the thing you want to set constant to get a system of 2 equations is any coefficient that is NOT zero.

7.9 Use of the cross product to find volumes

Definition. A parallelepiped is a 3D shape with 6 sides such that any 2 opposite sides are part of parallel planes. So it is like the 3d version of a parallelogram. Figure 23 shows an example.

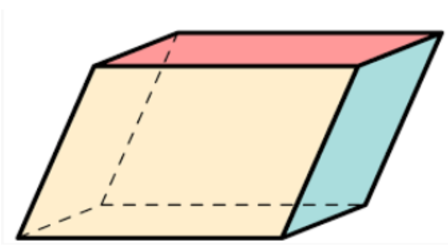

Figure 23

Proposition. Suppose we have 3 vectors with direction vectors a, b, c originating from the same point. Then the volume of the parallelepiped is given by $|a.\,(b \times c)|$, or equivalently $|b.\,(c \times a)|$ or $|c.\,(a \times b)|$.

Proof. Suppose the parallelogram spanned by b and c is at the base, then its area is $|b \times c|$ so the volume of the parallelepiped is $|b \times c|h$ where h is its height in the direction perpendicular to the base.

Its height is equal to $|a|$ times the cosine of the angle between a and the direction perpendicular to the base. Thus we have $|b \times c||a| \cos((b \times c)\, Oa) = |(b \times c)\, .a|$

By the same argument by putting a different thing as the base we get the other 2 equivalent forms.

□

Definition. A tetrahedron is a 3D shape with 4 triangle sides that looks like figure 24.

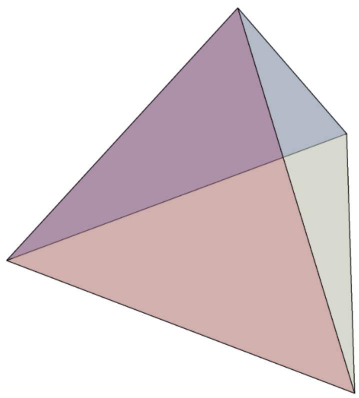

Figure 24

To find the volume of a tetrahedron, we have to do $\frac{1}{3}bh$ where b is the area of the base and h is the height. We will prove why this is the case in level 6.1. Since the area of the base is half the area of the parallelogram, we have the following fact:

The volume of a tetrahedron with edges a, b, c is $\frac{1}{6}$ the volume of the parallelepiped, i.e. $\frac{1}{6}\left|a.\left(b \times c\right)\right|$.

I'll find, given 4 points, the volume of the tetrahedron with those 4 points, as an example.

Example. Consider the 4 points

$$\begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 4 \\ 5 \end{pmatrix}$$

Then call these points A,B,C,D respectively. Then we want

$$\frac{1}{6}\left|\left(B - A\right).\left(\left(C - A\right) \times \left(D - A\right)\right)\right|$$

This is equal to

$$\frac{1}{6}\left|\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}.\left(\begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \times \begin{pmatrix} -1 \\ 3 \\ 6 \end{pmatrix}\right)\right|$$

$$= \frac{1}{6}\left|\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}.\begin{pmatrix} -3 \\ -11 \\ 5 \end{pmatrix}\right| = \frac{1}{6}\left|-36\right| = 6$$

Note that given 4 planes that enclose a tetrahedron, if we want to find the volume of the tetrahedron they enclose, we can convert it into a problem like the one above by finding the intersection between each of the sets of 3 planes by solving a system of equations. and finding the volume of the tetrahedron through those 4 vertices as we know how to do.

7.10 Projections of points onto planes

The projection of a point P onto a plane $r.n = d$ will be P moved by some multiple of the normal vector to the plane, i.e. $P + \lambda n$ for some λ. Thus we need to find λ such that $\left(P + \lambda n\right).n = d$. Ie, it will be the closest point between P and the plane.

Example. Consider the example above where we found the closest distance between the plane $3x + 12y - 4z = 7$ and the point $(3, 2, 5)$. Now I want to find the specific point on the plane that is closest to that point.

Actually, if you want you can use this method to find a point on a plane, as I demonstrated in the last section, by finding the point closest to the origin.

In this case, we want to solve for when

$$\left(\begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 12 \\ -4 \end{pmatrix}\right)$$

is on the plane, i.e. for when

$$\left(\begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 12 \\ -4 \end{pmatrix}\right) \cdot \begin{pmatrix} 3 \\ 12 \\ -4 \end{pmatrix} = 7$$

Ie, for when

$$(3 + 3\lambda)\, 3 + (2 + 12\lambda)\, 12 + (5 - 4\lambda)\, (-4) = 7$$

Ie, for when

$$(13 + 169\lambda) = 7$$

Thus $\lambda = -\frac{6}{169}$. Therefore the closest point will be

$$\begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} - \frac{6}{169} \begin{pmatrix} 3 \\ 12 \\ -4 \end{pmatrix}$$

We could simplify fractions or whatever to simplify this but you get the idea.

8 Conic sections

8.1 Ellipses

Definition. A conic section is a shape you get when you slice a cone.

Figure 25 shows the different possible conic sections.

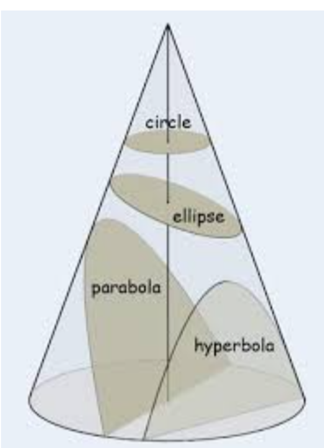

Figure 25

In this subsection we will talk about ellipses. We will give a bunch of random unmotivated definitions and state and not prove things about those definitions and defer the proofs to level 6.1, in order to keep in the spirit of A level textbooks. We will, of course, do the same for parabolas and hyperbolas.

Definition. An ellipse is basically a stretched out circle. An ellipse can be given by an equation like $ax^2 + by^2 = 1$, so it is like a circle with the x and y axes scaled. We will more conventionally write this as $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Essentially this gives an ellipse with a radius a along the x axis and b along the y axis.

Definition. Suppose that $a^2 > b^2$, in other words the ellipse is longer horizontally, in other words the x axis is the semimajor axis. Then the foci of the ellipse are the points at x coordinates $\pm\sqrt{a^2 - b^2}$

Definition. Suppose that the conditions above are satisfied. Then the eccentricity of the ellipse is defined as $\frac{\sqrt{a^2-b^2}}{a}$. We usually call this e. This is equivalent to $\sqrt{1 - \left(\frac{b}{a}\right)^2}$. This ranges between 0 and 1

The intuition for this definition is as follows. If it is a small number, i.e. close to 0, then the ellipse looks roughly like a circle. If it is a large number, i.e. close to 1, then the ellipse looks more elongated.

Proposition. If an ellipse is as above, then see figure 26: The claim is that as we move P around, the sum of the lengths of the lines connecting P to the foci is constant.

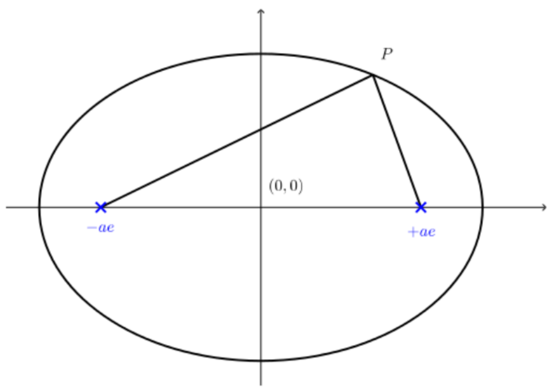

Figure 26

Proof. You really thought I was gonna prove this here? See level 6.1.

□

Assuming the proposition we actually do see that since the sum of the distances if $P = (a, 0)$ is $2a$ and the sum of the distances when $P = (0, b)$ must also be $2a$, it follows that $2\left(\sqrt{a^2e^2 + b^2}\right) = 2a$. Ie, $a^2e^2 + b^2 = a^2$, so we get the formula for e that we found above.

Oh don't worry we're not done with random definitions and non-proofs.

Definition. Again let an ellipse be as above with eccentricity e, then a directrix of the ellipse is one of the two lines $x = \pm\frac{a}{e}$.

Proposition. Let an ellipse be as above P be any point on the ellipse, F be a focus and l be the nearest directrix to F. Let x be the distance from F to P and y be the closest distance from P to l, then $x = ey$.

Proof. See level 6.1.

□

Note that since $b^2 = a^2\left(1 - e^2\right)$ it follows that the eccentricity is exactly the eccentricity of the ellipse

$$x^2 + \frac{y^2}{1 - e^2} = C$$

for any $C > 0$, noting that this can be written in "standard form" if we set $a = \sqrt{C}$.

Note that a point on an ellipse as above can be parametrized by $(a\cos(t), b\sin(t))$. It is easy to see from trigonometric identities that this satisfies $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

We can use the above definitions and some algebra to solve problems. I'll do an example.

Example. We will prove that at a point on an ellipse as above given by $(a\cos(\theta), b\sin(\theta))$, the normal to the ellipse has the equation

$$by = ax\tan(\theta) + \left(b^2 - a^2\right)\sin(\theta)$$

We will do this as follows. We will use a very useful idea that is important to know, that if a line passes through (x_1, y_1) and has slope m, then (it is easy to see why this is true), the equation for the line is

$$y - y_1 = m(x - x_1)$$

The slope of the normal to the ellipse will be (by properties of normal slopes discussed in levels 2 and 3)

$$-\frac{dx}{dy} = -\frac{dx/d\theta}{dy/d\theta} = -\frac{-a\sin(\theta)}{b\cos(\theta)} = \frac{a}{b}\tan(\theta)$$

So we have our m. We know that $(x_1, y_1) = (a\cos(\theta), b\sin(\theta))$. Therefore, by the general line equation above,

$$y - b\sin(\theta) = \frac{a}{b}\tan(\theta)(x - a\cos(\theta))$$

$$by - b^2\sin(\theta) = ax\tan(\theta) - a^2\tan(\theta)\cos(\theta)$$

Using $\tan(\theta)\cos(\theta) = \sin(\theta)$ and rearranging, the result follows.

Example. Determine the locus of complex numbers z such that $|z+1| + |z-1| = 4$

To do this we note that we want an ellipse with foci at 1 and -1, by properties of the focus.

To find the highest imaginary point on the ellipse, yi, we note that $|yi-1| + |yi+1| = 4$ so $2\sqrt{y^2+1} = 4$ so we find that $y = \sqrt{3}$. Similarly we find that the highest real point, x, must satisfy $|x-1| + |x+1| = 4$, and we can find, by a graph or otherwise, that $x = \pm 2$ satisfies this. Therefore we have an ellipse with $a = 2$ and $b = \sqrt{3}$, i.e. its equation is $\frac{x^2}{4} + \frac{y^2}{3} = 1$.

8.2 Parabolas

Definition. A parabola is the U shape you get from the graph $y = x^2$.

In this section, for reasons I don't understand, to keep in the spirit of the A level textbooks, we will consider parabolas on their side, i.e. with equations like $y^2 = 4ax$ for some constant a.

We say that the eccentricity of a parabola is 1. We can think of a parabola like an infinitely elongated ellipse in a sense.

Definition. The focus of a parabola $y^2 = 4ax$ is the point $(a, 0)$. The directrix of this parabola is the line $x = -a$.

Proposition. Let P be a point on the parabola $y^2 = 4ax$, then the distance from P to the focus equals the nearest distance from P to the directrix

Proof. It's not difficult, but since A level textbooks don't prove it, see level 6.1.

□

The parabola has the parametric equations $\left(at^2, 2at\right)$. We see that they indeed satisfy $y^2 = 4ax$.

We can solve problems similar to in the ellipse case.

8.3 Hyperbolas

Definition. A hyperbola is a shape you get from an equation like $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$. We could also have $\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1$ but we will just consider hyperbolas of the first type.

Definition. A rectangular hyperbola is a shape you get from an equation like $xy = c$ or $y = \frac{c}{x}$

To demonstrate what a hyperbola looks like, Figure 27 shows the curve with equation $x^2 - y^2 = 1$.

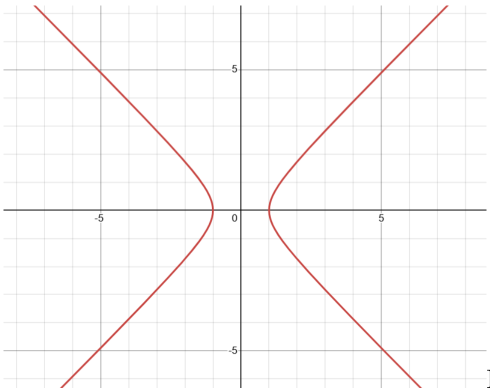

Figure 27

Ok here we go. Time for the usual unmotivated definitions and unproven theorems.

Definition. The eccentricity of a hyperbola, which always satisfies $e > 1$, is the positive number e such that the hyperbola can be written as $x^2 + \frac{y^2}{1-e^2} = C$ with $C > 0$, which is actually the same as for an ellipse, only now the denominators have different signs.

Definition. Let a be the unique positive number such that $(a, 0)$ is on the hyperbola, then the foci of the hyperbola are at $\pm ae$ and a directrix is one of the two lines $x = \pm\frac{a}{e}$.

Proposition. Let a hyperbola be as above P be any point on the ellipse, F be a focus and l be the nearest directrix to F. Let x be the distance from F to P and y be the closest distance from P to l, then $x = ey$.

Proposition. If a hyperbola is as above, then see figure 28: The claim is that as we move P around, the difference between the lengths of the lines connecting P to the foci is constant. This is unlike in the case of the ellipse where we were considering the sum of the lengths instead of the difference.

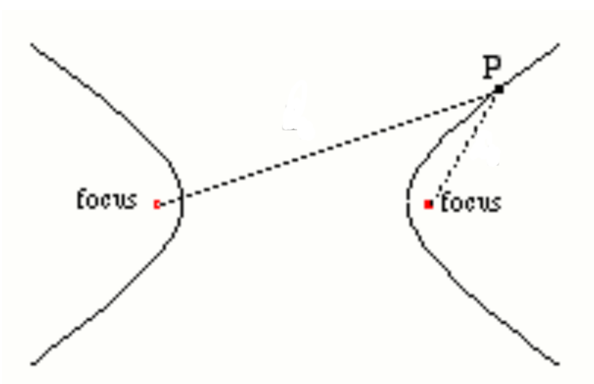

Figure 28

As usual see level 6.1 for the proof of both of these.

As for the parametric equations, we have the following:

- For a rectangular hyperbola if $xy = c^2$ then write $\left(ct, \frac{c}{t}\right)$

- For a hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ we can write either $(a\cosh(t), b\sinh(t))$ or $(a\sec(t), b\tan(t))$. Just be careful that if we use the former we will only get the positive x branch of the hyperbola unless we allow t to be complex. We ought to parametrize by $(\pm a\cosh(t), b\sinh(t))$

9 Solving recurrence relations

9.1 First order

Definition. A first order recurrence relation is when we are given $u_n = f(u_{n-1})$ and we want to solve for u_n. A second order recurrence relation would be something like $u_n = f(u_{n-1}, u_{n-2})$ and similarly for higher order.

We will give a few ways to solve first order recurrence relations.

If $u_n = Au_{n-1}$ then it is easy to see that $u_n = C * A^n$ where C is an arbitrary constant, and the value of C can be found if we are given an initial condition like the value of u_1.

If $u_n = u_{n-1} + f(n)$ it follows that

$$u_n = u_{n-2} + f(n) + f(n-1) = u_{n-3} + f(n) + f(n-1) + f(n-2) = ... = u_0 + \sum_{i=1}^{n} f(i)$$

In the most general case we may be given something like $u_n = Au_{n-1} + f(n)$. The obvious way to do this would be to divide everything through by A^n and reduce it to the previous case using the substitution $u_n = A^{-n}v_n$. if $A = 0$ we can't divide but then the solution $u_n = f(n)$ is right in your face so you're fine. But of course, the textbook does not do it that way, the textbook does it the hard way, so we will also do it the hard way.

The hard way is as follows. We are about to get PTSD from section 5 of this document. We write $u_n - Au_{n-1} = f(n)$. We find the complementary function, which is the general solution to $u_n - Au_{n-1} = 0$ which we know to be $C * A^n$. We then find a particular solution and we do this by guessing. If f is a polynomial or a function of the form $C * B^n$ we follow the same rule as in differential equations for our guess, and we also follow the magic-seemingly-made-up-if-you-haven't-seen-why rule that we guess that we multiply by n if the right hand side matches a term in the complementary function.

Since the same tricks to prove the methods for solving differential equations works for this, in level 6.1 there will not be much on this, just a brief discussion, as this actually has less technical details to worry about since we are working in the discrete world now.

The good news, though, is that it is easy to see that given u_0 and any solution that works, we can easily check it by induction, and we know it is the unique solution by the nature of the problem.

Example. Lets solve $u_n = u_{n-1} + n^2$ subject to $u_0 = 3$.

In this case, we do the sum rule. We have $u_n = 3 + \sum_{i=1}^{n} i^2 = 3 + \frac{n(n+1)(2n+1)}{6}$.

Example. Lets solve $u_n - 2u_{n-1} = -3$ subject to $u_0 = 4$.

To do this, we note that our complementary function is $C * 2^n$. We guess that our particular solution is a constant. We see that 3 works, so we have $C * 2^n + 3$. Given u_0 we can solve for C and we get that the solution is $u_n = 2^n + 3$.

Example. Lets solve $u_n - 2u_{n-1} = 2^n$ subject to $u_0 = 3$.

Our complementary function is as before. Since the right hand side term matches the complementary function we try $Cn * 2^n$. We substitute this in: We want $Cn * 2^n - 2C(n-1) * 2^{n-1} = 2^n$ so $2^n(Cn - C(n-1)) = 2^n$ so we see that $C = 1$. Therefore a particular solution is $n * 2^n$. So our general solution is $(n + A)2^n$. By the initial condition we find that $A = 3$ so the solution is $(n+3)2^n$.

9.2 Second order

We will focus on equations of the form $au_n + bu_{n-1} + cu_{n-2} = f(n)$, or equations that can be made into such an equation by a substitution.

We do this as follows. As usual we make up some magic rule that I say works because trust me bro. We let α, β be the roots of $ax^2 + bx + c = 0$. Then If $\alpha \neq \beta$ then our complementary function will be, for some arbitrary constants C and D that we would solve for given initial conditions, $C\alpha^n + D\beta^n$. If $\alpha = \beta$ we would instead have $(Cn + D)\alpha^n$. If α and β are a complex conjugate pair, we write that $\alpha = re^{i\theta}$ and $\beta = re^{-i\theta}$. Then our general solution will be $r^n\left(Ce^{in\theta} + De^{-in\theta}\right)$ which is actually just $r^n\left(C'\sin(n\theta) + D'\cos(n\theta)\right)$.

Example. We will solve $u_n - 4u_{n-1} + 8u_{n-2} = 0$. The roots of $x^2 - 4x + 8 = 0$ are $2 \pm 2i = 2\sqrt{2}\left(e^{\pm\frac{\pi i}{4}}\right)$. Therefore the general solution is

$$u_n = \left(2\sqrt{2}\right)^n \left(A\sin\left(\frac{n\pi}{4}\right) + B\cos\left(\frac{n\pi}{4}\right)\right)$$

If we impose that $u_0 = 3$ and $u_1 = 5$ we get the following equations:

If $n = 0$,

$$3 = u_0 = \left(2\sqrt{2}\right)^0 (B)$$

so $B = 3$. If $n = 1$ we get

$$5 = u_1 = \left(2\sqrt{2}\right)\left(A\sin\left(\frac{\pi}{4}\right) + 3\cos\left(\frac{\pi}{4}\right)\right) = \left(2\sqrt{2}\right)(A+3)\frac{\sqrt{2}}{2}$$

Therefore we have $2(A+3) = 5$ so $A = -\frac{1}{2}$ so the general solution is

$$u_n = \left(2\sqrt{2}\right)^n \left(-\frac{1}{2}\sin\left(\frac{n\pi}{4}\right) + 3\cos\left(\frac{n\pi}{4}\right)\right)$$

Example. We will solve $u_n - 2u_{n-1} + u_{n-2} = n$. Here it is tricky because by the "magic" rule - again see level 6.1 - our complementary function is $An + B$. Our particular solution matches a term on here, so we multiply by n and we try An^2 as a particular solution. We have to solve $A\left(n^2 - 2(n-1)^2 + (n-2)^2\right) = n$, but we will find that we want $2A = n$ which is a problem. What happens is we would try a particular solution $Cn + D$ since we have two collisions (one in the CF itself and one in the right hand term) we multiply by n^2 and try $Cn^3 + Dn^2$. It feels like nonsense magic, I know. We will do this and try solving

$$C\left(n^3 - 2(n-1)^3 + (n-2)^3\right) + D\left(n^2 - 2(n-1)^2 + (n-2)^2\right) = n$$

We get (with some algebra) $C(6n-6) + 2D = n$, so equating coefficients we get $C = \frac{1}{6}$ and $D = \frac{1}{2}$ so a particular solution that works is $\frac{1}{6}n^3 + \frac{1}{2}n^2$. Therefore our general solution is $\frac{1}{6}n^3 + \frac{1}{2}n^2 + An + B$. We could of course impose an initial condition yadda yadda you get the idea.

One cool application though is that we can prove that $\left(1+\sqrt{2}\right)^n$ is very close to an integer for large values of n. The recurrence relation $u_n = 2u_{n-1} + u_{n-2}$ with initial conditions $u_0 = u_1 = 2$ has the following

- The solution is $u_n = \left(1+\sqrt{2}\right)^n + \left(1-\sqrt{2}\right)^n$

- For all n, u_n is an integer.

Since $\left|1-\sqrt{2}\right| < 1$, it follows that $\left(1-\sqrt{2}\right)^n \to 0$ as n gets large, but since $\left(1+\sqrt{2}\right)^n + \left(1-\sqrt{2}\right)^n$ is an integer for every n it follows that $\left(1+\sqrt{2}\right)^n$ is a near-integer. We see this, where for example $\left(1+\sqrt{2}\right)^{10}$ is about 6725.99985132 which is indeed close to an integer, namely 6726.

9.3 An extra fun thing

This is not in A levels but it would be a crime to not show this.

Consider the recurrence relation $z_n = z_{n-1}^2 + c$ subject to $z_0 = 0$. We will allow c to be either real or complex. We note that if $c = 1$ then $z_1 = 1$, $z_2 = 2$, $z_3 = 5$, $z_4 = 26$, $z_5 = 677$ and so on and the sequence goes to infinity. On the other hand, if $c = -1$ then we have $z_1 = -1$, $z_2 = 0$, $z_3 = -1$ and we keep oscillating forever. In particular, the sequence stays bounded for $c = -1$ but goes to infinity for $c = 1$.

We then sketch on an Argand diagram the set of values of c for which this sequence stays bounded. When we do this, we get my favourite set of all time. It is called the mandelbrot set. An image of the set is shown in Figure 29. (For scale, the needle on the left goes until $c = -2$ and the crack on the right has the "critical point" at $c = \frac{1}{4}$). I'm so fascinated by the boundary of this thing.

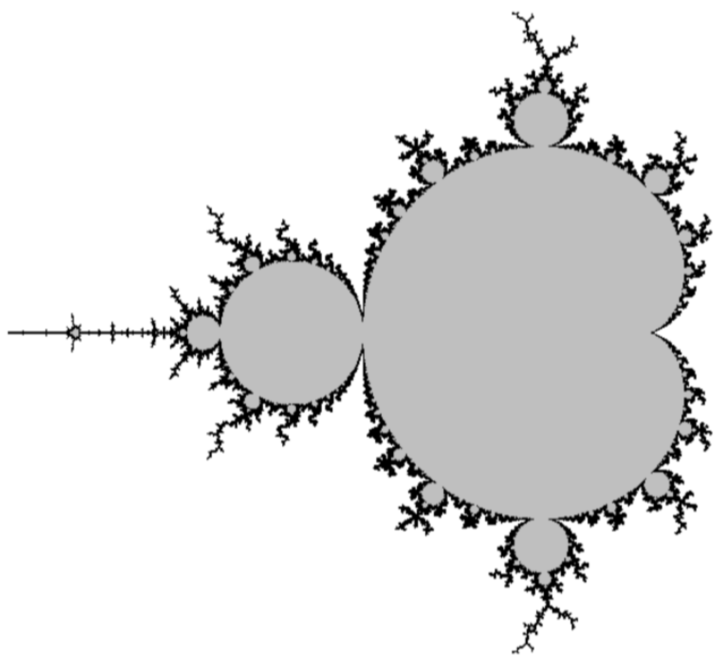

Figure 29

Also, remember the cardioid from section 4.4? It turns out that is exactly the shape of the main component.