

MAM1000W

Calculus Notes

By Dr Jonathan Shock



Calculus Notes by Dr Jonathan Shock licensed under a Creative Commons Attribution-Non-commercial 2.5 South Africa License. You are free to copy, distribute and make derivative works on condition that you give attribution to the author. To view this license, visit <http://creativecommons.org/license/by-nc/2.5/za/>.

Table of Contents

1. General statements	4
2. Basics – Integrating by substitution and by parts	5
2.1. The indefinite integral – The antiderivative	5
2.2. Integration by substitution	6
2.3. Integration by parts	8
3. Trig substitutions and trig integrals	10
3.1. Iteration formulae	13
3.2. Trig substitutions	13
4. Integration by partial fractions	15
5. Numerical integration	21
5.1. Integral approximations	21
5.2. The trapezoidal integral	22
5.3. Error estimates	23
5.4. Simpson's rule	25
5.5. The link between Simpson's rule and the fourth derivative of parabolic curves?	28
6. Improper integrals	30
6.1. Improper integrals of the first kind: Infinite integrals	31
6.2. Improper Integrals of the second kind: Discontinuous Integrand	35
6.3. The Comparison Theorem	37
7. Areas between curves	39
7.1. When to take the absolute area, and when the sign matters	47
7.2. Volumes	50
7.3. Breaking down a sphere into cylinders	52
7.4. Some more examples of volumes by cylindrical cross-sections	53
7.5. Volumes by cylindrical shells	60
7.6. Calculating the volumes of a torus using the method of cylindrical shells	63
7.7. Arc lengths	65
7.8. Average Value of a function	70
8. The binomial Theorem	74
8.1. Ordering objects	74
8.2. Choosing r objects from n objects	75
8.3. Pascal's identity	77
8.4. The Binomial Theorem	79
8.5. The multinomial expansion	82

9. Taylor and Maclaurin polynomials	83
9.1. Approximating a function with a power series	83
10. Complex numbers	97
10.1. A philosophical detour	97
10.2. A foray into a new number system	98
10.3. Plotting functions of complex numbers	105
10.4. Back to the algebra	106
10.5. More on the modulus/ argument form and exponential of a complex number	107
10.6. Another philosophical aside	110
10.7. Back to de Moivre	110
10.8. Calculating roots of complex numbers	115
10.9. Finding roots of polynomial equations	119
10.10. Complexity from complex numbers	121
11. Differential equations	128
11.1. Direction Fields and Euler's method	137
11.2. Separable differential equations	142
11.3. First order linear differential equations	147
11.4. Second Order differential equations	153
11.5. Summary	158
12. 3D geometry and vectors	160
12.1. Positions in three dimensions	161
12.2. Functions in three dimensions	162
12.3. Distances in three dimensions	163
13. Vectors	166
13.1. The scalar, or dot product	170
13.2. Scalar and vector projections	173
13.3. The vector, or cross product	175
13.4. Properties of the cross product	179
13.5. A multi-technique exercise	182

1 General statements, ideas and links

- These notes are for the second semester of MAM1000. They are neither complete nor exact and no responsibility is held for the accuracy within. That being said, I hope that they can be a useful resource in addition to the course textbook (Stewart) and additional online materials.
- I am always very grateful when people find mistakes in these notes. These may be in the form of spelling, grammar, calculational errors, typos in formulae, typesetting errors and anything else which doesn't seem to make sense. If an explanation is not clear, please contact me and I will do my best to explain it in another way. If you find errors, please email me at jonathan.shock@uct.ac.za and I shall make amendments to the notes as soon as I can.
- Many of the examples and some of the explanations, especially in the initial part of these notes come from the brilliant lectures which can be found and watched for free here: <http://www.centerofmath.org/video.html> I highly recommend watching them if anything here does not make sense.
- Take a look here: <http://www.bbc.co.uk/podcasts/series/maths> for a series of podcasts on the subject of mathematics. They are a lovely historical addition to the practical side of what we are doing here.
- This week we will start with a review of various integration techniques. You should already have studied these but they take a while and some practice to become natural. Believe me, that **with a total of a good few hours of practice they will be a breeze and you will have a good handle on not only how to get the answer, but also WHY and HOW they work - this is the difference between knowledge and understanding.**

In the next two lectures we will review:

- Integration by substitution
- Integration by parts
- Trigonometric integrals
- Trigonometric substitutions

2 Basics - Integrating by substitution and by parts

- There is an important distinction that we must make when looking at integrals. The two types that we will come across are **the indefinite and the definite integrals**. The former is in some way the most fundamental, and is used (**through the evaluation theorem**) to find the latter. The indefinite integral can be thought of as the opposite of differentiation - it is the antiderivative.
- The indefinite integral will allow you to make mathematical models of the way the world works, about how things vary and the dependences of one factor (age, price, strength, distance, area, risk...) on another, and then solve the equations that you come up with to find out how some system will behave. This will become clearer when we deal with differential equations.
- The definite integral allows you to calculate the area under a curve or the volume of a 3 dimensional (or higher!) object. This may not sound very interesting, but it allows you to calculate the total effect of something over time or space. It allows you to calculate the total population given varying population growth, or the total amount of money in a fund, given a certain behaviour over time of interest rates. It allows you to calculate the distance traveled given the speed, and it allows you to study volumes and surfaces which will have important effects on everything from economics to engineering, to energy consumption and much more besides.

2.1 The indefinite integral – The antiderivative

This is just a quick reminder. If you find **any** of this confusing, there is a very important trick for making it easier - practice, practice, practice! It doesn't take years to master this, it takes a few hours every week for a few weeks. You will become more and more familiar with the techniques and learn intuitively to know which technique to use in which situation.

Let's start with the very basics.

If $F(x) = x^2$ how do we solve:

$$f(x) = \frac{d(F(x))}{dx} \quad (1)$$

and find the answer is $f(x) = 2x$? We use our normal differentiation rules which you should now be very familiar with. How about if we didn't know $F(x)$ and wanted to find the answer to the equation:

$$2x = \frac{dF(x)}{dx} \quad (2)$$

The answer to this is $F(x) = x^2 + c$. When we differentiate the c we get zero. We write this as:

$$\int 2x dx = x^2 + c \quad (3)$$

If the derivative of $F(x)$ is $f(x)$ then the indefinite integral of $f(x)$ is $F(x) + c$.

You should be extremely familiar with the basic rules of integration by now. If not, practice over and over again until they are natural.

2.2 Integration by substitution

The trick is to become familiar enough with the examples to know what substitution will make the problem solvable. There are a fair number of tricks to know, which can't all be taught in class, but you can get them by studying the problem sheets in detail. It takes some practice but you will quickly understand WHY what works works. Example:

$$\int \cos(x^2)2x dx \quad (4)$$

What if we substitute $x^2 = u$:

$$\frac{du}{dx} = 2x \Rightarrow 2x dx = du \quad (5)$$

Let's substitute this into the original integral:

$$\int \cos(u) du = \sin(u) + c \quad (6)$$

we then substitute back again for x to get:

$$\sin(x^2) + c \quad (7)$$

We can note something important about the original integral:

$$\int \cos(x^2)2x dx \tag{8}$$

$2x$ is the derivative of x^2 which appears in the function $\cos x$. Let's say that $f(x) = \cos(x)$ and $g(x) = x^2$. Then we can write what we have as:

$$\int \cos(x^2)2x dx = \int f(g(x))g'(x) dx \tag{9}$$

But we know from the chain rule that:

$$\frac{dF(G(x))}{dx} = F'(G(x))G'(x) \tag{10}$$

We can integrate this equation to get:

$$\int \frac{dF(G(x))}{dx} dx = \int F'(G(x))G'(x) dx \tag{11}$$

but $\int \frac{dF(G(x))}{dx} dx = F(G(x)) + c$ which means that:

$$\int F'(G(x))G'(x) dx = F(G(x)) + c \tag{12}$$

or alternatively:

$$\int f(G(x))G'(x) dx = F(G(x)) + c \tag{13}$$

where $F(x)$ is the antiderivative (integral) of $f(x)$. In our example above. The antiderivative of $\cos(x)$ is $\sin(x) + c$ and so we have:

$$\int \cos(x^2)2x dx = \sin(x^2) + c \tag{14}$$

How about a definite integral by substitution:

$$\int_{e^1}^{e^2} \frac{1}{x(\ln x)^3} dx \tag{15}$$

This is really shorthand for:

$$\int_{x=e^1}^{x=e^2} \frac{1}{x(\ln x)^3} dx \tag{16}$$

Try $u = \ln x$ ($F'(x) = \frac{1}{x^3}$, $G(x) = \ln x$):

$$\frac{du}{dx} = \frac{1}{x} \quad (17)$$

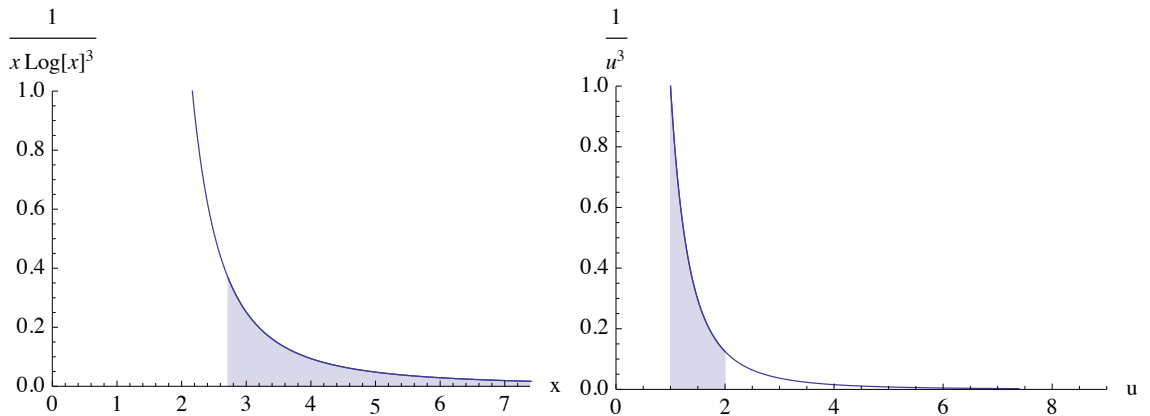
which means that:

$$\frac{1}{x} dx = du \quad (18)$$

We have to be careful that now we will be integrating over u and not x , so we have to change the limits too. When $x = e^1$, $u = 1$ and when $x = e^2$, $u = 2$. So the new integral is:

$$\int_1^2 \frac{1}{u^3} du = \left. \frac{-1}{2u^2} \right|_1^2 = \frac{3}{8} \quad (19)$$

Looking of the graphs written in both the x variable and the u variable they



look very different. The point is that the area is still the same. Both the integrand AND the limits much change correctly to get the right answer.

2.3 Integration by parts

We know the product rule for differentiation:

$$\frac{d(f(x)g(x))}{dx} = f'(x)g(x) + f(x)g'(x) \quad (20)$$

Integrating it gives us:

$$f(x)g(x) + c = \int f'(x)g(x)dx + \int f(x)g'(x)dx \quad (21)$$

We can then rearrange this to give:

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx \quad (22)$$

We have dropped the c because this will come also from doing the second integral. This is also sometimes written as:

$$\int u dv = uv - \int v du \quad (23)$$

Let's try a simple example:

$$\int x \ln x dx \quad (24)$$

If we take $f(x) = \ln x$ and $g'(x) = x$, then $f'(x) = \frac{1}{x}$ and $g(x) = \frac{x^2}{2}$. Then the equation becomes:

$$\int x \ln x dx = \frac{x^2}{2} \ln x - \int \frac{x^2}{2} \frac{1}{x} dx = \frac{x^2}{2} \ln x - \int \frac{x}{2} dx = \frac{x^2}{2} \ln x - \frac{x^2}{4} + c \quad (25)$$

A general rule, but not to be taken as sacred: When to use differentiation by parts: If you have a product of terms, one of which will simplify when you integrate it, and one of which will simplify when you differentiate it. The one that simplifies it when you differentiate it should be u , or alternatively $f(x)$ and the one that simplifies when you integrate it should be dv or alternatively $g'(x)$.

3 Trig substitutions and trig integrals

Trigonometric functions tell us about behaviour which happens in a periodic way. This might be related to the current flowing down an electrical circuit, the change in risk of certain weather patterns throughout the year or the position of a planet in the sky.

Here we're going to use a combination of simple integrals that we should all know by now and integration by substitution and parts along with trigonometric identities to solve certain integrals. Some of these integrals will be explicitly dependent on trigonometric functions and some of them will not, but trigonometric identities and integrals will come in very useful in solving them.

1. Let's look at an example. What is the solution to:

$$\int \tan \theta d\theta \quad (26)$$

Remember that $\tan \theta = \frac{\sin \theta}{\cos \theta}$ and so we can see that we have something that has the same pattern that we saw yesterday when looking at integration by substitution.

Let's try: $u = \cos \theta$ and so $du = -\sin \theta d\theta$ which gives us:

$$-\int \frac{1}{u} du = -\ln |u| + c \quad (27)$$

substitute back we get:

$$\int \tan \theta = -\ln |\cos \theta| + c = \ln |\sec \theta| + c \quad (28)$$

2. How about another example though a rather harder one:

$$\int \sec \theta d\theta \quad (29)$$

Look at the derivatives $(\tan \theta)' = \sec^2 \theta$ and $(\sec \theta)' = \sec \theta \tan \theta$. Now if you add them together you get:

$$(\sec \theta + \tan \theta)' = \sec^2 \theta + \sec \theta \tan \theta = \sec \theta (\sec \theta + \tan \theta) \quad (30)$$

It doesn't look very useful so far, but let's look at a u substitution of the form $u = (\sec \theta + \tan \theta)$, then $du = \sec \theta (\sec \theta + \tan \theta) d\theta$. This is

very interesting though because we can see that they differ only by a factor of $\sec \theta$:

$$\frac{du}{d\theta} = u \sec \theta \quad (31)$$

so $\frac{1}{u}du = \sec \theta d\theta$ performing this substitution we get:

$$\int \sec \theta d\theta = \int \frac{1}{u} du = \ln |u| + c = \ln |\sec \theta + \tan \theta| + c \quad (32)$$

3. Let's look at another integral:

$$\int \sec^3 \theta d\theta \quad (33)$$

whenever you see an odd power of a trigonometric function, look at splitting it off as follows:

$$\int \sec^3 \theta d\theta = \int \sec^2 \theta \sec \theta d\theta \quad (34)$$

we see a trig identity that $\sec^2 \theta = 1 + \tan^2 \theta$. So now we have:

$$\int \sec \theta (1 + \tan^2 \theta) d\theta = \int \sec \theta d\theta + \int \sec \theta \tan^2 \theta \quad (35)$$

but we already know the first part and the second part we will integrate by parts:

$$\ln |\sec \theta + \tan \theta| + \int (\tan \theta)(\sec \theta \tan \theta) d\theta \quad (36)$$

when we integrate by parts: $f(\theta) = \tan(\theta)$ ($f'(\theta) = \sec^2 \theta$) and $g'(\theta) = \sec \theta \tan \theta$. We know that the derivative of $\sec \theta$ is $\sec \theta \tan \theta$ so $g(\theta) = \sec \theta$. Now we get, for the integration by parts:

$$\int (\tan \theta)(\sec \theta \tan \theta) d\theta = \tan \theta \sec \theta - \int \sec^3 \theta \quad (37)$$

But note that the last term is what we started with, so we have:

$$\int \sec^3 \theta d\theta = \ln |\sec \theta + \tan \theta| + \tan \theta \sec \theta - \int \sec^3 \theta \quad (38)$$

so:

$$2 \int \sec^3 \theta d\theta = \ln |\sec \theta + \tan \theta| + \tan \theta \sec \theta + c \quad (39)$$

and:

$$\int \sec^3 \theta d\theta = \frac{1}{2} (\ln |\sec \theta + \tan \theta| + \tan \theta \sec \theta) + c \quad (40)$$

4. Another example:

$$\int \sin^5 \theta d\theta = \int \sin^4 \theta \sin(\theta) d\theta \quad (41)$$

We can use the identity $\sin^2 \theta + \cos^2 \theta = 1$ to get:

$$\int (1 - \cos^2 \theta)^2 \sin(\theta) d\theta \quad (42)$$

Now make the substitution: $u = \cos(\theta)$ which implies $-\sin(\theta)d\theta = du$, but we have something of this form already in the integrand:

$$\int -(1 - u^2)^2 du = \int -(1 - 2u^2 + u^4) du = -u + \frac{2}{3}u^3 - \frac{1}{5}u^5 + c \quad (43)$$

plug in the substitution again to get the final answer:

$$-\cos(\theta) + \frac{2}{3} \cos^3 \theta - \frac{1}{5} \cos^5 \theta + c \quad (44)$$

We can use this trick for all odd powers of $\sin(\theta)$ or $\cos(\theta)$ though for higher powers we will need to use an iterative expression that we can get from integrating by parts:

$$\int \sin^n x dx = -\frac{1}{n} \cos x \sin^{n-1} x + \frac{n-1}{n} \int \sin^{n-2} \quad (45)$$

for $n \geq 2$.

5. For even powers of \cos and \sin we can use the double angle formula:

$$\int \cos^2 \theta d\theta = \int \frac{1}{2}(1 + \cos 2\theta) d\theta \quad (46)$$

Substitute $2\theta = u, 2d\theta = du$ to get:

$$\int \frac{1}{4}(1 + \cos u) du = \frac{u}{4} + \frac{\sin u}{4} + c \quad (47)$$

substituting back in again, and we get:

$$\frac{\theta}{2} + \frac{\sin(2\theta)}{4} + c \quad (48)$$

If we had started with a higher even power of $\cos(\theta)$ we would simply have needed to use the double angle formula multiple times.

6. Another example:

$$\int \frac{\cos(\theta)}{\sqrt{1 + \sin(\theta)}} d\theta \quad (49)$$

Note that $\cos(\theta) = \frac{d\sin(\theta)}{d\theta}$ and try this yourself.

7. How about:

$$\int (1 + \tan \theta)^2 d\theta \quad (50)$$

Let's multiply it out:

$$\int 1 + \tan^2 \theta + 2 \tan \theta d\theta \quad (51)$$

We know the first part:

$$\int 1 + \tan^2 \theta d\theta = \tan \theta + c \quad (52)$$

If you don't know this, make sure that you understand it. The second part is:

$$\int \tan \theta d\theta = \int \frac{\sin(\theta)}{\cos(\theta)} d\theta \quad (53)$$

substitute $u = \cos(\theta)$ and you will get the final answer as:

$$\int (1 + \tan \theta)^2 d\theta = \tan \theta - 2 \log |\cos(\theta)| + c \quad (54)$$

3.1 Iteration formulae

You can use this to calculate the integral of the n^{th} power of \cos

$$\int \cos^n \theta d\theta = \frac{1}{n} \cos^{n-1} \theta \sin \theta + \frac{n-1}{n} \int \cos^{n-2} \theta d\theta \quad (55)$$

For $n \geq 2$. This comes from integrating by parts:

$$\int \cos^n \theta d\theta = \int \cos^{n-1} \theta \cos \theta d\theta \quad (56)$$

3.2 Trig substitutions

Sometimes we have an expression that doesn't appear to have anything to do with trig functions but actually we can solve them using trig substitutions

1. How about if we have an integral of the form:

$$\int \frac{1}{\sqrt{1-x^2}} dx \quad (57)$$

this will simplify greatly if we substitute $x = \sin u$ which gives $dx = \cos u du$. Note that the integral only makes sense between $-1 < x < 1$, therefore we will consider only u between $-\frac{\pi}{2} < u < \frac{\pi}{2}$. This will be important for taking the square root. Performing this substitution we get:

$$\int \frac{1}{\sqrt{1-\sin^2 u}} \cos u du = \int \frac{1}{\sqrt{\cos^2 u}} \cos u du \quad (58)$$

Because of the limits on u , $\cos u$ is always positive and so this is equal to:

$$\int 1 du = u + c = \arcsin(x) + c \quad (59)$$

Let's look at another example:

$$\int \frac{\sqrt{x^2-1}}{x} dx \quad (60)$$

In this case the appropriate substitution turns out to be $x = \sec u$ which means that $u = \arccos \frac{1}{x}$ and $dx = \frac{\sin u}{\cos^2 u} du$. Substituting this in we get:

$$\int \frac{\sqrt{\frac{1}{\cos^2 u} - 1}}{\frac{1}{\cos u}} \frac{\sin u}{\cos^2 u} du = \int \frac{\sin^2 u}{\cos^2 u} du = \int \tan^2 u du = \tan u - u + c \quad (61)$$

replacing the substitution we get:

$$\sqrt{x^2-1} - \arccos\left(\frac{1}{x}\right) \quad (62)$$

This isn't too obvious. Why is this last step correct?

2. Let's look at the integral:

$$\int \sqrt{a^2-x^2} dx \quad (63)$$

This is the integral along the arc of a circle. Circles have something to do with trig functions, so it isn't all that surprising that this might be related to a trig substitution. We remind ourselves that: $1 - \sin^2 \theta =$

$\cos^2 \theta$. If we can get the function to look like $1 - \sin^2 \theta$ things might simplify. Let's try a substitution: $x = a \sin \theta$, $dx = a \cos \theta d\theta$:

$$\int \sqrt{a^2 - x^2} dx = \int a \sqrt{1 - \sin^2 \theta} a \cos \theta d\theta = \int a^2 \cos^2 \theta d\theta \quad (64)$$

But now we can use the half angle formula to solve this: $\cos^2 \theta = \frac{1}{2}(1 + \cos(2\theta))$. We were able to let $\sqrt{\cos^2 \theta}$ because $-|a| \leq x \leq |a|$ if this is true then θ has to be between $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$ which means that $\cos \theta$ is positive and so $\sqrt{\cos^2 \theta} = \cos \theta$. Now:

$$\begin{aligned} \int a^2 \cos^2 \theta d\theta &= \int \frac{a^2}{2} (1 + \cos(2\theta)) = a^2 \frac{1}{2} \left(\theta + \frac{\sin(2\theta)}{2} \right) + c \\ &= \frac{a^2}{2} \left(\arcsin \frac{x}{a} + \frac{\sin(2 \arcsin(\frac{x}{a}))}{2} \right) + c \end{aligned} \quad (65)$$

but $\sin 2\theta = 2 \sin \theta \cos \theta$ thus $\sin 2 \arcsin \theta = 2\theta \sqrt{1 - \theta^2}$ so:

$$\int \sqrt{a^2 - x^2} dx = \frac{a^2}{2} \left(\arcsin \frac{x}{a} + \frac{x \sqrt{a^2 - x^2}}{a} \right) + c \quad (66)$$

There are many tricks to learn, but as you practice you will figure them out for yourselves.

4 Integration by partial fractions

Study appendix G in Stewart. There are more details about the current topic in there than we can do in a lecture.

1. What if you had to perform the following integral:

$$\int \frac{x+5}{x^2+x-2} dx \quad (67)$$

This looks hard, but what about:

$$\int \frac{2}{x-1} - \frac{1}{x+2} dx \quad (68)$$

This is much easier....but in fact it's exactly the same! Check that:

$$\frac{2}{x-1} - \frac{1}{x+2} = \frac{x+5}{x^2+x-2} \quad (69)$$

Splitting a fraction of polynomials into partial fractions makes things much more tractable. There are a number of steps to consider. We will study each one in detail by examples. Don't worry if you don't understand these yet, you should by the end of the lecture:

- (a) Convert the integrand into a proper rational form
 - (b) Factor the denominator into degree 1 terms $(x-r)$ (which may be repeated: $(x-r)^m$) and irreducible degree 2 terms (ax^2+bx+c) (which may be repeated $(ax^2+bx+cb)^n$).
 - (c) Take the irreducible degree two terms and complete the square to get: $((x-r)^2+b^2)^n$
 - (d) write out a sum of terms, one (or more if the factor is repeated) from each of the terms in the denominator
 - (e) solve the integrals in the sum one by one, either directly or by integration by substitution.
- Put your fraction into a proper rational form: The order of the polynomial on top must be smaller than the order of the polynomial on the bottom:

If this is not already the case:

$$\frac{P(x)}{Q(x)} = S(x) + \frac{R(x)}{Q(x)} \quad (70)$$

This step is done by long division of polynomial expressions. e.g.:

$$\frac{x^5+3x^2+6x-2}{x^3-7} = x^2 + \frac{10x^2+6x-2}{x^3-7} \quad (71)$$

Check this!

- Now we have a term that we have to integrate of the form $\frac{P(x)}{Q(x)}$ and we want to write it as a sum of terms of the form:

$$\left(\frac{A}{ax+b}\right)^i, \left(\frac{Ax+B}{ax^2+bx+c}\right)^i \quad (72)$$

eg:

$$\frac{5x-7}{x^2-7x+10} = \frac{5x-7}{(x-2)(x-5)} \quad (73)$$

Calculate the bottom with the quadratic formula. This is equal to:

$$\frac{A}{x-2} + \frac{B}{x-5} \quad (74)$$

Multiply this through and we get:

$$\frac{x(A+B) - 5A - 2B}{(x-2)(x-5)} \quad (75)$$

we want A and B such that the top = $5x - 7$ which is true for $A = -1$ and $B = 6$. So we have that the integral is:

$$\int \frac{5x-7}{x^2-7x+10} dx = \int \frac{-1}{x-2} + \frac{6}{x-5} dx \quad (76)$$

and this integral we can do.

2. How about something of the form:

$$\frac{Ax+b}{ax^2+bx+c} \quad (77)$$

where $b^2 - 4ac < 0$. At the moment we can't solve this in the form $(x+a_1)(x+a_2)$. Instead we complete the square. For example:

$$\int \frac{3x+2}{x^2+2x+2} dx \quad (78)$$

Here $b^2 - 4ac = -4 < 0$:

$$\int \frac{3x+2}{x^2+2x+2} dx = \int \frac{3x+2}{(x+1)^2+1} dx \quad (79)$$

Let $u = x + 1$:

$$\int \frac{3(u-1)+2}{u^2+1} du = \int \frac{3u}{u^2+1} - \frac{1}{u^2+1} du \quad (80)$$

We can solve this with trig substitutions. Show that:

$$\int \frac{3x+2}{x^2+2x+2} dx = \frac{3}{2} \ln |(x+1)^2+1| - \tan^{-1}(x+1) + c \quad (81)$$

3. How about if a factor is repeated?

$$\frac{3x - 6}{x^2(x - 1)^2} \quad (82)$$

For every repeated factor we write out every power of up to the power in the fraction:

$$\frac{3x - 6}{x^2(x - 1)^2} = \frac{A}{x} + \frac{B}{x^2} + \frac{C}{x - 1} + \frac{D}{(x - 1)^2} \quad (83)$$

Again, multiply this through to calculate the values of A , B , C and D . Check that the answer is:

$$\frac{-9}{x} - \frac{6}{x^2} + \frac{9}{(x - 1)} - \frac{3}{(x - 1)^2} \quad (84)$$

Again, we can do the integral without a problem.

4. For a repeated irreducible quadratic term of the form $((x - r)^2 + b^2)^m$, we can write this as a sum of contributions of the form:

$$\sum_{n=1}^m \frac{B_1(x - r) + c_1}{((x - r)^2 + b^2)^n} \quad (85)$$

Again, multiply this through to make sure that you can get any term of the form $\frac{\text{order 1 polynomial}}{((x-r)^2+b^2)^m}$

Let's check our original checklist:

1. Convert the integrand into a proper rational form
2. Factor the denominator into degree 1 terms $(x - r)$ (which may be repeated: $(x - r)^m$) and irreducible degree 2 terms $(ax^2 + bx + c)$ (which may be repeated $(ax^2 + bx + c)^n$).
3. Take the irreducible degree two terms and complete the square to get: $((x - r)^2 + b^2)^n$
4. write out a sum of terms, one (or more if the factor is repeated) from each of the terms in the denominator
5. solve the integrals in the sum one by one, either directly or by integration by substitution.

In the lecture I didn't explain one crucial point. Terms with first order polynomials on the bottom (to some power) in the partial fraction expansion, will always have a constant in the numerator:

$$\frac{A}{(ax + b)^m} \quad (86)$$

and it will be your task to figure out what A is by multiplying out the partial fraction and matching it to the original expression. On the other hand, terms with second order polynomials (ie. irreducible to the sum of two first order terms) in their denominator (to some power) in the partial fraction expansion will have a term linear in x in their numerator:

$$\frac{Ax + B}{ax^2 + bx + c} \quad (87)$$

This should be made clearer with some examples:

I will assume that when you are given a fraction where the numerator is of higher power (in x) than the denominator, you can do polynomial long division to get it into proper rational form. i.e.:

$$\frac{P(x)}{Q(x)} = R(x) + \frac{S(x)}{Q(x)} \quad (88)$$

If you don't feel comfortable with this, practice it!!!! You can easily come up with example problems of your own and test your technique.

Now that we have a proper rational form we can look at the possible techniques.

1. We have a quadratic function on the bottom which can be written as the product of two first order terms (ie. in the language of the quadratic equation, $b^2 - 4ac > 0$)

$$\frac{3x - 6}{x^2 + 5x + 6} = \frac{3x - 6}{(x + 2)(x + 3)} \quad (89)$$

This is an example of case I in Stewart appendix G Now split this up into the sum of two terms:

$$\frac{3x - 6}{(x + 2)(x + 3)} = \frac{A}{x + 2} + \frac{B}{x + 3} \quad (90)$$

multiply out the left hand side and match the powers of x and the powers of constant numbers to find A and B : $A + B = 3$ and $3A + 2B =$

–6. The solution to this is $A = -12$ and $B = 15$. Our final expression is:

$$\frac{-12}{x+2} + \frac{15}{x+3} \quad (91)$$

and if we need to integrate this we can use the normal techniques we have learnt up to now.

2. If $b^2 - 4ac < 0$ then we have to complete the square and use a substitution method to integrate. For instance

$$\frac{3x-6}{(x^2+6x+12)(x-3)(x+2)} = \frac{3x-6}{((x+3)^2+3)(x-3)(x+2)} \quad (92)$$

This is an example of case III in Stewart, appendix G. We write this as the sum of terms of the form:

$$\frac{3x-6}{((x+3)^2+3)(x-3)(x+2)} = \frac{A}{x-3} + \frac{B}{x+2} + \frac{Cx+D}{((x+3)^2+3)} \quad (93)$$

The last term can be integrated by a u substitution. Note that the quadratic term has not just a constant in the numerator, but a term $Cx+D$. This is very important to remember.

3. If we have a term which appears multiple times (ie. to a power greater than 1) in the denominator, we have to repeat this in the partial fraction expansion:

$$\frac{4x-3}{(x-3)(x+2)^2} \quad (94)$$

This is an example of case II in Stewart appendix G. Here $(x+2)$ appears twice (ie. it is to the second power). The partial fraction expansion of this is:

$$\frac{4x-3}{(x-3)(x+2)^2} = \frac{A}{x-3} + \frac{B}{x+2} + \frac{C}{(x+2)^2} \quad (95)$$

4. If we have multiple quadratic factors which can't be factorised to linear terms (ie. $b^2 - 4ac < 0$) then we must have multiple terms in the sum:

$$\frac{3x-6}{((x+3)^2+3)^2(x-3)(x+2)} = \frac{A}{x-3} + \frac{B}{x+2} + \frac{Cx+D}{((x+3)^2+3)} + \frac{Ex+F}{((x+3)^2+3)^2} \quad (96)$$

5 Numerical integration

5.1 Integral approximations

Most integrals are hard to solve exactly. This might be because:

- We don't know the antiderivative of the function
- We don't know the function exactly - it might come from experimental data and all we have are lines on a graph but not the function itself.

Computers may be better at approximations of the exact results. Sometimes it is best to get a computer to solve this problem, but it is important to understand what is happening when this is done and how we can understand the results.

We know that the integral is a limit of the Riemann sum:

$$\text{Riemann Sum } \xrightarrow[\Delta x \rightarrow 0]{\text{number of pieces} \rightarrow \infty} \text{definite integral} \quad (97)$$

In the same way, an approximation of an integral can be calculated as a Riemann sum with a finite number of rectangles. You have seen before that there are a number of choices that you can make for the rectangles. The point at which they intersect the line can be taken as their top left edge (LPA), their top right edge (RPA) or their midpoint at the top (MPA). For the midpoint approximation:

$$\int_a^b f(x)dx \approx M_n = \Delta x [f(\vec{x}_1) + f(\vec{x}_2) + \dots + f(\vec{x}_n)] = \Delta x \sum_{i=1}^n f(\vec{x}_i) \quad (98)$$

where \vec{x}_i is the midpoint of the i^{th} rectangle.

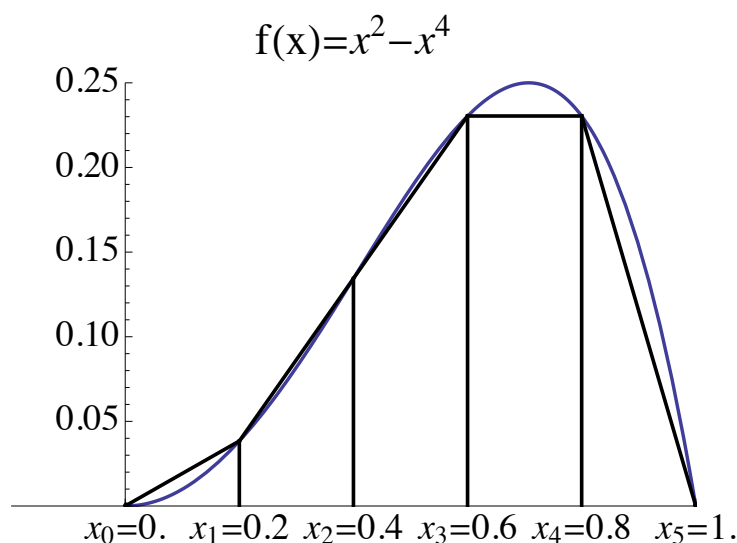
5.2 The trapezoidal integral

But maybe a rectangle isn't the best approximation we can make. When we are doing the left point approximation and the right point approximation, we are approximating only the value of the function at a certain point, and are not taking into account the slope. How about allowing the top of the rectangle to slope? This gives us a trapezoid. The area of a trapezoid is the length of the base times the average of the two side lengths. **Exercise: Why is the following the correct formula for the trapezoidal approximation:**

$$\int_a^b f(x)dx \approx T_n = \frac{\Delta x}{2} [f(x_0) + 2f(x_1) + 2f(x_2) + \dots + 2f(x_{n-1}) + f(x_n)] \quad (99)$$

Exercise: Find which out of the LPA, MPA, RPA and TPA give the best answer for:

$$\int_1^2 \frac{1}{x} dx \quad (100)$$



for $n=5$.

We know that the exact answer to this question is $\ln(2)$. The error of, for instance the trapezoidal integral is:

$$E_T \approx \ln(2) - T_5 \quad (101)$$

If what we are doing is correct, then the error should get less and less as we increase n . **Note, unless we know the actual value of the integral, we will never know what the error is. If we knew this then we wouldn't be doing the approximation in the first place. In the next section we will find that, even without knowing the exact value of the error, we can put a bound on its value, thus allowing us to chose n such that the error is definitely going to be lower than a certain value. For instance, if we want to approximate the integral of some curve so that the error is less than 0.00001, we can chose n such that this is true, without even knowing what the EXACT value of the error is.**

5.3 Error estimates (note that this is not examinable material, but it is interesting to understand)

We want to understand how the error in our approximation changes as we change the number of rectangles or trapezia. It should certainly go down as we increase the number and decrease their width, but by how much will it go down?

If the function we are trying to approximate does not curve, then the midpoint rule and the trapezoidal rule should give the exact answer.

Exercise: Why is this?

The curvature of the function (how much it deviates from being a straight line) makes the approximation inexact. For the left and right point approximations, even a straight line will not necessarily give you an exact answer when you do the LPA and RPA.

Can we approximate the function, in each region of the rectangle, or trapezoid, as a straight line, plus a quadratic term (a straight line plus something that is curving a bit).

$$f(x)_{x_{i-1}}^{x_i} \approx a + bx + cx^2 \quad (102)$$

If c is large, then the curve is very different from a straight line, if it's small, it's very close to a straight line. Because the trapezoidal and mid-point approximations give the exact answer if the function is a straight line, when $c = 0$ the error (the different between the approximation and the exact result) will be 0. If $|c|$ is large, it will give a large error. It is the cx^2 term which gives the error. You will understand exactly why you can make this approximation in a couple of weeks. Calculating the second derivative of the function gives $f''(x) = 2c$. We can make a statement about a bound on the error of the trapezoidal and midpoint rules, though we will not prove them here:

Error Bounds: If $|f''(x)| \leq K$ for $a \leq x \leq b$ then:

$$\begin{aligned} |E_T| &\leq \frac{K(b-a)^3}{12n^2} \\ |E_M| &\leq \frac{K(b-a)^3}{24n^2} \end{aligned}$$

You can see then that the bound on the midpoint approximation is stronger than that on the trapezoidal approximation. This doesn't mean that it's always better however.

For instance, for the function $f(x) = x^2 - x^4$, the second derivative is: $f''(x) = 2 - 12x^2$ but the absolute value of this is $|2 - 12x^2|$. The maximum value this takes is at $x = 1$ and is equal to 10. Thus we have an error bound on the trapezoidal integral which says that:

$$E_T \leq \frac{\max|f''(x)|(b-1)^3}{12n^2} = \frac{10(1-0)^3}{12n^2} = \frac{5}{6n^2} \quad (103)$$

$$\left(\begin{array}{cccccccc} n & 1 & 5 & 9 & 10 & 100 & 1000 \\ \frac{5}{6n^2} & 0.83 & 0.033 & 0.010 & 8.3 \times 10^{-3} & 8.3 \times 10^{-5} & 8.3 \times 10^{-7} \end{array} \right) \quad (104)$$

You can thus use this formula to work out how many pieces you need to break the function into to get a good approximation to the integral.

The reason that the error for the trapezium goes like the second derivative is because we are approximating the function to a sloped line. We are not taking into account how much it is curving, and the curving (away from the straight line) is encoded in the second derivative.

5.4 Simpson's rule (Also not examinable)

We can do better than approximating the curve by a series of trapezoids. We can actually try and follow some of the curvature of the function. This will reduce the error that we found in the last calculation. We need to approximate the curve by a series of parabolas rather than a series of trapezoids or rectangles (ie. functions of the form $f(x) = ax^2 + bx + c$).

How do we do this? First we split the region into an even number, n , of sectors and mark the points along the curve (p_0 to p_n). The width of the sectors are $h = \frac{b-a}{n}$. Let's give an example where $n = 6$ and $h = 2$ in figure 1:

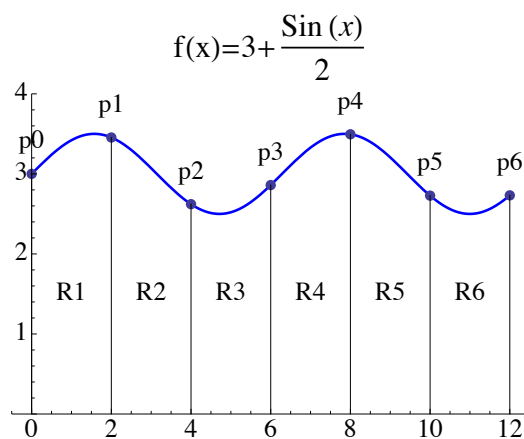


Figure 1: The graph of $f(x) = 3 + \frac{\sin(x)}{2}$ split into six regions and the seven points p_0 to p_6 along the curve

Now find a parabolic curve which goes through p_0, p_1 and p_2 , another curve which goes through p_2, p_3 and p_4 and another curve which goes through p_4, p_5 and p_6 . (**Exercise: See why we need an even number of regions for this to work**). If there were just two points in each group (eg. p_0 and

p_1 , p_1 and p_2 , etc.) then we could always find a straight line to go through them, but this time we have three points and thus we have to use a parabola. Any three points can always be joined by a parabola.

The first parabola passes through $((x_0, f(x_0)), (x_1, f(x_1)), (x_2, f(x_2))) = ((0, f(0)), (2, f(2)), (4, f(4)))$. So we need to find an a_1 , b_1 and c_1 such that:

$$\begin{aligned} a_1(0^2) + b_1(0) + c_1 &= f(0) \\ a_1(2^2) + b_1(2) + c_1 &= f(2) \\ a_1(4^2) + b_1(4) + c_1 &= f(4) \end{aligned} \tag{105}$$

We can solve this equation for a_1 , b_1 and c_1 . We can do this for all the parabolas and find the following three (green, yellow and red) which span the whole curve from 0 to 12. The green curve passes through p_0 , p_1 and p_2 , the yellow curve passes through p_2 , p_3 and p_4 and the red curve passes through p_4 , p_5 and p_6 . See figure 2:

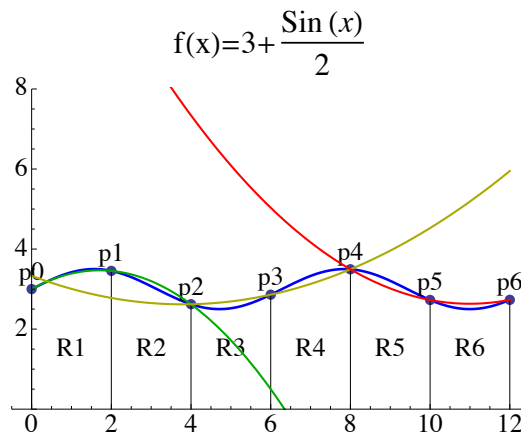


Figure 2: parabolas going through three points on the graph

You can see that the parabolas, in the regions between the points that they pass through (though definitely not outside those regions) approximate the curve pretty well. We can then integrate under the parabolas in those regions to get a better approximation to the exact integral, see figures 3, 4 and 5:

We can proceed in two different ways, first we can shift the integral to go over the origin as in figure 5. This allows for a simplification. The reason is that the integral under each parabola (for instance the green parabola between p_0 and p_2 given by $f(x) = a_1x^2 + b_1x + c_1$) is unchanged if we move the parabola to the left or right, the limits don't matter, only the difference

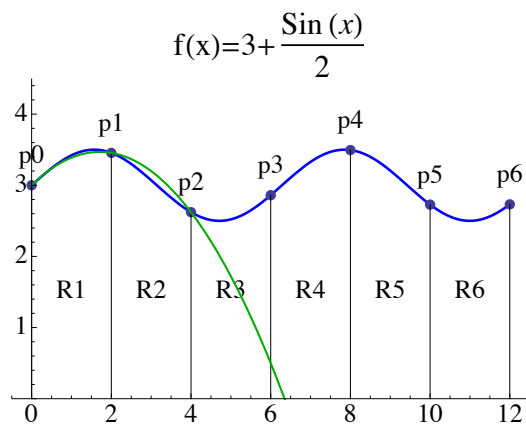


Figure 3: parabola going through (p_0, p_1) and p_2

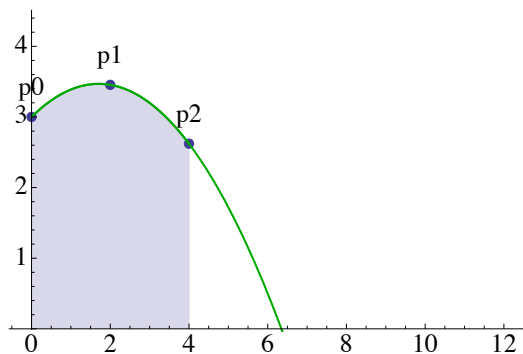


Figure 4: The parabola passing through points p_0 , p_1 and p_2

in the limits. In this case we will not do so, we will proceed directly without shifting the integral. (NB. in the lecture and in Stewart, the method of translation of the parabola is used).

We can then integrate this over the region between $x_0 \leq x \leq x_2$. You can show that this is equal to:

$$\int_{x_0}^{x_2} a_1x^2 + b_1x + c_1dx = \frac{2}{3}h(3x_0(2ah + ax_0 + b) + h(4ah + 3b) + 3c) \quad (106)$$

where $2h = x_2 - x_0$

However, by fitting the three points that the curve must go through, and thus solving for a_1, b_1 and c_1 , you can show that the integral is equal to

$$\frac{h}{3}(f(x_0) + 4f(x_1) + f(x_2)) \quad (107)$$

We can perform exactly the same calculation for all the other parabolas over their appropriate regions and find that the total integral is given by the

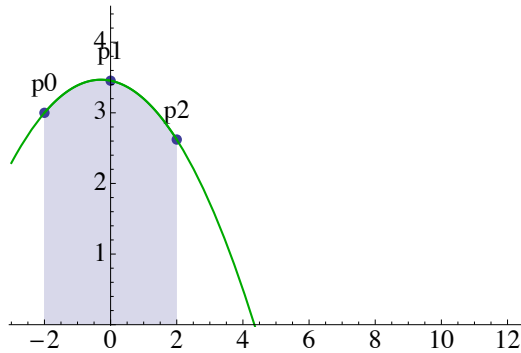


Figure 5: The same parabola as in 4 but shifted to straddle the origin (as done in Stewart and described in the lecture, but not performed here...)

expression known as **Simpson's Rule**:

$$\int_a^b f(x)dx \approx S_n = \frac{\Delta x}{3} [f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \dots 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)] \quad (108)$$

Notice that although we are approximating the curve with more complicated functions (parabolas rather than rectangles or trapezoids), the overall expression isn't much more complicated, but it does tend to be more accurate. The error from Simpson's rule goes like:

$$|E_S| \leq \frac{K(b-a)^5}{180n^4} \quad (109)$$

where $|f^{(4)}(x)| \leq K$.

5.5 Why does the error in Simpson's rule go like the fourth derivative if we are approximating a function with a series of parabolic curves?

I think the most intuitive understanding of this comes from looking at the MPA. here we are approximating the function by a straight line of slope zero for each interval (rather than a sloped line, as in the trapezium rule). The important point is that while the rectangle doesn't exactly approximate a sloped line, the errors above and below the line equal out and you end up with a result which is exact for sloped, linear functions, and has an error only at the second derivative of the function. See figure 6.

The same is true for the quadratic approximation which is used in Simpson's rule. While it may not approximate a cubic function exactly, the error

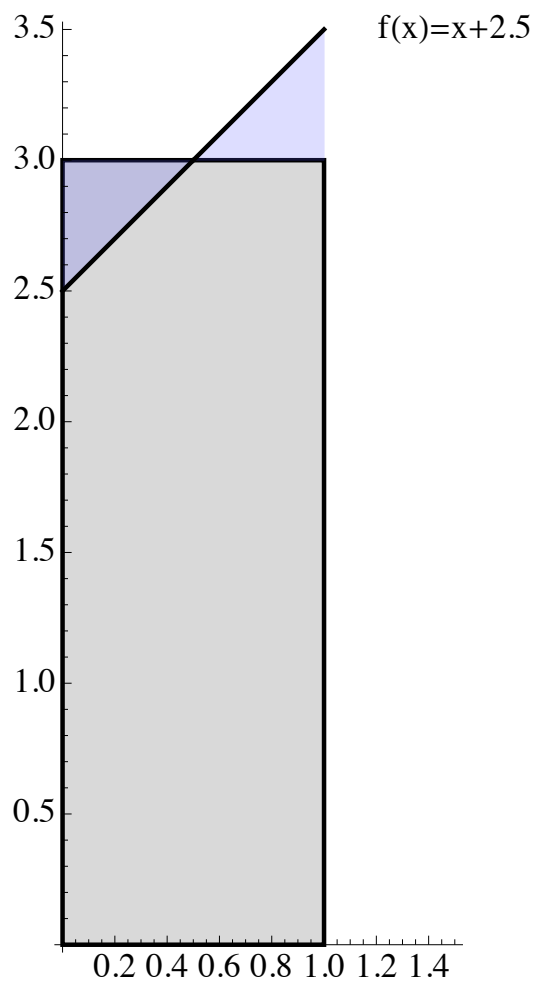


Figure 6: While the MPA doesn't exactly approximate the function, for constant slopes it gives an exact answer because the error above and below the line cancels out. This is not true for the LPA and RPA where there is an error even for a linear slope.

either side of the mid-point (if the parabolic curve goes through p_0 , p_1 and p_2 then the midpoint is p_1) cancels out and so Simpson's rule gives an exact answer if you are trying to approximate a cubic function. The error comes in if you are trying to approximate a fourth order or higher polynomial and thus the error goes like the maximum value of $|f''''(x)|$ and not $|f'''(x)|$

Please go through the examples in Stewart to get a good intuition of how these approximations work.

6 Improper integrals

When we defined the definite integral we gave some constraints. We can now integrate (either approximately or exactly):

$$\int_a^b f(x)dx \tag{110}$$

as long as $[a, b]$ is finite and as long as there are no infinite discontinuities in $a \leq x \leq b$. An infinite discontinuity means that $f(x)$ is not bounded at some point in $[a, b]$ (intuitively this means that the function goes to $\pm\infty$ at some value of x in $[a, b]$).

If we have an integral which does not abide by these constraints, we may still be able to calculate an answer for the area under the curve, but it will now be called an improper integral. The reason that these are defined as improper is because they will not themselves be well defined as Riemann sums, however, they will be limits of Riemann sums as we will soon see.

6.1 Improper integrals of the first kind: Infinite integrals

We can first ask about the integral over an open region, this will give us some intuition as to how to perform improper integrals. Can we perform some integral over the range $(0, 10]$ of x ? Certainly this doesn't make much sense if we are thinking about a Riemann sum. How much of the first rectangle do we include if we are supposed to be going down to, but not including 0? The answer is that we can think about this as a limit of a well defined Riemann integral:

$$\int_{(0,10]} x dx = \lim_{t \rightarrow 0} \int_t^{10} x dx = \lim_{t \rightarrow 0} \left. \frac{x^2}{2} \right|_t^{10} = \lim_{t \rightarrow 0} \frac{10^2}{2} - \frac{t^2}{2} = 50 \quad (111)$$

In this case the limit is equal to the regular Riemann integral over the closed region. Nothing very strange going on there, and it makes sense because we expect that just not including the point $x = 0$ shouldn't change the area under the curve by a finite amount.

1. We can however use this technique when we want to ask about more tricky integrals: those where there is an open range and the open end of the range is at ∞ , or $-\infty$ or both. We know how to integrate:

$$\int_1^4 \frac{1}{x^2} dx = \left. \frac{-1}{x} \right|_1^4 = \frac{-1}{4} - \frac{-1}{1} = \frac{3}{4} \quad (112)$$

This is just the area under the curve given in figure 7 How about if we wanted to integrate further? Can we go up to 100 rather than 4? Sure, this will give an answer of $1 - \frac{1}{100} = \frac{99}{100}$. But this is rather interesting. We have gone a lot further out, but the value of the integral has only change from 0.75 to 0.99. What about if we go even further out? We can see easily that as we go further and further out, the answer will

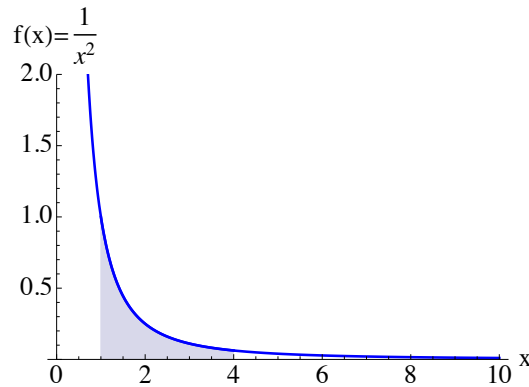


Figure 7: $\int_1^4 \frac{1}{x^2} dx$ gives a perfectly well defined area. Nothing funny going on here.

get closer and closer to 1. In fact we can define an integral, called an improper integral, where we integrate all the way up to ∞ . This integral is defined like:

$$\int_1^{\infty} \frac{1}{x^2} dx = \lim_{t \rightarrow \infty} \int_1^t \frac{1}{x^2} dx = \lim_{t \rightarrow \infty} \left. \frac{-1}{x} \right|_1^t = \lim_{t \rightarrow \infty} \left(\frac{-1}{t} - \frac{-1}{1} \right) \quad (113)$$

The limit of $\frac{-1}{t}$ as $t \rightarrow \infty$ is 0, and so the final answer is 1. It is the value that we were getting closer and closer to before we even knew what an improper integral was.

2. Let's look at another example

You may have guessed at first that if we have an integral that goes from $[0, \infty)$ that the area under the curve should be infinite, but this isn't necessarily so. Let's look at an example. We can certainly perform the following integral for finite t :

$$A(t) = \int_0^t \frac{1}{1+x^2} dx = \arctan x \Big|_0^t = \arctan t - 0 = \arctan t \quad (114)$$

What happens if we take $t \rightarrow \infty$:

$$\lim_{t \rightarrow \infty} \arctan t = \frac{\pi}{2} \quad (115)$$

Thus:

$$\int_0^{\infty} \frac{1}{1+x^2} dx = \lim_{t \rightarrow \infty} \int_0^t \frac{1}{1+x^2} dx = \frac{\pi}{2} \quad (116)$$

This example and the previous are set to be convergent. As $t \rightarrow \infty$ the value of the integral converges to some fixed value. Sometimes this

will not be so:

There was an excellent question in class about the above problem. I asked what value of x solved the equation $\lim_{t \rightarrow \infty} \tan x = t$, and indeed there are an infinite number of answers to this question. The real question however is what is $\lim_{t \rightarrow \infty} \arctan t$. The function $\arctan x$ is single valued and thus there is no ambiguity in saying that this limit gives $\frac{\pi}{2}$, whereas the way I stated the question, in terms of $\tan x$ does lead to an ambiguity. Great question and nice to think about these things!

3. For example, if we wanted to try something very very similar to the above examples:

$$\int_1^{\infty} \frac{1}{x} dx \quad (117)$$

This looks so similar to the first example, and we can see that as $x \rightarrow \infty$ the function itself goes to zero, but in this case it doesn't go to zero quite fast enough, as we integrate more and more of the function, we continue to get contributions and they never go to zero:

$$\int_1^{\infty} \frac{1}{x} dx = \lim_{t \rightarrow \infty} \int_1^t \frac{1}{x} dx = \lim_{t \rightarrow \infty} \ln |x| \Big|_1^t = \lim_{t \rightarrow \infty} (\ln |t| - 0) \quad (118)$$

But as t increases, $\ln |t|$ continues to increase, and so this doesn't give a finite answer. The result of this integral is ∞ and thus it is said to be divergent.

4. We can do exactly the same thing for an integral where the lower limit is $-\infty$ in exactly the same way:

$$\int_{-\infty}^{-1} \frac{1}{x^4} dx = \lim_{t \rightarrow -\infty} \int_t^{-1} \frac{1}{x^4} dx = \lim_{t \rightarrow -\infty} \frac{-1}{3x^3} \Big|_t^{-1} = \lim_{t \rightarrow -\infty} \left(\frac{-1}{3(-1)^3} - \frac{-1}{3t^3} \right) \quad (119)$$

In this case we can also take the limit which gives a convergent quantity and we get $\frac{1}{3}$.

5. On the other hand:

$$\int_{-\infty}^1 \frac{1}{\sqrt{-x}} dx \quad (120)$$

Is not convergent. ie. it is divergent.

6. In fact we can find out for what value of n , is:

$$\int_1^{\infty} x^n dx \quad (121)$$

is convergent. The answer is that for $n < -1$ the integral is convergent, and $n \geq -1$ it is divergent. We saw that if $n = -1$ it is divergent because of the logarithm (in fact this is called a logarithmic divergence). It should be very clear that for $n > 0$ this is divergent, but you can check for $n \leq 0$.

The definition, summarising the above, for an improper integral of type I can be found in Stewart page 414.

7. Sometimes we have to be a bit more clever about taking the limit as it won't be all that obvious. This often happens when you are integrating a product of functions, one of which diverges as $x \rightarrow \infty$ or $-\infty$ and one of which goes to zero in this limit. For instance:

$$\int_{-\infty}^0 xe^x dx \quad (122)$$

We can define this integral through the limit:

$$\int_{-\infty}^0 xe^x dx = \lim_{t \rightarrow -\infty} \int_t^0 xe^x dx \quad (123)$$

and then integrate by parts to get:

$$\lim_{t \rightarrow -\infty} (-te^t - 1 + e^t) \quad (124)$$

Clearly as $t \rightarrow -\infty$ -1 doesn't behave badly, and we know also that e^t goes to zero in this limit, but how about te^t ? We can use l'Hospital's rule to calculate this. This states:

$$\lim_{t \rightarrow -\infty} \frac{f(x)}{g(x)} = \lim_{t \rightarrow -\infty} \frac{f'(x)}{g'(x)} \quad (125)$$

This can be iterated as many times as necessary until you can see the answer. In this case we just need to take the derivative once:

$$\lim_{t \rightarrow -\infty} \frac{t}{e^{-t}} = - \lim_{t \rightarrow -\infty} \frac{1}{e^{-t}} = 0 \quad (126)$$

Thus, the overall integral is convergent and $= -1$.

8. The final extension of this is to have both the upper and lower limits be $+\infty$ and $-\infty$. In this case we can simply make a split and use the above definitions:

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^a f(x)dx + \int_a^{\infty} f(x)dx \quad (127)$$

where a is any finite number. If either of these terms is not convergent, then the whole integral is said to be divergent. Let's look at an example of that where it is not intuitive.

9. Let's look at an example that, done in the wrong way (but one which seems logical) will give the wrong answer:

$$\int_{-\infty}^{\infty} xdx \quad (128)$$

If we were to take this as a limit of $\int_{-t}^t xdx$ we would get that the answer is zero - it looks like the two sides of the odd function cancel each other out. But this is only because in this case we are taking the limits to $+\infty$ and $-\infty$ at the same rate. We can tend to these limits at different rates and find that the answer can be anything we want. Instead, the correct way to do this is to write:

$$\int_{-\infty}^{\infty} xdx = \int_{-\infty}^0 xdx + \int_0^{\infty} xdx = \left(\lim_{t \rightarrow -\infty} \int_t^0 xdx \right) + \left(\lim_{s \rightarrow \infty} \int_0^s xdx \right) \quad (129)$$

Both of these contributions do not converge, and so the whole thing diverges.

NB. The definition of divergent is simply that something is not convergent: It doesn't have to go to $\pm\infty$ to be divergent. For instance $\lim_{t \rightarrow \infty} \cos t$ oscillates, and so is not convergent, and so is said to be divergent.

6.2 Improper Integrals of the second kind: Discontinuous integrands

We have seen what happens when the x direction runs to ∞ or $-\infty$. Sometimes it gives a convergent value and we can find the integral, sometimes

the improper integral does not converge and it gives us an answer of $\pm\infty$. Now we are going to look at what happens if the function itself has an ∞ or $-\infty$ in it and we want to integrate over this region, or up to it. The formal definition of this can be found in Stewart page 418. Here we will look at the examples and examine what is happening.

1. The function $f(x) = \frac{1}{\sqrt{x-4}}$ is only defined on $(4, \infty)$ and looks like figure 8. Let's try and integrate it all the way up to the point 4, where there

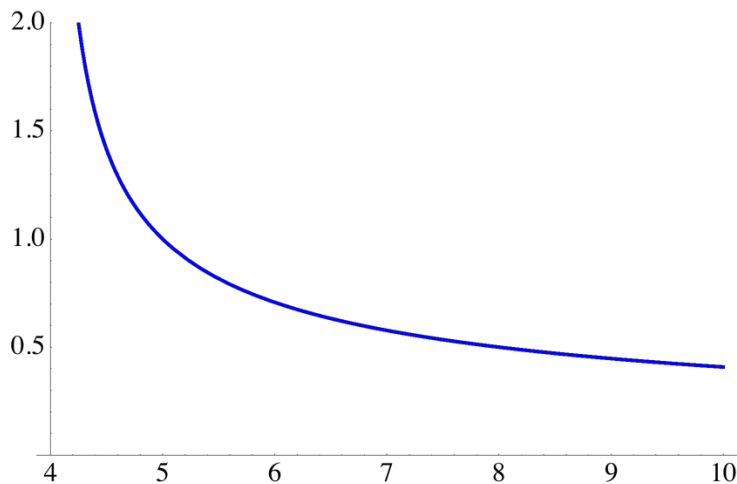


Figure 8: Plot of the function $f(x) = \frac{1}{\sqrt{x-4}}$. It seems like integrating from 4 to 10 is going to be difficult, but we'll see that it's not.

is the discontinuity...actually, we won't include the discontinuity but we will take the limit of the integral as it goes there:

$$\int_4^{10} \frac{1}{\sqrt{x-4}} dx = \lim_{t \rightarrow 4^+} \int_t^{10} \frac{1}{\sqrt{x-4}} dx = \lim_{t \rightarrow 4^+} 2\sqrt{x-4} \Big|_t^{10} \quad (130)$$

but, just as in the cases of some of the infinite integrals, we can take this limit and it will give us a convergent quantity:

$$\lim_{t \rightarrow 4^+} 2\sqrt{x-4} \Big|_t^{10} = 2(\sqrt{10-4} - \sqrt{4-4}) = 2\sqrt{6} \quad (131)$$

2. How about:

$$\int_0^1 \ln x dx \quad (132)$$

This diverges at $x = 0$ and so we can define the improper integral as:

$$\int_0^1 \ln x dx = \lim_{t \rightarrow 0^+} \int_t^1 \ln x dx = \lim_{t \rightarrow 0^+} -t \ln t - 1 + t \quad (133)$$

Clearly the last terms converge, but how about $t \ln t$. We know that t goes to zero and $\ln t$ goes to $-\infty$ so we have to be careful about this. We can invoke L'Hospital's rule again:

$$\lim_{t \rightarrow 0^+} t \ln t = \lim_{t \rightarrow 0^+} \frac{\ln t}{t^{-1}} = \lim_{t \rightarrow 0^+} \frac{\frac{1}{t}}{\frac{-1}{t^2}} = \lim_{t \rightarrow 0^+} (-t) = 0 \quad (134)$$

So the total answer is:

$$\int_0^1 \ln x dx = -1 \quad (135)$$

3. Let's look at a non-convergent example:

$$\int_0^1 \frac{1}{x^2} dx = \lim_{t \rightarrow 0^+} \int_t^1 \frac{1}{x^2} = \lim_{t \rightarrow 0^+} \left. \frac{-1}{x} \right|_t^1 \quad (136)$$

The limit does not converge (ie. it gives ∞ in this case).

4. For the time being, if we have an integral where any part of it is divergent, then it is said to be divergent as a whole. For instance:

$$\int_{-1}^1 \frac{1}{x^2} dx \quad (137)$$

can be split up as:

$$\int_{-1}^1 \frac{1}{x^2} dx = \lim_{t \rightarrow 0^-} \int_{-1}^t \frac{1}{x^2} dx + \lim_{t \rightarrow 0^+} \int_t^1 \frac{1}{x^2} dx \quad (138)$$

As the first term in this expression diverges, so the whole thing is said to be divergent. **Watch out for examples when neither of the limits have a problem, but the problem is somewhere in the range $[a, b]$**

6.3 The Comparison Theorem

For two continuous functions, $f(x)$ and $g(x)$, if $f(x) \geq g(x) \geq 0$ for $x \geq a$ then:

1. If $\int_a^\infty f(x) dx$ is convergent, then so is $\int_a^\infty g(x) dx$.
2. If $\int_a^\infty g(x) dx$ is divergent, then so is $\int_a^\infty f(x) dx$.

Let's look at an example. If we have the integral:

$$\int_0^{\infty} e^{-x^2} dx \quad (139)$$

We may not know how to do this integral. However, we can still say something about whether it converges or diverges. We use a trick for this and split the integral into two regions:

$$\int_0^{\infty} e^{-x^2} dx = \int_0^1 e^{-x^2} dx + \int_1^{\infty} e^{-x^2} dx \quad (140)$$

Certainly the first part is finite because it is continuous and the limits are finite. The question is whether the second part is finite. The reason that we split the integral up like this is because there is a simple function which is always greater than e^{-x^2} between 1 and ∞ (but not between 0 and 1 which is why we made the split). We can show that:

$$e^{-x} \geq e^{-x^2} \quad (141)$$

for $x \geq 1$. We can then ask whether the integral:

$$\int_1^{\infty} e^{-x} dx \quad (142)$$

converges, and indeed we can show with the methods that we've learnt above, that it does. Thus, we can use the comparison theorem and say that because e^{-x} is always greater than or equal to e^{-x^2} in the region between 1 and ∞ and that the integral over the former converges, that the integral over e^{-x^2} must also converge.

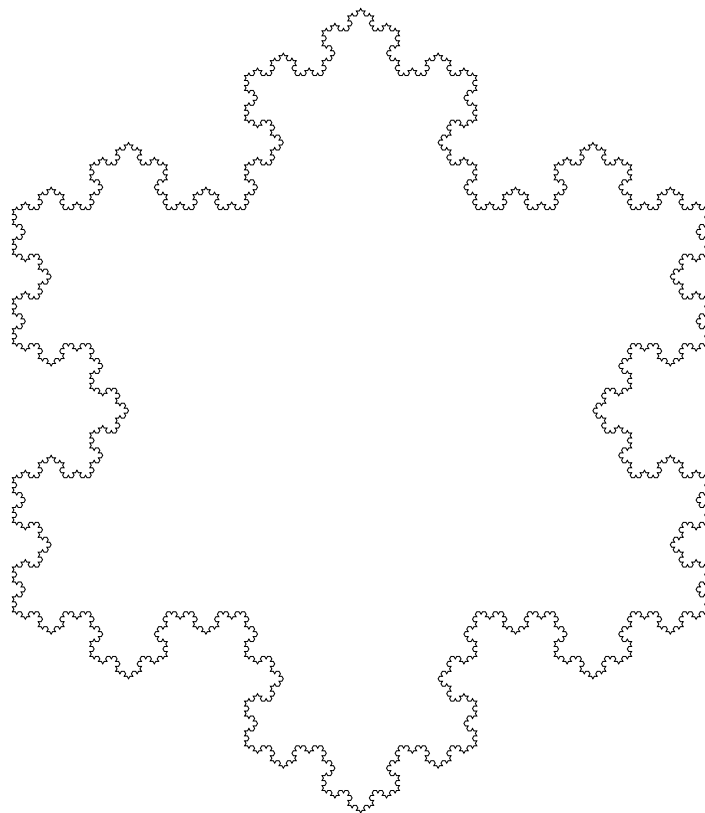


Figure 9: Sometimes things which seem finite are not. The length of the line in a Koch snowflake is infinite. This is not related to the course, but is just a nice example of a surprising result!

7 Areas between curves

We know how to calculate the area between a curve and the horizontal axis of a graph. We have learnt a number of very sophisticated techniques which allow us to get an analytic answer, or, if the antiderivative of the function is not known, an approximate answer to any accuracy we want, if the function is well-behaved enough.

It is a fairly simple extension to study not the area between a curve and the horizontal axis, but between two curves. You can see here the link between the two, as shown in figure 10

We can see in figure 11 how we can approximate the area by a sum of rectangles. In this case we have used the left point approximation. Each rectangle here has $\Delta x = 0.02$ and the heights are $f(x_i) - g(x_i)$ where the x_i are the left points of each rectangle, in this case $= 0.02(i - 1)$. So $x_1 = 0$,

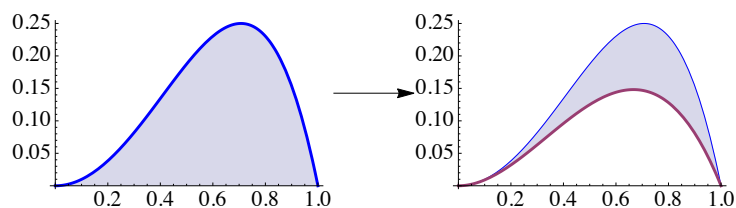


Figure 10: The area between a curve and the horizontal axis, given by the definite integral, and the area between two curves. We will soon see how to calculate the latter

$x_2 = 0.02$, etc.

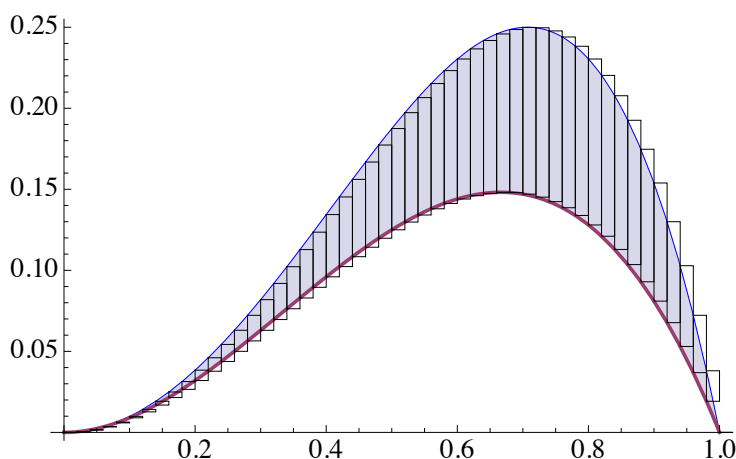


Figure 11: The area approximated by a sum of rectangles.

We started off studying integration by taking the limit of a Riemann sum. We can do exactly the same thing here, taking the simplest possible Riemann sum, with rectangular regions which stretch between the two curves (see the diagrams on page 448 in Stewart) and then taking the limit to find an expression in terms of integrals. You will not be surprised by the final answer! If we have two functions $f(x)$ and $g(x)$ where, in some region $x \in [a, b]$ $f(x) \geq g(x)$, we can write the area between them as the Riemann sum:

$$\text{Area between curves } f(x) \text{ and } g(x) \approx \sum_{i=1}^n \Delta x (f(x_i^*) - g(x_i^*)) \quad (143)$$

where x_i^* are the positions in the rectangles that go into making up the area. This can be chosen as the left point, right point or mid-point, or any other carefully chosen point. Taking the limit of this sum gives a difference in

integrals, or alternatively, the integral of the difference in the values of the functions:

$$A = \int_a^b [f(x) - g(x)] dx \quad (144)$$

This is clearly the same as the area under $f(x)$ minus the area under $g(x)$ because $\int_a^b f(x) - f(g)dx = \int_a^b f(x)dx - \int_a^b g(x)dx$.

1. Sometimes you will be asked to find the enclosed area between two curves. This means that you need not just to find the integral, but also to find the upper and lower limits of the integral (a and b in the notation $\int_a^b f(x)dx$). The limits will be given by the intersection points of the curves. Here is a simple example:

Find the area between the curves $f(x) = -x^2 + 6$ and $g(x) = x^2 + 3$. First of all we need to find the intersection points. These two curves intersect when:

$$-x^2 + 6 = x^2 + 3 \quad (145)$$

The solutions of this equation are $x = \pm\sqrt{\frac{3}{2}}$, so what we have to calculate is the area shown in figure 12. The area between the two

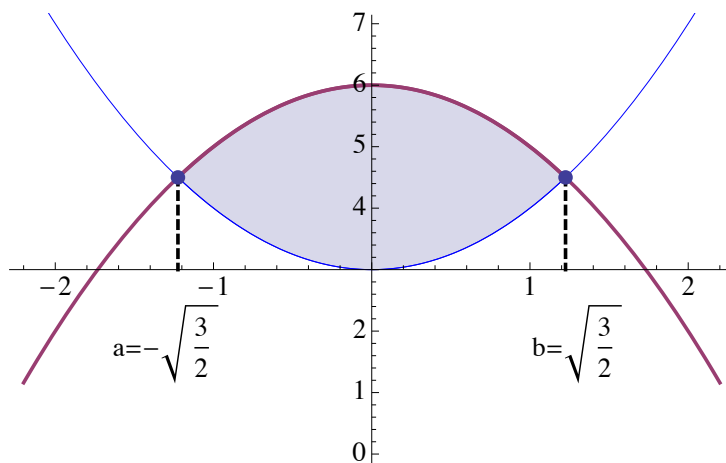


Figure 12: The intersection points of the two functions are given by $x = \pm\sqrt{\frac{3}{2}}$

curves is thus:

$$A = \int_{-\sqrt{\frac{3}{2}}}^{\sqrt{\frac{3}{2}}} (-x^2 + 6) - (x^2 + 3) = \int_{-\sqrt{\frac{3}{2}}}^{\sqrt{\frac{3}{2}}} (-2x^2 + 3) = -\frac{2}{3}x^3 + 3x \Big|_{-\sqrt{\frac{3}{2}}}^{\sqrt{\frac{3}{2}}} \quad (146)$$

We can evaluate this and it gives $A = 2\sqrt{6}$. Note that if you are doing everything correctly, then the area between the curves should always be positive. A negative definite integral makes sense, but a negative area between two curves does not.

- Here is an example where you have to find the area between two curves and you are given the region of integration. The curves $f(x) = x$ and $g(x) = -x^2 - 3$ do not intersect, but we can still ask what the area between them in the region between, let's say -4 and 3 is. This is shown in figure 13 The area is simply:

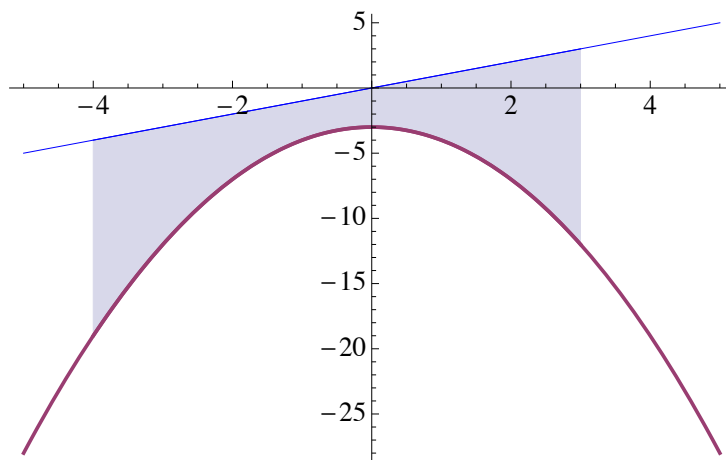


Figure 13: These two curves never intersect but we can ask for the area between them in the range -4 to 3. Think of this as having four curves. The two functions, and two lines which go vertically along the line $x = -4$ and $x = 3$. There is an area enclosed between these four lines.

$$\int_{-4}^3 (x - (-x^2 - 3)) dx = \frac{x^2}{2} + \frac{x^3}{3} + 3x \Big|_{-4}^3 = \frac{287}{6} \quad (147)$$

- Sometimes you won't know exactly the intersection points, but you can calculate them using Newton's method. What is the area between the curves: $f(x) = \sin x$ and $g(x) = x^2 - 2$? First we have to find the intersection points, and this is done using Newton's method. The answer is that they intersect at $x \approx -1.06$ and $x \approx 1.73$. This is shown in figure 14

In order to use Newton's method in this context we have to write down the equation we want to find the zeros of. We know that the intersection points are at $f(x) = g(x)$ and thus $\sin x = x^2 - 2$. We can

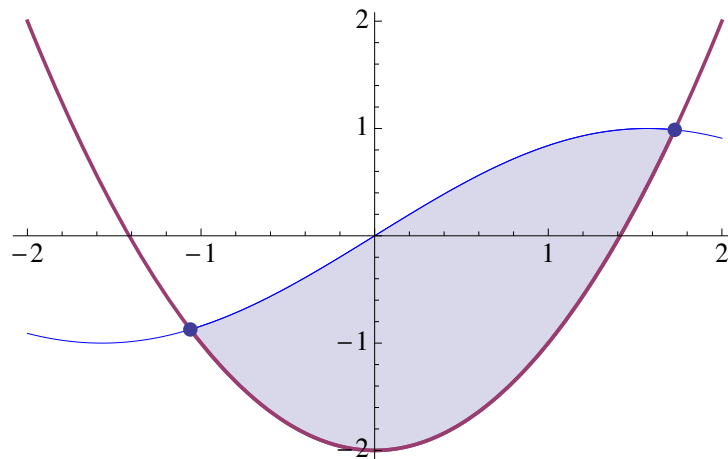


Figure 14: $f(x) = \sin x$ and $g(x) = x^2 - 2$ and their intersection points, found by Newton's method.

write this as: $\sin x - x^2 + 2 = 0$. We thus want to find the zeros of $p(x) = \sin x - x^2 + 2$. First sketch this. In fact it looks like figure 15 In

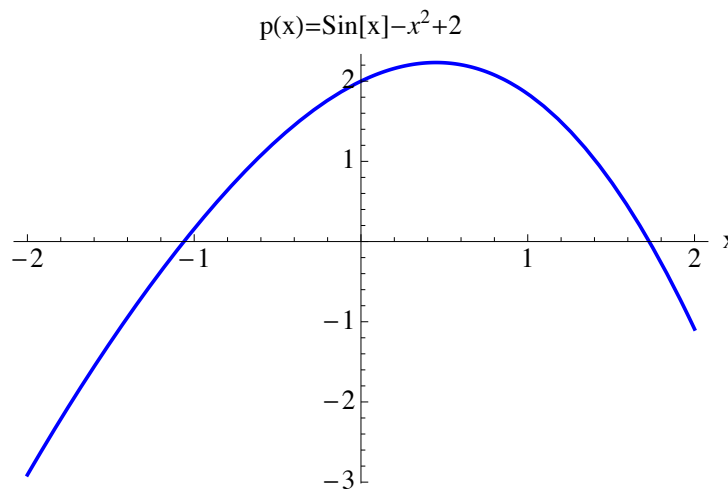


Figure 15: $\sin x - x^2 + 2$. The intersection points of $f(x)$ and $g(x)$ are given by the zeros of this function. There are clearly two of them.

order to use Newton's method we have to start with an approximate value. In this case we can start with $x = -1$ and we will zoom in to the left intersection point. We can also use the point $x = 1$ and we will zoom into the right point. Figure 16 shows the iteration procedure to get the right intersection point.

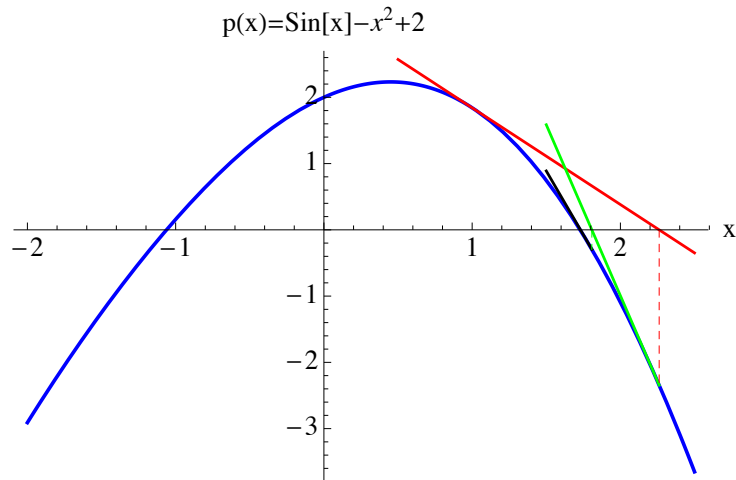


Figure 16: Three iterations of Newton's method starting at $x = 1$ to find the intersection point of the function $p(x)$ with the x-axis. The same thing can be done for finding the left intersection point, but starting with the initial guess of $x = -1$ for Newton's method.

Thus the area between the curves is:

$$\int_{-1.06}^{1.73} \sin(x) - (x^2 - 2) = -\cos x - \frac{x^3}{3} + 2x \Big|_{-1.06}^{1.73} \approx 4.10 \quad (148)$$

4. We are used to thinking always of some function being dependent on x , but we can actually write it the other way around, i.e. $x = f(y)$. x and y are just labels and so, as long as the function is single valued, we can write a function however we want. For instance, the lines $y = x - 1$ and $x = \frac{y^2 - 6}{2}$ (which would not be single valued if we wrote the curve as $y(x)$) look like figure 17: If we want to perform the integral in x variables we would have to do:

$$Area = \int_{-3}^{-1} \sqrt{2x + 6} - (-\sqrt{2x + 6}) dx + \int_{-1}^5 \sqrt{2x + 6} - (x - 1) \quad (149)$$

Alternatively, we can perform the integral in y variables as:

$$\int_{-2}^4 (y + 1) - \left(\frac{1}{2}y^2 - 3\right) dy \quad (150)$$

The two will give the same answer but the latter calculation is clearly easier.

Exercise: Make sure that you understand the above calculation.

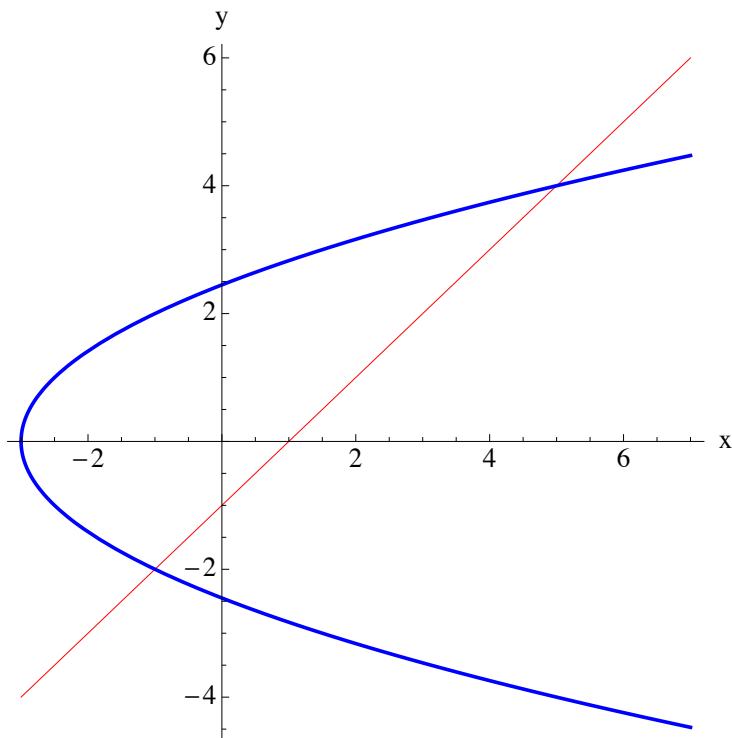


Figure 17: If we want to find the area between the two lines using x coordinates we will have to do so in two steps, in y coordinates we can do it in a single integration.

- Let's look at a slightly more complicated region. If we have the curves $y = \sin 2x$ and $y = \cos x$ and ask for the region between the two curves in the range $0 \leq x \leq \frac{\pi}{2}$ we will find that there are actually two contributions to this. See figure 18.

Recap:

There are several situations that we have to consider:

- We are given two functions, $f(x)$ and $g(x)$ and are asked to find the area between them from a to b . If the functions do not intersect in this region then find which function is larger. If $f(x) > g(x)$ in the region from $x = a$ to $x = b$ calculate: $\int_a^b f(x) - g(x) dx$. If $g(x) > f(x)$ in the region then calculate $\int_a^b g(x) - f(x) dx$. Remember the area should always be positive, so this is a good check.
- We are given two functions $f(x)$ and $g(x)$ which intersect at two points and we are asked to integrate the area between them. First calculate

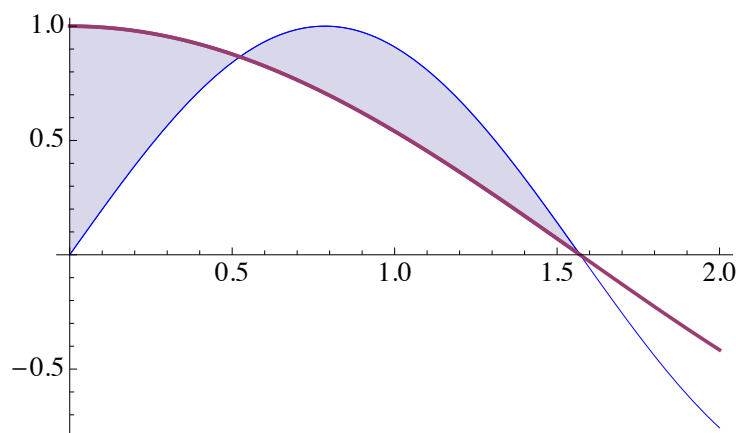


Figure 18: There are two regions in between these two graphs so the integral has to be done in two parts.

the intersection points, either by solving the equations simultaneously or by using Newton's method then perform the above integral, making sure that we come out with a positive answer

3. We are given two functions $f(x)$ and $g(x)$ which may intersect at more than two points. Calculate these points, then treat the integrals between the sets of points separately. See example 5 above.
4. We are given a function or functions which are multivalued in x . See if they are single-valued in y . If they are, then see if the problem will be easier to solve when we write the functions as $f(y)$ and $g(y)$ rather than as functions of x .

Important point: If you are confused as to how best to perform the integral, think first of approximating the area between the curves as a Riemann sum then convert this to an integral, or integrals by taking the limit of the Riemann sum that the number of rectangles goes to ∞ .

Note that it will not always be the case that you can add the two areas together as we have done above. If you are asked simply to find the areas between the curves then this will be the case. If however you have a problem which has a physical interpretation (for instance, see Example 4 on page 434 of Stewart) then you will have to be more careful about adding together areas.

7.1 When to take the absolute area, and when the sign matters

Let's take two functions: $f(t) = \cos t$ and $g(t) = -\cos t$. They look like figure 19:

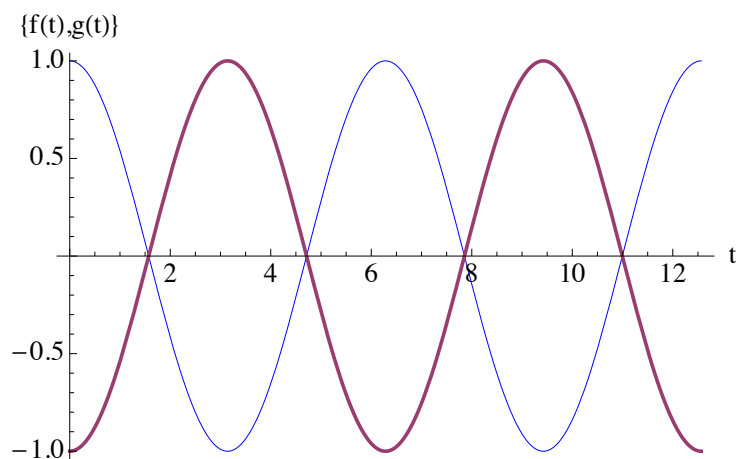


Figure 19: $\cos t$ and $-\cos t$.

We may be asked to simply find the area between the curves, between two points on the t -axis. Let's say between 0 and some point t' . We know that when we are calculating the area between two curves we must have the larger function minus the smaller function, so we have to split this integral up into segments because, between 0 and $\frac{\pi}{2}$ $f(t)$ is greater than $g(t)$ but then between $\frac{\pi}{2}$ and $\frac{3\pi}{2}$, it's the other way around. So the area between the curves, let's say up to 2π is:

$$\begin{aligned}
 \text{Area} &= \int_0^{\frac{\pi}{2}} \cos x - (-\cos x) dx + \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} (-\cos x) - \cos x dx + \int_{\frac{3\pi}{2}}^{2\pi} \cos x - (-\cos x) dx \\
 &= \int_0^{\frac{\pi}{2}} 2 \cos x dx - \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} 2 \cos x dx + \int_{\frac{3\pi}{2}}^{2\pi} 2 \cos x dx \\
 &= 2(\sin \frac{\pi}{2} - \sin 0) - (2(\sin \frac{3\pi}{2} - \sin \frac{\pi}{2})) + 2(\sin 2\pi - \sin \frac{3\pi}{2}) \\
 &= 8
 \end{aligned} \tag{151}$$

So the total area under the curve between these points is 8. In fact we can even plot what the area looks like as we integrate up to some point t' and it looks like figure 20

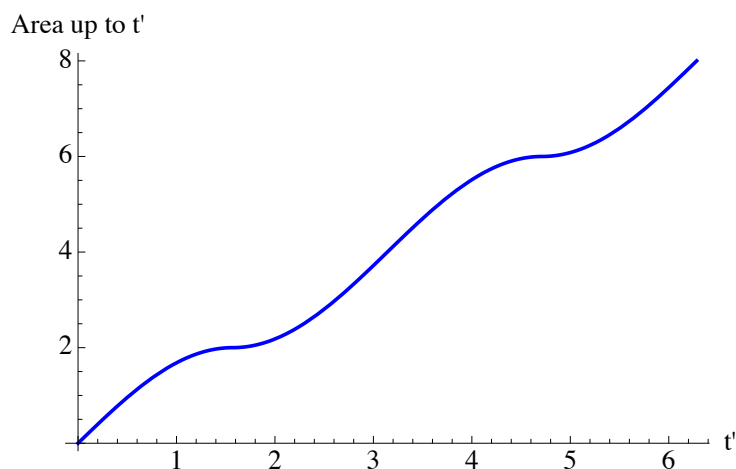


Figure 20: The area between the graphs up to t' .

We can however ask a different question related to the same graph. If we say that the graph shows a physical measurement, like velocity, then we have to be more careful with what it means to find the area between the curves.

In fact the area between the curves is equal to the displacement between two objects moving at a particular velocity. Let's say that the blue curve corresponds to object 1 and the red curve corresponds to object 2. We can see that we start off with object 1 moving with a positive velocity (let's define this to be in the positive x direction). At some point it starts to slow down and by time $t = \frac{\pi}{2}$ it has stopped moving completely (its velocity is zero), then it starts to move with a velocity in the opposite direction, before coming to a stop and then moving the other way again. This movement is an oscillatory movement exactly as you would have with a pendulum moving back and forth. The red curve is exactly the same thing, but starting with a velocity in the opposite direction. Think of two pendula swinging from the same pivot but in opposite directions. They will move away from each other, then towards each other, then pass each other, then do the same thing in reverse. We can envision what is happening in figure 21

Now, from a physical point of view, the area between the curves is telling us the distance between the two objects. In the part where the blue is higher than the red, to start with, we can say (if they both start off at the same position) that the objects are moving away from each other, but if they then stop and turn back (when they cross over the axis) they will start to move towards each other. If they are allowed to move towards each other then the distance between them must be allowed to decrease and so we can have negative contributions to this. In particular the distance between them no

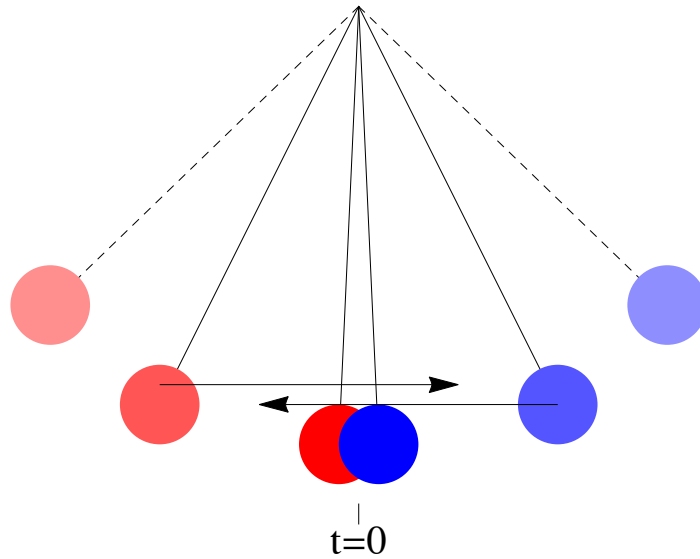


Figure 21: Two pendula moving back and forth. We define time $t = 0$ to be a time when they are passing each other in the middle, moving with opposite velocities. They slow, then change the direction of their velocities, then pass each other. Notice that the graph we have in figure 18 is only about the velocity. We have to integrate under the graph to find the distance (and in this case integrate between the graphs to find the displacement of the two objects from one another).

longer needs to be split up so that each area is positive. Some of the areas will give positive contributions (when they are moving apart) and some will give negative contributions to the distance between them (when they are moving together). In fact some of the time the distance between them (when measured with respect to a given direction) can be negative. In this case the integral is given by:

$$\int_0^{t'} \cos t - (-\cos t) dt = \int_0^{t'} 2 \cos t dt = 2 \sin t \Big|_0^{t'} = 2 \sin(t') \quad (152)$$

We thus see that the maximum distance between them is 2 and there is a time also when they are at displacement -2. This means that their positions have been reversed. ie. the object 1 is on the left hand side and object 2 is on the right hand side.

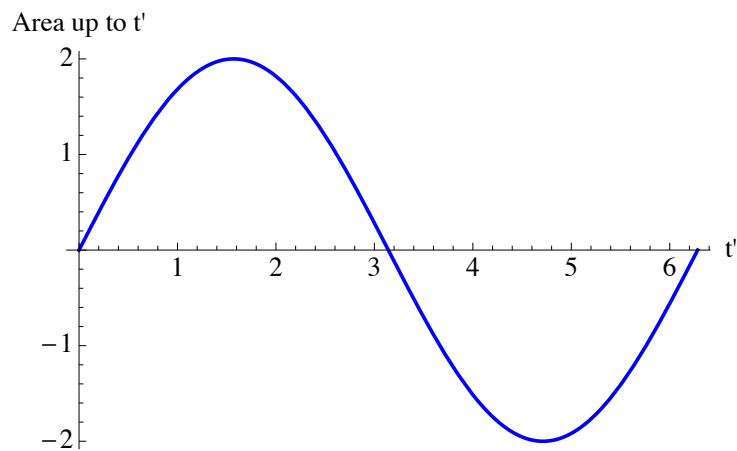


Figure 22: The distance between the two objects which are each undergoing oscillatory behaviour.

7.2 Volumes

Having understood how to calculate the areas between two curves as a limit of the Riemann sum we can start to think about how to approximate not areas but volumes. If we can calculate an area by adding together small rectangles, perhaps we can calculate volumes by adding together small boxes, or other shapes. In fact we will use not boxes but cylinders, and perhaps not cylinders as we normally think of them.

You probably think of a cylinder as a tube with a circular cross-section, but in fact that is a particular type of cylinder called a circular cylinder. The general idea of a cylinder is a three dimensional object with a constant cross section. Some examples of this are:

- A toblerone packet has a constant triangular cross-section. See figure 23

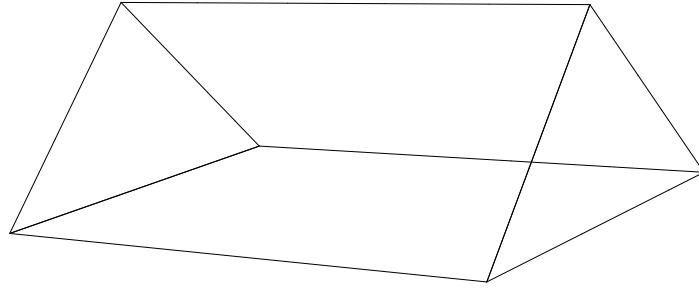


Figure 23: A shape with a constant triangular cross section.

- A box of matches has a rectangular cross section, thus this is a rectangular cylinder
- The cross section of a toilet roll is an annulus (a disk with a hole in it).

Definition:

A cylinder is a shape which is bounded by two identical surfaces (the two ends of the toblerone box, the toilet roll or the circular cylinder, etc.). The two surfaces are called B_1 and B_2 . The cylinder's length is h . The cylinder is the volume made by joining all lines perpendicular to the two endpoints. If the area of the base is A then the volume of the cylinder is:

$$V = Ah \tag{153}$$

For a circular cylinder the volume is:

$$V_{\text{circular cylinder}} = h \times \pi r^2 \tag{154}$$

For an equilateral triangle cylinder (figure 23):

$$V_{\text{equilateral triangle cylinder}} = h \times \frac{l^2\sqrt{3}}{4} \tag{155}$$

We can also work out the cylinder whose face is defined by some curve, or the intersection of some curves. How about if we have a cylinder whose face is defined by the area between the two parabolic curves:

$$f(x) = x^2 - 2, \quad \text{and} \quad g(x) = -x^2 + 2 \tag{156}$$

We now how to find the area between two curves. The intersection points of these two curves are $x = \pm\sqrt{2}$ and so the area between them is:

$$\int_{-\sqrt{2}}^{\sqrt{2}} -x^2 + 2 - (x^2 - 2)dx = \frac{16\sqrt{2}}{3} \quad (157)$$

So the volume of a cylinder with this face is:

$$V_{\text{cylinder with parabolic intersection face}} = h \times \frac{16\sqrt{2}}{3} \quad (158)$$

7.3 Breaking down a sphere into cylinders

Ok, so now we know what a cylinder is, how can we use this to start approximating the volumes of solids? We can start off by trying to calculate the volume of a sphere. We could approximate a sphere by a single circular cylinder. The smallest circular cylinder that we could fit around a sphere of radius r would have height r and a circular face with radius r . This is shown in figure 24 We will call this approximation to the volume of the sphere $V_{1 \text{ cylinder}}$ and it is given by $r \times \pi r^2 = \pi r^3$. This is clearly not a very good approximation, so we might try and approximate it by more cylinders of varying sizes. This is shown in figure 25 for an approximation with 4, 8 and 16 cylinders. So we can draw, these, but can we write down the volumes as the sum of cylindrical volumes? We can certainly take a limit where the widths of the cylinders are very thin and work out a general expression for a cylinder at a particular position within the sphere. In figure 26 we see a single, very thin (small h) cylinder at a particular radius x taken through the sphere. In this particular example the sphere is radius 1, but we can think about a general sphere of radius r . If we take a slice at position x then the radius of the slice is going to be the value of y shown in the plot, but because x , y and r form a right-angled triangle, we can see that $y = \sqrt{r^2 - x^2}$. The area of the slice is thus $\pi\sqrt{r^2 - x^2}^2 = \pi(r^2 - x^2)$ and the volume (where the thickness we now call Δx which we will shortly let go to an infinitesimal size), is:

$$V_{\text{Thin slice at position } x} = \Delta x \pi (r^2 - x^2) \quad (159)$$

So if we have a lot of these slices at different positions, then the total volume of them is:

$$V_{\text{total volume of thin slices}} = \sum_i \Delta x \pi (r^2 - x_i^2) \quad (160)$$

where the x_i are the positions of the centres of the disks (thin cylinders). We know how to take the limit of this, and we find that the total volume when

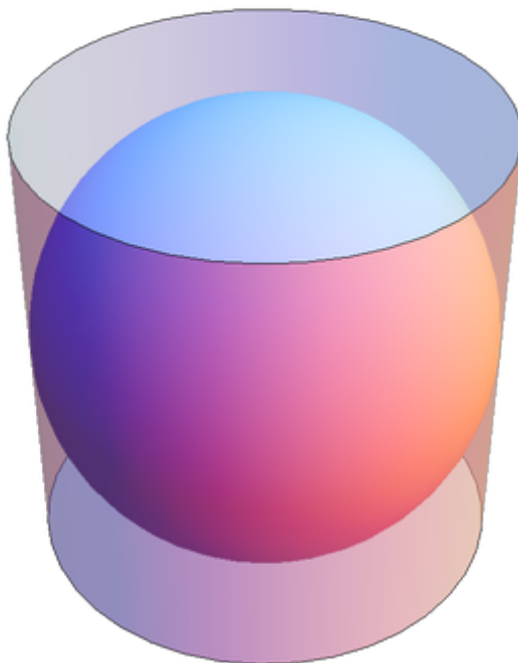


Figure 24: A sphere of radius r approximated by a circular cylinder of height r and radius r .

we take the number of disks to infinity (their widths to zero size):

$$V_{sphere} = \int_{-r}^r \pi(r^2 - x^2)dx = \pi\left(rx - \frac{x^3}{3}\right)\Big|_{-r}^r = \frac{4}{3}\pi r^3 \quad (161)$$

So miraculously we used exactly the same trick that we used previously with the Riemann sum for a two dimensional area to calculate a three dimensional volume!

7.4 Some more examples of volumes by cylindrical cross-sections

We are very familiar with spheres, but we can define shapes very simply by giving a function and rotating that function around an axis. Let's look at an example.

1. We can define an area by the region between the function $y = \sqrt{x}$ and the x -axis, and a line at $x = 1$. This looks like the left hand side of

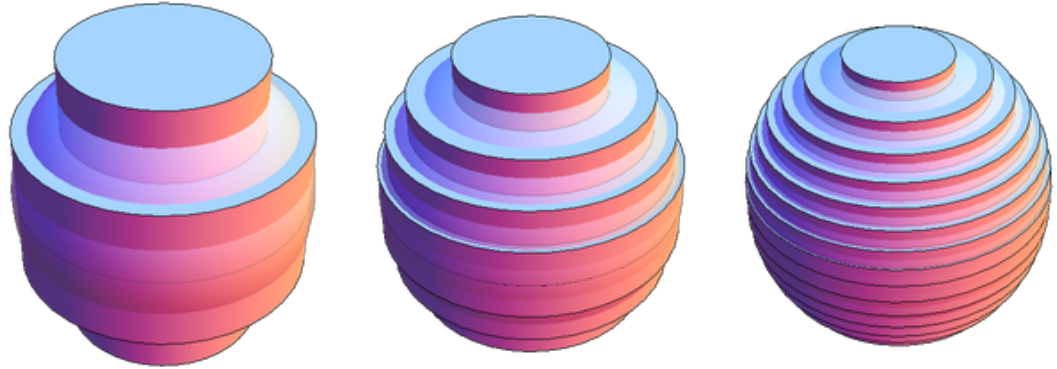


Figure 25: A sphere of radius r approximated 4 cylinders, 8 cylinders and 16 cylinders

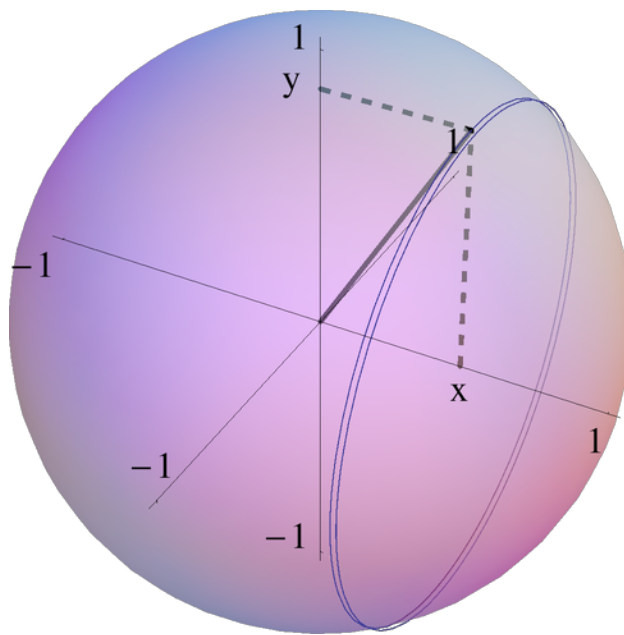


Figure 26: A single thin cylinder taken as a slice through the sphere

figure ???. We can then imagine taking this wedge and letting it rotate around the x -axis. If it does this then it will sweep out a volume in the (x, y, z) plane which looks like the right hand side of figure 27 Again, just as we did in the case of the sphere, we can think about taking

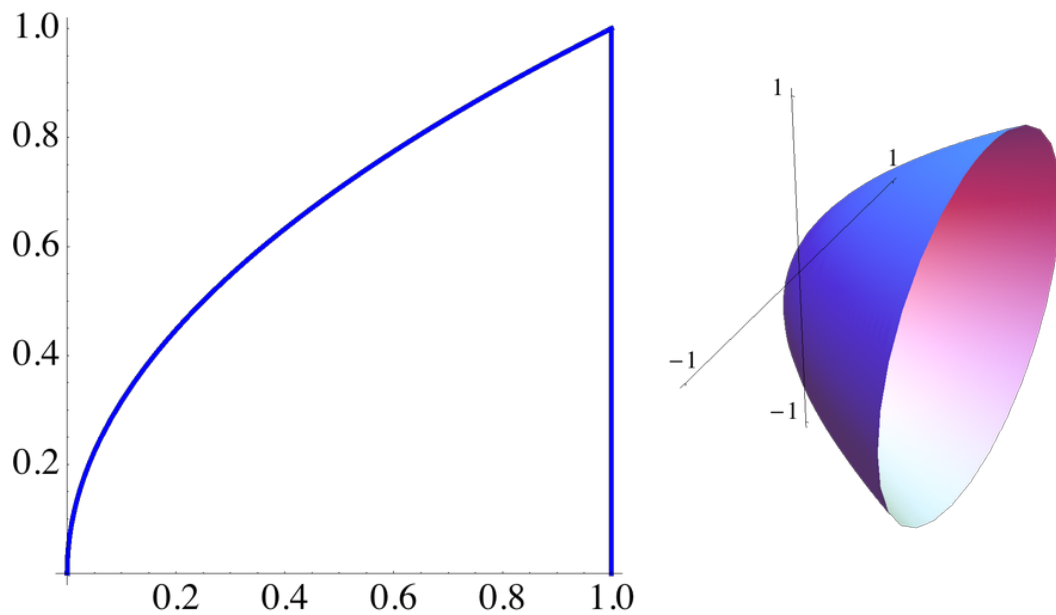


Figure 27: On the left, the function $y = \sqrt{x}$ and on the right the volume swept out by the wedge defined by this function. We have to remember that while the line $y = \sqrt{x}$ will only sweep out the surface, we are interested in the whole volume inside, caused by sweeping out the area underneath the curve.

slices through this shape. Again, the slices will be circular disks. At a given position x , the radius of the disk will be the height that the function goes, which is \sqrt{x} and thus the area of a disk at position x will be $\pi(\sqrt{x})^2$ and the volume will be $\Delta x \pi \sqrt{x}^2 = \Delta x \pi x$. We can again take the sum of the volumes of all the disks going from $x = 0$ to $x = 1$ and take the limit that their thickness goes to zero, while their number goes to infinity, to get:

$$Volume_{\sqrt{x} \text{ swept around } x \text{ axis}} = \int_0^1 \pi x dx = \frac{\pi}{2} \quad (162)$$

2. We made a very particular choice in the previous example, to sweep the shape around the x axis. We could have chosen any axis to rotate it around. let's try and rotate it around the y axis instead. This will give us a shape given in figure 28 This time we want to take horizontal slices through the figure. The thickness of the disks will now be Δy because the thickness is in the y direction. We now have to ask what the shape of the disks will be. We can imagine that the disks will be annuli (disks

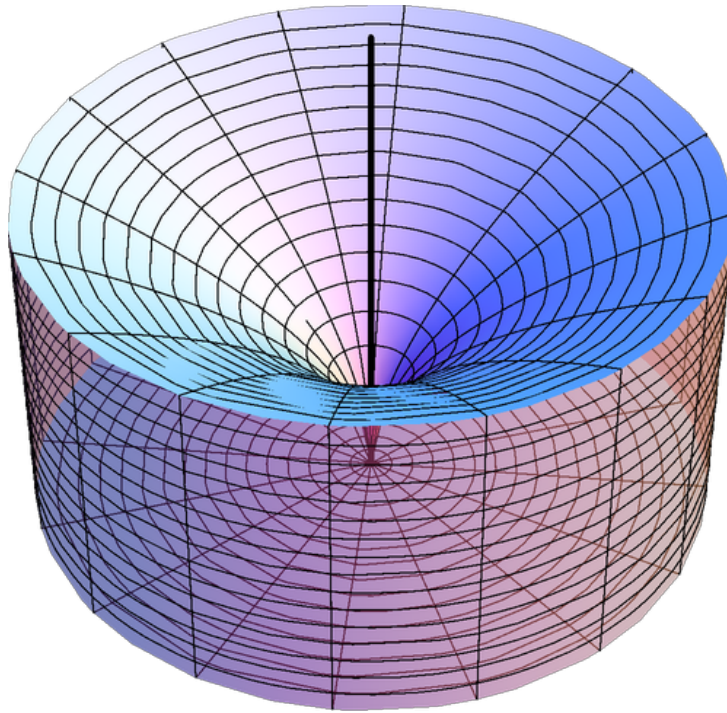


Figure 28: The function $y = \sqrt{x}$ rotated about the y axis which is shown by the vertical black line.

with holes in the middle). For y near the top of the shape, they annuli will be very thin, but towards the bottom they will be very thick, with only a small hole in the centre. Figure 29 shows what would happen if we approximated the shape by five disks and then pulled them apart for effect

For a disk at a given height up the y axis, the inner radius of the annulus is going to be y^2 because $y = \sqrt{x}$ and so $x = y^2$. The outer radius is always going to be 1. So the area of of one of the annuli at position y is going to be the area of a disk of radius 1 minus the area of a disk of radius y^2 . This gives: $area = \pi(1^2 - (y^2)^2) = \pi(1 - y^4)$. The volume of one of these disks is then: $\Delta y \pi(1 - y^4)$ and thus we can add them all up in a Riemann sum and take the limit, to get:

$$V_{\sqrt{x} \text{ rotated about the } y \text{ axis}} = \int_0^1 \pi(1 - y^4)dy = \pi\left(1 - \frac{1}{5}\right) = \pi\frac{4}{5} \quad (163)$$

3. In fact we can chose a different axis again to rotate our shape around. This time let's chose a different shape. If we chose the shape defined by

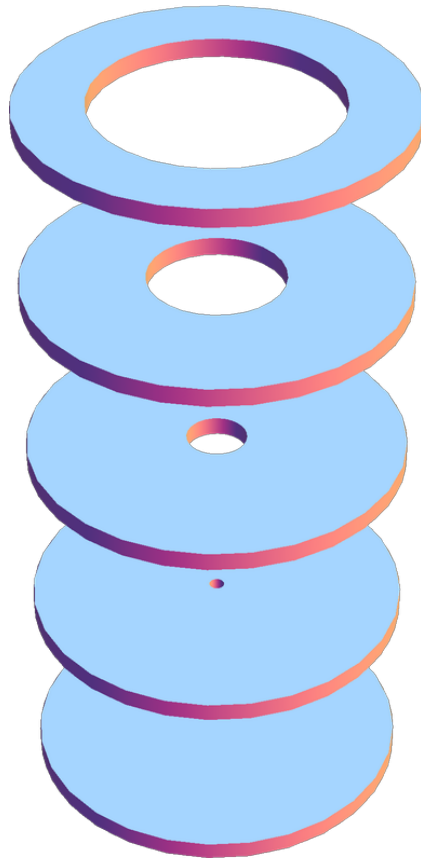


Figure 29: The cylindrical segments of figure 28 if we approximated it by five disks and pulled them apart.

the area in between the functions $y = x^2$ and $y = x$ we get the shape on the left hand side of figure ???. Let's chose to rotate this around the line $x = -1$ - again, a vertical line, but now displaced. This will give the figure shown on the right hand side of figure ??. Again, we have to ask what the cylinder is going to look like at height y (ask yourself what a thin slice of the shape would look like at height y). This time both the inner and the outer radii of the annulus are going to change as we change y . Now the inner radius is going to be $1 + \sqrt{y}$ and the outer radius is going to be $1 + y$. You can see this from the cross section in figure 31 Thus we can see that the annulus at height y is going to have area $\pi((1 + \sqrt{y})^2 - (1 + y)^2)$ (the 1 is because we are going from the point $x = -1$ to the point y and \sqrt{y}). The volume is just this multiplied by Δy . Thus, again, adding up all the pieces, we get a total

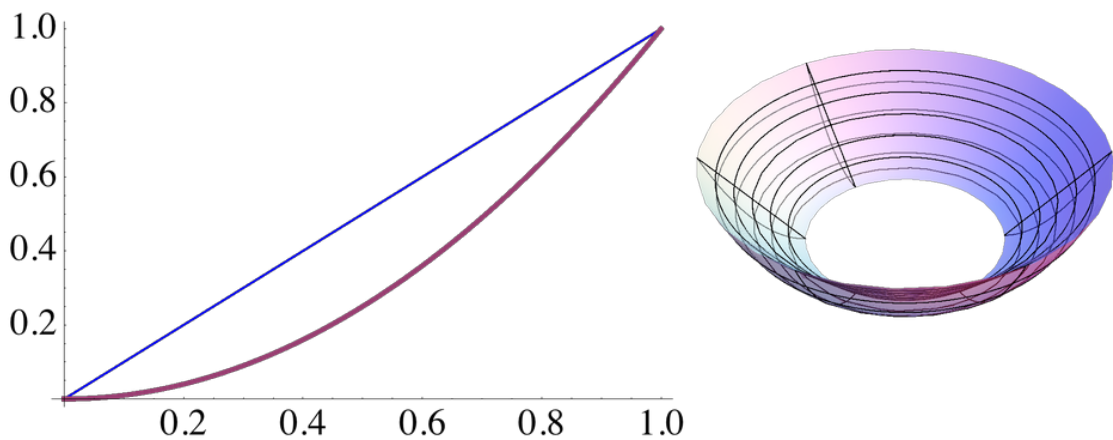


Figure 30: The area between the lines $y = x$ and $y = x^2$ rotated about the line $x = -1$.

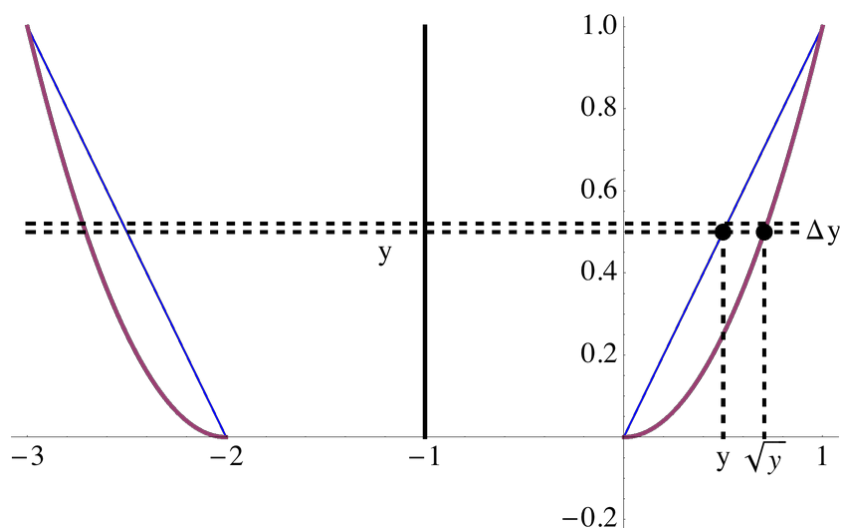


Figure 31: A cross-section through the shape in figure 30 showing the inner and outer radii of the annuli

volume of:

$$V_{\text{shape in figure 30}} = \int_0^1 \pi((1 + \sqrt{y})^2 - (1 + y)^2)dy = \frac{\pi}{2} \quad (164)$$

4. We have come up with some examples, but let's try and come up with a general expression for this method, when we rotate a function about the x axis and when we rotate it about the y axis.

(a) Let's first consider the case when we revolve a function $f(x)$ in the

range between 0 and x' about the x axis. We can take cylindrical slices through it in the (z, x) plane (see figure 26). The radius of a shell at x is $f(x)$ and so its area is $\pi f(x)^2$. Thus, its volume is $\Delta x \pi f(x)^2$ and then, after adding up all disks and taking the limit that their number goes to infinity gives:

$$Volume = \int_0^{x'} \pi f(x)^2 dx \quad (165)$$

If we can calculate this integral, then we can work out the volume.

- (b) How about if we rotate about the y -axis. In this case we take disks at a given position in the y direction and thickness Δy . If you are taking the shape to be the area below the function, swept out around the x axis you will have a set of annuli, each with outer radius x' and inner radius given by $f^{-1}(y)$. The slices will go from 0 to $f(x')$ (as we are taking them in the y direction. Thus, the total volume of the shape is:

$$Volume = \int_0^{f(x')} \pi(x'^2 - f^{-1}(y)) dy \quad (166)$$

We may be able to do this, but we can only do so if we know the inverse function. Note also that the above is only true when we have a monotonically increasing function in the region of interest. If we do not then the inverse function is not single valued and so you have to be more careful with the general definition.

Calculating the inverse function is not always easy and so sometimes it is easier to perform a calculation by taking the cylinders in another way. This is the method of cylindrical shells.

7.5 Volumes by cylindrical shells

Sometimes it is easier to calculate the volume of a solid (formed by the revolution of a surface about an axis of revolution) by dividing it up in a different way. We will still use cylinders, but this time they will be cylindrical shells, where the height of the cylindrical shell mimics the function. Figure 32 illustrates how this works for a particular function (in this case $f(x) = \sqrt{1 - x^2}$ between $x = 0$ and $x = 1$).

Let's first think about the volume of a single shell which has inner radius r_1 , outer radius r_2 and height h as shown in figure 33. The volume of such a figure is just going to be the height times the area of the cylindrical face (which in this case is an annulus). The area of the annulus is:

$$Area_{annulus} = \pi(r_2^2 - r_1^2) \tag{167}$$

so the volume of this object is $h\pi(r_2^2 - r_1^2)$. We can rewrite this in a way

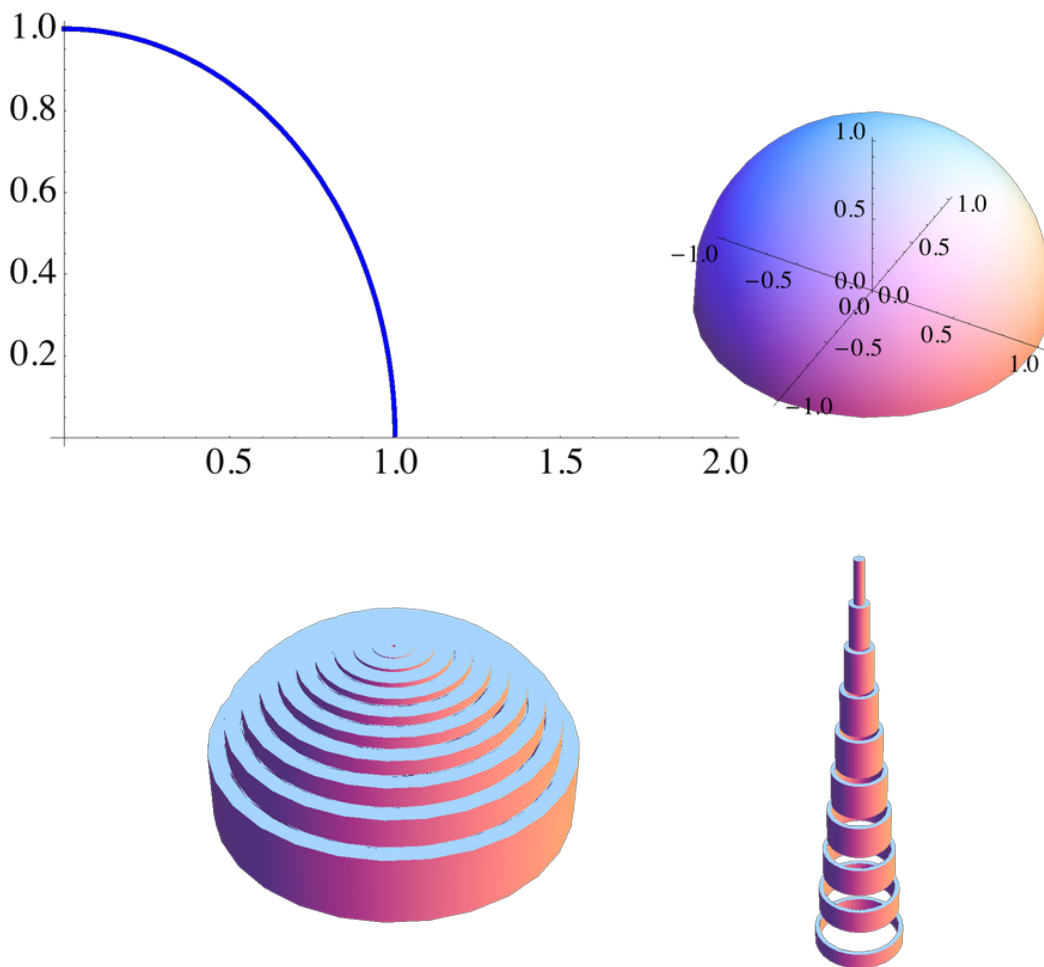


Figure 32: Top Left image: the function $\sqrt{1-x^2}$. Top right image: The volume formed by revolving this about the y axis. Bottom left image: Breaking this shape up into cylindrical shells (think of taking a wire loop of varying diameters and slicing vertically through the function. Bottom right image: The cylindrical shells when pulled apart (this is just for illustration purposes). In this case the central shell is taller than the outer shells because in the centre of the shape the function is highest and towards the edges it is lowest.

which will become much more useful shortly:

$$h\pi(r_2^2 - r_1^2) = h\pi 2\frac{r_2 + r_1}{2}(r_2 - r_1) \quad (168)$$

We can see then that $\frac{r_2+r_1}{2}$ is the radius up to the middle of the shell's edge and $r_2 - r_1$ is the thickness of the shell. Eventually we are going to want to

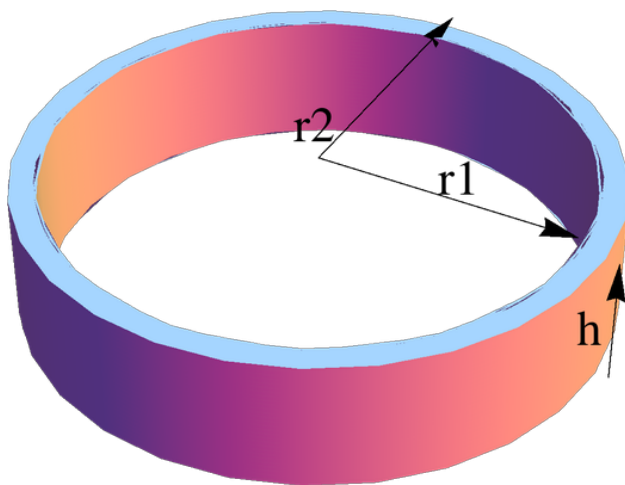


Figure 33: A cylindrical shell of height h , inner radius r_1 and outer radius r_2

take thin shells and add them together, so we can think of $r_2 - r_1$ as Δx . Also, when the shells are thin, we can just call the distance to the middle of the shell's edge x . If we want to mimic the shape made by rotating a given function around the y -axis, the height of the cylindrical shell will then just be $f(x)$ (see figure 32). Thus we can write the volume of a given shell which is of radius x as:

$$V_{\text{cylindrical shell}} = \Delta x \pi 2x f(x) \quad (169)$$

If we add all of the shells which are different radii x then we will end up with a Riemann sum and we can take the limit of the Riemann sum. This gives the exact answer to the volume as:

$$\int_a^b 2\pi x f(x) dx \quad (170)$$

where a and b are the inner and outer radii of the shape of interest (in the case of figure 32 this will be $x = 0$ and $x = 1$ but you have to be careful with this as we will see in the next example.

The important thing to note about this expression is that when we used the method of cylindrical cross-sections we needed to know $f^{-1}(y)$ which is not always possible. Now we just need to know how to integrate $x f(x)$ which may be easier.

Let's look at an example:

7.6 Calculating the volume of a torus using the method of cylindrical shells

A torus is created by taking a circular cross section in the (x, y) plane and revolving it around the y axis, as shown in figure 34. It will actually be easier

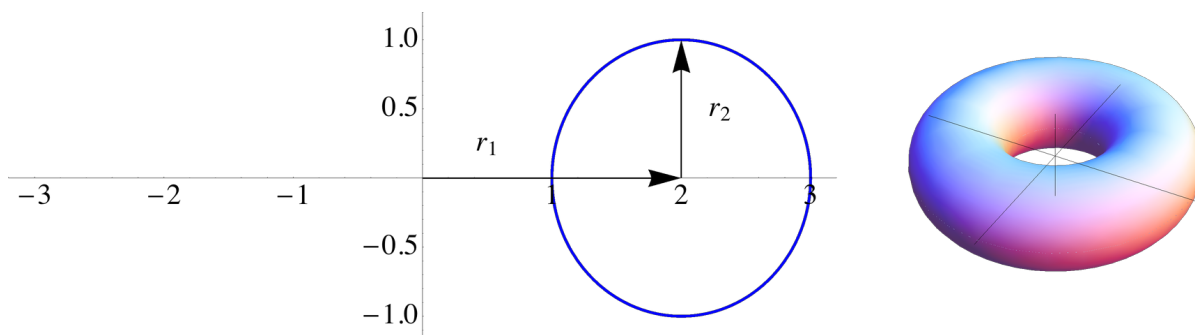


Figure 34: A torus formed by revolving a circle around the y axis.

to work with half of this torus (the top half or the bottom half) because the function is then single valued. If we chose the top half, then the equation that defines the half circle is $y = \sqrt{r_2^2 - (x - r_1)^2}$. We want to slice this function into cylindrical pieces which will look like figure 35 if we take just five pieces. Of course we will take the limit that we have an infinite number of pieces, but for now this is an illustrative approximation. The torus starts at $x = r_1 - r_2$ and ends at $x = r_1 + r_2$. We can thus write down the formula for the integration by the method of singular shells as (remembering that we are just looking at the half torus, so to get the full torus volume we multiply by two):

$$2 \int_{r_1 - r_2}^{r_1 + r_2} 2\pi x \sqrt{r_2^2 - (x - r_1)^2} dx \quad (171)$$

Now we do a u -substitution of the form $x - r_1 = ur_2$. This gives the integration limits in the u variable as ± 1 , so the integration is then:

$$2 \int_{-1}^1 2\pi(ur_2 + r_1)r_2\sqrt{1 - u^2}r_2 du \quad (172)$$

We can split this up into two integrals:

$$2 \int_{-1}^1 2\pi(ur_2)r_2\sqrt{1 - u^2}r_2 du + 2 \int_{-1}^1 2\pi(r_1)r_2\sqrt{1 - u^2}r_2 du \quad (173)$$

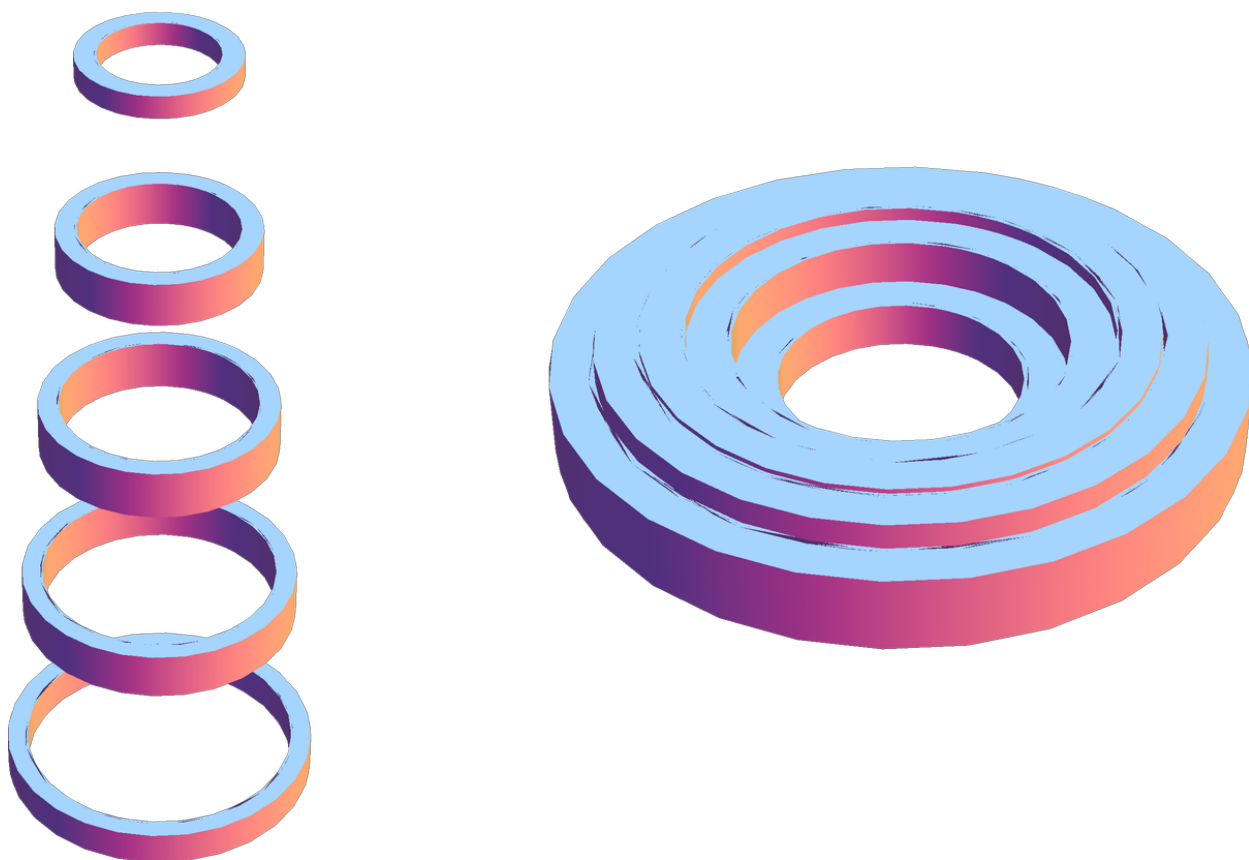


Figure 35: The cylindrical shells making up the torus when it is split into just five pieces. The left image is a pull-apart of the right image. You can see that the outer ring and the inner ring have the minimum height, and the third ring has the maximum height because it is making up the middle of the torus where it is heights.

The first integral is an odd function which, when integrated from -1 to 1 necessarily gives 0 . The second term is an even function so we can write:

$$Volume_{torus} = 4 \int_0^1 2\pi(r_1)r_2^2\sqrt{1-u^2}du \quad (174)$$

and the integral of $\int \sqrt{1-u^2}du$ is $\frac{1}{2}(\sqrt{1-u^2}u + \sin^{-1}(u))$. The first term vanishes at both ends of the limits (both at 0 and at 1) and the second term has a contribution only at $x = 1$, ($\sin^{-1} 1 = \frac{\pi}{2}$). Thus the final result for the integration gives:

$$Volume_{torus} = 2\pi^2r_1r_2^2 \quad (175)$$

This in itself is a rather lovely result because it can be written as $2\pi r_1 \times \pi r_2^2$ which is the perimeter of a circle of diameter r_1 times the area of a circle of radius r_2 , which is roughly (but not quite) what a torus is. If you imagine a cylinder that is $2\pi r_1$ in length and r_2 in radius then this is the volume that it represents. You can imagine then curling this round to make a torus (take a tube and join the ends). There is a rather lovely explanation here: <http://whistleralley.com/torus/torus.htm> as to why this is the case. It might seem obvious at first glance, then it seems surprising on further inspection, and finally it should seem very elegant!

7.7 Arc lengths

We have now learnt a great deal about various properties of curves (ie. functions). We can study their gradients, the areas under them and between them, we can even study properties of functions which have discontinuities and which extend all the way to ∞ . We can look at the revolution of a curve about an axis and study the volume enclosed using cylinders of various forms and we have a good understanding of the link between finite sums of pieces which make up an area or volume and the limit that these pieces become infinitesimally small and thus how we end up with an integral from a Riemann sum.

There is one piece of the puzzle left, which is to know the length of a curve

between two points. Let's say we have some function $f(x)$ and we want to know the length of the curve (ie. how long a piece of string would be that went along the curve and stretched from some point a to another point b on the curve). The most naive guess would be to lie the string straight between the points $f(a)$ and $f(b)$ and calculate the length like that. Looking at figure 36 we can see that this will give:

$$L = \sqrt{(b - a)^2 + (f(b) - f(a))^2} \quad (176)$$

This is clearly a pretty bad approximation however. Clearly a better approx-

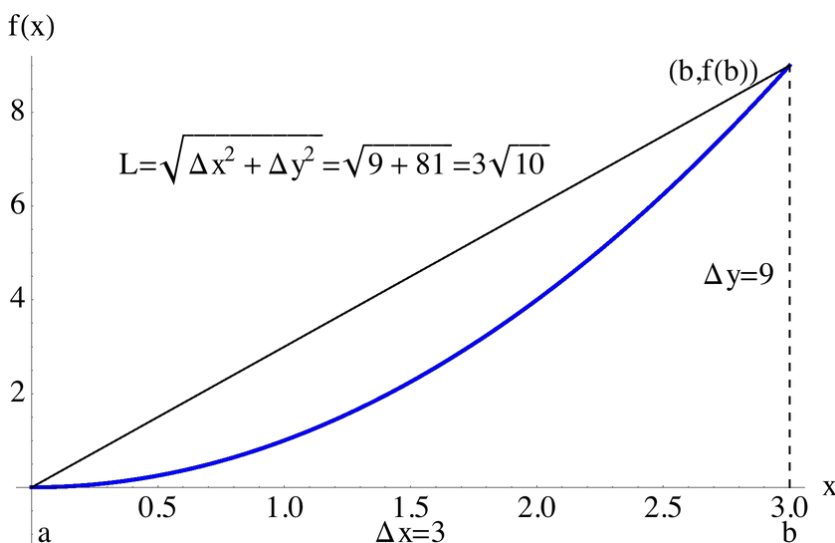


Figure 36: The most naive guess we could make for the length of our curve would be to approximate it by a straight line running from one end to the other.

imation would be to split the line up into many parts and approximate each of these by a straight line. Already alarm bells should be ringing and you should be thinking of integration. You are going to add up lots of pieces and hopefully take the limit that each of these pieces will be very small. This sounds like the beginnings of an integral problem.

In figure 37 we split the curve up into 3 parts, each with the same Δx but now Δy is a function of the position of the i^{th} point that we are looking at. We can write down a general expression for the i^{th} line approximation. The length of it will be:

$$L_i = \sqrt{\Delta x^2 + (f(x_i + \Delta x) - f(x_i))^2} \quad (177)$$

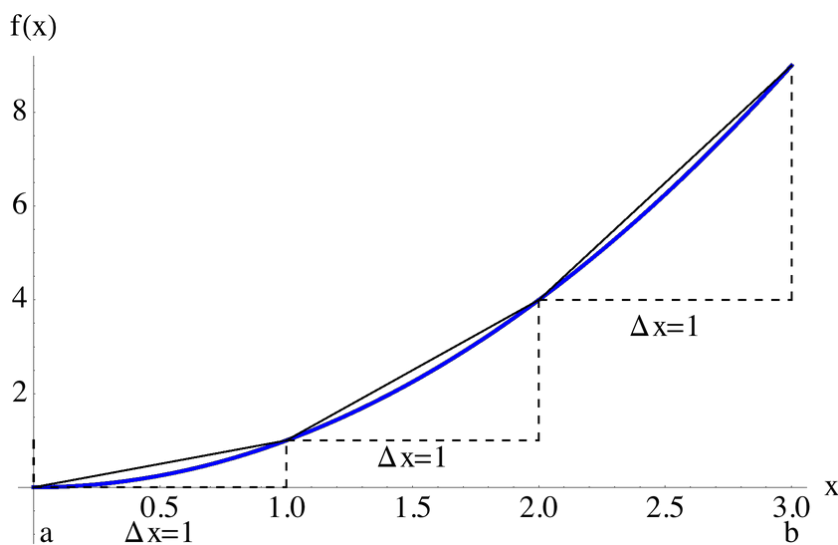


Figure 37: The same curve as in figure 36 but this time split into 3 sections

Let's pull out a factor of Δx to make this look more like a term in a Riemann sum:

$$L_i = \Delta x \sqrt{1 + \frac{(f(x_i + \Delta x) - f(x_i))^2}{\Delta x^2}} \quad (178)$$

So the total length is going to be:

$$L \approx \sum_{i=0}^{n-1} \Delta x \sqrt{1 + \frac{(f(x_i + \Delta x) - f(x_i))^2}{\Delta x^2}} \quad (179)$$

Clearly, the more lines we draw along the curve, and the shorter they are, the better is going to be the approximation. We want to take the limit that these lines go to zero size. We can look at the term inside the square root and see that the expression $\frac{(f(x_i + \Delta x) - f(x_i))}{\Delta x}$ as the Δx goes smaller (ie. the number of pieces goes larger) will get closer and closer to our definition of the gradient of the slope, ie. the derivative of the function at x_i . That is:

$$\begin{aligned} L &= \lim_{\Delta x \rightarrow 0} \sum_{i=0}^{n-1} \Delta x \sqrt{1 + \frac{(f(x_i + \Delta x) - f(x_i))^2}{\Delta x^2}} = \lim_{\Delta x \rightarrow 0} \sum_{i=0}^{n-1} \Delta x \sqrt{1 + \left(\left. \frac{df(x)}{dx} \right|_{x=x_i} \right)^2} \\ &= \int_a^b \sqrt{1 + \left(\frac{df(x)}{dx} \right)^2} dx \end{aligned} \quad (180)$$

where we first took the limit of the terms inside the square root, then we looked at the whole thing and took the limit that the number of pieces that

we were studying went to infinity. This isn't strictly the best way to do it, but it's rather more intuitive to see that as we take the number of pieces smaller and smaller, the expression inside the square root looks more and more like the derivative, and then we have an infinite number of such pieces in our Riemann sum which add together to give the integral expression.

And that's it! That's the expression for the length of a curve from a to b . Now let's look at some examples:

1. Let's first work out the circumference of a circle using this method. A circle can be written as two curves: $f(x) = \pm\sqrt{r^2 - x^2}$ where the positive part is the half circle above the x axis, and the negative part is below the x axis. We can simply say then that the circumference of a circle is twice the length along the curve $f(x) = \sqrt{r^2 - x^2}$ from $x = -r$ to $x = r$ (see figure 38). We simply plug it into our equation for the

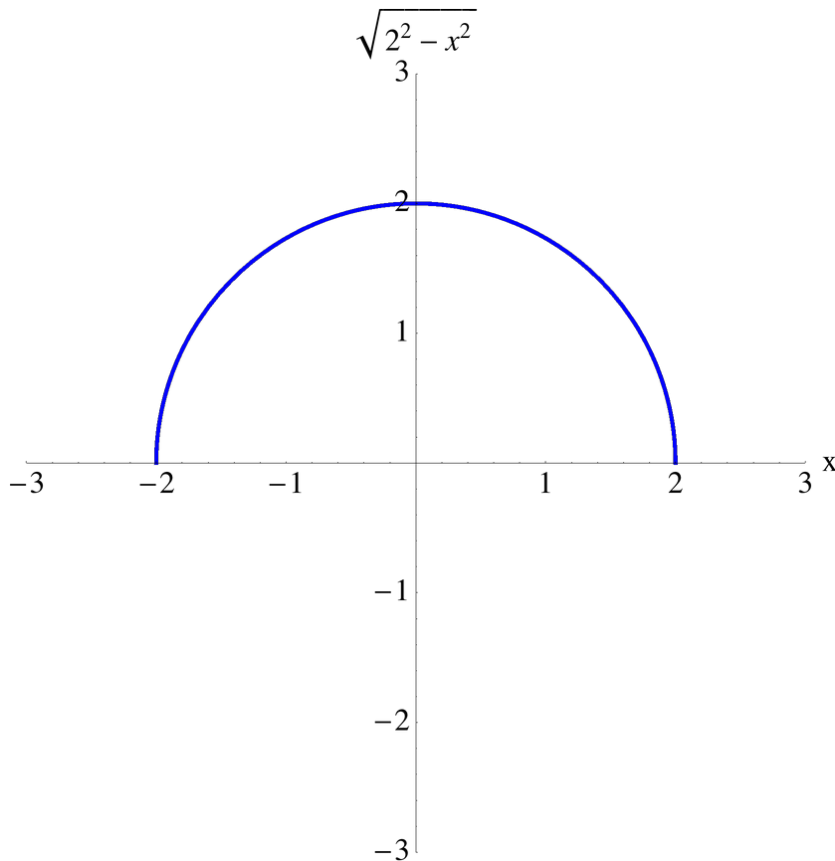


Figure 38: The top half of a circle. The circumference of the circle will simply be twice this arc length.

length of a curve and see what we get. First we need the derivative $\frac{dy}{dx}$

which is $\frac{-x}{\sqrt{r^2-x^2}}$. Then the length of the curve is given by:

$$L = \int_{-r}^r \sqrt{1 + \frac{x^2}{r^2 - x^2}} dx = \int_{-r}^r \frac{r}{\sqrt{r^2 - x^2}} dx \quad (181)$$

we use a substitution $x = r \sin \theta$ to get:

$$L = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r d\theta = r\pi \quad (182)$$

So this is the length of the half circle, so the perimeter of a full circle is twice this = $2\pi r$ as we know that it should be.

2. Let's find an example which at first sight looks rather complicated. We will find the length of this arc between 0 and $\frac{\pi}{4}$:

$$\begin{aligned} y &= \ln(\sec x) \\ \frac{dy}{dx} &= \frac{\sec x \tan x}{\sec x} = \tan x \\ &= \int_0^{\frac{\pi}{4}} \sqrt{1 + f'(x)^2} dx = \int_0^{\frac{\pi}{4}} \sqrt{1 + \tan^2 x} dx = \int_0^{\frac{\pi}{4}} \sqrt{\sec^2 x} dx \\ &= \int_0^{\frac{\pi}{4}} \sec x dx = \ln |\tan x + \sec x| \Big|_0^{\frac{\pi}{4}} \\ &= \ln(\sqrt{2} + 1) \end{aligned} \quad (183)$$

Note that we were able to equate $\sqrt{\sec^2 x}$ with $\sec x$ only because we were interested in this function at 0 and $\frac{\pi}{4}$ at which points it is positive.

3. Exercise: Calculate the arc length of the function $y = \sqrt{x - x^2} + \arcsin \sqrt{x}$ between 0 and π .

It is important to note that Stewart introduces arc length using the formalism of parametric equations. We do not do this here and while it will not be examined, it is a useful technique to understand.

7.8 Average Value of a function

If we have a function in some range $[a, b]$ we can ask what the average value of the function is. We are going to do this in a very intuitive way. We split up the curve into n points and take the average of those points, and then ask what happens to that number as we take the number of points to infinity. Let's consider the curve $y = x^2 - x^3$ between -3 and 3 . If we take seven points along the curve (ie. at $x = -3, -2, -1, 0, 1, 2, 3$) We will get the function values: $f(x) = 36, 12, 2, 0, 0, -4, -18$ as can be seen in figure 39 The average value of the function when we take just seven sample points is thus: $\frac{36+12+2+0+0-4-18}{7} = 4$. Figure 40 shows what happens when we take more and more points. We see that the mean value converges to a fixed number, in this case 3. But there is a better way to do this and again, it involves integration. What we have written down so far is that for $n + 1$ sample points, which have a distance Δx in between them, starting at $x = a$,

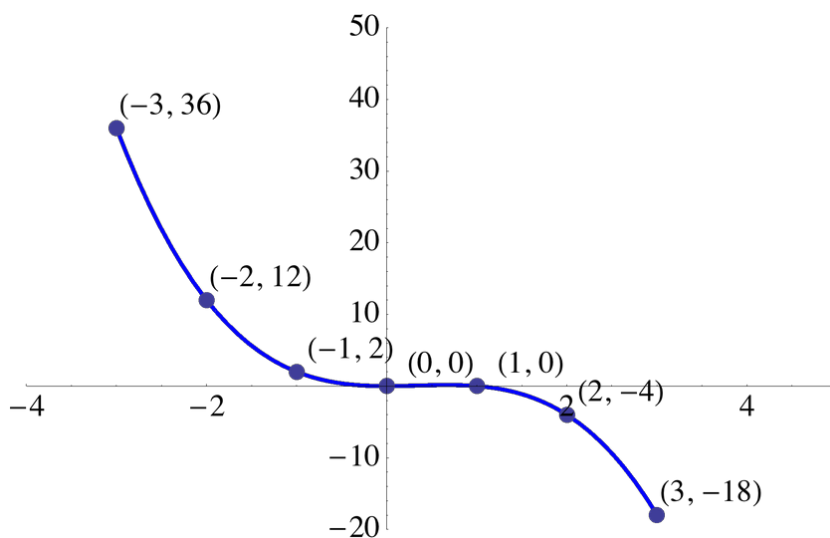


Figure 39: $y = x^2 - x^3$ between $x = -3$ and 3 with 7 sample points.

Mean value

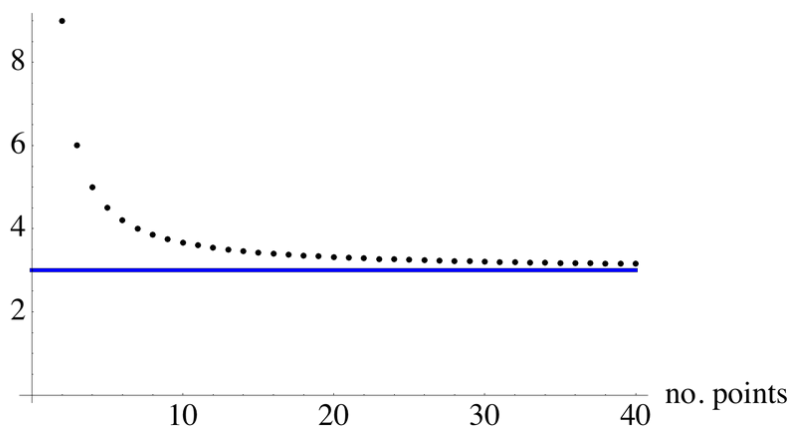


Figure 40: The mean value of the average of a set of n sample points in $y = x^2 - x^3$ between $x = -3$ and 3 .

the mean value of the function is:

$$Mean = \frac{\sum_{i=0}^n f(x_i)}{n} \quad (184)$$

If we multiply and divide this by Δx we get:

$$Mean = \Delta x \frac{\sum_{i=0}^n f(x_i)}{\Delta x n} \quad (185)$$

but Δxn is just $b - a$ and so we can take the limit of this as a Riemann sum and get:

$$Mean = \frac{\int_a^b f(x)dx}{(b - a)} \quad (186)$$

This makes intuitive sense. Just as the mean of a sample of items is just the sum of the items divided by the number of them, so the mean value of a function is just the integral of the function (the limit of a Riemann sum) divided by the total distance in between the limits. In fact we can think about this in another way. If the mean of the function is M and the interval is $b - a$ then $M(b - a)$ as the area of a rectangle between b and a with height M and this is equal to the total integral of the function - ie. the area under the curve. This says that we can write the area under the curve as the mean value of the function times the integral, which again makes intuitive sense. Taking the example from figure 39, we find that the mean value is:

$$Mean = \frac{\int_{-3}^3 x^2 - x^3}{6} = \frac{\left. \frac{x^3}{3} - \frac{x^4}{4} \right|_{-3}^3}{6} = 3 \quad (187)$$

and thus the area under the curve is equal to 3 times the interval (6)=18. In this example you have to be a bit careful because there is both a positive and negative contribution to the area under the curve but you can see from figure 41 that the total area under the curve (both the negative and positive contributions added together) equals the area under the mean value line (the red line) which is at $y = 3$. We can go further than this and make a statement about the mean value itself. This is given by the Mean Value Theorem for integration:

- **The Mean Value Theorem for integrals:** If f is a continuous function on $[a, b]$, then there exists at least one number c in $[a, b]$ such that:

$$\int_a^b f(x)dx = f(c)(b - a) \quad (188)$$

This says that the mean value of the function must be a value of the function which is within the upper and lower bounds of f between a and b . In other words, the line $y = \text{Mean Value of } f$ must intersect f at least once between $x = a$ and b . We can see that this is true in figure 41 where the intersection of the red line (the mean line) with the function is at roughly $(-1.175, 3)$. The value c in the mean value theorem is found by saying that, in this case: $x^2 - x^3 = 3$. This has only one real solution which is at $x \approx -1.175$.

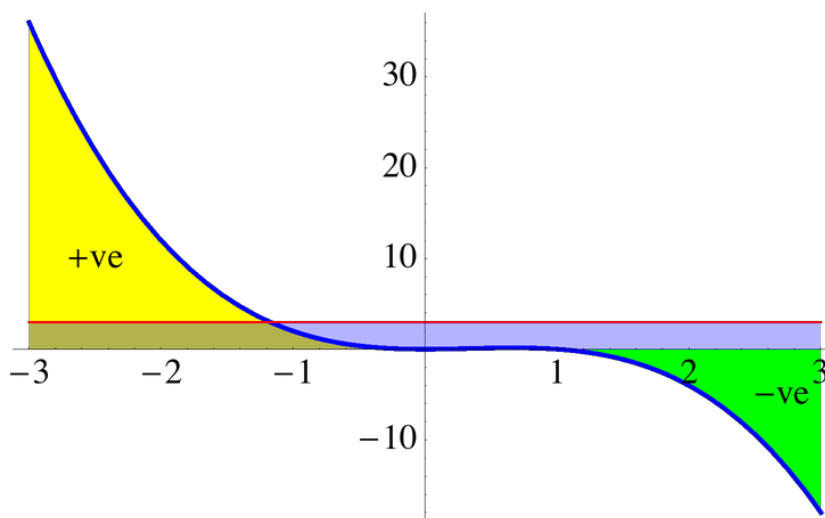


Figure 41: The integral between the curve and the x axis is equal to the integral under the red line which itself is the mean value of the function between -3 and 3 .

Let's look at another example:

If $f(x) = \cos^4 x \sin x$ and we want to know the mean value of this function between 0 and π , this is given by:

$$Mean = \frac{\int_0^\pi \cos^4 x \sin x dx}{\pi} = \frac{-\frac{\cos^5 x}{5} \Big|_0^\pi}{\pi} = \frac{2}{5\pi} \quad (189)$$

Now we can ask what the value of c is for which:

$$\int_0^\pi \cos^4 x \sin x dx = f(c)\pi \quad (190)$$

Which is the same as asking for what value of c is:

$$f(c) = Mean = \frac{2}{5\pi} \quad (191)$$

In fact there are four solutions to this equation for c between 0 and π as can be seen in figure 42. These values are approximately $(0.13, 0.88, 2.26, 3.01)$

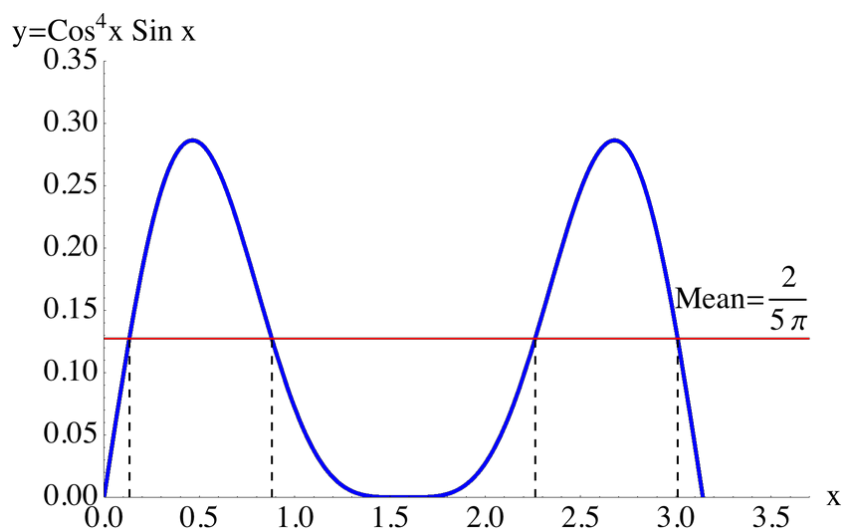


Figure 42: There are four values of c in between 0 and π for which $\int_0^c \cos^4 x \sin x dx = f(c)\pi$.

8 The Binomial Theorem

The aim of this section will be that you can understand: The combinatorics of lottery tickets and poker hands, understand why there are likely to be several people in a small room with the same birthday, understand the combinatorics of the binomial expansion $(a+b)^n$ and understand how this links in with Pascal's triangle. You should also be able to manipulate combinatorial expressions.

8.1 Ordering objects

How many ways can we order n objects? Let's start with just 2: a and b can be ordered as ab and ba . With 3 objects we have abc , acb , bac , bca , cab , cba , ie. 6 combinations. With 4 objects we can start to see a pattern. The first object a can go in four possible positions. Having chosen these positions there are then 3 possible positions for b , then there are two left for c and

finally there will only be a single place for each of these that you can put d . So, there are $4 \times 3 \times 2 \times 1$ possible ways of ordering 4 objects. This reasoning leads to the general statement that to order n objects there are:

$$n \times (n - 1) \times (n - 2) \times (n - 3) \dots \times 2 \times 1 = n! \quad (192)$$

ways. This is called n factorial (or sometimes n shriek). $4! = 4 \times 3 \times 2 \times 1 = 24$. How many ways are there of ordering 0 objects? Well there is one way, and it is to simply have no objects, so we define $0!$ to be 1. To ask how many ways are there of order 3.5 objects is clearly absurd and so $3.5!$ will be meaningless. We can see also that $n! = n \times (n - 1)!$. For $r < n$:

$$\frac{n!}{r!} = \frac{n(n - 1)(n - 2)(n - 3) \dots (r + 1)r!}{r!} = n(n - 1)(n - 2)(n - 3) \dots (r + 1) \quad (193)$$

so $\frac{6!}{4!} = 6 \times 5 = 30$. It is also true that as n tends to *infy*:

$$n! \sim n^{\frac{1}{2}} \left(\frac{n}{e}\right)^n \quad (194)$$

Factorials get large very very quickly: $10! = 3628800$. This is how many ways there are of ordering ten objects, which doesn't sound like a very big thing but you can see that there are over 3 million ways to do it.

8.2 Choosing r objects from n objects

Let's say that we have 6 objects ($\{a, b, c, d, e, f\}$) and we want to know the number of ways of picking 4 of them (where the order doesn't matter, so $\{a, b, c, d\}$ is no different from $\{a, b, d, c\}$). We can imagine writing down all of the possible permutations of the 6 objects (of which there are $6!$ as we know) and then just taking the first four items. For instance, these are the first few permutations and the items in brackets are the ones that we pick as

our 4:

(abcd)ef
 (abcd)fe
 (abce)df
 (abce)fd
 (abdc)ef
 (abde)cf
 (abde)fc
 (abdc)ef
 (adb)ef
 (adb)fe
 ...etc....

If we pick the items in the brackets then clearly we are going to get many repetitions of the same thing. For instance, the first two items are the same (both (abcd)). Also many of the items in the brackets are the same, up to a reordering, and we said that we didn't care about the order, so (abcd) is the same as (abdc). We know that in total there are 6! items in the long list of permutations, but we have over-counted if we want to know how many ways there are of choosing 4 objects from 6. We've double counted because we have got multiple permutations of the last two objects (i.e.. (abcd)ef and (abcd)fe are the same) and we have multiply counted because we have many ways of writing the same items in the brackets. In fact we have over counted by 2! due to the last items and overcounted by 4! because of the permutations of the first items. Thus the total number of unique ways of picking 4 items from 6 is not 6! but $\frac{6!}{2!4!} = 15$.

In fact this generalises completely. If we have n items and we want to choose r from them then we have:

$${}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!} \quad (195)$$

where we read "n choose r" and both ${}_nC_r$ and $\binom{n}{r}$ are just two different ways of writing this.

1. How many different possible lottery tickets are there?

Clearly the order of the numbers doesn't matter, and we can chose 6 numbers from 49. This is ${}_{49}C_6$ which is:

$$\frac{49!}{6!(49-6)!} = \frac{49 \times 48 \times 47 \times 46 \times 45 \times 44}{6!} = 13983816 \quad (196)$$

If you buy one lottery ticket, then the chances of you winning are 1 in 13983816, which is pretty terrible!

2. How about if you toss a coin 6 times, how many ways are there of getting 2 heads? We can start to enumerate them:

HHTTTT

HTHTTT

HTTHTT

HTTTHT

HTTTTH

THHTTT

etc.

We can think of labelling the positions of the heads. The first in the list above will be (1,2), the second will be (1,3), etc. We can see that what we are doing is choosing 2 numbers from 6, but we know that this is just ${}_6C_2 = \frac{6!}{4!2!} = 15$. Note that choosing r from n gives the same number as choosing $n - r$ from n . There is a nice symmetry because of the denominator in ${}_nC_r$.

There are some more important properties of ${}_nC_r$. How many ways are there of choosing n objects from n : The answer is always 1, whatever the value of n . How about how to choose 0 objects from n . Again, the answer is 1. There is also a nice symmetry of this. How many ways are there of choosing 4 objects from 6 is the same as choosing 2 objects from 6 (it's like choosing not to choose the 4 objects!). In general: ${}_nC_r = {}_nC_{n-r}$.

8.3 Pascal's identity

There is a very nice identity, which at first site might not seem very useful, but in fact it will help a lot with working out expressions of the form ${}_nC_r$. This is Pascal's identity:

$$\binom{n}{r} + \binom{n}{r+1} = \binom{n+1}{r+1} \quad (197)$$

We can prove this by starting with the left hand side:

$$\begin{aligned}
\binom{n}{r} + \binom{n}{r+1} &= \frac{n!}{r!(n-r)!} + \frac{n!}{(r+1)!(n-(r+1))!} \\
&= \frac{n!}{r!(n-r)(n-(r+1))!} + \frac{n!}{(r+1)r!(n-(r+1))!} \\
&= \frac{n!}{r!(n-(r+1))!} \left[\frac{1}{n-r} + \frac{1}{r+1} \right] \\
&= \frac{n!}{r!(n-(r+1))!} \left[\frac{r+1+n-r}{(n-r)(r+1)} \right] \\
&= \frac{n!}{r!(n-(r+1))!} \left[\frac{n+1}{(n-r)(r+1)} \right] \\
&= \frac{(n+1)n!}{(r+1)r!(n-r)(n-(r+1))!} \\
&= \frac{(n+1)!}{(r+1)!(n-r)!} = {}_{n+1}C_{r+1} \tag{198}
\end{aligned}$$

So we have an identity which tells that that to work out a term related to choosing from $n + 1$ objects, we have to add together two terms related to choosing from n objects. This is encoded entirely in Pascal's triangle:

$$\begin{array}{rcccc}
n = 0: & & & & 1 \\
n = 1: & & & 1 & 1 \\
n = 2: & & 1 & 2 & 1 \\
n = 3: & 1 & 3 & 3 & 1 \\
n = 4: & 1 & 4 & 6 & 4 & 1
\end{array}$$

Pascal's triangle is the complete enumeration of all ${}_nC_r$. In fact, the above can be written as:

$$\begin{array}{rcccccc}
n = 0: & & & & & & {}_0C_0 \\
n = 1: & & & & & & {}_1C_0 & & & & {}_1C_1 \\
n = 2: & & & & & & {}_2C_0 & & & {}_2C_1 & & & {}_2C_2 \\
n = 3: & & & & & & {}_3C_0 & & & {}_3C_1 & & & {}_3C_2 & & & {}_3C_3 \\
n = 4: & & & & & & {}_4C_0 & & & {}_4C_1 & & & {}_4C_2 & & & {}_4C_3 & & & {}_4C_4
\end{array}$$

Check for yourself that the ${}_nC_r$ expressions that you calculate match up exactly with the numbers in Pascal's triangle. Now, the rather mysterious looking Pascal's identity makes a lot more sense. It is precisely the algorithm that you use to construct Pascal's triangle. For instance Pascal's identity tells

you that ${}_4C_3 = {}_3C_2 + {}_3C_3$ which is the same as seeing that the number 4 is given by $3 + 1$ in Pascal's triangle. In order to work out ${}_nC_r$ then we just need to write out Pascal's triangle down to the $(n + 1)^{th}$ layer (the first layer is for ${}_0C_0$).

8.4 The Binomial Theorem

What is $(x + y)^{34}$? This seems like an almost impossible task; you'd have to write out $(x + y)(x + y)(x + y)\dots$ 34 times, then multiply them all out and it would become incredibly messy. However, everything we've done in the last section will allow us to see precisely what this, and any other expression of this form, is given by. Let's start with a simpler example:

Example

How about $(x + y)^4$? This is a simpler example and one that we could think of doing by hand, but we will show that there is a very general way to get any expression of this form. We start by writing this out in long form:

$$(x + y)^4 = (x + y)(x + y)(x + y)(x + y) \quad (199)$$

We are now going to label the x 's and y 's, though the labels (which will be indices) will just be dummy labels and we will remove them in the end. We write out the above expression as:

$$(x_1 + y_1)(x_2 + y_2)(x_3 + y_3)(x_4 + y_4) \quad (200)$$

Now let's start to expand this and see what it looks like:

$$\begin{aligned} & x_1x_2x_3x_4 \\ & + x_1x_2x_3y_4 + x_1x_2y_3x_4 + x_1y_2x_3x_4 + y_1x_2x_3x_4 \\ & + x_1x_2y_3y_4 + x_1y_2x_3y_4 + x_1y_2y_3x_4 + y_1x_2x_3y_4 + y_1x_2y_3x_4 + y_1y_2x_3x_4 \\ & + x_1y_2y_3y_4 + y_1x_2y_3y_4 + y_1y_2x_3y_4 + y_1y_2y_3x_4 \\ & + y_1y_2y_3y_4 \end{aligned} \quad (201)$$

This looks pretty horrible, but of course the indices that we put on the x 's and y 's were dummy indices and we can remove them. When we remove them we find that the first line in the equation above is just x^4 , the second line gives something times x^3y , the third gives something times x^2y^2 , the fourth gives something times xy^3 and the fifth gives y^4 . In fact it gives precisely:

$$x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4 \quad (202)$$

But how did we get these numbers 1, 4, 6 and 4, the coefficients in front of each term. Let's look at the third line in the formula above. We see that there are terms of the form $x_1x_2y_3y_4$, etc. What we have here are all the different ways of choosing 2 x 's from the four possible ones. We can choose 1 and 2, 1 and 3, 1 and 4, 2 and 3, 2 and 4 or 3 and 4. In fact we know how to calculate the number of ways of doing this and it is ${}_4C_2 = 6$. For the second line as well, we have all the ways of choosing a single x from the four x 's, so this number is ${}_4C_1 = 4$. In fact, we can write the above expression as:

$${}_4C_4x^4 + {}_4C_3x^3y + {}_4C_2x^2y^2 + {}_4C_1xy^3 + {}_4C_0y^4 = \sum_{r=0}^4 {}_4C_r x^r y^{4-r} \quad (203)$$

Check it, but note that on the left hand side we are starting at $r = 4$ and working backwards to $r = 0$ whereas on the right hand side we sum the other way around. This doesn't matter because $a + b + c + d = d + c + b + a$.

In fact we have, by constructing the expression in this way, found a much more general result. This is called the Binomial Theorem and it tells us that:

$$(x + y)^n = \sum_{r=0}^n {}_n C_r x^r y^{n-r} \quad (204)$$

The ${}_n C_r$ are called the binomial coefficients. This is a very powerful result and allows us to calculate otherwise very complicated expressions. We can also note that because Pascal's triangle is made up precisely of the numbers of the form ${}_n C_r$ we can use it to find the coefficients of the binomial expansion. You don't need to calculate all of the ${}_4 C_r$ terms to find $(x + y)^4$. Just look at the $(4 + 1)^{th}$ row of Pascal's triangle: (1, 4, 6, 4, 1) and these are the coefficients of x^4 , x^3y , x^2y^2 , xy^3 , y^4 . Let's look at some more examples:

1.

$$(x+y)^6 = \sum_{r=0}^6 {}_6 C_r x^r y^{6-r} = x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^4 + 6xy^5 + y^6 \quad (205)$$

Note that we could take the numbers (1, 6, 15, 20, 15, 6, 1) from the 7th layer of Pascal's triangle.

It is also important to note that it doesn't matter whether we write $x^r y^{n-r}$ or $x^{n-r} y^r$ because ${}_n C_r = {}_n C_{n-r}$.

2. We can also ask about just a particular term in an expansion. In an expansion of the form $(a + b)^n$, the coefficient of the term $a^r b^{n-r}$ will

be ${}_n C_r$. For example:

$$(x + 3y)^6 = \sum_{r=0}^6 {}_6 C_r x^r (3y)^{6-r} \quad (206)$$

The coefficient of the term $x^4 y^2$ will come with a factor ${}_6 C_4$ but we have to be careful because it will also come with a factor 3^2 . So it will be $\frac{6!}{4!2!} 3^2 = 135$

3. How about a much more complicated looking example:

$$(76y + 35x)^{322} \quad (207)$$

What is the term with the 163^d power of x ? We know that:

$$(76y + 35x)^{322} = \sum_{r=0}^{322} {}_{322} C_r (76y)^r (35x)^{322-r} \quad (208)$$

so the term that goes like x^{163} is ${}_{322} C_{322-163} (76y)^{322-163} (35x)^{163} = \frac{322!}{163!159!} (76y)^{159} (35x)^{163}$. You wouldn't be expected to calculate 322! but you would be expected to be able to work out the above expression for this term.

4. How about looking at an expression of the form:

$$\left(x + \frac{1}{x}\right)^{10} \quad (209)$$

We can use the binomial theorem to write this as:

$$\left(x + \frac{1}{x}\right)^{10} = \sum_{r=0}^{10} {}_{10} C_r x^r \left(\frac{1}{x}\right)^{10-r} \quad (210)$$

If we were asked to find the coefficient of the x^4 term we would have to work out for which r does this go like x^4 , so we write out the above as:

$$\left(x + \frac{1}{x}\right)^{10} = \sum_{r=0}^{10} {}_{10} C_r x^{r+r-10} \quad (211)$$

The power of x is 4 when $r = 7$, so we have to work out ${}_{10} C_7 = 120$. So the term that goes like x^4 is $120x^4$.

8.5 The multinomial expansion

We will just touch on it here, but it is possible to write out the expansion of something with multiple terms (more than 2) using the same machinery as above. For instance, what is the expansion of $(a + b + c)^n$?. We start by writing it as $(a + (b + c))^n$ and using the binomial expansion:

$$(a + b + c)^n = (a + (b + c))^n = \sum_{r=0}^n {}_n C_r a^r (b + c)^{n-r} \quad (212)$$

but then the second part of this is just a binomial expression and so we can write that out too:

$$(b + c)^{n-r} = \sum_{s=0}^{n-r} {}_{n-r} C_s b^s c^{n-r-s} \quad (213)$$

and so the whole expression becomes:

$$(a + b + c)^n = \sum_{r=0}^n {}_n C_r a^r \sum_{s=0}^{n-r} {}_{n-r} C_s b^s c^{n-r-s} \quad (214)$$

If needed, you could work this out using the binomial coefficients, ${}_n C_r$, as normal.

9 Taylor and Maclaurin polynomials

An important note: In this section we will be dealing with both polynomials and series. A polynomial will be defined as an expression of the form:

$$\sum_{i=0}^n a_i x^i \quad (215)$$

For some set of constants a_i and some finite n . A series on the other hand will be more general and we can think of it as an infinite polynomial:

$$\sum_{i=0}^{\infty} a_i x^i \quad (216)$$

9.0.1 Recap:

We saw when we studied the binomial theorem how we could take an expression like: $(x + y)^n$ where n was some positive integer, and express this in a very compact form as a polynomial:

$$(x + y)^n = \sum_{r=0}^n {}_n C_r x^r y^{n-r} \quad (217)$$

where ${}_n C_r$ are the binomial coefficients, which are very easy to calculate, either using factorials or by constructing the appropriate layer of Pascal's triangle. It was clear why we should be able to write this as a series of terms of higher and higher powers in x . In particular, if we let $y=1$ then we can say that a function $f(x) = (1 + x)^n$ has a series expansion given by the above equation with y set to 1. It turns out that we can construct such a series for much more general functions than just binomial expressions. We start by trying to understand how we might approximate a function.

9.1 Approximating a function with a power series

Let's take the function shown in figure 43 (It happens to be $2 \sin x + \frac{\cos 3x}{2}$ but that doesn't matter for this discussion). Let's say that we want to be able to write down an approximate expression for this function that's going to give a pretty good approximation of it near some particular point. Let's say that the point that we want to approximate the function is at $x = 2.5$.

What is the simplest approximation we can imagine for this function, $f(x)$, close to $x = 2.5$? Well, how about we approximate it by a constant $f(2.5)$? This looks almost nothing like the function, but at least very very

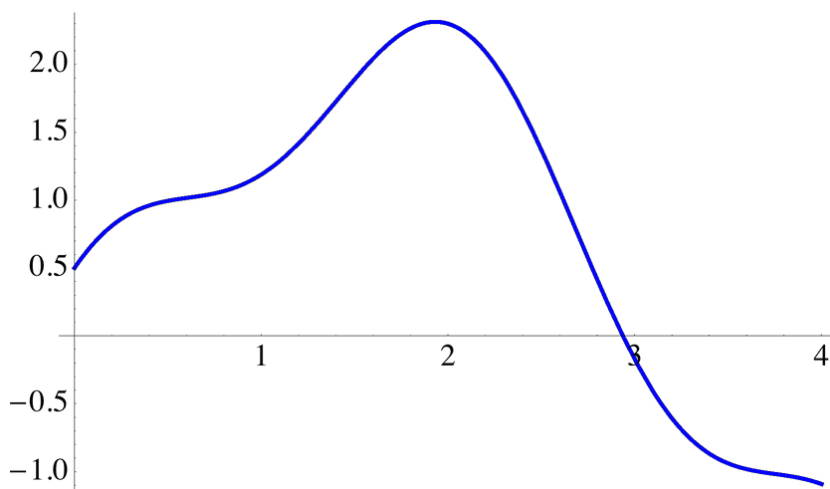


Figure 43: A function which we would like to approximate by a polynomial.

close to $x = 2.5$ it is something like the function. At least it gets the value of the function right at $x = 2.5$ by definition! This approximation is shown in figure 44.

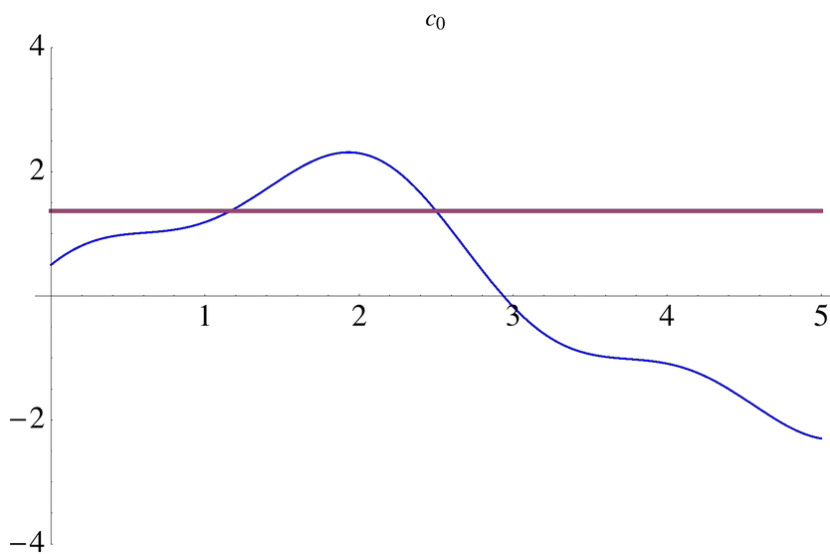


Figure 44: If we want to approximate the function close to $x = 2.5$ by a polynomial, the crudest thing that we could do would be to have a constant. At the point $x = 2.5$ this would be correct, but anywhere away from there it is a terrible approximation.

How about we try and approximate it a little better. What about the

tangent line to the function at $x = 2.5$? at least then the gradient of our approximation as well as the value itself would be right at $x = 2.5$. In figure 45 we show a straight line which has the same value, and same first derivative of our function $f(x)$ at $x = 2.5$. This line is given by $f(2.5) + c_1(x - a)$ where soon we will see exactly how to calculate c_1 but for now it's just a constant that we fix so that this line shares the same gradient as our function at the point $x = 2.5$.

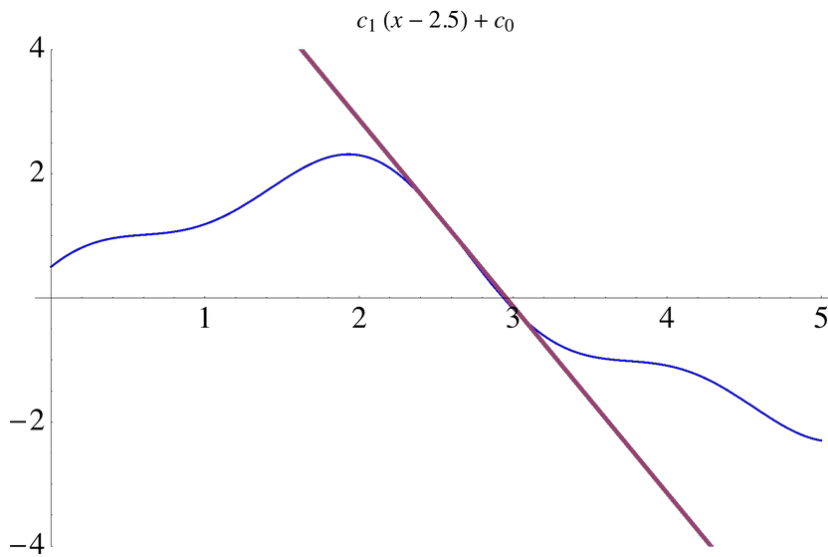


Figure 45: Here we try and do a little better. We have approximated the function, close to $x = 2.5$ by a constant plus a linear term (we will show soon how to calculate these c_i). We see that it does a better job at approximating the function close to $x = 2.5$ but still it isn't great..

What about a slightly better approximation? How about if we try and find a quadratic function that matches the value, and the first derivative and the second derivative of the function at $x = 2.5$? This is shown in figure 46. The parabolic line has equation $f(2.5) + c_1(x - a) + c_2(x - a)^2$ where again, c_1 and c_2 are chosen so that the zeroth, first and second derivatives of the line match that of our function at $x = 2.5$.

How about a cubic expression that get's the zeroth, first, second and third derivatives of the function correct? Figure 47 has the cubic polynomial which best approximates the function.

We could clearly go on. In fact we can write down a general polynomial expression:

$$\sum_{i=0}^n c_i(x - a)^i \tag{218}$$

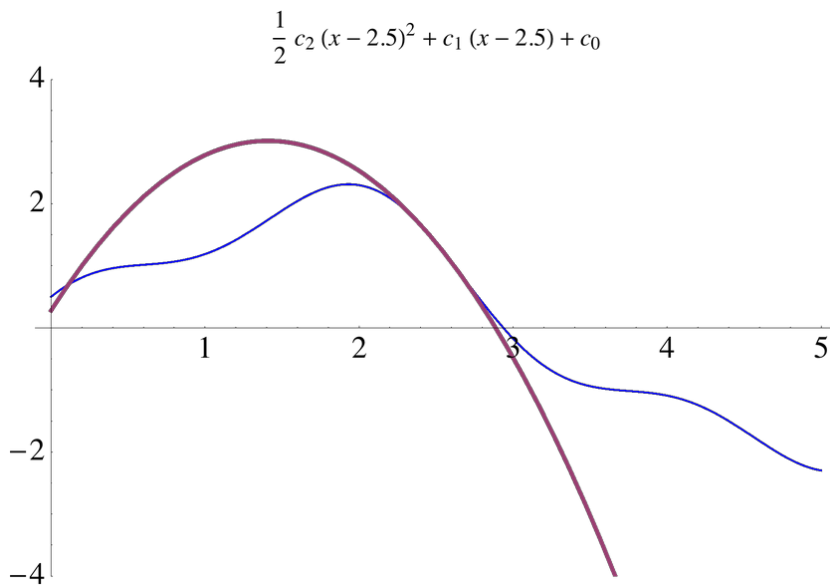


Figure 46: Now we are trying a quadratic function. We fix c_0 , c_1 and c_2 such that the quadratic polynomial has the same value, first derivative and second derivative as the function, at $x = 2.5$. We can see that this is starting to look a bit more like the function itself than did our first two terrible approximations but anywhere far from $x = 2.5$ it's still pretty awful.

where the c_i are constants and in this case $a = 2.5$. The c_i will be chosen so that the i^{th} derivative of this polynomial matches the i^{th} derivative of our function. In figure 48 we see what happens if we go up to a fourteenth order polynomial. We see that it is really starting to approximate the function pretty well close to $x = 2.5$. Note that what we are doing is adding together shapes with weights given by the c_i 's and getting something which is a better and better approximation of the function itself.

In fact what we are doing is a general procedure. We can say that we can approximate a function, about some point a by a polynomial:

$$f(x) \approx \sum_{i=0}^n c_i (x - a)^i \quad (219)$$

We have $(x - a)^n$ so that close to $x = a$ the higher order terms will become less and less important and we can safely truncate the series at some n to get a reasonable approximation to the function. At $x = a$ all that is left is c_0 which is just the value of the function at $x = a$. The higher the value of n (the higher the order of the polynomial) the better should be our approximation to the actual function. Now we just have to work out how to calculate the

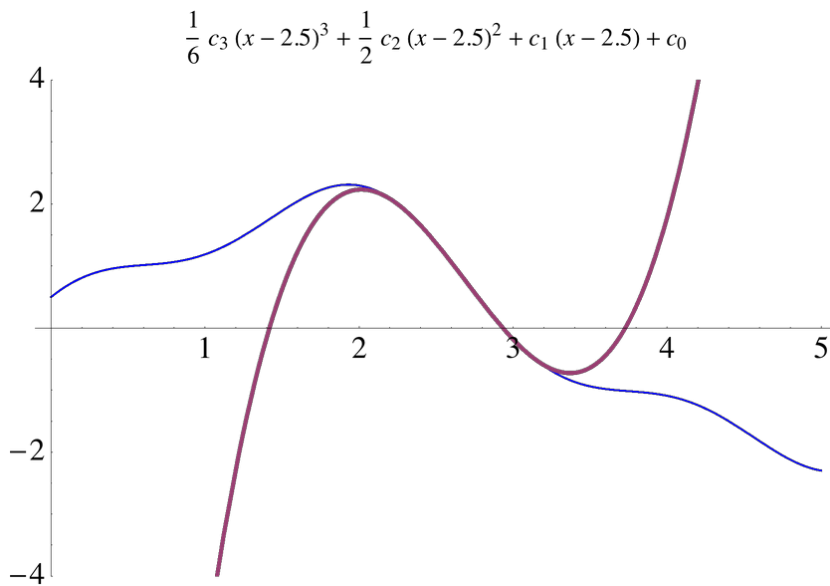


Figure 47: Approximating the function by a third order polynomial. We see that the approximation is matching the function close to $x = 2.5$ better and better, the higher order the polynomial we choose.

constants c_i . The first thing to do is to plug in the value a to both sides of equation 219. This will clearly give:

$$f(a) = c_0 \tag{220}$$

so this has already fixed c_0 for us. Next we can take a derivative of both sides and then, again, plug in $x = a$. Check for yourself that this will give:

$$f'(a) = c_1 \tag{221}$$

Now differentiate again and plug in $x = a$. This will give:

$$f''(a) = 2c_2 \tag{222}$$

ie. $c_2 = \frac{f''(a)}{2}$ Doing the same again gives:

$$f'''(a) = 3 \times 2c_3 \tag{223}$$

We can carry on with this procedure and we will quickly realise that there is a very general expression for the c_i which is:

$$c_i = \frac{f^{(i)}(a)}{i!} \tag{224}$$

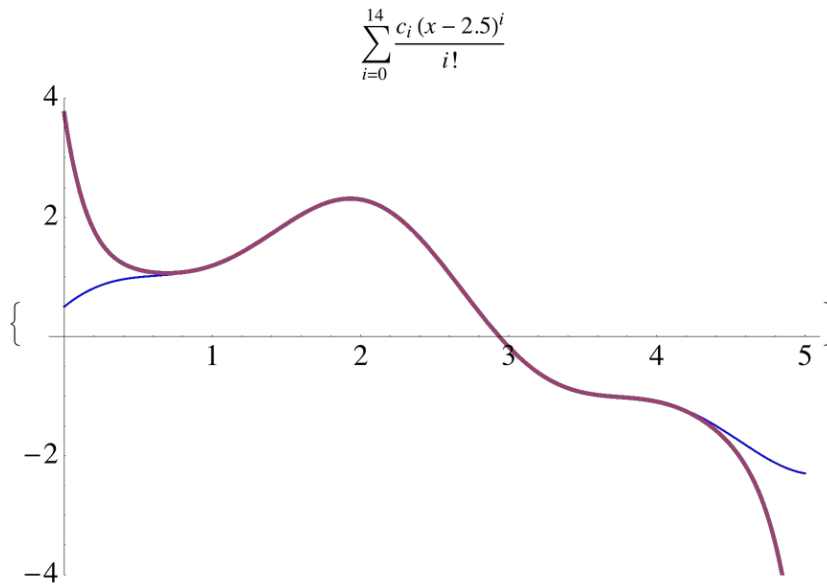


Figure 48: Here we have approximated the function by the fourteenth order polynomial. It's a very good fit close to the point $x = 2.5$ and matches the first fourteen derivatives of the function at that point too. If you go far from $x = 2.5$ it starts to diverge from the actual function. For instance out at $x = 5$ already it gives a pretty bad approximation to the function.

where $f^{(i)}$ is the i^{th} derivative of the function and we evaluate it at $x = a$.

So, we can plug this back into the expression we wrote down for the polynomial:

$$f(x) \approx \sum_{i=0}^n \frac{f^{(i)}(a)}{i!} (x - a)^i \quad (225)$$

This is known as a Taylor polynomial. If we were to take an infinite number of terms, it becomes what is known as a Taylor series. There are subtleties with taking the number of terms to ∞ , but the examples that we will cover will not have these problems. Clearly, as we take more and more terms in the Taylor Polynomial we will get a better and better approximation to the function that we are interested in as long as we are close to the point about which we are expanding. This last point is important. If we expand around $x = a$ and then ask for the behaviour of the polynomial close to $x = a$, let's say at some $x = b$ which is close to $x = a$, then we will have terms of the form $(b - a)^i$ where $b - a$ will be close to zero. Small terms to high powers get smaller and smaller and so we can happily ignore higher order terms when we are close to $x = a$. For instance, if we expand around $x = 1$ and we ask for the value of the Taylor polynomial at $x = 1.1$ then we will have a series

which has terms like $(1.1 - 1)^i$. Clearly the tenth term will go like 0.1^{10} which is a tiny number and can probably be ignored. When we are close to the point about which we are expanding, we can thus safely ignore the very high order terms and can 'truncate' the polynomial at some finite n , happy in the fact that higher order terms in the polynomial will contribute very little. On the other hand, If we expand about $x = a$ but then try to ask what the value of the approximation is far from $x = a$ (let's say at $x = b$), we will end up with terms of the form $(b - a)^i$ where $b - a$ is large. When we take higher and higher powers of these terms, they will contribute more and more and so we have to be very careful with their coefficients. Sometimes the coefficients (of the form $\frac{f^{(i)}(a)}{i!}$) can overcome the high value of $(b - a)^i$ and so we are safe, but sometimes you have to be very careful.

You might ask why we are doing this, but we will soon see that to be able to write out a polynomial expression for a function can be very useful. Let's start by looking at some examples. In fact, the first examples we will look at we will want to approximate the function about the point $x = 0$, so in the above expression $a = 0$. When you approximate a function about $x = 0$ a Taylor series is known as a Maclaurin series. Let's look at a few examples of Maclaurin series:

1. Let's look at the function $f(x) = e^x$ and write this as a Taylor series about the point $x = 0$ (ie. a Maclaurin series). In order to do this we have to calculate the values of the derivatives of the function at $x = 0$.

We set up a table:

$$\begin{pmatrix} i & f^{(i)}(x) & f^{(i)}(0) \\ 0 & e^x & 1 \\ 1 & e^x & 1 \\ 2 & e^x & 1 \\ 3 & e^x & 1 \\ 4 & e^x & 1 \\ 5 & e^x & 1 \end{pmatrix} \quad (226)$$

In this case it is incredibly simple to work out the derivatives of the function and the value of the derivative at $x = 0$ because they are all the same. Now we can plug this into the expression for the Taylor polynomial and find:

$$e^x \approx \sum_{i=0}^n \frac{x^i}{i!} \quad (227)$$

where this is only an approximation of e^x if we take a finite number of

terms. In this case, if we let $n \rightarrow \infty$ we get an exact expression for e^x :

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!} \quad (228)$$

In figure 49 you can see how the first six terms in the Taylor polynomial add up to get a function which is a better and better approximation of the exponential function. It is also very important to note that

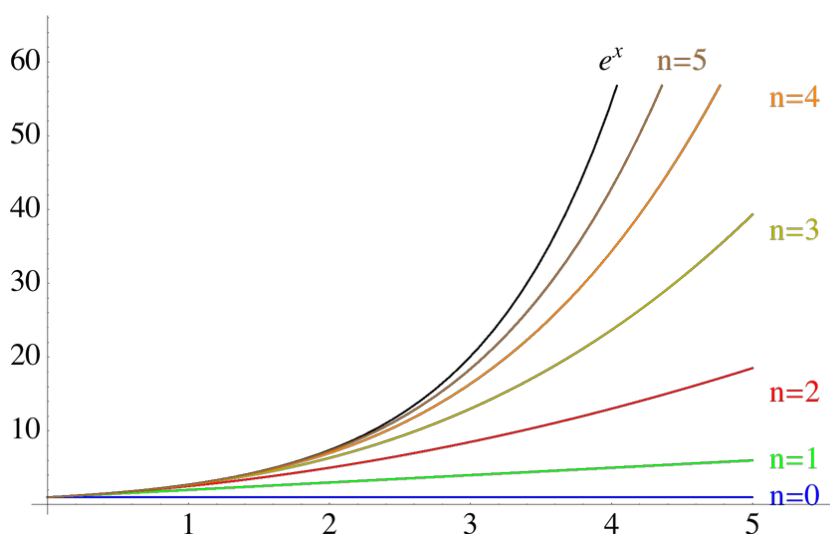


Figure 49: The Taylor polynomial approximation of e^x with order n up to the fifth order Taylor polynomial expanded around $x = 0$ (ie. the Maclaurin polynomial). You can see that as you add more and more terms the approximations to the exponential get better and better.

the further away from 0 (which was where we expanded the function about) we are, the worse is our approximation, independent of n . For $x \approx 0$ even $n = 1$ will give a reasonable approximation to the value of e^x . The further from x we go, the higher the n we need to get a good approximation for the function value. We can see from the equation itself that there is a balance at play. For x close to zero, the higher order terms in the polynomial die off quickly because x^i for small x and larger and larger i get smaller and smaller. However, for larger x , (let's say 2), the value of x^i increases as i gets larger and so we have to wait until the $i!$ in the denominator get large enough to start to make the terms $\frac{x^i}{i!}$ get small. Of course it's not quite as simple as that as we are not just looking at individual terms but we are adding them together.

The subject of how quickly these terms in the polynomials die off and can be ignored is a subject in and of itself.

What can we do with this? Well, the simplest thing that we can do is to get a value of e in an algebraic way. By putting in $x = 1$ we find that:

$$e^1 = \sum_{i=0}^{\infty} \frac{1}{i!} = 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \dots \quad (229)$$

You can thus calculate e by hand to arbitrary precision.

2. The exponential function was particular easy because its derivative is equal to the function itself every time. Let's look at a slightly more involved example: $f(x) = \sin x$ about $x = 0$. Again, we start with the table of derivatives:

$$\begin{pmatrix} i & f^{(i)}(x) & f^{(i)}(0) \\ 0 & \sin(x) & 0 \\ 1 & \cos(x) & 1 \\ 2 & -\sin(x) & 0 \\ 3 & -\cos(x) & -1 \\ 4 & \sin(x) & 0 \\ 5 & \cos(x) & 1 \end{pmatrix} \quad (230)$$

Now the values of the derivatives are not always the same. They are zero every other one, and they change in sign when they are not zero. This leads to a very elegant expression for the sin function expanded around $x = 0$:

$$\sin x = \sum_{i=0}^{\infty} \frac{(-1)^i x^{2i+1}}{(2i+1)!} \quad (231)$$

The important point is that here the terms get smaller and smaller as you take more and more of them, so if, for instance, you want to know the value of $\sin 2.4$ you can plug it into the right hand side, take a finite number of terms and you will get an approximation. Let's look at exactly what this means. Plugging in 2.4 (as a random example), gives:

$$\sum_{i=0}^{\infty} \frac{(-1)^i 2.4^{2i+1}}{(2i+1)!} \quad (232)$$

but we can't practically take an infinite number of these terms, so let's see how they go as we take a finite number. If we write:

$$T_n(\sin x) = \sum_{i=0}^n \frac{(-1)^i x^{2i+1}}{(2i+1)!} \quad (233)$$

where this is the n^{th} order Taylor Polynomial of x , then we find that as we increase n , the polynomial expression quickly tends to a fixed value as you can see in figure 50

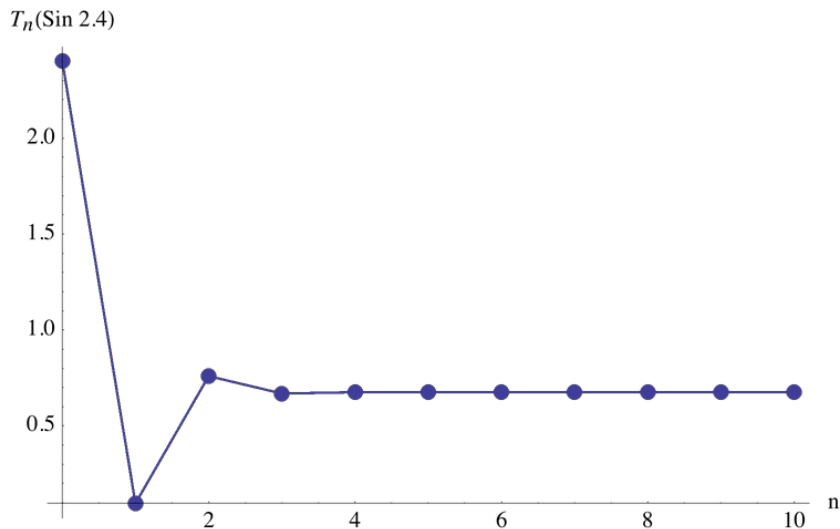


Figure 50: The value of the Taylor polynomial as we increase n , the number of terms in the polynomial for $\sin(2.4)$ expanded about $x = 0$. We can see that for $n > 3$ or so, we get a pretty good approximation to the actual value ($\sin 2.4 = 0.675$)

- Let's look at expanding around a different point. We have, in these two examples just looked at Maclaurin polynomial (a Taylor polynomial) expanded around $x = 0$). Now let's look at the exponential function expanded about $x = 2$. Remember that the general form for the Taylor polynomial is:

$$\sum_{i=0}^n \frac{f^{(i)}(a)}{i!} (x - a)^i \quad (234)$$

so we have to calculate the derivative of e^x at $x = 2$. Again, this is pretty simple in this case because the derivative is always the same and so the value of the any derivative (including the zeroth derivative) at $x = 2$ is e^2 . So the n^{th} order Taylor polynomial for e^x at $x = 2$ is:

$$\sum_{i=0}^n \frac{e^2}{i!} (x - 2)^i \quad (235)$$

Just as we did in figure 49 let's see what this looks like for different values of n . Again, if we were to take an infinite number of terms (a

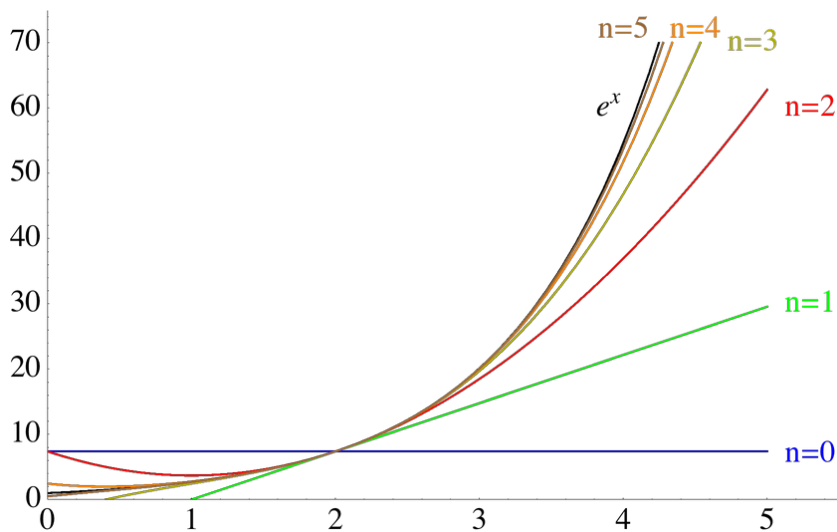


Figure 51: The Taylor polynomial approximation of e^x with order n up to the fifth order Taylor polynomial expanded about $x = 2$. Here it is clear that close to $x = 2$ even the low order polynomials get a reasonable approximation to the function value but further away from $x = 2$ we need to take more and more terms to approximate the function.

Taylor series) we would get a perfect match with the function. In fact they are equal:

$$e^x = \sum_{i=0}^{\infty} \frac{e^2}{i!} (x - 2)^i \quad (236)$$

4. **Exercises:** Find the expansion for $\cos x$ around $x = 0$. Find the expansion for $\sin x$ around $x = \frac{1}{2}$.
5. **Harder, but you can do it!:** Find the expansion for $\arctan x$ about $x = 0$. Hint: You know that $\int \frac{1}{1+x^2} dx = \arctan x + c$ and you can do the expansion of $\frac{1}{1+x^2}$.
6. Now we are going to return to our friend, the binomial expansion and show that we get the same answer as we got using our combinatoric method by making a Taylor polynomial for our expression. Let's say we take $(x+y)^6$ and we want to work out the Taylor expansion for this about $x = 0$. NB. Here we just treat y as a dummy variable - do not think that $y(x)$, it's just a number here.

We go through exactly the same procedure as before, writing down a list of derivatives of $(x+y)^6$ and plugging in $x = 0$ (because that is

what we are expanding about here.

$$\begin{pmatrix} i & f^{(i)}(x) & f^{(i)}(0) \\ 0 & (x+y)^6 & y^6 \\ 1 & 6(x+y)^5 & 6y^5 \\ 2 & 30(x+y)^4 & 30y^4 \\ 3 & 120(x+y)^3 & 120y^3 \\ 4 & 360(x+y)^2 & 360y^2 \\ 5 & 720(x+y) & 720y \\ 6 & 720 & 720 \\ 7 & 0 & 0 \\ 8 & 0 & 0 \\ 9 & 0 & 0 \\ 10 & 0 & 0 \end{pmatrix} \quad (237)$$

We notice here that the derivatives all vanish after $i = 6$. This means that if we include all terms up to 6, we will get not just an approximation to our function, but an exact equality. We should note however that the expression in the table above can be simplified somewhat and written in factorial form as follows:

$$\begin{pmatrix} i & f^{(i)}(x) & f^{(i)}(0) \\ 0 & \frac{6!}{6!}(x+y)^6 & \frac{6!}{6!}y^6 \\ 1 & \frac{6!}{5!}(x+y)^5 & \frac{6!}{5!}y^5 \\ 2 & \frac{6!}{4!}(x+y)^4 & \frac{6!}{4!}y^4 \\ 3 & \frac{6!}{3!}(x+y)^3 & \frac{6!}{3!}y^3 \\ 4 & \frac{6!}{2!}(x+y)^2 & \frac{6!}{2!}y^2 \\ 5 & \frac{6!}{1!}(x+y) & \frac{6!}{1!}y \\ 6 & \frac{6!}{0!} & \frac{6!}{0!} \\ 7 & 0 & 0 \\ 8 & 0 & 0 \\ 9 & 0 & 0 \\ 10 & 0 & 0 \end{pmatrix} \quad (238)$$

Check that you understand the above. We can thus write the Taylor polynomial as:

$$(x+y)^6 = \sum_{i=0}^6 \frac{6!}{(6-i)!i!} y^{6-i} x^i \quad (239)$$

But we recognise the factor $\frac{6!}{(6-i)!i!}$ as being the binomial factor ${}_6C_i$, so we have just rederived the binomial expression for $(x+y)^6$ using a very different method.

Well, this doesn't seem like we've done anything very useful. We already knew how to expand $(x+y)^6$, but you can see that the derivation can generalise a huge amount. When we looked at binomial expressions and we expanded out the brackets, it was very important that there were an integer number of brackets. It didn't make sense to ask how many ways there were to chose i x 's from a non-integer number of x 's. Now, however, we've seen how to derive the same expression using a Taylor expansion, but nowhere in the definition of the Taylor expansion did we say that the power needed to be an integer. Let's look at an example which we couldn't possibly hope to answer using a combinatoric argument.

7. Expand $(1+x)^{\frac{9}{2}}$ in a Maclaurin Polynomial. Again, we take derivatives and plug in $x = 0$:

$$\begin{pmatrix} i & f^{(i)}(x) & f^{(i)}(0) \\ 0 & (x+1)^{9/2} & 1 \\ 1 & \frac{9}{2}(x+1)^{7/2} & \frac{9}{2} \\ 2 & \frac{63}{4}(x+1)^{5/2} & \frac{63}{4} \\ 3 & \frac{315}{8}(x+1)^{3/2} & \frac{315}{8} \\ 4 & \frac{945\sqrt{x+1}}{945} & \frac{945}{945} \\ 5 & \frac{16}{32\sqrt{x+1}} & \frac{16}{32} \\ 6 & -\frac{64(x+1)^{3/2}}{2835} & -\frac{64}{2835} \\ 7 & \frac{128(x+1)^{5/2}}{14175} & \frac{128}{14175} \\ 8 & -\frac{256(x+1)^{7/2}}{99225} & -\frac{256}{99225} \\ 9 & \frac{512(x+1)^{9/2}}{893025} & \frac{512}{893025} \\ 10 & -\frac{893025}{1024(x+1)^{11/2}} & -\frac{893025}{1024} \end{pmatrix} \quad (240)$$

ok, so we see two things here. First we see that the numbers are pretty horrible, but that's ok. Secondly we see that the factors don't vanish at any i as they did in the case of an integer power. This is because in the case of an integer power, at some point we had taken enough derivatives that we got to a constant and then the derivative of that was zero. In this case we never hit zero, but keep on going into negative powers. Now the Taylor polynomial is given by an expression that can be written as:

$$\sum_{i=0}^n \frac{1}{i!} \left(\prod_{j=0}^i (4.5 - j) \right) x^i \quad (241)$$

where here the notation $\prod_{j=0}^i (4.5 - j)$ is similar to the *sum* notation which stands for "add these together" and means "multiply all these

together". So, for instance:

$$\prod_{j=0}^5 (4.5-j) = (4.5-0) \times (4.5-1) \times (4.5-2) \times (4.5-3) \times (4.5-4) \times (4.5-5) \quad (242)$$

This is similar to having a factorial ($5!$) but is much more general because we can have non-integer values as well.

In the case of an integer value for k in $(1+x)^k$ we had a finite number of terms. With a non-integer value we can take an arbitrary number of terms to try and get closer and closer to the real value of the function. Again, for this approximation, if we take x away from the point that we are expanding about (here $x=0$) we may get a bad answer. In fact, in this case, for subtle reasons, this will only work for $x \leq 1$ and if we try and look at x far away from 1, we will get worse and worse answers, the more terms we take. This is covered using something called "The radius of convergence" but we will not cover that topic here.

8. Let's look at another binomial expression which we will use to calculate a rather harder looking series: $f(x) = (1-x^2)^{-\frac{1}{2}}$

What you should understand from this topic:

- That you can approximate functions close to some point by a polynomial.
- That better and better approximations will accurately give you higher and higher derivatives of the function at the point about which you are expanding.
- That there is a general formalism for expanding a function about an arbitrary point.
- That in general expanding with more and more terms will give you a better approximation to the function.
- That you can use these expansions for getting analytic expressions for functions that you can otherwise only calculate numerically (for instance, as we saw in the exponential function).
- You should know how to calculate the Taylor polynomial for a given function, expanded around an arbitrary point.
- You should understand how to derive the binomial theorem using a Maclaurin polynomial.

10 Complex numbers

These will be an addition to the notes already on Vula on complex numbers. Please refer to that document as well as I will be taking a slightly alternative approach on occasion.

10.1 A philosophical detour

Before we get on to talking about imaginary numbers and complex numbers, let's try and break down our preconceptions about numbers in general. We look at the world around us and see many things which we categorise. We see a computer, a piece of paper, we see other people, we see our hands. These are labels that we use to categorise the world around us, but these objects seem very physical and very real. We rarely question their existence, though if one wants to take the Cartesian view, we should also question the reality we are in. We are not going to go that far, but let's try and ask about the existence of numbers. I have definitely seen five pieces of paper, but I have never seen a five. I've seen the number written down, but I can write down anything I want and it doesn't necessarily mean that it exists. I can write down a erga[oeiave21 but that doesn't suddenly bring a erga[oeiave21 into existence. How about a -5? I've definitely never seen a -5 though I understand perfectly well what it means. The integers seem to be very good ways of describing, or more specifically counting objects and the negative numbers are a good way of keeping track the transfer of objects from one place to another. I can also ask you to give me 3 coffees, and here I am really asking you to apply 3 as an operation to the object coffee. 3 is acting almost more like a verb than it is a noun. When I describe that there are 30 people in a class, I am really thinking about this as a description, or an adjective. So in the real world, somehow numbers feel like verbs and adjectives. I certainly wouldn't say that 'heavy' exists, but certainly a book which has been described as heavy does.

However, there is a world in which numbers really do seem to be more like nouns than they do in the world around us, and that is in the abstract world of mathematics. In the universe of equations, numbers somehow feel much more concrete and I can manipulate them and transmogrify them from one form to another using a set of mathematical operations which become more and more finely tuned and specialised as we learn more and more mathematics. I can take a 5 and I can apply a sin function to it to give me another number. I think of this rather as taking an object and putting it in a machine which turns it into another object. Here 5 is very real, but so is -5 and so is $\frac{\pi}{4}$ and so are all the numbers that we've ever used. They are simply the objects

which are manipulated by our mathematical machinery. Whether or not they exist as objects around us isn't very important for our use of them in the mathematical universe.

Incidentally, I have here separated the real universe from the more abstract, platonic, mathematical one, but it is fair to say that we have found mathematics as the best language with which to accurately describe the real universe. All of our models and precise descriptions of the universe are built using mathematics, and it acts as an incredible way of describing the laws of nature. Which came first, the mathematics or the universe? That is not a question I am going to get onto here, but it's certainly a profound one!

10.2 A foray into a new number system

OK, so we have a mathematical world of numbers and we can manipulate them. Thus, we should be perfectly happy to have some more ingredients in that world, that don't have such an obvious mapping to the things in the world around us. We will discover that actually they help us enormously in the things that we can do with the mathematical machinery. It's like having a powerful car but not the right fuel to really take it up to top speed. We are about to find out what that fuel is and push the limits of what our car can do!

Previously, if we set up a certain type of quadratic equation and plugged it into our machinery to find a solution, the machinery would jam and we wouldn't get an answer out. This was a real shame because it didn't seem to be that much more of a complicated equation than any other that we had studied. We are perfectly happy with solving an equation like:

$$x^2 - 1 = 0 \tag{243}$$

You can plot the graph of the left hand side and see that it equals 0 at two points $x = \pm 1$. That's fine, our mathematical machinery can deal with that fine, but when we ask to solve something so similar:

$$x^2 + 1 = 0 \tag{244}$$

our traditional machinery comes juddering to a halt and we get an error message on the screen. In fact when we plot the graph it's clear that it doesn't cross the x axis, so it can't have a solution...can it? Maybe we're not looking hard enough. Maybe our machinery is fine, but we've just fed it the wrong ingredients. In fact, we can find the solution just fine. The solutions are:

$$x = \pm\sqrt{-1} \tag{245}$$

You might look at this and go "Absolutely not!" You can't take the square root of a negative number, but if you plug that into the equation, it works just fine and is a perfectly good solution. What is not true is that $\sqrt{-1}$ is like the normal numbers that we are used to using. In fact, let's give this solution a name. We'll call it i :

$$\sqrt{-1} = i \tag{246}$$

What is i ? It's one of the solutions to the equation above. Plug it in, you'll see that it works. There's no funny business going on here. So what if it doesn't correspond to a number in the world you see around you, nor does -76 but we don't have a problem with using that number, do we? i stands for imaginary. So we call i the imaginary number, but in fact it is no more imaginary than most other numbers, it's just a little harder to understand it because we are used to things which represent a size. Numbers which are not imaginary are called real, but again, this name is probably not a very good name as -32 is no more "real" than is i .

Once you have defined this new type of number - a number that squares to a negative number, we open up so many new possibilities. Things that previously would have driven our machinery to a halt are now very easily accessible. With this new number we've just upgraded our mathematical machinery so that it can handle so many more problems than it could before.

It might seem that i wouldn't have anything to do with the real world and it's true that any measurement of the real world will give us one of the real numbers that we are used to, but using i makes many things much more natural than they would be without it. This is true in many many fields of science and so having i at hand is absolutely indispensable when you want to describe the real world.

However, before we start to use our shiny new ingredient, we have to understand about some basic properties of i . We know that $\sqrt{-1} = i$ by definition. How about $\sqrt{-2}$. Well, we use the fact that $\sqrt{ab} = \sqrt{a}\sqrt{b}$ to split this into $\sqrt{-2} = \sqrt{2}\sqrt{-1} = i\sqrt{2}$. Simple, huh! So, now we know how to take the square root of any real number, be it positive or negative. How about adding two imaginary numbers together. Again, it works just as you would expect:

$$ai + bi = (a + b)i \tag{247}$$

If you've got 2 lots of i and 7 lots of i then you've got a total of 9 lots of i . Again, nothing funny to see here!

What about adding an imaginary number to a real number: what if we have $3 + 4i$. Well, real numbers and imaginary numbers are a bit like apples and pears. You can add them together, but they don't give you

pomegranates. You are just left with apples and pears, so $3 + 4i$ is as simple as you can get it. A number that has a real part and an imaginary part is called a complex number. We can write any complex number as $a + bi$ where a and b are both real numbers. With this ingredient you can now solve ANY quadratic equation! Previously half of all quadratic equations would have been impossible to solve because, when you used the magic quadratic solution formula, you would have found the square root of a negative number. Now we know what to do with that. Let's look at a simple example:

$$x^2 + 2x + 8 = 0 \tag{248}$$

We plug it into the normal formula for solving quadratic equations and get:

$$x = \frac{-2 \pm \sqrt{4 - 4 \times 8}}{2} = \frac{-2 \pm \sqrt{-28}}{2} = -1 \pm i\sqrt{7} \tag{249}$$

Plug those two solutions back in and you'll see that this is the correct solution to that previously unsolvable equation!

Ok, we're doing good, we now know what a complex number is and we know how to add and subtract complex numbers. Multiplying is just as easy, as long as you remember that because $i = \sqrt{-1}$, $i^2 = -1$. So, multiplying two complex numbers together goes like this:

$$(a+bi)(c+di) = ac+adi+bic+bidi = ac+i(ad+bc)-bd = (ac-bd)+i(ad+bc) \tag{250}$$

Notice how we put the real part in a bracket, and the imaginary part in a separate bracket.

We have notation for taking just these parts. If we have a complex number $z = a + bi$ then $Re(z) = a$ and $Im(z) = b$. Don't be fooled, the imaginary part is the coefficient of i , and doesn't include the i .

So, a complex number is made up of two parts, an imaginary part and a real part. We are very used to dealing with numbers lying along the real line, from $-\infty$ to ∞ . When we want to denote a real number on the x axis of a graph, we just put a point there. Is there anything equivalent for complex numbers? The difference is simply that real numbers have only one dimension - ie. one thing that can change, whereas complex numbers have two things that can change (their real part and imaginary part), and thus they live in a two dimensional space. But this is lucky, because paper is in general more or less two dimensional and so we can still denote a complex number as a point on a piece of paper. Now though we have to draw two lines, and not just one (as we had to for the real line). Now we have the real axis, and the imaginary axis. In figure 52 we plot a few complex numbers

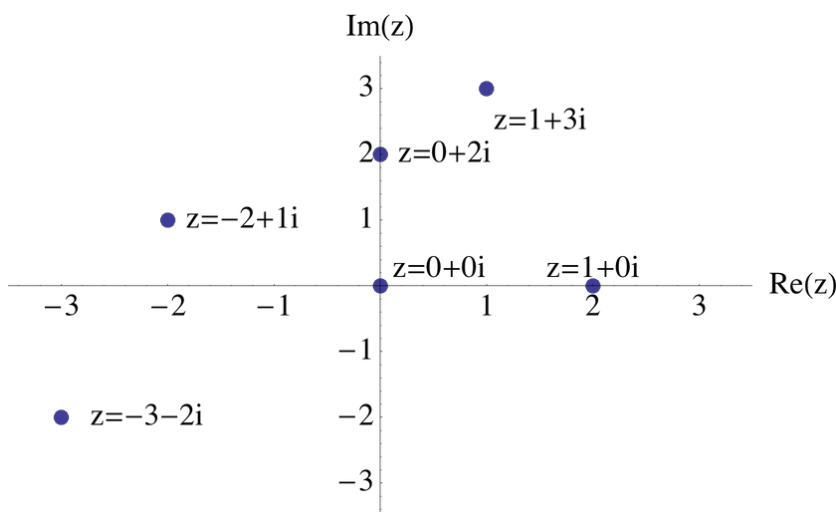


Figure 52: A few complex numbers plotted in the complex, or Argand plane. The "x" position is the real part and the "y" position is the imaginary part. Normally we would call $z = 0 + 0i$ just $z = 0$ and, for instance $z = 1 + 0i$ to be $z = 1$ but here we are including the imaginary or real part multiplied by zero explicitly to show that they are living in the complex plane. **NB. there is an error in the label at (2, 0).**

and show where they lie in the complex plane, also known as the Argand plane.

There is another way that we could think about writing these complex numbers, and that is simply by their coordinates. This is called ordered pair notation and means that instead of writing $a + bi$ we just write (a, b) . This makes it even more explicit that they live in a two dimensional space with two degrees of freedom.

There's actually another perfectly good way that we can specify the position of a complex number in the complex plane. You can see that you could draw a line from a complex number to the origin, and that number would have a length, and there would be an angle between that line and the positive x axis. Giving these two values (the length and the angle) would completely specify the complex number. So, if I give you (a, b) position in the complex plane, you could come back to me with (r, θ) . We show this in figure 53

There is a simple relationship between a and b and r and θ (if a and b are positive, see below otherwise):

$$\begin{aligned}
 r &= \sqrt{a^2 + b^2}, \quad \theta = \arctan \frac{b}{a} + 2n\pi \mid n \in \mathcal{Z} \\
 a &= r \cos \theta, \quad b = r \sin \theta
 \end{aligned}
 \tag{251}$$

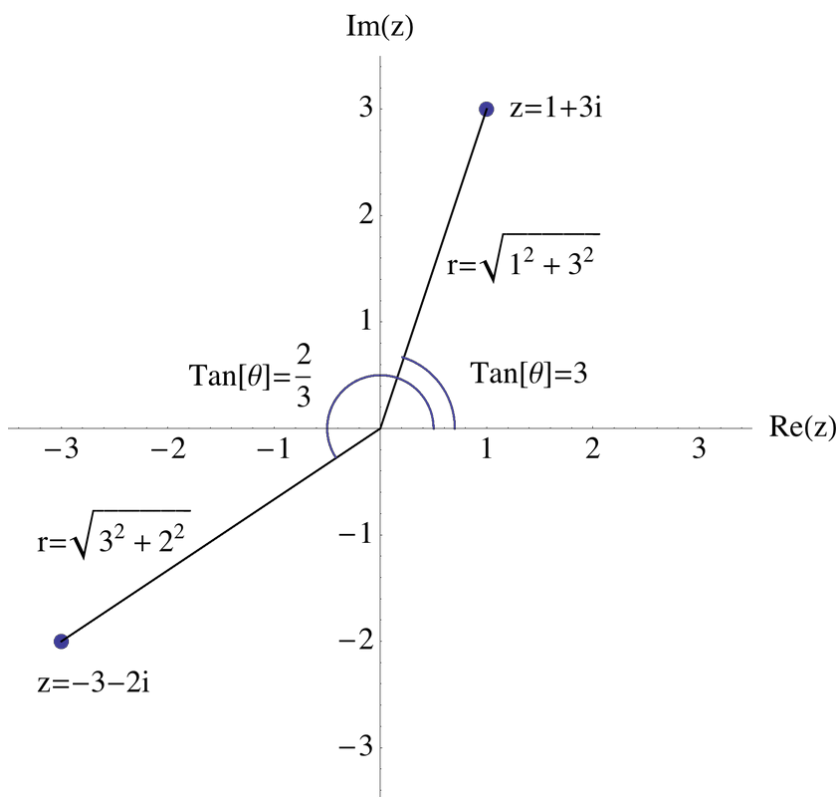


Figure 53: The angle and radius going out from the origin to the point of the complex number also uniquely identifies the point. Note however that there are an infinite number of angles, linked by the addition of integer multiples of 2π which will take you to the same number. Note also that you have to be careful with simply taking the arctan of $\frac{b}{a}$. You have to work out which solution you want. Solutions to $\arctan \frac{2}{3}$ can give 0.588 or 3.7296 depending on which branch of the tangent function you chose to invert. This just means that you have to be careful about a factor of π when you're solving such a system. It should be pretty obvious if the answer you have calculated is off by a factor of π just by looking at whether the angle should be between 0 and π or π and 2π .

Writing the complex number in terms of a and b is known as the cartesian form and in terms of r and θ is known as the modulus/argument form (the modulus is the size of the complex number (the length of the radial line) and the angle θ is known as the argument of the complex number). The modulus of a complex number is denoted as $|z|$ for a complex number z .

Note that we have actually been very sloppy with the above. We have to be more careful when finding θ . It's true that when both b and a are positive,

this formula is true, but when one of them is negative, this will not give the correct answer. In fact, when we are in the top left quadrant of the Argand plane - ie. when a is negative and b is positive, then $\theta = \pi + \arctan \frac{b}{a} + 2n\pi$. When we are in the bottom left quadrant we have $\theta = \pi - \arctan \frac{b}{a} + 2n\pi$ and when we are in the bottom right quadrant we have $\theta = 2\pi + \arctan \frac{b}{a} + 2n\pi$. You should be able to see this by plotting the position of the complex number in the Argand plane and finding the angle subtended between the line which goes between the origin and that number, and the positive real axis.

Note that there are an infinite number of θ 's which specify the same complex number. This is because you can go round the circle as many times as you want (ie. add integer multiples of 2π) and you will get back to the same angle you started at. The other way to think about this is that in the definition of a and b above, you can add on 2π to the θ inside the sin function and it will give you the same answer.

Note then that we can write $a + bi$ as $r(\cos \theta + i \sin \theta)$. It's exactly the same thing, but the radial length and the angle have been written explicitly. For instance: $1.5 + 4i = \sqrt{1.5^2 + 4^2}(\cos \arctan \frac{4}{1.5} + i \sin \arctan \frac{4}{1.5}) = 4.272(\cos 1.212 + i \sin 1.212)$.

Note importantly that if you see a complex number which is just $\cos \theta + i \sin \theta$ then the modulus, or in other words, size, of the complex number is $\cos^2 \theta + \sin^2 \theta = 1$ because $|a + bi|^2 = a^2 + b^2$. So the modulus of any complex number of the form $\cos \theta + i \sin \theta$ is just 1. You can thus see that when written in this particular form, the angular part and the magnitude part are very explicitly separated.

So, we can take a complex number, multiply it by another, add and subtract and write it in two different forms (akin to cartesian and circular coordinates for a regular vector in two dimensions). There's a simple algebraic operation which we can't yet do, which is to divide. In fact division is easy if you know one more special operation on complex numbers. This is known as complex conjugation. Complex conjugation flips the sign of the imaginary part and the sign for the complex conjugate of a complex number z is \bar{z} . If $z = a + bi$ then $\bar{z} = a - bi$. That seems pretty innocuous but it's going to be very powerful. In particular we should note that in the complex plane, the position of \bar{z} is the reflected position of z in the x -axis as shown in figure 54.

One thing to note immediately is that the modulus of a complex number, which we write as $|z| = |a + bi| = \sqrt{a^2 + b^2}$ can now be written as $|z| = \sqrt{z\bar{z}}$. ie. if you multiply a number by its complex conjugate it gives you a purely real number - the imaginary stuff cancels. Now if we want to think about dividing by a complex number we can employ a famous mathematical trick, which is to multiply by 1. Let's try and work out what $\frac{1}{z}$ is in terms of a real

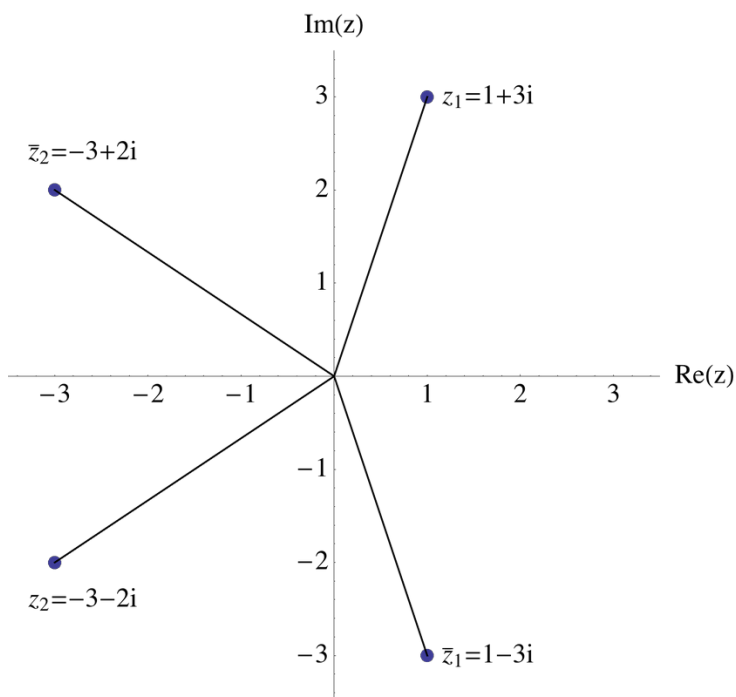


Figure 54: Two complex numbers, labeled z_1 and z_2 and their complex conjugates.

part and an imaginary part. If $z = a + bi$ then we can multiply $\frac{1}{z}$ by $\frac{\bar{z}}{\bar{z}}$:

$$\frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{\bar{z}}{|z|^2} = \frac{a - bi}{a^2 + b^2} \quad (252)$$

So we can always take a fraction of a complex number and turn it into a form of a real part and an imaginary part:

$$\frac{1}{a + bi} = \frac{a - bi}{a^2 + b^2} \quad (253)$$

10.3 Plotting functions of complex numbers (Not examinable)

Just to get a bit of a picture of what dividing by a complex number means, we can play a bit of a game. Normally we think of functions as going from a real number to another real number. $\sin(x)$ takes a real number x and gives you another real number. We can plot this on a graph by plotting a two dimensional set of data which tells you about the value that $\sin(x)$ takes for every x along the real line. We are very used to this idea of a function. However, a function of a complex number is more difficult to visualise. Complex numbers themselves live in 2 dimensions (they have a real part and an imaginary part) and when you apply a function to them, very often the result is another complex number which also lives in a 2 dimensional space. So trying to plot this function would require us to draw a three dimensional surface in four dimensions (cf. a one dimensional line in 2 dimensions for purely real functions). This we can't do, but what we can do is to ask what a set of complex numbers looks like when we apply a function to them, by having two complex planes next to each other.

Let's think about the complex numbers which lie along a point of fixed real part but varying imaginary part. Let's pick the real part to be 2 for now. These numbers are thus of the form: $z = 2 + ib$. As we vary b we trace out a vertical line which cuts through the real axis at 2. What happens to these points when we apply to them the function $f(z) = \frac{1}{z}$. We know that they will give: $\frac{2-ib}{2^2+b^2}$. As we vary the value b these will trace out not a straight line, but in fact a circle of radius 2. This is not immediately obvious until you've played around quite a bit with complex numbers, so don't be put off if you don't see it. However, we can plot this using Mathematica and find that vertical lines are mapped to circles of the form shown in figure 55 while horizontal lines are mapped to circles of a slightly different form as can be seen in figure 56

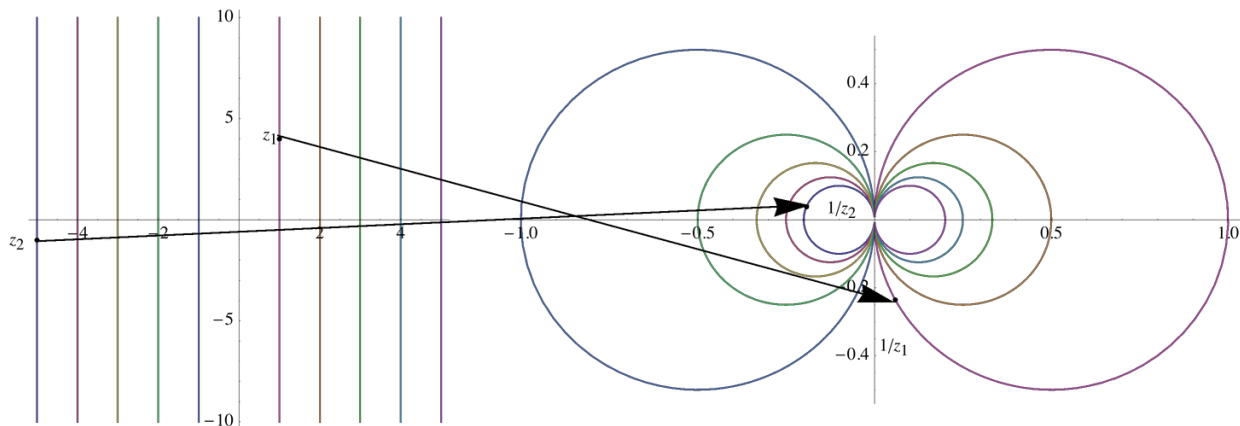


Figure 55: Vertical lines in the complex plane get mapped to circles when we apply the function $f(z) = \frac{1}{z}$ to them. Think of the line as being made up of a load of points in the complex plane, each of those points get mapped to another set of points but if we were to plot both the original line and the mapped image in the same graph it would look very confusing, so instead we draw the image in a plane next to the original. The arrow shows two example points on lines which get mapped to two points on circles.

10.4 Back to the algebra

OK, so that was a nice diversion which allows us to understand how to visualise functions of complex numbers - we do it by seeing how a whole line in the complex plane is mapped to another line in the complex plane under the action of a function. Let's get back to the main action.

We now know how to do all of the usual algebraic operations on complex numbers. Having learnt how to do $\frac{1}{z}$ we can also calculate one complex number divided by another. The reason we want to be able to do this is so that we can always convert a given complex number into cartesian form.

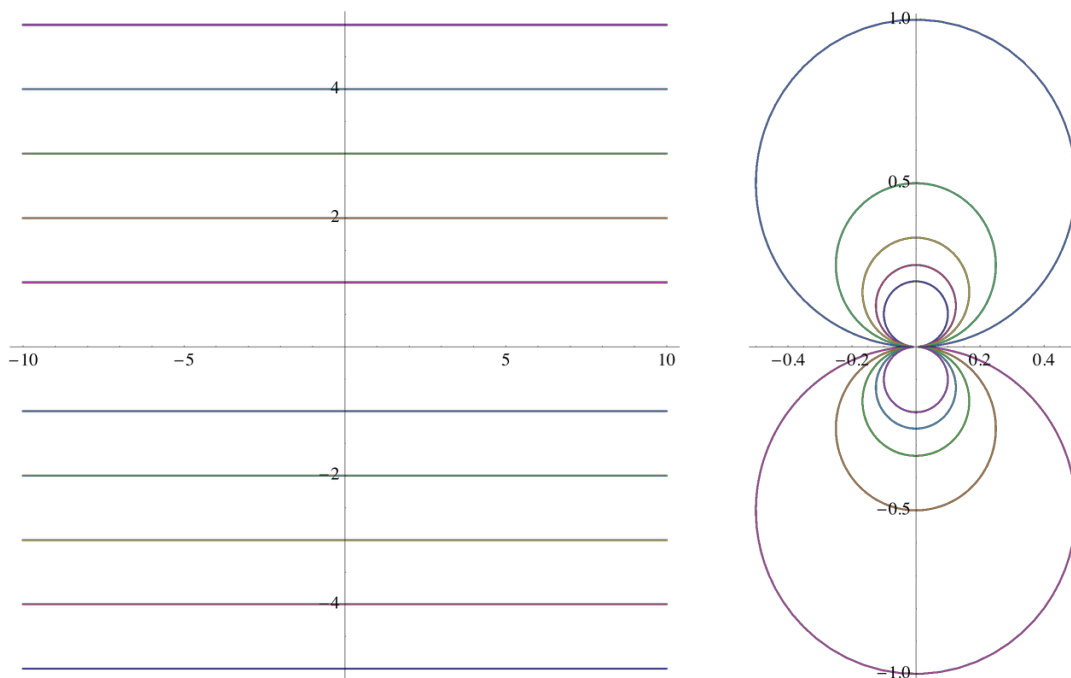


Figure 56: The equivalent mapping as figure 55 using the function $f(z) = \frac{1}{z}$ but for points along horizontal lines.

Let's take a completely generic example so that we will be able to do it in all situations. If we are given:

$$\frac{a + bi}{c + di} \tag{254}$$

and we want to convert it to normal cartesian form, we note that this is just $z_1 \frac{1}{z_2}$. We can then multiply top and bottom by the conjugate of z_2 which will give us: $\frac{z_1 \bar{z}_2}{|z_2|^2}$. In this case, this gives:

$$\frac{(a + bi)(c - di)}{c^2 + d^2} = \frac{ac + bd + i(bc - ad)}{c^2 + d^2} = \frac{ac + bd}{c^2 + d^2} + i \frac{bc - ad}{c^2 + d^2} \tag{255}$$

Because the coefficients a , b , c and d are real numbers, this expression is clearly in cartesian form, so we've done what we set out to do.

10.5 More on the modulus/argument form and the exponential of a complex number

We saw previously how the modulus and argument give another way of describing a complex number, though a given complex number will have an

infinite number of possible arguments (angles) which describe it. We are going to find shortly that this form is, for many operations on complex numbers, a more powerful form to use. We start by reminding ourselves that for a complex number $a + bi$ where the length of the line between the origin of the complex plane and the number in the complex plane is r and the angle subtended between that line and the positive real axis is θ , we can, through some simple trigonometry, show that:

$$a + bi = r \cos \theta + ir \sin \theta = r(\cos \theta + i \sin \theta) \quad (256)$$

Sometimes this particular combination of \cos and \sin is called "cis": $\text{cis} \theta = \cos \theta + i \sin \theta$ but we shan't use that notation very often. It's going to turn out that this particular combination of \cos and \sin has an even simpler way of being written. We note this by reminding ourselves of the Maclaurin series of \cos and \sin .

$$\cos \theta + i \sin \theta = \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n+1}}{(2n+1)!} \quad (257)$$

It turns out (show this) that this can be written in a much more compact form:

$$\cos \theta + i \sin \theta = \sum_{n=0}^{\infty} i^n \frac{\theta^n}{n!} = \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!} \quad (258)$$

but this last equation is almost exactly the same as the Maclaurin expansion of e^x except now it's ix instead of x . But that's great! That means that it's the Taylor expansion of e^{ix} . But thus just means that:

$$\cos \theta + i \sin \theta = e^{i\theta} \quad (259)$$

That's pretty simple! θ is just a real number. How about if we try and take the exponential of a complex number? We've never had to do that before, but now we know how to take the exponential of a purely imaginary number ($i\theta$) and we already know how to take the exponential of a real number, so it can't be that hard, can it? If we want to take the exponential of z where $z = a + bi$, this will be given by:

$$e^{a+bi} = e^a e^{bi} \quad (260)$$

But the first bit is just a real number and the second bit we saw how to do above, so:

$$e^a e^{bi} = e^a (\cos b + i \sin b) \quad (261)$$

But that's even better, because we know how to calculate $\cos b$ and $\sin b$, so we can calculate this on our calculators. We can thus calculate the exponential of any complex number - nice, huh! In fact, there's something very special going on here. The magnitude of this exponential is $|e^z| = e^a$ and the argument: $\arg(e^z) = b$. This means that when we exponentiate a complex number, the magnitude of the result is purely related to the real part of the complex number and the argument is purely related to the imaginary part. As an exercise, write down a list of ten complex numbers and then calculate their exponent. We can play the same game that we played before with seeing how vertical and horizontal lines get mapped under the exponential function in the complex plane. The result is shown in figure 57

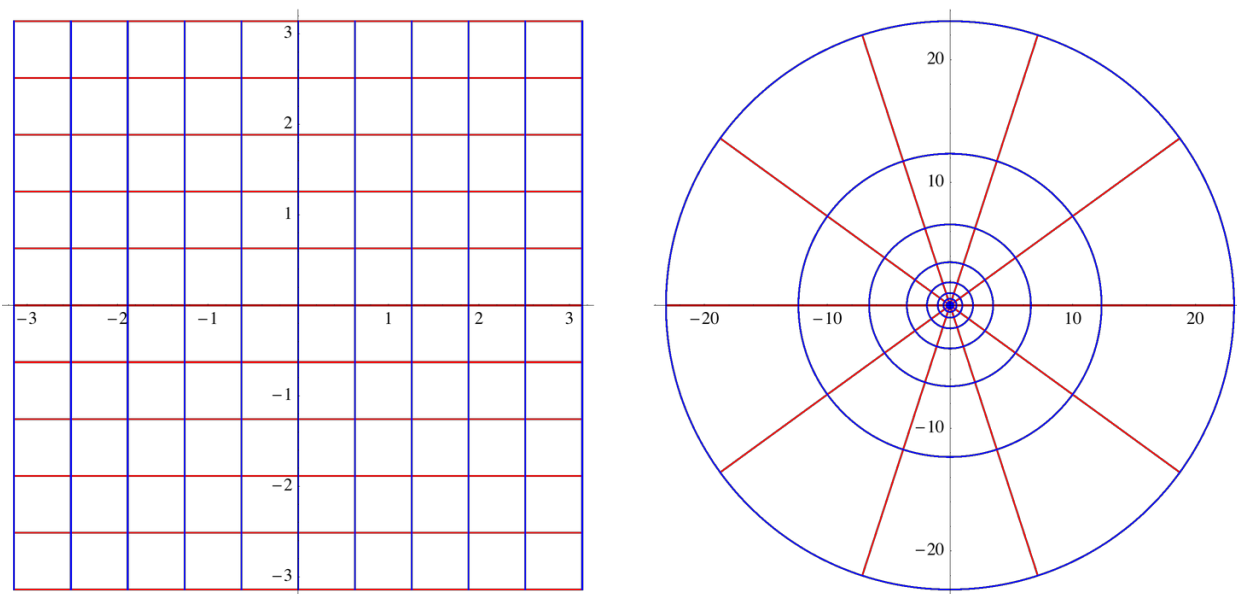


Figure 57: The complex plane mapped under the exponential function. You see that horizontal lines get mapped to radial lines, while vertical lines get mapped to circles where the radius is the exponential of the real value of the points on the line.

Now that know that $r(\cos \theta + i \sin \theta) = re^{i\theta} = e^{\ln(r)+i\theta}$ (this is what we wrote previously but now we have written r explicitly on the left and have simply put it in the exponent on the right), it is a trivial step to show a very important theorem related to complex numbers and trigonometric functions. In fact it becomes trivial once we have that previous identity. De Moivre's theorem states that:

$$(\cos \theta + i \sin \theta)^n = (\cos n\theta + i \sin n\theta) \quad (262)$$

for integer n . This would not have been obvious previously, but now when we write the complex number in exponential form it is clear. Show that this is true!

10.6 Another philosophical aside - or why complex numbers have just given us something very powerful!

We should step back for a moment and note something highly non-trivial but very important. We have shown that there is a deep link between exponentials and trigonometric functions. This link would not have been at all clear had it not been for complex numbers. Previously we saw the exponential function as a function which was always increasing, was its own derivative and had certain specific derivative behaviour at $x = 0$, and trigonometric functions were periodic functions which were related to triangles and circles and all sorts of nice geometric operations. By the power of complex numbers we have shown that the two are actually two faces of the same coin - and that is a very powerful thing to have shown!!

As a general, but not strict rule, the addition of new elements in mathematics shows deep links between different areas which otherwise would not be visible. We have just been a powerful highway between trigonometry and exponential growth using the tarmac of complex numbers. The link was always there before, but not visible until we had added this new ingredient.

10.7 Back to de Moivre...

So we have seen that if we raise $\cos\theta + i\sin\theta$ to an integer power, we end up with an expression where the argument θ is just multiplied by the integer power. It is time now to look at what multiplying a complex number by $\cos\theta + i\sin\theta$ or equivalently by $e^{i\theta}$ does.

Let's take some complex number $a + bi$ and multiply it by some $e^{i\theta}$. We know that we can write $z = a + bi$ as $|z|e^{i\arg(z)}$ (if this isn't obvious then take the left hand side and show that this is true). Now if we multiply z by $e^{i\theta}$ we clearly get $ze^{i\theta} = |z|e^{i(\arg(z)+\theta)}$. What do we notice about this complex number? Well, the modulus of it is the same (just $|z|$) and the angle has been changed from $\arg(z)$ to $\arg(z) + \theta$, it's as if we took the line in the complex plane linking the origin to z and we simply rotated it by an angle θ . You can see this explicitly in figure 58.

It is very important to note that while we have made it explicit here that θ is the argument of $e^{i\theta}$, we could just have well have written e^{ib} or any

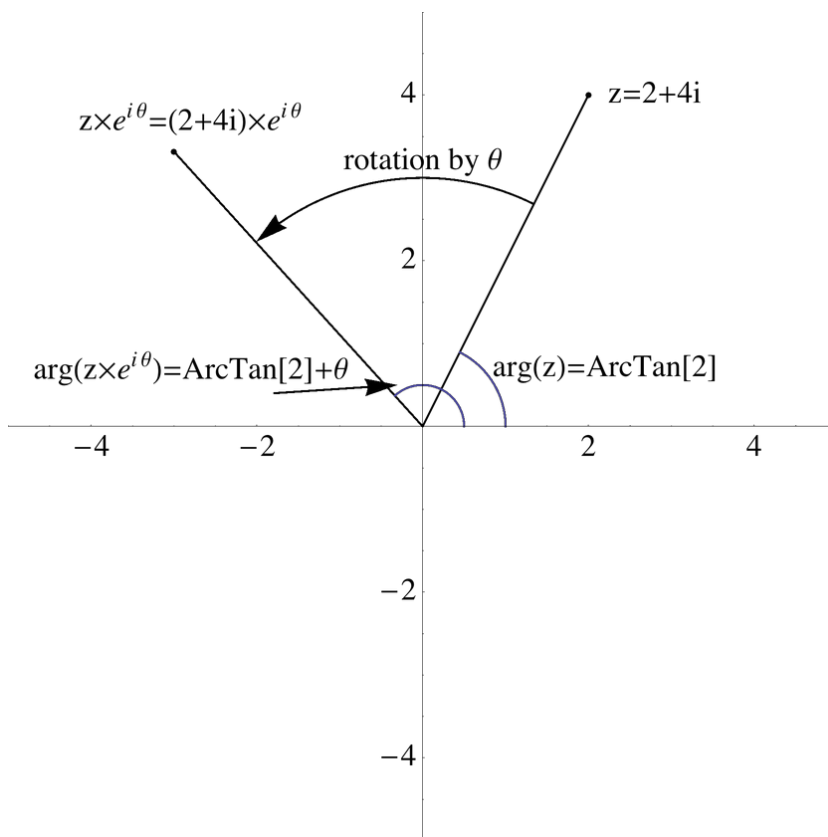


Figure 58: Multiplying a number by $e^{i\theta}$ is like rotating it anticlockwise in the complex plane by the angle θ .

other letter, and of course it would have been the same - we simply use θ in this case to make it clear that it's an angle. So, we see that if we multiply a complex number by something of the form e^{bi} we simply rotate it by an angle b . If we multiply a complex number by another complex number of the form e^{a+bi} then we will increase its magnitude by a factor e^a and change its angle (also called 'phase') by b . When we have the exponential of a complex number, the real part in the exponent will always tell us about magnitude and the imaginary part in the exponent will tell us about phase (or angle). You can see that multiplying by the exponential of a complex number allows us to transform complex numbers in very obvious ways. In the same way, if we want to add complex numbers together, then it's more clear what we are doing when we use the cartesian form.

de Moivre's theorem allows us to do some pretty powerful computations very easily. What if we wanted to know what $(1 - \sqrt{3}i)^{31}$ was? Let's call this z^{31} and let's first find the modulus and argument of z . Clearly $|z| =$

$\sqrt{1+3} = 2$ and $\arg(z) = \arctan(-\sqrt{3}) = -\frac{\pi}{3}$ (note that really the argument is the whole set of numbers you get when you add integer multiples of 2π to this but here we are happy to find a single solution - try, and make sure that the final result for the real and imaginary parts do not change when you add on multiples of 2π , this will be because $31 \times 2\pi$ is also an integer multiple of 2π and so when you take the sin and cos of this you will get the same number). So we can write:

$$z = 2(\cos(-\frac{\pi}{3}) + i \sin(-\frac{\pi}{3})) \quad (263)$$

now we can use de Moivre to write z^{31} as:

$$z^{31} = 2^{31}(\cos(-\frac{\pi}{3}) + i \sin(-\frac{\pi}{3}))^{31} = 2^{31}((\cos(-\frac{31\pi}{3}) + i \sin(-\frac{31\pi}{3}))) \quad (264)$$

but in fact $(\cos(-\frac{31\pi}{3}) + i \sin(-\frac{31\pi}{3})) = \cos(-\frac{\pi}{3}) + i \sin(-\frac{\pi}{3}) = \cos(\frac{\pi}{3}) - i \sin(\frac{\pi}{3})$ so:

$$(1 - \sqrt{3}i)^{31} = 2^{31}(\cos(\frac{\pi}{3}) - i \sin(\frac{\pi}{3})) \quad (265)$$

and we never even had to invoke the binomial theorem! Again, see what complex numbers have given us. If we didn't know how to manipulate these sorts of expressions using de Moivre we would have had to do this using the binomial theorem and we would have a horrible set of 32 terms that we'd have to somehow resum to give us something sensible. None of that was necessary now that we know about the special behaviour of complex numbers.

Note that we can think of this again in the complex plane. The number we start off with has magnitude 2 and angle with the positive real axis $-\frac{\pi}{3}$. If we multiply this by itself 31 times, we will end up with a number of magnitude 2^{31} and the angle will be 31 times the original angle (thanks to de Moivre). In figure 59 we give an example of different powers of z where $z = (1 - \sqrt{3}i)$

So, we know how to take the exponential of any complex number now. We do it by converting the exponential into the exponential of the real and imaginary parts separately, and then use the relationship between e^{ia} and the cos and sin functions to write everything in terms of functions of real numbers, which we know how to deal with. How about the trigonometric functions applied to complex numbers? Well, we have a pretty good hint already from how we got from the exponential of complex numbers to trigonometric functions of real numbers. In fact we're just going to give the answer, but you can work it out using Taylor series as well. For a complex number z :

$$\begin{aligned} \cos z &= \frac{e^{iz} + e^{-iz}}{2} \\ \sin z &= \frac{e^{iz} - e^{-iz}}{2i} \end{aligned} \quad (266)$$

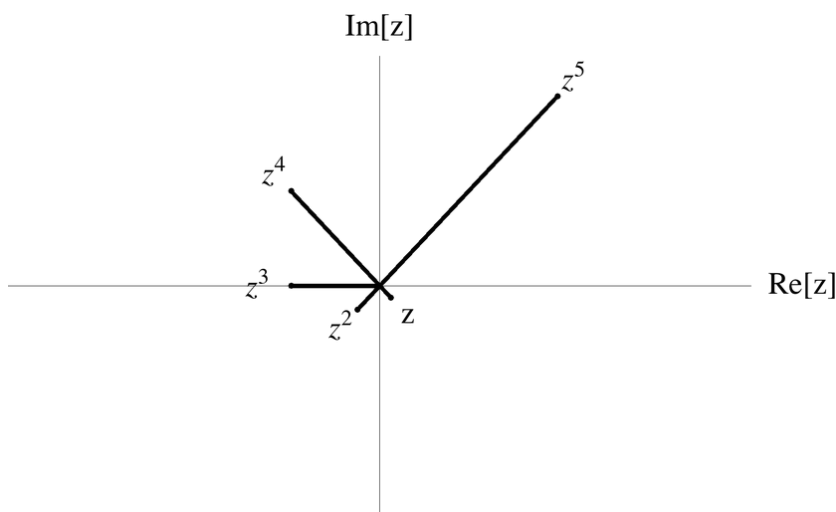


Figure 59: $z = 1 - \sqrt{3}i$ to different powers. Because the magnitude of this complex number is 2, the magnitudes of the different powers are just 2 to that power, and the angles add together each time simple to get $n \times \arg(z)$. The numbers on the axes have been left off here to keep the figure cleaner.

The first thing to check is that this is true when z is a real number. It looks pretty strange at first site, especially the definition of \sin because there's an i sticking out in the denominator like a sore thumb! How can this give us something real, as we know that it must? Let's take $z = \theta$ where θ is real. Then:

$$\begin{aligned} \cos \theta &= \frac{e^{i\theta} + e^{-i\theta}}{2} = \frac{\cos \theta + i \sin \theta + \cos(-\theta) + i \sin(-\theta)}{2} = \cos \theta \\ \sin \theta &= \frac{e^{i\theta} - e^{-i\theta}}{2i} = \frac{\cos \theta + i \sin \theta - \cos(-\theta) - i \sin(-\theta)}{2i} = \sin \theta \end{aligned} \quad (267)$$

So that holds. How about trying to understand the behaviour of the trigonometric function of a purely imaginary complex number? Well, we know for sure that trig functions are always between -1 and 1, right? Well, it turns out that this is only true if θ is purely real. In fact if it's complex then the trig functions can give you any complex number you like. We'll see that shortly. Let $z = ib$ where b is a real number, then:

$$\begin{aligned} \cos ib &= \frac{e^{-b} + e^b}{2} \\ \sin ib &= \frac{e^{-b} - e^b}{2i} \end{aligned} \quad (268)$$

Then we can see that $\sin ib$ is purely imaginary and $\cos ib$ is purely real. The

other thing that we can note is that for large positive or negative b , both of these numbers are going to be large in magnitude. In fact these functions have names $\cos ib = \cosh b$ which is known as hypergeometric cosine, and $i \sin ib = \sinh b$ which is hypergeometric sine. We can plot both of these functions in figure 60 and see that they don't behave at all like the normal sine and cosine functions that we are used to. There are hyperbolic trig

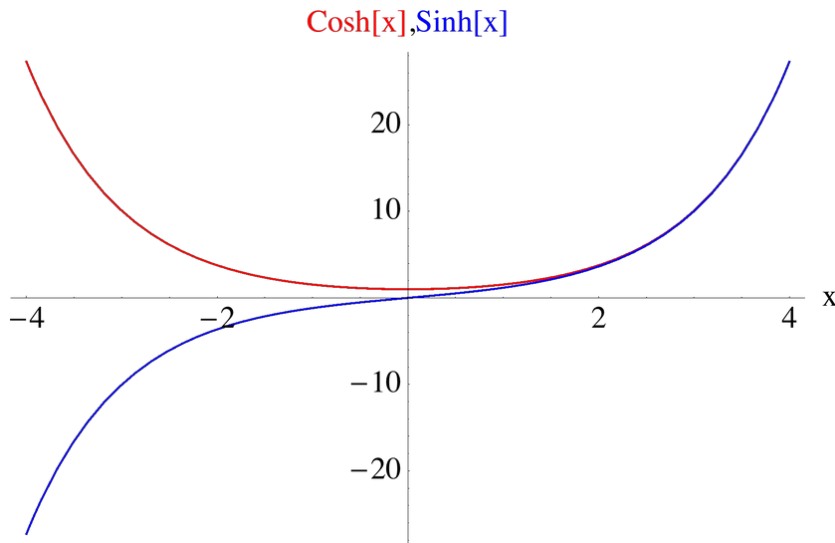


Figure 60: The hyperbolic sine and cosine functions. They are neither periodic nor bounded as are the normal sine and cosine functions.

identities just as there are for normal trig functions, but unlike the case where $\sin^2 x + \cos^2 x = 1$, now $\cosh^2 x - \sinh^2 x = 1$. In fact it is precisely this hyperbolic identity which gives the functions their name. Unlike the sine and cosine functions which describe circular behaviour, so the hyperbolic versions describe the behaviour of things moving along hyperboloids. Note that none of this is talk of hypergeometric functions is examinable, but it is interesting.

OK, so let's look at a very general example to see that actually we can have trig functions giving any result we want, not just numbers between -1 and 1 as we had when we were restricted to the real numbers. Let's ask a very general question. Find z when:

$$\sin z = w \quad (269)$$

We start by writing the trig function in exponential form and then multiply both sides of the equation by $2ie^{iz}$. This gives:

$$e^{2iz} - 1 = 2wie^{iz} \quad (270)$$

Now we note that we can write $e^{iz} = p$ and we can happily now write the equation as a normal quadratic equation:

$$p^2 - 2wip - 1 = 0 \quad (271)$$

But we can solve this using the normal quadratic recipe (You can prove that this still holds, even for complex numbers, simply by using it and then plugging the result back into the quadratic equation). So:

$$p = \frac{2wi \pm \sqrt{-4w^2 + 4}}{2} = wi \pm \sqrt{1 - w^2} \quad (272)$$

So, we now have that:

$$e^{iz} = wi \pm \sqrt{1 - w^2} \quad (273)$$

How can we calculate what z must be? Let's start by saying that $z = a + bi$ so:

$$e^{ia-b} = e^{-b}(\cos a + i \sin a) = wi \pm \sqrt{1 - w^2} \quad (274)$$

We can then solve this by taking the modulus and argument of both sides. First the modulus:

$$|e^{-b}(\cos a + i \sin a)| = e^{-b} = |wi \pm \sqrt{1 - w^2}| \quad (275)$$

and the arg:

$$\arg(e^{-b}(\cos a + i \sin a)) = a = \arg(wi \pm \sqrt{1 - w^2}) \quad (276)$$

so now, given a w , we can calculate precisely the z such that $\sin z = w$. You can see that you would be able to do this for any w you are given, because any w has both a modulus and an argument (though you'd have to be careful about $w = 0$). Of course we have to be a little more careful than this because we can add integer multiples of $2\pi i$ to z and it will still fulfil the criteria. to be a solution to the equation, thus, the full solution is:

$$z = \arg(wi \pm \sqrt{1 - w^2}) - i \ln |wi \pm \sqrt{1 - w^2}| + 2\pi in \quad n \in \mathbb{Z} \quad (277)$$

10.8 Calculating roots of complex numbers

So, we can do basic algebra with complex numbers, powers of them, trig and exponential functions of them. There isn't much left that we might want to do, but taking powers of them is very important and also pretty easy. Let's say we wanted to calculate the solution to the equation:

$$z^2 = -1 \quad (278)$$

Well, we know what z is for this, because that's what got us into this mess in the first place! We know that there are two solutions and they are $\pm i$. That is to say that if you multiply either of these numbers by themselves, you get -1, by definition. How about $z^2 = i$. The first thing to do whenever you have to take the root of a complex number, or a real number (here the second root of i) is to convert that number into modulus argument form - it will be infinitely easier like that. We know that $i = re^{i\theta}$ where $r = 1$ and $\theta = \frac{\pi}{2} + 2n\pi$. So we can write:

$$z^2 = e^{\frac{i\pi}{2} + 2n\pi i} \quad (279)$$

or just:

$$z = e^{\frac{i\pi}{4} + n\pi i} \quad (280)$$

n again goes over the integers, but we have to ask whether it's going to make a difference to what z is. When $n = 0$ this number is $e^{\frac{i\pi}{4}}$ and you can work out that this is $\frac{1+i}{\sqrt{2}}$, when $n = 1$ we have $\frac{-1-i}{\sqrt{2}}$ which is different, but when $n = 2$ we get back to the same answer as when $n = 0$, and for $n = 3$ we get the answers for when $n = 1$. We can see very quickly that there are actually only two unique solutions to this equation which are $\pm \frac{1+i}{\sqrt{2}}$. Plot these in the Argand plane and you will see how they relate to i . It is pretty clear that when you multiply these numbers by themselves that they will give you i (remember, multiplying two complex numbers together is done by adding the arguments and multiplying the magnitudes. In this case the magnitudes are 1 and the phases add up to give $\frac{\pi}{4}$ (modulo 2π in both cases)).

How about a slightly more complicated question. How about:

$$z^5 = (1 + i) \quad (281)$$

Again, we start by converting from cartesian to mod/arg form which gives:

$$z^5 = \sqrt{2}e^{\frac{i\pi}{4} + 2n\pi i} \quad (282)$$

so:

$$z = 2^{\frac{1}{10}}e^{\frac{i\pi}{20} + \frac{2n}{5}\pi i} \quad (283)$$

You can see quickly that when $n = 0, 1, 2, 3$ and 4 you will get unique results, but $n = 5$ will give the same as $n = 0$, and $n = 6$ will give the same as $n = 1$, etc. We can plot these in the Argand plane and see what they look like in figure 61. It's always nice to have a general procedure when we're trying to solve a set of problems, and in this case we can find a very simple general procedure. You should have noticed one thing by now, which is that the n^{th} root of a number (be it real or imaginary) will have n solutions. That might

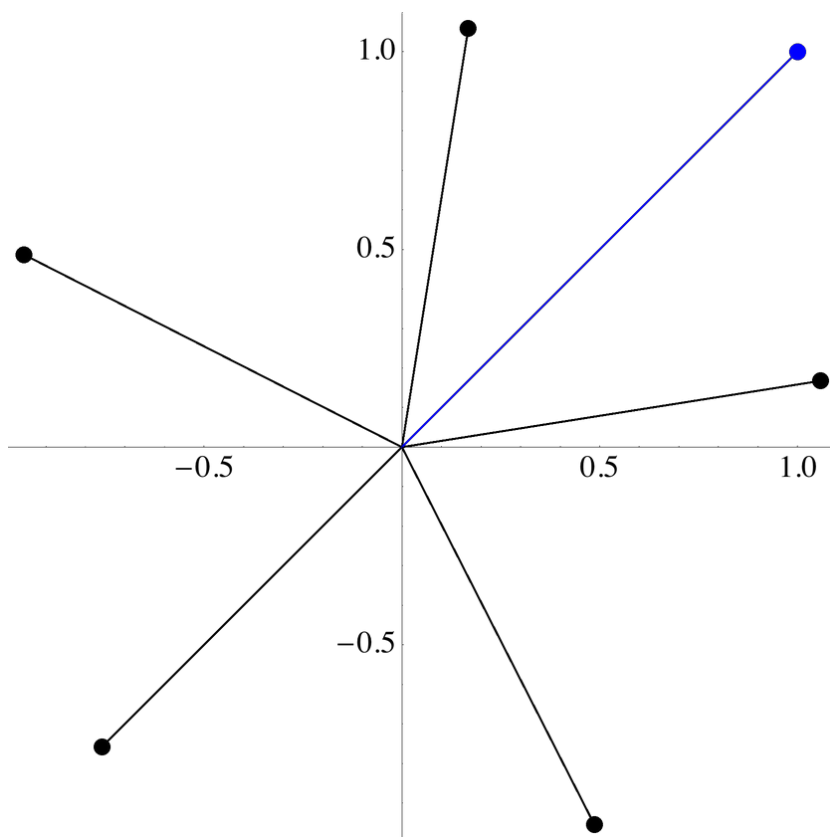


Figure 61: The five fifth roots of $1 + i$. The point $1 + i$ is shown in blue, the five roots are shown in black. You can see that when you multiply each one of them by itself five times you will get $1 + i$ - this is the definition of the fifth root.

seem strange that there are 7 solutions to the equation $z^7 = 1$ but check it - it is true, and this is because there are seven angles, that when you add them to themselves 7 times, give you integer multiples of 2π - think about it...

So, what if we want to solve:

$$z^n = q \tag{284}$$

We start by writing q in mod/arg form, calling its magnitude ρ and its angle ϕ :

$$z^n = \rho e^{i\phi + 2k\pi i} \tag{285}$$

now:

$$z = \rho^{\frac{1}{n}} e^{\frac{i\phi}{n} + 2\frac{k}{n}\pi i} \tag{286}$$

where now $k = 0, 1, \dots, (n - 1)$. so again we have n solutions for an n^{th} root.

We should note something rather important here. When we write $\rho^{\frac{1}{n}}$ this really is the n^{th} root of a real number, whereas when we write $z^n = \rho$, the two are not inverse of each other. You can see this by the fact that you write $x^2 = 1$ which has two solutions but writing $x = \sqrt{1}$ is just one of them, the negative solution is encoded in the angular part of the solution. ie. in this case you would write $x = \sqrt{1}e^{in\pi}$ where $n = 0, 1$ which is just the statement that ± 1 are the two solutions to the equation $x^2 = 1$.

10.9 Finding roots of polynomial equations

When we were playing around with partial fractions we appeared to make a bit of an assumption which was that the only forms that we had to deal with in the denominator of a fraction could always be written as a factor of either linear parts $(a + bx)$ or quadratic parts which we could factor into linear parts $(ax^2 + bx + c)$ where $b^2 - 4ac < 0$, and of course multiple powers of these, for instance we could have terms like $(a + bx)^3$ in the denominator. How do we know that we can always split a polynomial up into these factors where the coefficients are real? Couldn't it be for instance that if I gave you a cubic polynomial that all the roots were complex and so I couldn't factor it in a way that every factor came out with real coefficients? It turns out that the answer is no, but we need a couple more ingredients to prove this.

We said that we were dealing with ratios of polynomials, eg. $\frac{P(x)}{Q(x)}$ but a polynomial where you only have linear and quadratic factors multiplied together doesn't seem to be very general. How do we know that we can always write a polynomial in this way? Well, it turns out that using complex numbers we can prove this very powerful fact. The statement is:

Theorem: Every polynomial with real coefficients factors into a product of factors which are either real or quadratic irreducible.

Quadratic irreducible means that it can't be split into a product of two linear factors with real coefficients. However, in order to prove this we need to look at a couple of important things. The first is another theorem which we will prove:

Theorem: The non-real roots of a polynomial equation with real coefficients occur in complex conjugate pairs.

This says that if some complex number q is a root of a polynomial, then so is \bar{q} . This sounds rather strange, but let's prove it in a few simple lines.

Let our polynomial with real coefficients be:

$$f(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 \quad (287)$$

where by definition, the a_i are all real. If α is a root then, by definition $f(\alpha) = 0$ (this is what a root means in the context of a polynomial equation). Then we want to show that $f(\bar{\alpha})$ is also a root which simply means that the complex conjugate of a root is also a root. What is $f(\bar{\alpha})$? Well, we just plug it in:

$$f(\bar{\alpha}) = a_n \bar{\alpha}^n + a_{n-1} \bar{\alpha}^{n-1} + \dots + a_1 \bar{\alpha} + a_0 \quad (288)$$

but we know that the conjugate of a product of complex numbers is the same as the product of the conjugate (ie. $\bar{a}^n = \bar{a}^n$). See we can rewrite the above as:

$$f(\bar{\alpha}) = a_n \bar{\alpha}^n + a_{n-1} \bar{\alpha}^{n-1} + \dots a_1 \bar{\alpha} + a_0 \quad (289)$$

but we also said that the coefficients were real, and the conjugate of a real coefficient is just that coefficient, so we can "expand" the conjugation:

$$f(\bar{\alpha}) = a_n \bar{\alpha}^n + a_{n-1} \bar{\alpha}^{n-1} + \dots a_1 \bar{\alpha} + a_0 \quad (290)$$

and furthermore, the sum of a set of conjugates is the conjugate of the sum:

$$f(\bar{\alpha}) = a_n \bar{\alpha}^n + a_{n-1} \bar{\alpha}^{n-1} + \dots a_1 \bar{\alpha} + a_0 \quad (291)$$

but this is just the same as:

$$f(\bar{\alpha}) = f(\alpha) = \bar{0} = 0 \quad (292)$$

So we have shown that if α is a root, then so is $\bar{\alpha}$. This means that any time you find a complex root of an equation, it must have another complex root which is the conjugate of the first root. In fact, for the case of a quadratic equation we can see this immediately because if a , b and c are real, then the only place that an imaginary number can come into the root is from the part which is $\pm\sqrt{b^2 - 4ac}$ which means that the two solutions, if $b^2 - 4ac < 0$, are going to be conjugate of each other (the + part and the - part). We have just shown that this is true also for higher order polynomials. In fact this is true for any polynomial at all.

There's a second crucial theorem that we need to show that we can factor any polynomial into a product of linear and irreducible quadratic factors. This is the Fundamental Theorem of Algebra:

Theorem Every polynomial (with real OR complex coefficients) which is of degree at least one, has a zero.

This says that there is always a solution to any polynomial (equalling zero) which is order more than 0. A zeroth order polynomial is just a constant and so it's clear that this will not have a solution which is equal to zero, unless the constant is zero.

We will not prove the Fundamental Theorem of algebra but we will use it to prove our initial statement about factoring polynomials.

Let's take a polynomial $P(z)$. by the Fundamental Theorem, it has at least one root, let's call it r_1 . That root can either be real or complex. If it's real, then we know that the Polynomial can be written as $(z - r_1)\tilde{P}(z)$

where $\tilde{P}(z) = \frac{P(z)}{(z-r_1)}$. We can always divide a polynomial by another to get another polynomial as long as the order of the polynomial on top is greater than that on the bottom.

The other option for our root is that it is complex. If it is complex then its conjugate is also a root. This means that $(z - r_1)(z - \bar{r}_1)$ is a factor of the polynomial, but this is purely real. We can see this by noting that if $r_1 = a + bi$ then $(z - r_1)(z - \bar{r}_1) = (z - a - bi)(z - a + bi) = ((z - a) - bi)((z - a) + bi) = ((z - a)^2 + b^2)$ which is an irreducible quadratic (ie. we can't write it as the product of linear factors). In this case we can write: $P(z) = ((z - a)^2 + b^2)\tilde{P}(z)$.

Then we can take $\tilde{P}(z)$ and play exactly the same game. In fact we can do this until we are left with just a constant and this clearly can't be written as the product of linear and quadratic factors - it's as simple as it can be.

So we have proved our initial claim using both the fundamental theorem of algebra and the fact that roots always come in conjugate pairs. This fact allows us to play with partial fractions as we did before, so it's really a very powerful statement.

Let's just see where that gets us. Now if I give you the polynomial $2z^3 - 9z^2 + 14z - 5$ and tell you that it has a zero at $z = 2 - i$, then you can immediately factor this cubic. The first thing you know for sure is that $z = 2 + i$ is also going to be a root. Now you can take this function and divide by $(z - (2 - i))(z - (2 + i)) = ((z - 2)^2 + 1)$ This gives:

$$\frac{2z^3 - 9z^2 + 14z - 5}{((z - 2)^2 + 1)} = (2z - 1) \quad (293)$$

and so:

$$2z^3 - 9z^2 + 14z - 5 = ((z - 2)^2 + 1)(2z - 1) \quad (294)$$

We have factored the polynomial into a linear factor and an irreducible quadratic factor. Now we know the solutions to this equation.

10.10 Complexity from complex numbers - The beauty of the Mandelbrot set (non-examinable)

We are about to show that you can get incredible structure from the simplest algorithm when we use complex numbers.

The equation we are going to look at is an iterative equation:

$$z_{i+1} = z_i^2 + C \quad (295)$$

with $z_0 = 0$. You simply get the next z_i from plugging in the previous one, squaring it and adding a number C . I'm going to give you a value for C , then

you're going to iterate this equation and see what happens. For instance, if I give you the number $C = 3$:

$$\begin{pmatrix} i & z_i + C^2 & |z_{i+1}| \\ 0 & 0^2 + 3 & 3 \\ 1 & 3^2 + 3 & 12 \\ 2 & 12^2 + 3 & 147 \\ 3 & 147^2 + 3 & 21612 \end{pmatrix} \quad (296)$$

You can see that this number is just going to keep on increasing without end if we keep applying the algorithm. How about a smaller number, let's say $C = 0.1$:

$$\begin{pmatrix} i & z_i + C^2 & |z_{i+1}| \\ 0 & 0^2 + 0.1 & 0.1 \\ 1 & 0.1^2 + 0.1 & 0.1121 \\ 2 & 0.1121^2 + 0.1 & 0.112566 \\ 3 & 0.112566^2 + 0.1 & 0.112671 \end{pmatrix} \quad (297)$$

It looks like this is tending to some value. In fact it has come to a fixed point where $z = z^2 + 0.1$. There are actually two solutions to this equation but one of them is 0.112702 which is where we are tending towards. I can clearly perform this procedure with C being any number and I can ask whether the z_i diverges as $i \rightarrow \infty$, or whether it always stays small. In fact I can be a bit more strict. I can ask whether the magnitude of the z_i always stays less than 2 or whether it becomes greater than 2. Let's ask this for all numbers which are real.

It turns out that this is a pretty easy problem. If $-2 < z < 0.25$ then the magnitude of z always stays below 2 and if it's outside this range, then it blows up. As an example Let's look at the value of the z_i as we iterate for numbers just below and just above $C = 0.25$. This is shown in figure 62

But we've only looked at values of C on the real line, and frankly the results were not very interesting. Why should it be any more interesting if we let C be any complex number? It turns out that the answer is both surprising and beautiful.

We'll work out what is the set of complex numbers for which z_i always stays below 2 using a technique of picking numbers in the complex plane at random and testing them - ie. plugging them into the iterative algorithm and seeing what happens to z_i . We're only going to run the iterative algorithm for a maximum number of times, but if during that time the magnitude of z goes above 2 we will quite the algorithm and say that that value of C that we used is not in the set that we are interested in. In fact to get the set accurately we would have to be more sophisticated than this and run the

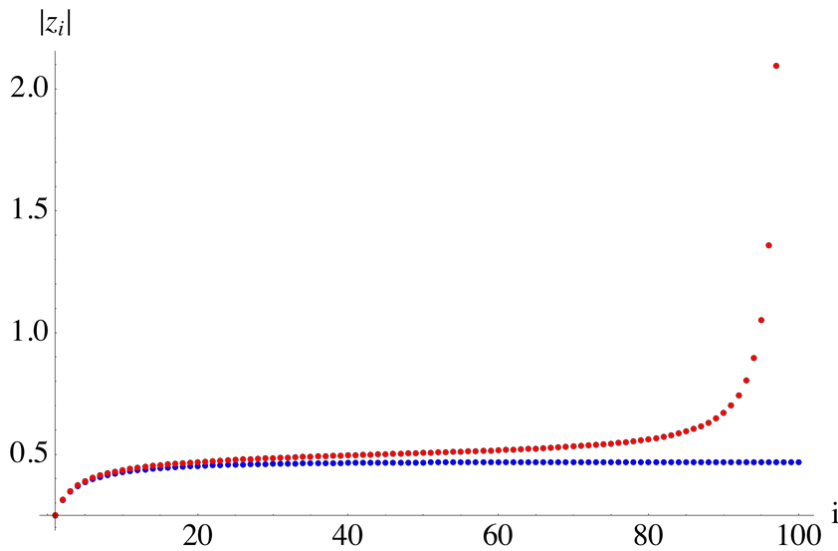


Figure 62: Values of z_i for $C = 0.249$ (blue points) and $C = 0.251$ (red points). We see that for the blue points the value of z_i converges to 0.5, whereas for $C = 0.251$ it diverges. It is true that for all points between -2 and 0.25 z_i converges and for all points outside of this set of numbers, it diverges.

algorithm in a more intelligent way so that we're not including points in the set which should not be there. For instance, in figure 62 had we only run the algorithm 50 times, we would have thought that 0.251 was in the set, whereas we see that eventually, the magnitude gets above 2 and so it's not in the set. However, for the purposes of this explanation, running the algorithm 100 times will be sufficient to see the complexity emerge.

Let's try with a random complex number. Let's take $C = \frac{1+i}{4}$. This, together with $C = \frac{1+i}{2.5}$ is shown in figure 63

We can see immediately that the behaviour looks somewhat different. For real numbers the numbers either converged, or diverged fairly obviously, but for complex numbers they seem to oscillate before they converge or diverge and it's not altogether clear what will happen if you look at the values at any particular i .

Let's now take random points in the complex plane (values of C) and if the value of z_i converges, we'll put a blue point there and if it diverges (within 100 iterations of the algorithm) we'll put a red point there. What will this look like? This is shown in figure 64 for 100,000 random points.

You can see that it looks rather like a random splotch of paint. In fact it turns out that this really isn't a very detailed view of the set at all (the word

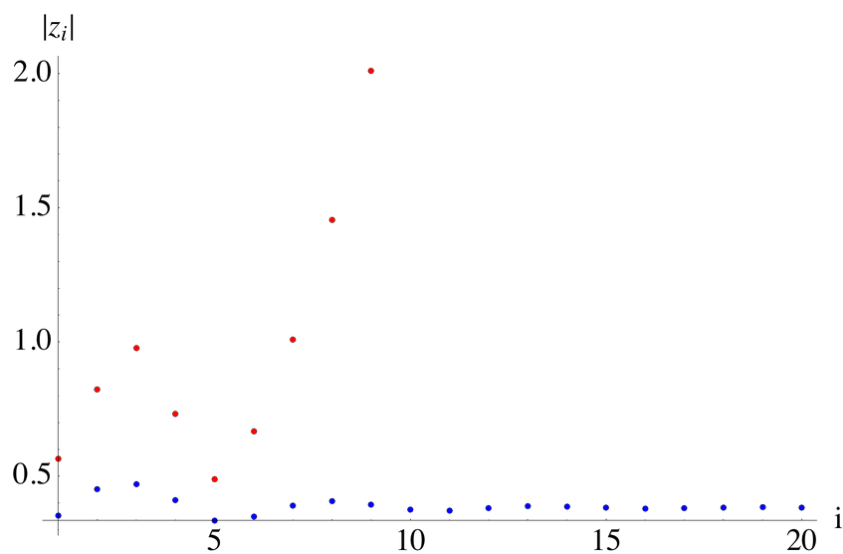


Figure 63: Values of z_i for $C = \frac{1+i}{4}$ (blue points) and $C = \frac{1+i}{2.5}$ (red points). For the blue points the values fluctuate but get closer and closer to a fixed value which is less than 2, whereas for the red points, they quickly diverge.

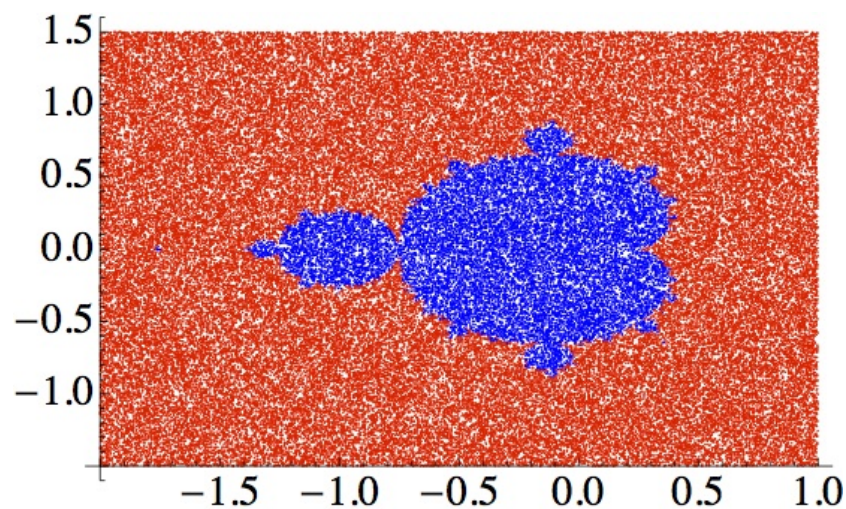


Figure 64: Graph of those values of C (in the complex plane) for which the iterative values of z_i diverge within 100 iterations (in red) and those for which they don't (in blue). This is a sample of around 100,000 points.

set in 'Mandelbrot set' is because we are looking for the set of numbers with a particular behaviour, in this case the blue points are in the set, and the red points are not). In fact it turns out that there is an infinite amount of detail

to be seen. What does this mean. Well, let's say that we take a small patch of figure 64 and try and look at it in more detail. Let's take a region of the complex numbers C with real parts between 0.2 and 0.5 and imaginary parts 0.25 and 0.8. We sample in this smaller region with 100,000 points and find the image in figure 65

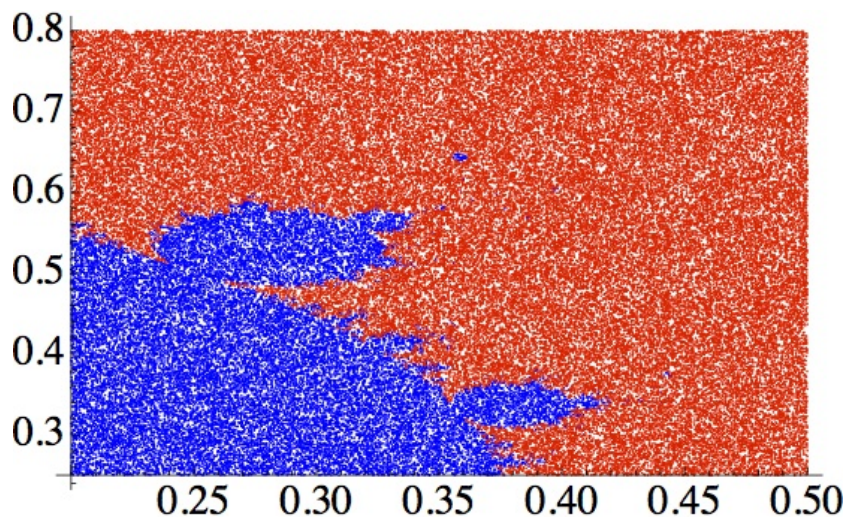


Figure 65: We zoom in on a small region of this graph to try and get down to the lowest level of detail. We've taken a small part of figure 64 and sampled 100,000 points in that small region.

But it looks like there might be more structure at even smaller levels. Let's zoom in on an even smaller region here. In figure 66 we look at a small region in figure 65 of complex numbers C with real parts between 0.3 and 0.325 and imaginary part between 0.55 and 0.62.

Again, we zoom in on a tiny region in figure 66, again sampling 100,000 points, but this time in the tiny region of C with real part between 0.315 and 0.316 and imaginary part between 0.576 and 0.578. This is shown in figure 67

In fact we could keep doing this for every and the images would never become smooth and we would never stop getting more detail. This is a fractal and you can zoom in infinitely far and keep seeing more structure. Think of this like a coastline. If you take a picture of a country from above, you will get a certain detail of the coastline, but as you zoom in more and more you will start to see the structure of individual beaches, then further in you will see structure of individual curves in the beach, then you will start to see structure in the rocks in the beach, then you will start to see structure in the sand, then in smaller and smaller particles. However, you can only

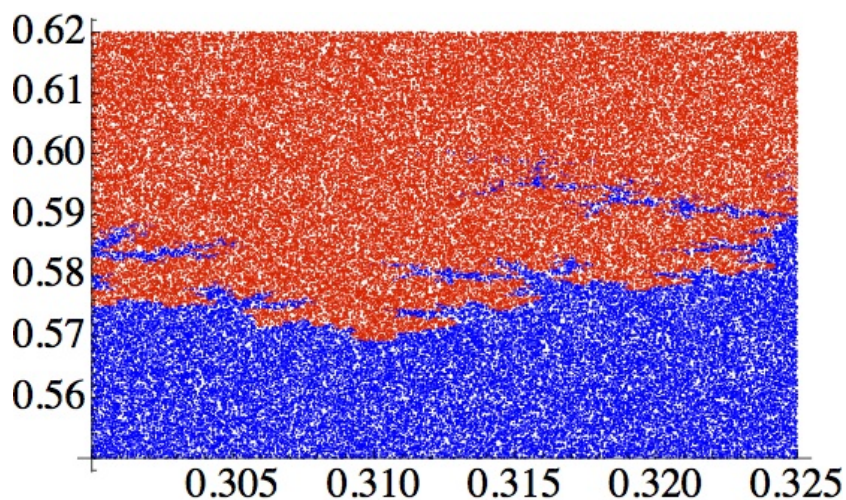


Figure 66: We now take a small region of figure 65 and zoom in on that, again sampling a very small area with 100,000 random points.

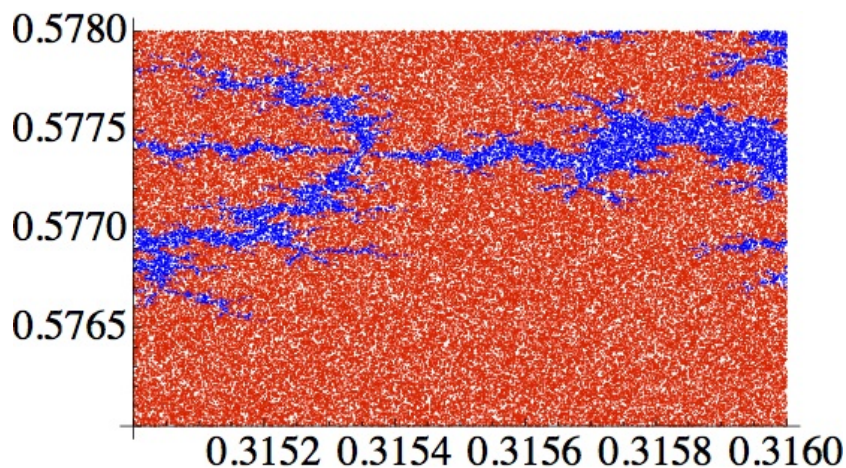


Figure 67: Sampling a small region in figure 66 with another 100,000 points. We find that we can zoom in further and further and there will always be more detail at every level. This goes on for an infinite number of zooms!

zoom in so far because there seems to be a fundamental limit to the scale of the universe. However, a fractal is different. In the mathematical world you can zoom in more and more and always get more detail....always!!! This is a pretty amazing fact for such a simple equation and its because of the wonderful behaviour of complex numbers that we get this complexity!

In fact using a relatively small number of sample points (only 100,000) doesn't give us a very good picture of the intricacy of the image. In figure 68

there is an image taken from wikipedia (http://en.wikipedia.org/wiki/File:Mandelset_hires.png) which shows in much better detail the mandelbrot set.

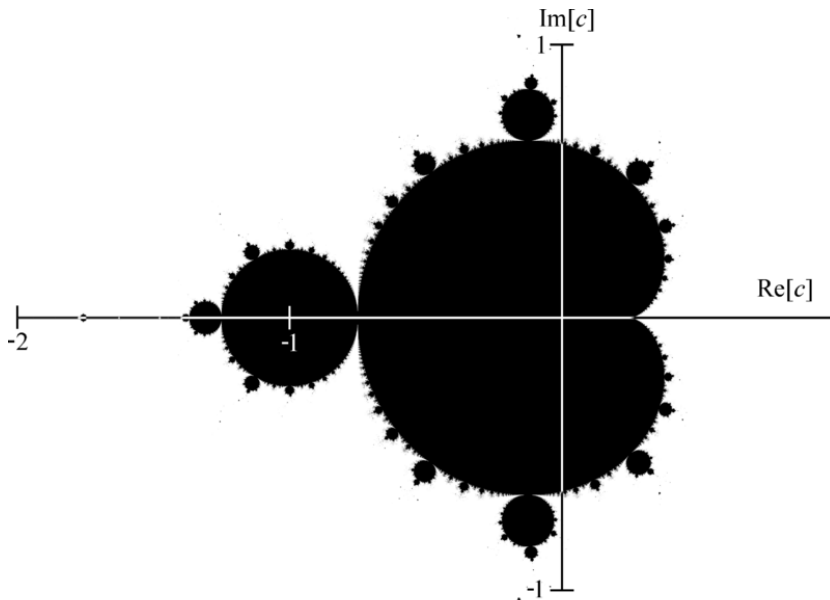


Figure 68: A much better rendering of the Mandelbrot set, though it's not as obvious here how it is found as with our discrete sampling technique. Anything in black is in the set, anything in white is not in the set. Each of the nodes can be zoomed into infinitely far and you will never stop seeing new structure.

What we have done here is very crude, we have only zoomed in a tiny bit (in fact, what does tiny mean when we can zoom in infinitely far, in theory - isn't anything tiny compared to that?). Have a look here: <https://www.youtube.com/watch?v=0jGai087u3A> for a video of a zoom into a Mandelbrot set by over 200 orders of magnitude (zooming in by a magnification of 10 over 200 times). Note that the colours are related to those numbers which are not in the set but they can be colour coded by the number of iterations it takes for the algorithm to give you a z_i whose magnitude is greater than 2.

11 Differential equations

We're now going to take everything that you learnt about integration and turn it into a way to model and understand the world around us. This is a very powerful statement and indeed differential equations are without a doubt the most powerful mathematical tool we have to understand the behaviour of everything from fundamental particles to populations, economies, weather, flow of wealth, heat, fluids, the motion of planets, the life of stars, the flight of an aircraft, the trajectory of a meteor, the way a pendulum swings, the way a ponytail swings (see paper on this: <http://epubs.siam.org/action/showAbstract?page=2667&volume=70&issue=7&journalCode=smjmap>), the way fish move, the way algae grow, the way a neuron fires, the way a fire spreads...and so much more.

So what is a differential equation? It is an equation which contains one or more derivatives of a function. Let's look at a very simple example, of population growth. We might want to ask the question, how fast does a population grow? We will take the most naive model we can think of...which will be horribly wrong, but it will give us an idea of how to deal with differential equations and what they can tell us.

What might the rate of change of a population be related to? Well, presumably if you already have more members of a population then the population can grow faster. If you only have two members of a population then it can't grow very fast but if you have 100 members, there will be more babies added to the population more regularly. Let's say that a couple will have a baby every 30 years, on average - of course in reality the rate won't be so smooth, but when we have a large sample, this won't matter. If we have a single couple then the rate of change of the population will be one every 30 years (or $\frac{dP(t)}{dt} \approx 1/30$ - the rate of change of the population is 1 every 30 years. Note that we can choose to use whatever units we want, ie. years, seconds, millennia, as long as we are consistent throughout and use the same unit for every measurement). If we have 10 couples, then on average there will be one baby every 3 years (Note that we're making a gross approximation here that we are mixing the number of couples and the number of individuals but to get a general idea of what's going on, this won't matter very much - basically factors of 2 are not very important for this discussion). We're not yet taking into account the fact that these babies will grow and start having babies of their own. This will be the crucial step in a moment. So it seems that the higher the population, the higher the rate of baby production. It seems reasonable that as we double the population, we'll double the rate of

change of the population so let's postulate that:

$$\frac{dP(t)}{dt} \sim P(t) \quad (298)$$

That is, the rate of change of the population at time t is proportional to the population that you have at time t . There is a constant of proportionality, and that is the rate of baby production per individual. Let's call that constant k . It will depend on the type of population, be it people, or rabbits, or bacteria. They clearly have very different reproductive rates. So now we have:

$$\frac{dP(t)}{dt} = kP(t) \quad (299)$$

This is our first differential equation - it's an equation, and it has a differential in it. It's called first order because it contains the first derivative. If we had the second derivative anywhere in the equation it would be called a second order differential equation, and so on. The aim is going to be to find a function $P(t)$ that satisfies this equation. You'll see what that is when we've found it! Let's rearrange the equation a little bit :

$$\frac{dP(t)}{P(t)} = kdt \quad (300)$$

This links the way $P(t)$ changes (i.e. $dP(t)$) with small changes in time (dt). We can now simply put an integral sign on both sides (as long as we do the same thing to both sides, we're good!). This gives us:

$$\int \frac{dP(t)}{P(t)} = \int kdt \quad (301)$$

But that's fine, we know how to work with this because we've spent an age looking at integrals by now! Let's just integrate both sides to get:

$$\ln |P(t)| + C_1 = kt + C_2 \quad (302)$$

we see here that we have two constants of integration, but they are just constants, so let's add them together to get:

$$\ln |P(t)| = kt + C \quad (303)$$

where C is a new constant, having absorbed C_1 into C_2 . Whenever we have two constants of integration like this we can just add them together to get a single constant. Soon you will see exactly WHY we have constants of integration and what they really mean. They are very very important, and

very much related to the problem we are looking at. Everything you did before will suddenly become oh so very clear!

Ok, let's do a bit more rearrangement but first we're going to postulate that the population will always be positive and so we can remove the modulus about $P(t)$. The equation we have now rearranges to:

$$P(t) = e^{kt+C} = e^C e^{kt} \quad (304)$$

Now e^C is also just a constant, so let's rename this whole thing in a rather suggestive way. Let's call this the constant P_0 :

$$P(t) = P_0 e^{kt} \quad (305)$$

remember, P_0 is just a constant, and k is also a constant which we have to give because it tells us about the average rate of baby production per member of the population. This type of behaviour is called exponential growth and models many things, from populations of bacteria, to compound interest, to radioactive decay.

What does P_0 represent? Well, let's put in $t = 0$ into the equation and we get:

$$P(0) = P_0 \quad (306)$$

P_0 is the size of the population at time $t = 0$. It is the initial condition of our population. Note that we are also free to choose when we start to measure time, but once we decide upon it, we must always measure time relative to that point. That point will be $t = 0$. Let's say that at the start of our experiment the population of bacteria in our dish is 1000, then $P_0 = 1000$. If we look at the population of people in South Africa, and we chose to start measuring from now, when the population is around 50 million, then $P_0 = 5 \times 10^7$. Note that there is something really really important that we've missed out. We are not taking into account that there is another way for the population to change than simply new people being born/bacteria being produced. That is death. This is a fairly important part of the way a population works, but for now we are leaving it out. We will see that because we have left it out, we will get some pretty crazy results, but we will at least get an idea of what a differential equation can give us.

Let's work our way through an example. Let's take rabbits. If you look it up online, you will find that for rabbits k (the rate of production of rabbits per rabbit) is somewhere around 13 per year. So we have:

$$P(t) = P_0 e^{13t} \quad (307)$$

Let's say we start with 2 rabbits. What happens after a year? Well, according to our calculations this will give us:

$$P(1year) = 2e^{13 \times 1} = 884827 \quad (308)$$

That's already a lot of rabbits, but let's see what happens after 10 years...

$$P(10years) = 2e^{13 \times 10} \approx 5.7 \times 10^{56} \quad (309)$$

damn, that's a lot of rabbits!!! How much rabbit is that? Well, let's say that a rabbit is about 2 kilos. This is 10^{57} kilos of rabbit. The sun weighs around 10^{30} kilos, so this is around 10^{27} suns, which is more than there are stars in all the galaxies in the observable universe. After 10 years we already have a universe of rabbits. Let's keep going with this calculation, just because it's fun. Let's say that our population of rabbits which is increasing rapidly in size is in the form of a giant ball of rabbits, all reproducing and making more rabbits (it'd be strange if they started making lemurs). Let's also say that you can fit roughly 100 rabbits into a cubic meter. This is probably an overestimate in normal conditions, but in the centre of the sphere of rabbits the pressure will be pretty high, so this is probably a vast underestimate. If that's the case then the volume of the population is $\frac{P(t)}{100} m^3$. What's the radius of this sphere of rabbits? $V_{sphere} = \frac{4}{3}\pi r^3$ so: $r = \left(\frac{3V_{sphere}}{4\pi}\right)^{\frac{1}{3}}$ so the radius is:

$$r_{\text{sphere of rabbits}}(t) = \left(\frac{3\frac{P(t)}{100}}{4\pi}\right)^{\frac{1}{3}} \quad (310)$$

let's plug in our known rabbit population growth and we have that the radius at time t is given by:

$$r_{\text{sphere of rabbits}}(t) = \left(\frac{3\frac{2e^{13t}}{100}}{4\pi}\right)^{\frac{1}{3}} = \left(\frac{3 \times 2}{100 \times 4\pi}\right)^{\frac{1}{3}} e^{\frac{13t}{3}} \quad (311)$$

assuming that we start with 2 rabbits at time $t = 0$. And what's the rate of change of the radius over time? Well, we can think of this as the velocity of the surface of the sphere:

$$velocity_{\text{surface of sphere of rabbits}} = \frac{dr(t)}{dt} = \left(\frac{3 \times 2}{100 \times 4\pi}\right)^{\frac{1}{3}} \frac{13}{3} e^{\frac{13t}{3}} \quad (312)$$

The speed of light is roughly 9.5×10^{15} meters per year. So we can ask when the surface of our shell of rabbits starts expanding faster than the speed of light:

$$9.5 \times 10^{15} = \frac{dr(t)}{dt} = \left(\frac{3 \times 2}{100 \times 4\pi}\right)^{\frac{1}{3}} \frac{13}{3} e^{\frac{13t}{3}} \quad (313)$$

Solving this we get that this happens at roughly $t = 8.5$ years. If we took an idealised rabbit population that reproduced in proportion to its population then after 8.5 years this ball of rabbits would be expanding faster than the speed of light. In fact we've not taken a few crucial things into account, including the fact that the pressure of the inner part of the sphere of rabbits would start nuclear fusion and our sphere would become star, which might make reproduction hard, and even if it didn't, the sphere would have become a black hole far before its outer limb accelerated past the speed of light.

Phew! ok, that was all very silly but quite fun and we have learnt several things:

- We have solved our first differential equation
- The assumptions that went into our initial differential equation probably weren't very realistic.
- If you ever find perfectly idealised rabbits, don't let them breed.
- Never let a mathematical physicist with an overactive imagination set your tut questions.

What we did learn that is very important is that starting off with knowledge of the rate of change of a quantity, we were able to find out how it increased over time.

Let's try and make this example at least a little more realistic. Let's suppose that the environment only supports a fixed number of rabbits, let's call that fixed number M . This is the maximum number of rabbits that we can stably have. It turns out that there is a very very important equation which will model this sort of behaviour very well, and it shows up all over the place. This is called the logistic equation and looks like:

$$\frac{dP(t)}{dt} = kP(t) \left(1 - \frac{P(t)}{M} \right) \quad (314)$$

We have pulled this equation out of thin air, so rather than deriving where it comes from, we will simply motivate that it seems to have the right sort of behaviour of what we want.

For very small populations (much less than the stable equilibrium population M) $\frac{P(t)}{M} \ll 1$, so the term in brackets can be safely ignored. For very small population then the equation looks like our original equation which gave us uncontrolled exponential growth. However, this is only true for small populations, so this means that when we start out, we'll get exponential growth, but after a while as $P(t)$ stops being much less than M we will have to take

into account the term in brackets and this will save us from the ball of rabbits problem that we encountered before.

What will happen then when $P(t)$ is a little bit less than M . We see that the term in brackets is small but positive. This means that the rate of change of the population will be small and positive. As we get closer and closer to $P(t) = M$ we will find that the rate of change slows to zero. For populations just above M we will find that the rate of change of the population is actually negative. This means that the population is getting less, and as we get closer to M from above, the rate of change slows to zero again. How about for very large populations? Well, in this case $\frac{P(t)}{dt} \sim -\frac{kP(t)^2}{M}$ and so the population is going to go down in proportion to the square of the population - ie. it's going to go down very fast for large populations. This seems a bit more controlled. We can see that we are never going to get the growing sphere of rabbits problem (or GSOR problem). There is some control going on. For large populations, the population will get smaller until we get closer and closer to the equilibrium population M and for small population it will grow and slowly asymptote to M . In fact it turns out that the solution to this equation is given by:

$$P(t) = \frac{Me^{CM+kt}}{e^{CM+kt} - 1} \quad (315)$$

where again C is our integration constant. We are not going to show how to find this, but we are going to show that this is indeed a solution.

Note that while this is a good solution to the Logistic equation it is not the solution you will often find in textbooks. The reason for this is that C here can actually be complex which makes things a bit more tricky. Normally, you will be given the solution:

$$P(t) = \frac{MCe^{kt}}{Ce^{kt} + 1} \quad (316)$$

where $A = \frac{P_0}{M-P_0}$. In what follows we will use the above, somewhat less recognised form but you should be able to work out a solution given the initial condition either way.

How can we check that a given function is a solution to the equation of motion? We simply plug it in, just as we would with a normal equation.

Let's look at the left hand side:

$$\begin{aligned}\frac{dP(t)}{dt} &= \frac{d\left(\frac{Me^{CM+kt}}{e^{CM+kt}-1}\right)}{dt} \\ &= \frac{kMe^{CM+kt}}{e^{CM+kt}-1} - \frac{kMe^{2CM+2kt}}{(e^{CM+kt}-1)^2} = -\frac{kMe^{CM+kt}}{(e^{CM+kt}-1)^2}\end{aligned}\quad (317)$$

How about the right hand side?

$$\begin{aligned}kP(t)\left(1 - \frac{P(t)}{M}\right) &= k\left(\frac{Me^{CM+kt}}{e^{CM+kt}-1}\right)\left(1 - \frac{\left(\frac{Me^{CM+kt}}{e^{CM+kt}-1}\right)}{M}\right) \\ &= -\frac{kMe^{CM+kt}}{(e^{CM+kt}-1)^2}\end{aligned}\quad (318)$$

so the left hand side and the right hand side of the equations are equal - well, that sounds like this function satisfies the equation of motion then! We still have the integration constant and it would be nice to be able to write this in terms of an initial population again. We will start by redefining e^{CM} as \tilde{C} . C is a constant and so we can rewrite it in whatever way is most convenient for us:

$$P(t) = \frac{\tilde{C}Me^{kt}}{\tilde{C}e^{kt}-1}\quad (319)$$

Now define P_0 as the population at $t = 0$ and plug into the solution $P(t = 0) = P_0$ and $t = 0$. This gives:

$$P_0 = \frac{\tilde{C}M}{\tilde{C}-1}\quad (320)$$

Now we can solve this equation for \tilde{C} to get:

$$\tilde{C} = \frac{P_0}{P_0 - M}\quad (321)$$

Now we can reinsert this into our solution to get:

$$P(t) = \frac{\frac{P_0}{P_0-M}Me^{kt}}{\frac{P_0}{P_0-M}e^{kt}-1} = \frac{MP_0e^{kt}}{P_0(e^{kt}-1)+M}\quad (322)$$

Check that at $t = 0$ $P(t) = P_0$. Now we have everything we need. I can give you a value for k (population growth for a small population) and M , a maximum sustainable population, and a P_0 the initial population, and you

can work out how this population will vary over time. Let's go back to our rabbits and say that $k = 13$ and say that in a given environment 10,000 rabbits can be sustained. Let's look at what happens over time as we start with different numbers of rabbits. The graph in figure 69 shows what happens when you start with 2, 100, 1000, 5000, 200000 and 500000 rabbits. You can

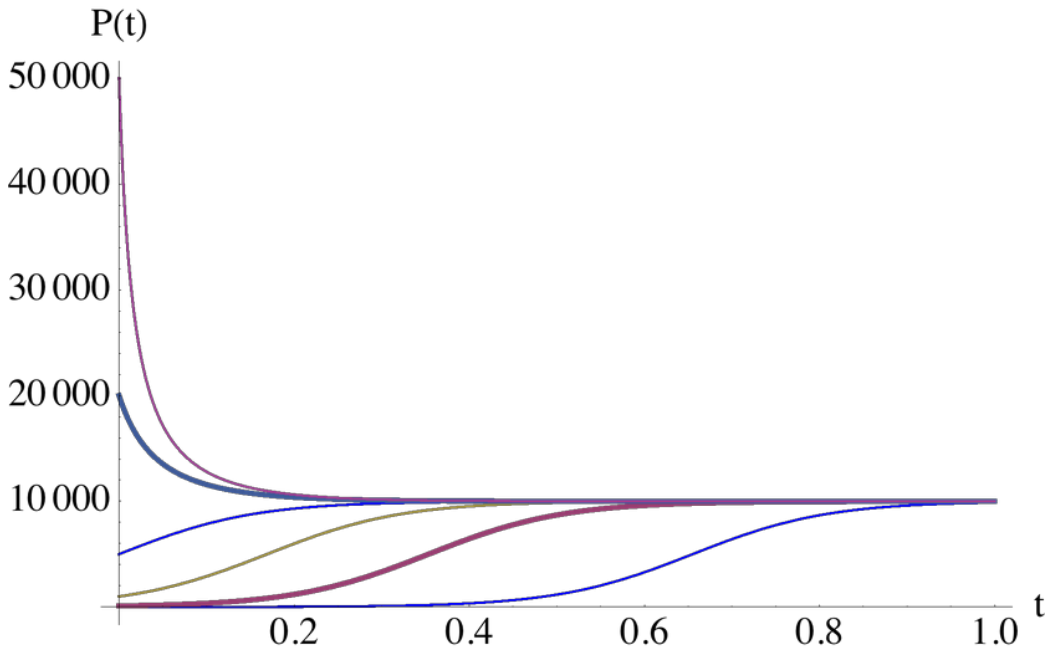


Figure 69: Solutions to the logistic equation modelling population growth of a rabbit population where the equilibrium population (ie. the population that the environment can sustainably support) is 10,000 rabbits, with a growth rate of 13 rabbits per rabbit per year and various starting populations (values of P_0).

see something remarkable happening here (not that remarkable because it is exactly what we had predicted, but a bit remarkable because we're about to remark on it!). Independent of the initial population (except if $P_0 = 0$, if that's the case then you will always have a population of zero - spontaneous rabbit production is not built into our equations), over time the population will always tend to the value that the environment can sustain (10,000). This isn't all that surprising because we built the logistic equation to have precisely this behaviour, but it's nice to see that indeed it gives the expected result when we plot the solutions.

Incidentally, the solution to the logistic equation is called the sigmoidal function and shows up all over the place.

What you should know by now about differential equations so far:

- The basic definition of a differential equation
- That differential equations are a powerful way of modelling the world
- That sometimes the simplest equation you might guess for a given situation needs to be modified to get realistic behaviour
- How to take a solution to a differential equation and show that it is indeed a solution
- That the integration constants we found before are really related to the freedom to set the initial conditions in a system described by a differential equation

Incidentally, a system where you are given a differential equation and an initial condition (eg. population at some particular time) is called an initial value problem. Note also that I could have given you the population at time $t = 0.5$ and you would also have been able to calculate the solution. You would have plugged in this population and $t = 0.5$ and this would have fixed \tilde{C} for us. In fact, let's take an example. Let's say that at time $t = 0.5$ the population is 2000. Then:

$$2000 = \frac{\tilde{C}M e^{0.5k}}{\tilde{C}e^{k0.5} - 1} \quad (323)$$

gives:

$$\tilde{C} = \frac{1.0005e^{-0.5k}}{M} \quad (324)$$

This can then be plugged back into the equation and you can work out the population for any time in the past or the future of that point.

11.1 Direction Fields and Euler's method

We haven't yet studied any general ways to solve differential equations. In the first case of exponential growth we found an easy way to solve the equation, but for the logistic equation we just gave the solution and showed that it indeed satisfied the equation. Here we are going to look at some methods for finding not the exact solution, but approximations of the solutions. The first method is the method of Direction Fields and it will give us a good idea of what the solutions are going to look like. The second method, Euler's method will give us an approximation to a single solution and we will be able to improve it to get arbitrarily good solutions to any differential equation (so long as there aren't particularly nasty pathologies in the differential equation).

11.1.1 Direction Fields

Let's take a differential equation:

$$\frac{dy}{dx} = x + y \quad (325)$$

We know then that given a point x and y we can find the derivative there. Let's say we pick the point $(3, 2)$. At this point, a function which solves the differential equation and goes through there will have gradient $3 + 2 = 5$. We can think about drawing in a tiny bit of the function, let's say a short line segment there with gradient 5. We know that close to the point $(3, 2)$ the function will still have gradient roughly 5, so drawing in a straight line (ie. one where the gradient is constant) is going to be a reasonable approximation to whatever the function will really do there. Of course at $(3.1, 2.1)$ the gradient will be a little more than 5, but that is the approximation we're making.

We can go through a whole array of points in the (x, y) plane and work out what the gradients of lines going through them will be. What we will be doing is actually mapping out all of the solutions to the differential equation. Given an initial condition, of course we won't go through all points in the plane, but we are now mapping out the flow of all possible initial conditions. Let's do this for a grid of in the (x, y) plane and work out the derivative of

the function at those points:

$y \ x :$	-3	-2	-1	0	1	2	3
-3	-6	-5	-4	-3	-2	-1	0
-2	-5	-4	-3	-2	-1	0	1
-1	-4	-3	-2	-1	0	1	2
0	-3	-2	-1	0	1	2	3
1	-2	-1	0	1	2	3	4
2	-1	0	1	2	3	4	5
3	0	1	2	3	4	5	6

(326)

We can then draw a plot in the (x, y) plane of small lines, with these gradients, at these points. This will look like figure 70. If we want to see this in more

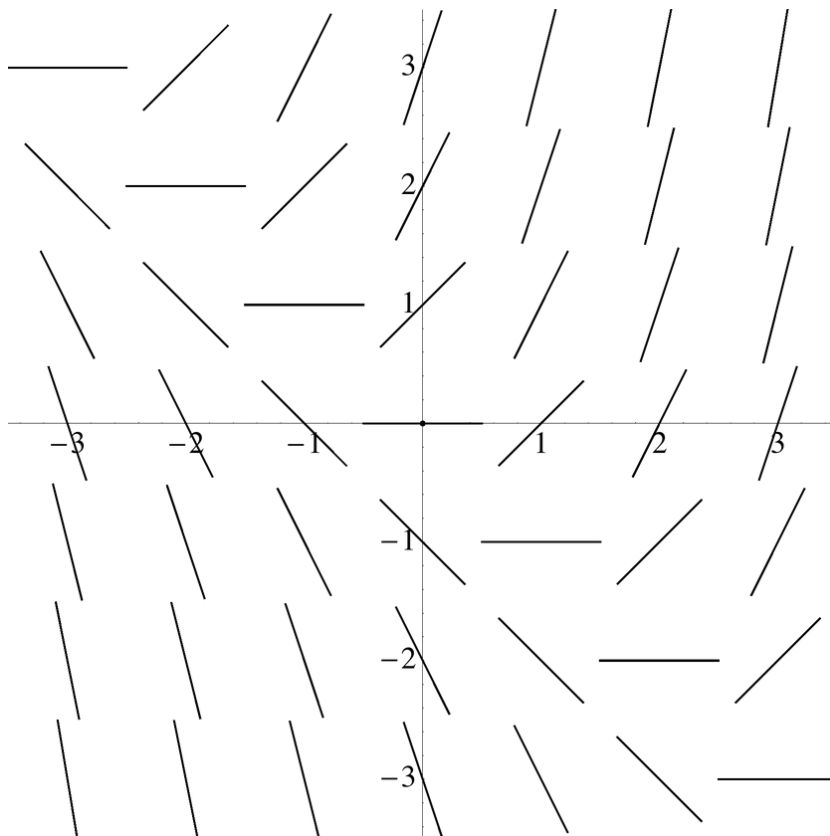


Figure 70: Direction fields for the equation $\frac{dy}{dx} = x + y$ for points at integer values of x and y .

detail then we can look at a larger number of grid points and we will find, for instance, the direction field in figure 71. You should get a general sense

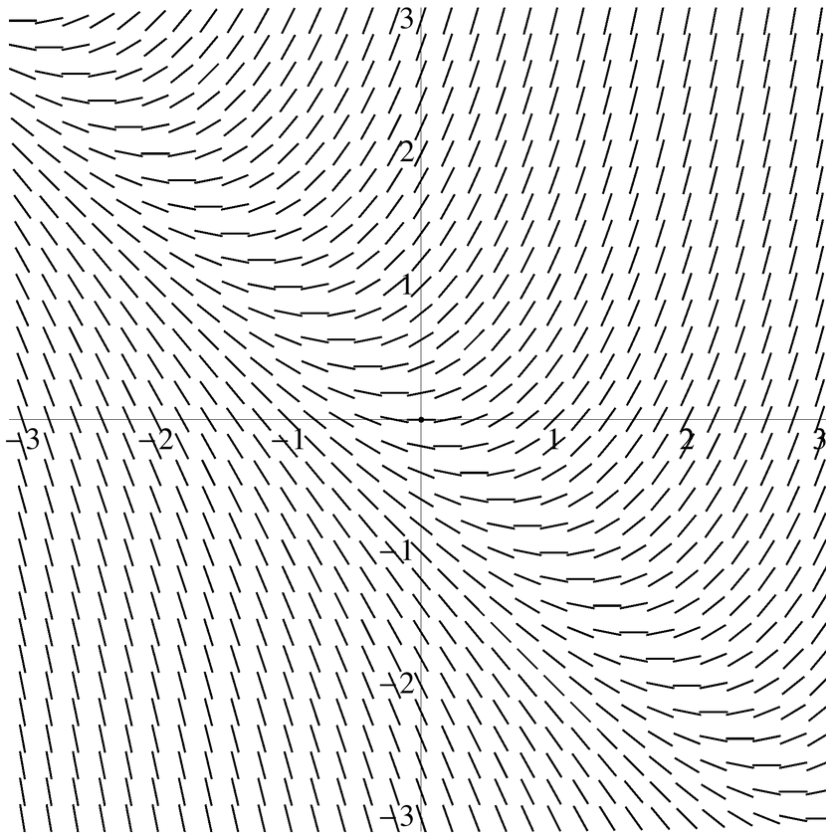


Figure 71: Direction fields for the equation $\frac{dy}{dx} = x + y$ for points at smaller intervals in the x and y directions than in figure 70.

of the "flow" in these diagrams. You can imagine that a particular function is going to follow along a set of flow lines. In fact, the solution of this equation is known and is $y(x) = -1 - x + Ce^x$. Four different solutions with different initial conditions (ie. different values of c) are shown in figure 72 on top of the direction fields. You can thus see that given any first order differential equation, you can figure out the direction field by simply plugging in the x and y values. This will give the general behaviour of all solutions of the differential equation.

11.1.2 Euler's method

The method of Direction Fields was a useful way to get the general behaviour of all solutions to a differential equation. Euler's method is going to allow us to find an approximation of a particular solution given an initial condition.

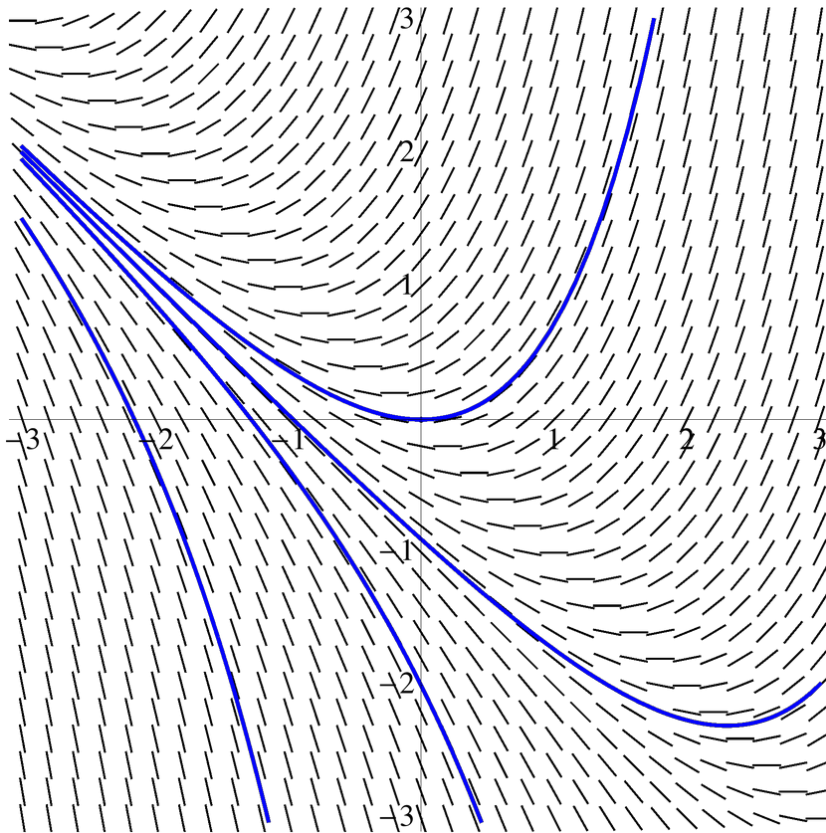


Figure 72: Four particular solutions to the differential equation with four different initial conditions plotted on top of the direction field. You can see that the direction field gives the flows of all possible solutions to the equation. Starting at a particular point (an initial condition) allows for the behaviour of that solution to be understood.

Let's take the same equation that we had in the previous example:

$$\frac{dy}{dx} = x + y \quad (327)$$

but now we are going to ask what the solution which has initial condition $(0, -0.5)$ (ie. at $x = 0$, $y = -0.5$) looks like. We know that close to this point, the gradient of the graph is going to be -0.5 (ie. $0 + (-0.5)$), so we can draw a short line at that point with gradient -0.5 . Let's draw a line of extension 1 in the x direction from the point $(0, -0.5)$ with gradient -0.5 . This is going to take us to the point $(1, -1)$ At this point we again look at the differential equation and it tells us that the gradient at that point $(1, -1)$ should be 0 (ie. $1 + (-1)$), so we draw a line with gradient 0 another unit

in the x direction, which takes us to the point $(2, -1)$. Again, we ask what the gradient should be at this point, and clearly it should be 1. We draw in a line of gradient 1 of length 1 which will take us to the point $(2, 0)$. We continue this again and again until we have got as far in the x direction as we want. Of course we chose to go 1 unit in the x direction each time but in fact this is a very crude approximation. We know that really the gradient will be changing continuously but we are letting it change only discretely. We could shorten the lines that we use, say to 0.5 long, or 0.1 long, or 0.01 long and each time it would give us a better and better approximation of the real solution to the differential equation. Various length sections for Euler's method are shown in figure 73

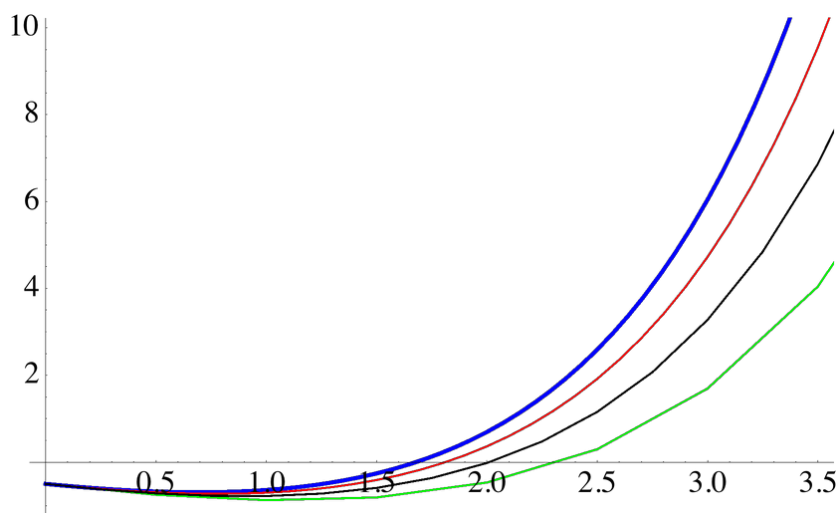


Figure 73: The blue curve shows the exact solution to the differential equation in eq.327. The green curve shows the approximation found with Euler's method when the step size is 1, the black is when the step size is 0.5 and the red is when the step size is 0.1. As the step size decreases we get closer and closer to the exact solution.

Please, if you read the part on Euler's method and it doesn't make sense, please email me and I'll write in some more details

11.2 Separable differential equations

In some ways these are the easiest differential equations to solve in theory, though in practice the final step (that of integrating) may be difficult or impossible. A separable differential equation is one of the form:

$$\frac{dy}{dx} = \frac{f(x)}{g(y)} \quad (328)$$

where $f(x)$ and $g(y)$ are any functions of x and y respectively. For instance:

$$\frac{dy}{dx} = xy \quad (329)$$

is of this form where $f(x) = x$ and $f(y) = \frac{1}{y}$. The reason that these equations are simple in theory is because we can rearrange them to be:

$$g(y)dy = f(x)dx \quad (330)$$

ie. we have all the x stuff on one side and all the y stuff on the other and then we can integrate both sides:

$$\int g(y)dy = \int f(x)dx \quad (331)$$

and that's it. As long as you can do the integrals, you can get a function y in terms of x . let's look at some examples:

$$\frac{dy}{dx} = xy \quad (332)$$

gives the following integral:

$$\int \frac{1}{y} dy = \int x dx \quad (333)$$

and so:

$$\ln |y| + c_1 = \frac{x^2}{2} + c_2 \quad (334)$$

here we have one constant of integration from each side of the interval, but because they are just constants, we can put them into one constant and call it c :

$$\ln |y| = \frac{x^2}{2} + c \quad (335)$$

we can then rearrange this to give:

$$|y| = e^{\frac{x^2}{2} + c} = e^c e^{\frac{x^2}{2}} \quad (336)$$

We can then call e^c just a constant, and let's call it y_0 because we can see that when x is zero, $|y|$ is just going to be given by this constant:

$$|y| = y_0 e^{\frac{x^2}{2}} \quad (337)$$

This has two solutions (one where y is positive, and one where it is negative), so we can choose one of them, depending on the initial condition for y . If we are told that for $x = 0$ $y = 3$ then:

$$y = 3e^{\frac{x^2}{2}} \quad (338)$$

If we are told that for $x = 0$, $y = -2$ then:

$$y = -2e^{\frac{x^2}{2}} \quad (339)$$

so we have found the general set of solutions (and here, two particular solutions) to the initial differential equation.

Given that you can already integrate, basically that is all you need to know about solving separable differential equations, but let's take another couple of examples:

$$\frac{dy}{dt} = \frac{4 \sin 2t}{y} \quad (340)$$

given the initial condition that at $t = 0$, $y = 1$, ie. $y(0) = 1$. We rearrange the equation to get the integral:

$$\int y dy = \int 4 \sin 2t dt \quad (341)$$

This gives (taking the two integration constants and writing them as one:

$$\frac{y^2}{2} = -2 \cos 2t + c \quad (342)$$

so $y = \sqrt{c - 4 \cos 2t}$, where we have absorbed a factor of 2 into c . Now we want to find the value of c given our initial condition, so we set $t = 0$ and $y = 1$:

$$1 = \sqrt{c - 4} \quad (343)$$

so $c = 5$ and so the particular solution is:

$$y = \sqrt{5 - 4 \cos 2t} \quad (344)$$

In figure 74 we have plotted this solution, along with a number of solutions with other initial conditions. figure 73 Lets look at another example:

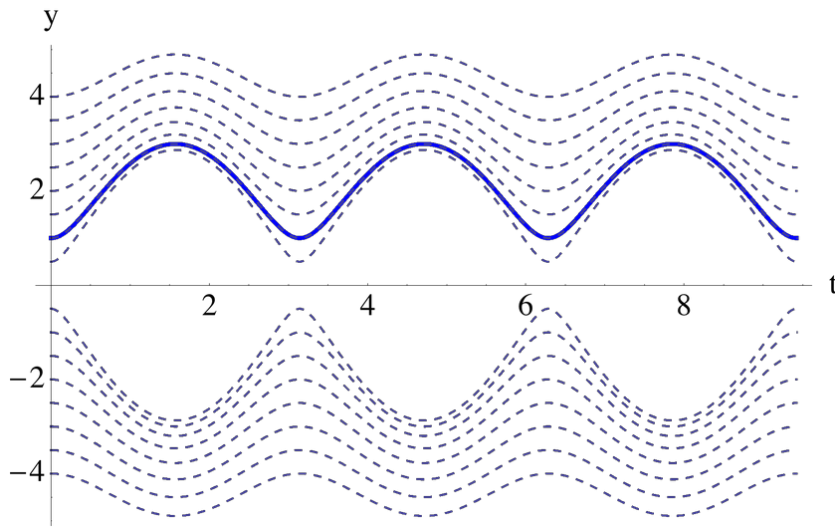


Figure 74: Solutions to the differential equation $\frac{dy}{dt} = \frac{4 \sin 2t}{y}$ with different initial conditions (dashed lines) and with the particular initial condition $y(0) = 1$ in the thick blue line.

$$\frac{dy}{dx} = e^{-y}(2x - 4) \quad (345)$$

with initial condition $y(5) = 0$. Again, we rearrange the equation to be an integral:

$$\int e^y dy = 2x - 4dx \quad (346)$$

and so $e^y = x^2 - 4x + c$ giving the solution:

$$y = \ln(x^2 - 4x + c) \quad (347)$$

we plug in $x = 5$ and $y = 0$ to get:

$$0 = \ln(5^2 - 4 \times 5 + c) \quad (348)$$

which is solved by $c = -4$. So the particular solution that we are interested in is:

$$y = \ln(x^2 - 4x - 4) \quad (349)$$

Note that this only makes sense for $x^2 - 4x - 4 > 0$ which has two solutions: $2 + 2\sqrt{2} < x$ and $x < 2 - 2\sqrt{2}$. Only the first of these contains the point $x = 5$ and so this is the branch that we are interested in, so the solution is given above and is valid for $2 + 2\sqrt{2} < x$. The particular solution along with several other solutions are shown in figure 75.

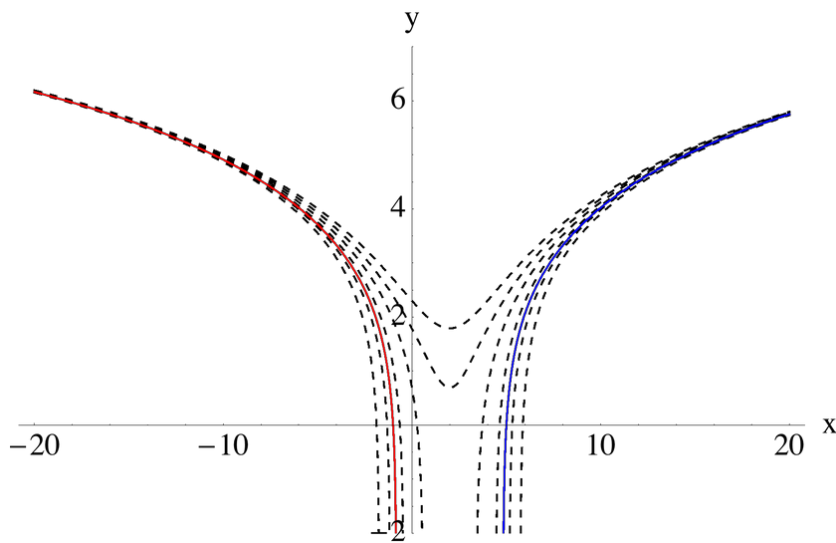


Figure 75: Solutions to the differential equation $\frac{dy}{dt} = e^{-y}(2x-4)$ with different initial conditions (dashed lines) and with the particular initial condition $y(5) = 0$ in the thick blue line. The thick red line has the same value of c as the particular solution we are interested in, but it is disjoint from that solution as it never reaches $x = 5$.

How does the integration tie in with Euler's method. Well, in Euler's method we are slowly adding up more and more change as we move along the x direction and it is telling us the total y that we have got to at any one value of x . Integration is just adding up and so Euler's method is just the adding up of the little bits in discrete steps rather than as a continual function as in an interval.

11.2.1 A way of thinking about these equation

Differential equations seem a bit mysterious at first, but we really can take them as extensions of algebraic equations. When you are given the equation:

$$x^2 = x + 4 \tag{350}$$

You are being asked to find that number (or those numbers) whose square is equal to 4 plus itself. Similarly when you are given then differential equation:

$$\frac{dy}{dx} = x^2 + y \tag{351}$$

You are being asked to find the family of functions y whose gradients are equal to their value plus the square of the x value at any one point. You

should be able to go anywhere on one of these curves and the relationship between x , y and the gradient of the line at that point should satisfy the equation. How do you know that it does? Well, just plug in the function you believe to be a solution into the differential equation and if it does indeed give you zero then you've found a solution.

11.3 First order linear differential equations

We are now going to deal with another subset of first order differential equations which in some ways are easier than the previous and in other ways more complicated. These are linear first order differential equations. The general form of a first order linear differential equation is:

$$\frac{dy}{dx} + P(x)y = Q(x) \quad (352)$$

where $P(x)$ and $Q(x)$ are any functions of x . Sometimes you will be given an equation which is not obviously in this form but it can be transformed to this form. For instance:

$$\frac{1}{y} \frac{dy}{dx} = x^2 + \frac{\sin x}{y} \quad (353)$$

This can easily be transformed into the canonical form for a linear first order DE. While in its present form, it is not clear how to solve this, there is another equation which can be solved very easily. This is:

$$\frac{d(yR(x))}{dx} = S(x) \quad (354)$$

Note that we leave the explicit x dependence off y here but you should always bear in mind that y is a function of x . If we were keeping everything explicit this would be $\frac{d(y(x)R(x))}{dx} = S(x)$. This is easy to solve because we can rearrange the d terms to get:

$$d(yR(x)) = S(x)dx \quad (355)$$

and then integrate both sides:

$$\int d(yR(x)) = \int S(x)dx \quad (356)$$

The left hand side looks rather confusing, but we can simply relabel $yR(x) = u$ and then we have $\int du$ which we know to be $u + c = yR(x) + c$ so in fact the left hand side is really trivial. The solution to this is then:

$$y = \frac{1}{R(x)} \int S(x)dx \quad (357)$$

so as long as we can integrate $S(x)$ we can find the solution to this equation.

Why does this help us? We seem to have solved a completely unrelated differential equation. However, the point is that if we can transform the

original differential equation in eq 352 into the form in eq 354 then we know how to solve this.

This means that somehow we want to take:

$$\frac{dy}{dx} + P(x)y = Q(x) \quad (358)$$

and transform it to look like a single derivative on the left hand side. Let's try and transform this into the form of eq 354 but with the right hand side being $S(x) = R(x)Q(x)$. This is equivalent to trying to transform our equation into:

$$\frac{1}{R(x)} \frac{d(yR(x))}{dx} = Q(x) \quad (359)$$

How can we do this? Well, let's take the left hand side of eq 359 and use the product rule on the derivative. It gives:

$$\frac{1}{R(x)} \frac{d(yR(x))}{dx} = \frac{dy}{dx} + \frac{y}{R(x)} \frac{dR(x)}{dx} \quad (360)$$

Which means that $\frac{1}{R(x)} \frac{d(yR(x))}{dx}$ is the same as $\frac{dy}{dx} + yP(x)$ as long as $P(x) = \frac{1}{R(x)} \frac{dR(x)}{dx}$. So let's see if we can find what $R(x)$ has to be for this to be true. We can use a rather nice trick which is that:

$$\frac{1}{R(x)} \frac{dR(x)}{dx} = \frac{d(\ln R(x))}{dx} \quad (361)$$

so we have $\frac{d(\ln R(x))}{dx} = P(x)$. Let's bring up the dx from the bottom and integrate, to give:

$$\ln R(x) = \int P(x)dx \quad (362)$$

and thus:

$$R(x) = e^{\int P(x)dx} \quad (363)$$

This looks pretty strange, but in fact it's fine. If we can integrate $P(x)$ then we can certainly take the exponential of the integral. So, we are claiming that:

$$\frac{1}{R(x)} \frac{d(yR(x))}{dx} = \frac{dy}{dx} + yP(x) \quad (364)$$

if $R(x) = e^{\int P(x)dx}$. We also know that if $\frac{1}{R(x)} \frac{d(yR(x))}{dx} = Q(x)$ then the solution is:

$$y = \frac{1}{R(x)} \int R(x)Q(x)dx \quad (365)$$

So we have done it...without knowing what we were doing...We can plug back in what $R(x)$ has to be in terms of $P(x)$ and we have found the solution to the equation. The complete solution is then:

$$y = \frac{1}{e^{\int P(x)dx}} \int e^{\int P(x)dx} Q(x) dx \quad (366)$$

which might look horrible, but there's no part of that which you can't do...as long as the integrals themselves aren't too complicated.

Let's look at an example of this where we will follow the same reasoning but it will all be a bit more clear when we have specific P 's and Q 's and R 's. Let's take a really simple example:

$$\frac{dy}{dx} + \frac{y}{x} = x \quad (367)$$

so now we have $P(x) = \frac{1}{x}$ and $Q(x) = x$. We thus have that $R(x) = e^{\int P(x)dx} = e^{\ln x + c} = e^c e^{\ln x} = e^c x$. We are missing off the absolute value because in fact the complex part of $\ln x$ for negative x can be absorbed into the c . So apparently the left hand side of the equation should be the same as:

$$\frac{1}{R(x)} \frac{d(yR(x))}{dx} \quad (368)$$

Check this by plugging in $R(x)$:

$$\frac{1}{e^c x} \frac{d(ye^c x)}{dx} = \frac{1}{x} \frac{d(yx)}{dx} = \frac{dy}{dx} + \frac{y}{x} \quad (369)$$

which is indeed the left hand side of the original equation. So now we have that:

$$\frac{1}{x} \frac{d(yx)}{dx} = x \quad (370)$$

which is just $\frac{d(yx)}{dx} x = x^2$. We can then solve this to get:

$$\begin{aligned} yx &= \frac{x^3}{3} + c \\ y &= \frac{x^2}{3} + \frac{c}{x} \end{aligned} \quad (371)$$

Check that this is a solution by plugging it into the original differential equation that we were trying to solve. You'll see that it is a general solution to this equation. As always c can be fixed to a particular value by being given an initial condition.

When you are faced with an equation which is linear and first order there are thus two ways to go about solving it. You can manually try and put the left hand side into a form which is a single derivative on a combined function, or you can just take the values of $P(x)$ and $Q(x)$ and plug them into the general formula for the solution. It's up to you, but it's definitely a good idea to run through examples in the former way first in order to understand what's going on under the bonnet of this formalism.

Let's take another example. This time it corresponds to a falling body accelerating under gravity but with air resistance:

$$\frac{dv}{dt} = 9.8 - 0.196v \quad (372)$$

The 9.8 is the acceleration due to gravity. The 0.196 is a number which represents the amount of air resistance, but in fact it's chosen here to make the numbers come out nicer in the end. Let's start by putting it into the canonical form for a first order linear DE:

$$\frac{dv}{dt} + 0.196v = 9.8 \quad (373)$$

The left hand side can then be rewritten to give:

$$\frac{1}{e^{0.196t}} \frac{d(v e^{0.196t})}{dt} = 9.8 \quad (374)$$

Now we rearrange, and integrate to give:

$$\int d(v e^{0.196t}) = \int 9.8 e^{0.196t} dt \quad (375)$$

which gives, after integration:

$$v e^{0.196t} = 50 e^{0.196t} + c \quad (376)$$

or:

$$v = 50 + c e^{-0.196t} \quad (377)$$

If we state that at $t = 0$ the velocity is 0 then we get:

$$0 = 50 + c e^0 \quad (378)$$

so $c = -50$:

$$v = 50 - 50 e^{-0.196t} \quad (379)$$

which corresponds to a body accelerating under gravity (like being dropped out of a plane) and the acceleration being slowed by air resistance until it reaches a terminal velocity of 50ms^{-1} .

Let's just make sure that we could have used our formalism to get the answer straight away without the juggling we did of the left hand side: Calculating:

$$y = \frac{1}{e^{\int P(t)dt}} \int e^{\int P(t)dt} Q(t) dt \quad (380)$$

with $P(t) = 0.196$ and $Q(t) = t$ you'll find that you do indeed get the same solution we found above.

Let's look at another example:

$$y' - \frac{y}{2} = 4 \sin(3t) \quad (381)$$

I'm jumping about between different forms of notation because in general you will come across all of these and it's useful for you to be able to translate between them. The dash is just the usual derivative and we can see here that y is supposed to be a function of a variable t . It should always be clear what the functional dependence will be just by looking at the equation.

Try running through this example like we did with the last one by getting the left hand side into a term which has a single derivative on y times some function of t . We will however simply take $P(t)$ and $Q(t)$ and plug it into the formalism. here $P(t) = -\frac{1}{2}$ and $Q(t) = 4 \sin(3t)$ so $e^{\int P(t)} = e^{-\frac{t}{2}+c_1}$ and thus:

$$y = \frac{1}{e^{-\frac{t}{2}+c_1}} \int 4 \sin(3t) e^{-\frac{t}{2}+c_1} dt = -\frac{8}{37} (6 \cos(3t) - \sin(3t)) + ce^{\frac{t}{2}} \quad (382)$$

A reiteration

I think that it's important to understand how we're solving these first order linear differential equations. The idea is to make a rearrangement of them into a form which is very easily to solve. In order to do that we have to find a special function $R(x)$ which allows for this transformation, but once we've found it and made the rearrangement, we can solve the equation with ease (modulo being able to do the integrals).

We said at the very beginning of this section that in some ways these equations are more complicated than the separable equations we first looked at, and in other ways are more complicated. The method to solve them was certainly more involved than when we solved the separable equations but it turns out that the linearity of these equations imbues them with some very special properties. In general solutions to linear equations can be combined in ways which make finding the combined effects of different solutions very

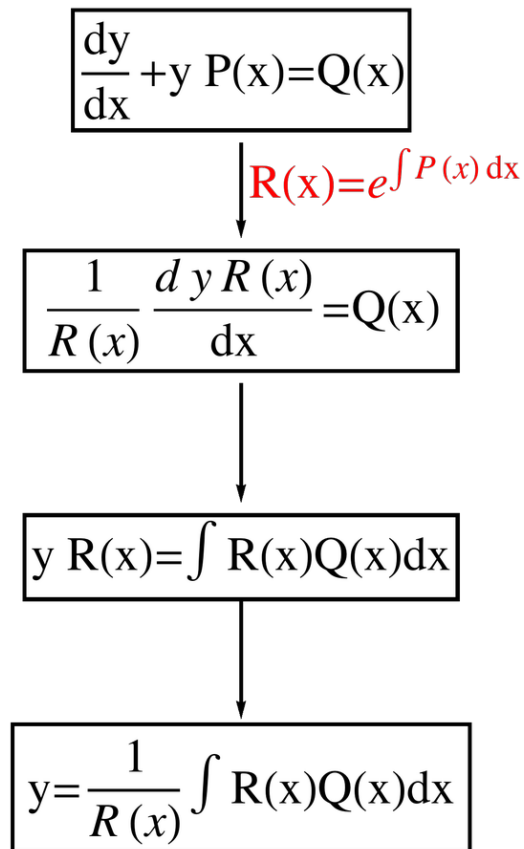


Figure 76: A flowchart showing the formulation of the general solution to a linear first order differential equation.

easy. We will not have a chance to go into this, but suffice it to say that in general, finding solutions to non-linear equations can be incredibly complicated, but they can have some amazing behaviour.

11.4 Second Order differential equations

We are only going to look at a particular subset of all possible second order differential equations (that is, equations which contain at most second derivatives) but these particular equations are absolutely ubiquitous across every field of science. The particular subset we are going to look at are linear, homogenous second order differential equations with constant coefficients. These can be written in general as:

$$\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0 \quad (383)$$

It is linear because it contains at most (and in this case at least) a single power of y in each term. It is homogenous because there is no term which has no powers of y (ie. the right hand side is not a constant), and the coefficients b and c are any real numbers (though you can extend this to having complex numbers very easily). We will see that depending on the relationship between these numbers (b and c) we can have very different behaviour of the equation. Let's first of all rearrange this equation and see what it says:

$$\frac{d^2y}{dx^2} = -b\frac{dy}{dx} - cy \quad (384)$$

This says that we are looking for a function whose second derivative at any point is linked only to its first derivative and value. In fact we can guess a solution to this equation because we already know a function whose derivatives look very much like the function without any derivatives. These are exponential functions. Let's take the original equation and plug in a trial solution. Let's guess that the solution might look something like $y = e^{rx}$ for some value of r . First we plug this into the original equation and see if this is a solution:

$$\frac{d^2e^{rx}}{dx^2} + b\frac{de^{rx}}{dx} + ce^{rx} = 0 \quad (385)$$

This can be rewritten as:

$$e^{rx}(r^2 + rb + c) = 0 \quad (386)$$

$r^2 + rb + c$ is known as the characteristic equation for the differential equation. Check that you understand where this comes from. We hoped to find a solution to the equation, but clearly this will only be a solution if either $e^{rx} = 0$ or $(r^2 + rb + c) = 0$. The former can't be true for all x except if $r = -\infty$ and that's not a very interesting solution. The latter however is a good constraint and says that:

$$r = \frac{-b \pm \sqrt{b^2 - 4c}}{2} \quad (387)$$

This means that the solution is:

$$y = e^{\frac{-1 \pm \sqrt{b^2 - 4c}}{2}x} \quad (388)$$

Check that this is indeed a solution. In fact we can see that in general there will be two solutions (except when $b^2 = 4c$). Let's take an example equation. Let's look at:

$$\frac{d^2y}{dx^2} - 2\frac{dy}{dx} - 3y = 0 \quad (389)$$

Check that the solutions to this equation are $y = e^{3x}$ and $y = e^{-x}$.

Now the linearity of the equation comes into effect. If $y = e^{3x}$ is a solution then so is $y = ke^{3x}$ for any k . This wouldn't be true if the equation wasn't linear or homogeneous.

Theorem: If $y_1 = e^{r_1x}$ and $y_2 = e^{r_2x}$ are two solutions to a differential equation $\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$, then so is the combination $y = k_1y_1 + k_2y_2$ for any constants k_1 and k_2 .

This means that if we find two solutions to such a differential equation we can combine them in a linear fashion with whatever coefficients we want. In fact the two coefficients will turn out to be the free constants which we can fix with two initial conditions. For all of the differential equations that we are dealing with, the number of boundary conditions are equal to the order of the differential equation. A second order differential equation is fixed with two initial conditions.

A good example of a second order differential equation is the equation which governs projectile motion, like throwing a ball. A second order linear differential equation will tell you the trajectory of the ball after you let go. However, in order to predict the trajectory, you need to know two things. The initial position, and the initial velocity of the projectile.

In general you will fix the two constants for a second order differential equation by being given an initial value for y and also an initial gradient of y . Or you can fix the two constants by being given the value of y at two different values of x . Having the two constants k_1 and k_2 will be enough to fix everything. However, before we understand about completely general solutions and what we need for that, we have to understand about independent solutions.

Two solutions to a differential equation are said to be independent if one of them is not a constant multiple of the other. ie. $y = 3e^x$ and $y = e^x$ are not independent but $y = 3e^x$ and $y = 5e^{2x}$ are independent.

It is a general theorem of these differential equations that ALL solutions to the equation can be written in the form $y = k_1y_1 + k_2y_2$ where y_1 and

y_2 are two independent solutions. It doesn't matter which two independent solutions, you can always add together two independent solutions in two different ways to get two more independent solutions. For instance, let's say we have found two independent solutions y_1 and y_2 . Then $y_3 = y_1 + y_2$ and $y_4 = y_1 - y_2$ are also independent solutions. Then any solution can be written as $y = k_1y_1 + k_2y_2$ or $y = k_3y_3 + k_4y_4$ for particular choices of k_1 and k_2 or k_3 and k_4 .

Now, this means that as long as we can find two independent solutions, then we have found all of the solutions. This is the same statement that for a first order linear equation, as soon as we've found the general solution which includes a single integration constant, then we've found all of the solutions. It turns out that there are three different possible scenarios, depending on the particular constant coefficients in the original differential equation. We will deal with them all separately.

We are interested in finding the general solutions to the differential equation:

$$y'' + by' + cy = 0 \tag{390}$$

The three scenarios are:

1. $b^2 - 4c > 0$

Then there are two solutions to the characteristic equation, call them r_1 and r_2 , and the general solution to this equation is:

$$y = k_1e^{r_1x} + k_2e^{r_2x} \tag{391}$$

where k_1 and k_2 are fixed by initial conditions.

2. $b^2 - 4c = 0$

Then there is a single solution to the characteristic equation, call it r . It turns out however, that in this case, and this case only, there is another type of solution to this differential equation, which is of the form xe^{rx} . Thus, the general solution to this class of equations is:

$$y = k_1e^{rx} + k_2xe^{rx} \tag{392}$$

3. $b^2 - 4c < 0$

This is perhaps the most interesting class of solutions. Again, there are two solutions, but now they will be complex. In fact they will be

complex conjugates of one another. Let's call them $r_1 = \alpha + i\beta$ and $r_2 = \alpha - i\beta$. So, we can write the two general solution as:

$$\begin{aligned} y &= k_1 e^{(\alpha+i\beta)x} + k_2 e^{(\alpha-i\beta)x} \\ &= e^{\alpha x} (k_1 e^{(i\beta)x} + k_2 e^{(-i\beta)x}) \\ &= e^{\alpha x} (k_1 (\cos \beta + i \sin \beta x) + k_2 (\cos \beta - i \sin \beta x)) \end{aligned} \tag{393}$$

Where we chose to write the complex exponential in sin and cos form. If we chose k_1 and k_2 each to be $\frac{1}{2}$ then we can see that there is a solution which is $y = e^{\alpha x} \cos \beta x$. If we let $k_1 = \frac{1}{2i}$ and $k_2 = -\frac{1}{2i}$ then there is also a solution which is $y = e^{\alpha x} \sin \beta x$. These are two good, independent solutions to the original equation (you can check this) and thus we can write down the general solution with two new integration constants:

$$y = k_3 e^{\alpha x} \cos \beta x + k_4 e^{\alpha x} \sin \beta x \tag{394}$$

Thus the solution to this particular form is an exponential growth or decay (depending on α multiplied by periodic functions, the frequency of which is determined by β).

Let's look at an example of each of these three classes of solution:

1.

$$y'' - 2y' - y = 0 \tag{395}$$

Thus $b^2 - 4c = 5$ and so we are in the first class. The solutions to the characteristic equation are $r = 2 \pm \sqrt{5}$ and so the general solution to this equation is:

$$y = k_1 e^{(2+\sqrt{5})x} + k_2 e^{(2-\sqrt{5})x} \tag{396}$$

If we are given the initial condition $y(1) = 2$ and $y'(1) = -3$ then we can simply write:

$$2 = k_1 e^{(2+\sqrt{5})1} + k_2 e^{(2-\sqrt{5})1} \tag{397}$$

and, after taking the derivative of the solution:

$$-3 = k_1 (2 + \sqrt{5}) e^{(2+\sqrt{5})1} + k_2 (2 - \sqrt{5}) e^{(2-\sqrt{5})1} \tag{398}$$

These two equations can be solved simultaneously and give the coefficients:

$$\begin{aligned} k_1 &= \frac{1}{4} (4 - 5\sqrt{2}) e^{-\sqrt{2}-1} \\ k_2 &= \frac{1}{4} (4 + 5\sqrt{2}) e^{\sqrt{2}-1} \end{aligned} \tag{399}$$

2.

$$y'' - 2y' + y = 0 \quad (400)$$

With the same boundary conditions as in the previous example. Show that the following is the solution:

$$y = k_1 e^x + k_2 x e^x \quad (401)$$

with $k_1 = 7e^{-1}$ and $k_2 = -5e^{-1}$

3.

$$y'' + y' + y = 0 \quad (402)$$

Thus the solutions to the characteristic equation are $r = \frac{-1 \pm i\sqrt{3}}{2}$. This means that in our previous formulation, $\alpha = -\frac{1}{2}$ and $\beta = \frac{\sqrt{3}}{2}$. So the general solution is:

$$y = k_1 e^{-\frac{1}{2}x} \cos \frac{\sqrt{3}}{2}x + k_2 e^{-\frac{1}{2}x} \sin \frac{\sqrt{3}}{2}x \quad (403)$$

If we set the boundary conditions as $y(0) = 3$ and $y'(0) = -2$ then the particular solution is:

$$\frac{1}{3} e^{-x/2} \left(9 \cos \left(\frac{\sqrt{3}x}{2} \right) - \sqrt{3} \sin \left(\frac{\sqrt{3}x}{2} \right) \right) \quad (404)$$

In figure 77 we plot a few different particular solutions to this equation with different boundary conditions.

4. Finally we will consider an equation which only has imaginary solutions to the characteristic equation. For instance:

$$y'' + y = 0 \quad (405)$$

The characteristic equation has $r = \pm i$ and therefore the solution has $\alpha = 0$ and $\beta = 1$ giving:

$$y = k_1 \cos x + k_2 \sin x \quad (406)$$

Check that indeed this is the solution to this equation. This is a very important equation because it produces purely periodic behaviour.

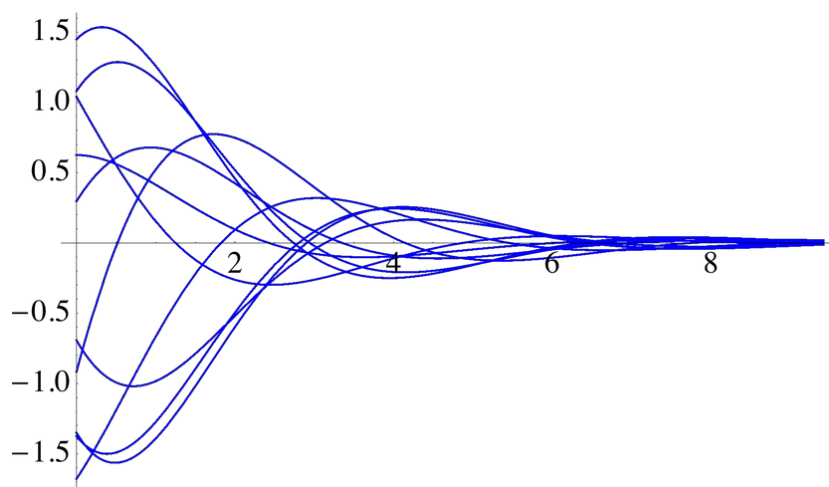


Figure 77: Solutions to the equation $y'' + y' + y = 0$ with different boundary conditions. All of the solutions are trigonometric functions multiplied by decaying exponentials but the initial position and initial gradient affect the overall solution.

11.5 Summary

We've now taken a tour through some of the most important differential equations that we can solve at this level. We have built up a lot of powerful machinery to tackle many problems and being able to manipulate such systems will give you control over how you model the real world. You should now:

- Understand what a differential equation is and understand how it can reflect physical aspects describing a particular system.
- Know how to solve a separable differential equation to get the general solution (one with a single integration constant).
- Know how to fix integration constants when you are given initial conditions.
- Know how to solve first order linear differential equation.
- Know how to solve second order linear differential equations with constant coefficients.
- Know how to plot Direction fields given a first order differential equation.

- Know how to use Euler's method given a first order differential equation.

12 3D geometry and vectors

A lot of the following is going to be rather intuitively clear, but we need to build up a framework where we are all speaking the same language to develop the powerful tools that we are going to find over the coming sections. We will be dealing here specifically with three dimensional space but we will discuss along the way the extension of these concepts to higher dimensional spaces. The higher dimensional stuff is not examinable but I think that sometimes it helps to understand the things which are special about three dimensions, and the things which are not.

In particular, I can recommend having a look at the web page of John Baez who discusses the regular polytopes in different numbers of dimensions here: <http://math.ucr.edu/home/baez/platonic.html>

It's clear that to define where you are in three dimensional space you need to set up a few key ingredients first. What you need is first of all an origin. You will use this to define all relative positions. Everything will be defined in terms of distances and directions from the origin. The other thing that you need to specify are the axes of your space. Generally we will call them the x , y and z axes. The important point about them is that we want them to be orthogonal - ie. all at right angles to each other. Clearly in four you would need four axes, and in 11 dimensions you would need 11 orthogonal axes.

In three dimensions the three axes define for us three orthogonal planes, which we denote the xy -plane, the xz -plane and the yz -plane. You might have thought that if you have three of these in three dimensions then in some other number d of dimensions you will have d planes, but this isn't the case. To define a plane we need to pick two axes. In three dimensions this is choosing two directions from three, or ${}_3C_2$, which is three. In four dimensions you pick two directions from four. ie. ${}_4C_2 = 6$. If we label the axes in four dimensions as x, y, z, r for instance, then we have the xy , the xz , the xr , the yz , the yr and the zr planes. In 11 dimensions there are ${}_{11}C_2 = 55$ planes. We're not going to enumerate them! In fact there is a simpler way of writing ${}_dC_2$ which is that this is always going to be given by $\frac{d(d-1)}{2}$.

In two dimensions there are said to be four quadrants. where both the x and y values are positive, two when one is positive and one negative and one when they are both negative. In three dimensions there are eight of them. We can denote them by $+++$, $++-$, $+ - +$, $+ - -$, $- + +$, $- + -$, $- - +$, $- - -$. We call these octants. This means that in each sector we have to choose either positive or negative for each of the directions. This gives 2^d sectors in d dimensions.

12.1 Positions in three dimensions

A point in three dimensions is denoted $P(a, b, c)$ and is a distance a from the (y, z) plane, a distance b from the (x, z) plane and a distance c from the (x, y) plane (Note that sometimes we will call it the (x, y) plane and sometimes the xy plane. Learn to go seamlessly between these notations. We can see from figure 78 that the point defines for us a rectangle with the

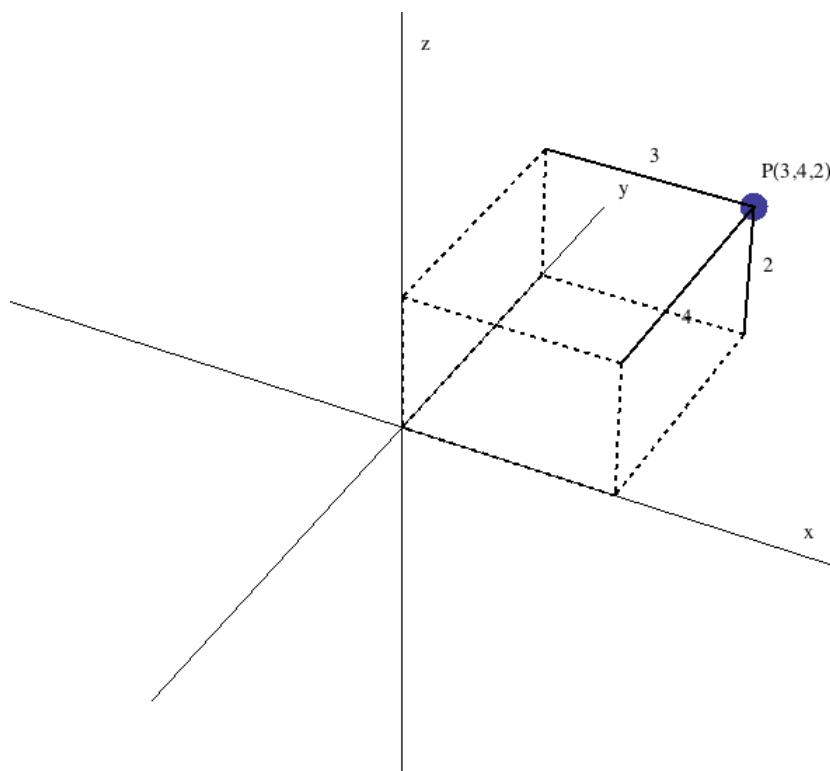


Figure 78: A point $P(3, 4, 2)$ and its distances from the various coordinate planes.

origin as the extreme vertex. In particular we can project the point onto the coordinate planes. If we project the point $P(3, 4, 2)$ onto the xy coordinate plane, this is like shining a light from above and seeing where the shadow of the original point lies on the xy plane. The answer is another point (call it Q) which is at $Q(3, 4, 0)$. Similarly we can project onto the other planes and we get $R(3, 0, 2)$ and $S(0, 4, 2)$. These can similarly be projected onto the coordinate axes. If we take Q and project it onto the x axis we get a point at $(3, 0, 0)$, etc.

Projections can be very useful for looking at straight lines in 3 dimen-

sions. If you are given the 3 points $(3, 7, -2), (2, 4, 2)$ and $(1, 3, 3)$ and asked whether they form a straight line, then it's not clear by drawing them as a two dimensional picture (which all pictures are), whether or not they form a straight line. We also don't quite know yet how to find gradients in higher dimensions. However, one thing we can be sure of is that a straight line in 3 dimensions will still look like a straight line when we project it onto the coordinate planes. In order to check that it does form a straight line we'll have to project it onto all three planes. First, let's project it onto the xy -plane. The points project to: $(3, 7, 0), (2, 4, 0)$ and $(1, 3, 0)$. Now we can calculate the gradient between the first and the second point (call it g_1) and between the second and the third point (call it g_2): $g_1 = \frac{4-7}{2-3} = \frac{3}{1} = 3$ and $g_2 = \frac{3-4}{1-2} = 1$. So the gradients are different and thus this can't be a straight line. If the projection isn't a straight line then the original line can't be straight either. Had we found that this projection was a straight line then we still would have had to check the other projections to see whether they gave straight lines. Only if all three projections give straight lines can we be sure that the line in three dimensions is straight.

12.2 Functions in three dimensions

A point in three dimensional space is given by three real numbers, thus the set of all points in three dimensional space is $\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3 = \{(x, y, z) | x, y, z \in \mathbb{R}\}$. In two dimensions, a curve is a set of points in \mathbb{R}^2 . In 3D, an equation involving x, y and z is a surface in \mathbb{R}^3 . For instance, in 2D we could look at the set:

$$\{(x, y) | y = x^2\} \quad (407)$$

or

$$\{(x, y) | y = 3\} \quad (408)$$

The first gives the set of points along the curve $y = x^2$ and the second gives the set of points on the straight line at $y = 3$. How about in three dimensions? The set of points described by:

$$\{(x, y, z) | z = 3\} \quad (409)$$

Involves all points such that $z = 3$, so for instance $P(4, 15, 3)$ is such a point, and $P(-12, 2, 3)$ is such a point. In fact all points, independent of the x and y values but with $z = 3$ satisfy the requirements to be in this set. This means that the set of points is just the flat surface at $z = 3$. This is shown in figure 79 Let's look at a slightly less trivial function. How about:

$$\{(x, y, z) | x^2 + y^2 = 1\} \quad (410)$$

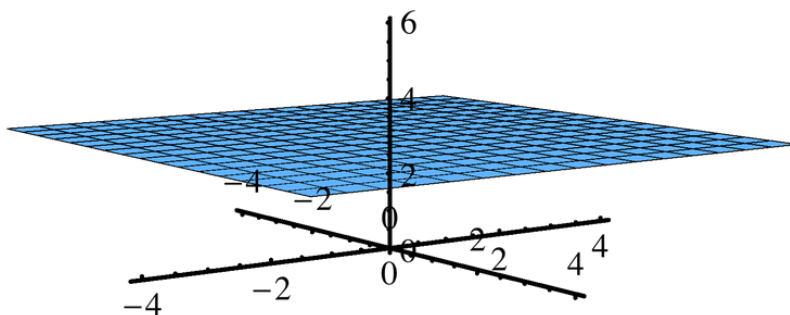


Figure 79: The function $z = 3$ plotted in three dimensions.

Of course we can just write this as "the surface in three dimensions specified by the function $x^2 + y^2 = 1$ " rather than writing it in the language of the set of points, but they are equivalent. This surface is shown in figure 80. We can see that because z is not constrained, this is just a cylinder. ie. a circle defined at all values of z . We could ask for an intersection of two surfaces, for instance $z = 3$ and $x^2 + y^2 = 1$ would give us a horizontal circle at height $z = 3$. The function $x^2 + y^2 + z^2 = 1$ is of course a sphere of radius 1. We can define a sphere of any radius, centred at any point (x_0, y_0, z_0) with: $(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2$.

Note that any function which is at most linear in the coordinate directions will be a flat surface. Anything which is at least quadratic will not be flat.

The definition of regions in three dimensions is also relatively intuitive. For instance: the region defined by $1 \leq x^2 + y^2 + z^2 \leq 4$ and $z \leq 0$ is the lower half of the region in between the spheres of radius 1 and 2 centred on the origin.

12.3 Distances in three dimensions

We can ask about distances in three dimensions as a simple extension of two dimensions. The distance between two points at (a_x, a_y, a_z) and (b_x, b_y, b_z) is just $\sqrt{(a_x - b_x)^2 + (a_y - b_y)^2 + (a_z - b_z)^2}$ as you would expect. You can see this by looking at the distances of the projections of these points onto the coordinate axes.

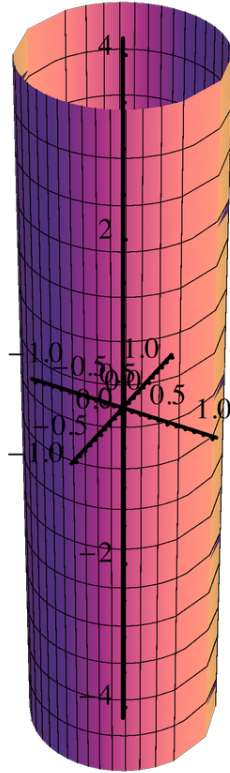


Figure 80: The function $x^2 + y^2 = 1$ plotted in three dimensions.

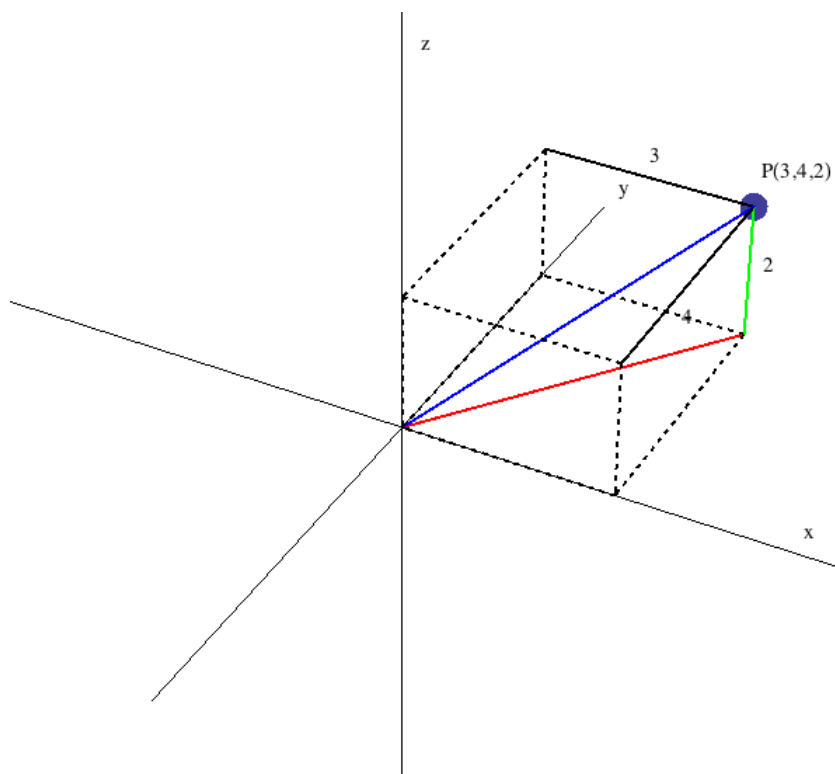


Figure 81: The distance from the origin to the point $P(3, 4, 2)$ is given by using Pythagoras with the red, green and blue lines. ie. $distance^2 = (\sqrt{3^2 + 4^2})^2 + (2)^2$.

13 Vectors

Vectors are quantities which have both magnitude (ie. size) and direction. The most common examples of these are velocity ($3ms^{-1}$ to the right) and force (10 Newtons pointing vertically down). The easiest way to describe such a quantity is an arrow, where the magnitude gives the length and the direction is given by, well, the direction of the arrow. The important point about this is that the position of the vector itself doesn't matter. In figure 82 we place the same arrow in several different places and they are all the same vector.

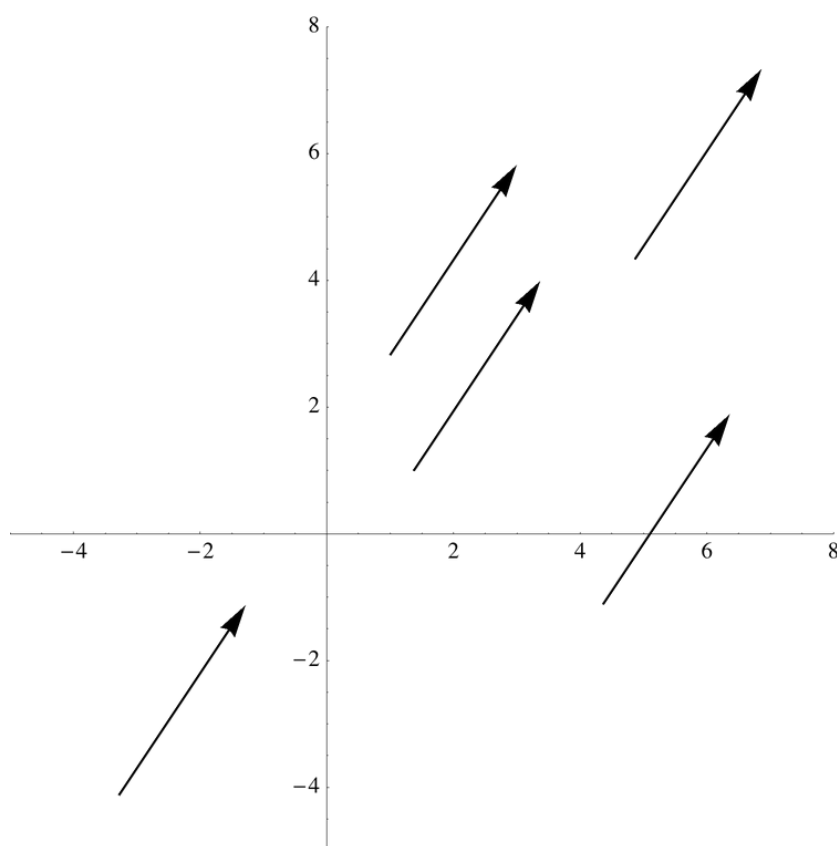


Figure 82: A vector, with magnitude given by its length and direction given by the direction of the arrow, placed at different points in the plane. Note that the position of the start of the arrow is now important, just the relationship between the start and the end of the arrow.

We could define a vector by the length and the angle that it makes with the horizontal axis, but in general we define it by how much it goes in the

horizontal direction and how much it goes in the vertical direction, that is, how much it goes in the x -direction and how much it goes in the y -direction. You can see that in figure 82 the vectors all go along 2 units in the x -direction and up 3 units in the y -direction. We write this vector (which we will call \vec{v}) as $\vec{v} = \langle 2, 3 \rangle$. The numbers 2 and 3 are said to be the components of the vector \vec{v} . All of the arrows in the figure correspond to precisely this vector. The starting position of the vector is unimportant. All that is important is how much it goes along and how much it goes up. This can easily be extended to higher dimensions. We can think of the amount that an arrow goes along in the x , y and z directions and this would be labelled by three numbers, for instance $\langle 6, -2, 1 \rangle$ where the -2 just means that it goes backwards two units in the y direction.

The magnitude of a vector is just the size, without the direction, of the vector, so the magnitude of a vector is a scalar quantity. It is just the same as the Pythagorean length of the vector. So in the case of the vector in figure 82 the magnitude is $\sqrt{2^2 + 3^2} = \sqrt{13}$. We denote the magnitude of a vector as $|\vec{v}|$ and so $|\vec{v}| = \sqrt{13}$. The magnitude is always equal to the square root of the sum of the squared components of the vector. I could write a vector, \vec{p} , in 7 dimensions, for instance (not that I could picture it) as $\vec{p} = \langle 1, 6, -2, 5, 2, 7, -5 \rangle$ and the magnitude would be: $\sqrt{1^2 + 6^2 + (-2)^2 + 5^2 + 2^2 + 7^2 + (-5)^2} = 12$.

Now that we can write down a vector using the angle brackets we can ask what happens when we add together two vectors. We can think of adding together their descriptions. For instance, if we have a vector which is $\vec{v} = \langle 3, 4 \rangle$ and one which was $\vec{w} = \langle 5, 7 \rangle$, then $\vec{v} + \vec{w}$ corresponds to "go 3 to the right and 4 up, then 5 to the right, then 7 up", which is the same as 8 to the right, then 11 up which we can write as: $\vec{v} + \vec{w} = \langle 8, 11 \rangle$. So, adding together vectors is done by simply adding them component by component. If $\vec{v} = \langle v_1, v_2, \dots, v_m \rangle$ and $\vec{w} = \langle w_1, w_2, \dots, w_m \rangle$ are vectors in an m -dimensional space where v_i and w_i , $i = 1 \dots m$ are just numbers, ie. the components of the vectors then:

$$\vec{v} + \vec{w} = \langle v_1 + w_1, v_2 + w_2, \dots, v_m + w_m \rangle \quad (411)$$

It is clear then that if this is how we add vectors together, that it doesn't matter which one comes first, and so $\vec{v} + \vec{w} = \vec{w} + \vec{v}$. We can also view this pictorially in figure 83

It turns out that if we know how to add together two vectors then we can work out how to multiply a vector by a scalar. The answer is very simple but we can think about it first just in terms of adding a vector to itself, ie. multiplying it by the scalar 2. If $\vec{v} = \langle v_1, v_2 \rangle$ then $2\vec{v} = \vec{v} + \vec{v} = \langle 2v_1, 2v_2 \rangle$.

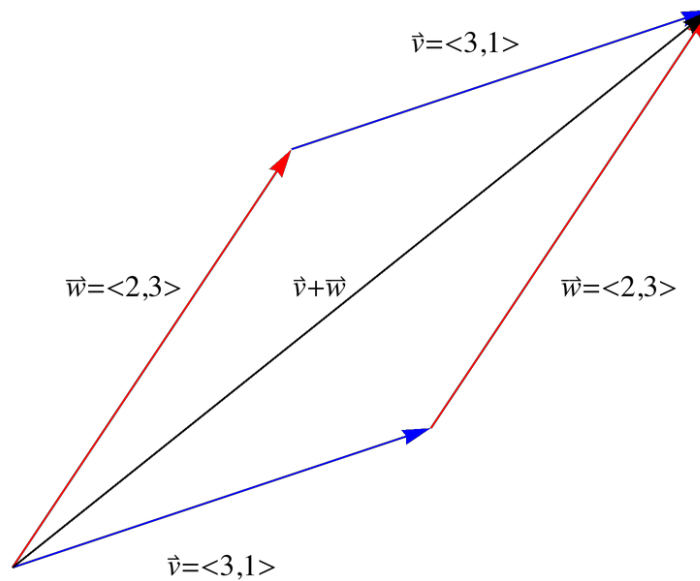


Figure 83: The addition of two vectors. We don't need to put axes here because the position of the vectors are unimportant. When we add them together we can first move along \vec{v} and then \vec{w} or vice versa. The black line is the addition of the two which gives $\vec{v} + \vec{w} = \langle 2 + 3, 3 + 1 \rangle = \langle 5, 4 \rangle$

In fact:

$$s\vec{v} = \langle sv_1, sv_2 \rangle \tag{412}$$

where s is any scalar. This includes negative numbers and so the negative of a vector is just a vector of the same length pointing in the opposite direction. Thus if you are asked to find the difference between two vectors, this is the same as adding one to the negative of the other. This is shown in figure 84

Keep this in mind at all times: A vector doesn't have a particular starting point, it is an arrow and can be moved around to any position. It is only the length and direction of the vector which is important.

So, we've shown how to find the length of a vector, and so it's very easy to create a unit vector in a given direction. Let's say that we have a vector $\vec{v} = \langle 5, 6, 7 \rangle$ and we want to find a unit vector in this direction. We can do this by simply dividing the vector by its magnitude, which in this case is $\sqrt{5^2 + 6^2 + 7^2} = \sqrt{110}$. So a unit vector in the same direction as \vec{v} is:

$$\vec{u} = \frac{1}{\sqrt{110}} \langle 5, 6, 7 \rangle \tag{413}$$

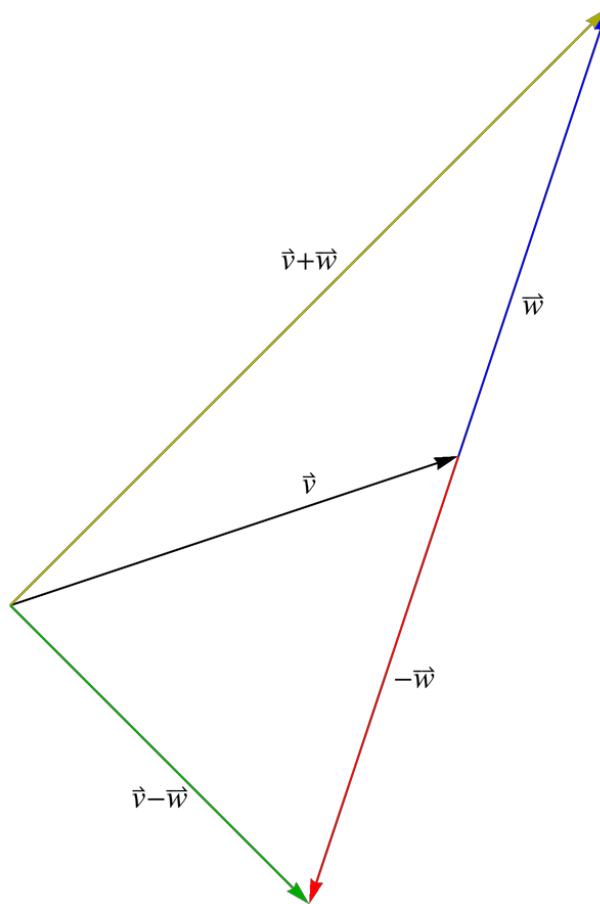


Figure 84: The addition and subtraction of vectors. Subtraction should be thought of as the addition of the negative of the vector. Remember: The position of the vector is unimportant. For addition you simply lay them end to end in whatever order you want.

Check for yourself that this new vector has magnitude 1. There is another unit vector with the same direction which is the vector pointing in the opposite direction - i.e. $-\frac{1}{\sqrt{110}} \langle 5, 6, 7 \rangle$. In general a unit vector in the direction of a vector \vec{v} is:

$$\vec{u} = \frac{\vec{v}}{|\vec{v}|} \tag{414}$$

You can show in general that this will have magnitude 1.

So far we have always used the angled bracket notation to write down our vectors but we can write them in another way. To do this we have to define what are called the standard basis vectors. They are normally given by the

vectors \vec{i} , \vec{j} and \vec{k} where:

$$\begin{aligned}\vec{i} &= \langle 1, 0, 0 \rangle \\ \vec{j} &= \langle 0, 1, 0 \rangle \\ \vec{k} &= \langle 0, 0, 1 \rangle\end{aligned}$$

Of course, if we are dealing with higher dimensions then we need to be a bit more creative with the names, but for now, this will do just fine. We can now write any vector as a sum of these 3 standard basis vectors. For instance:

$$\langle 1, -2, 6 \rangle = \vec{i} - 2\vec{j} + 6\vec{k} \quad (415)$$

and in general if we have a vector $\vec{a} = \langle a_1, a_2, a_3 \rangle$ then we can write this as: $\vec{a} = a_1\vec{i} + a_2\vec{j} + a_3\vec{k}$. As an exercise, we can also see that a unit vector in the direction of \vec{a} is:

$$\vec{u} = \frac{(a_1\vec{i} + a_2\vec{j} + a_3\vec{k})}{\sqrt{a_1^2 + a_2^2 + a_3^2}} \quad (416)$$

Of course this can be generalised to any number of dimensions.

13.1 The scalar, or dot product

We have seen now how to add together vectors and how to multiply them by scalars, but we haven't seen how to multiply two vectors together. In fact it's not all that obvious what it means to multiply two vectors together. A vector has a magnitude and a direction, how do you multiply directions? The answer is that there are two different ways to multiply together vectors. The first way which we will explore now is the scalar, or dot product. This will take two vectors and the product of them using this rule will give us a scalar. We definitely want something that is linear in both of the magnitudes of the vectors. That is to say that we want some way of multiplying together vectors so that when we double the magnitude of one of the vectors, we double the product. We will express the scalar or dot product of two vectors as: $\vec{v} \cdot \vec{w}$. For the above property to hold, we must have that:

$$\vec{v} \cdot \vec{w} \sim |\vec{v}| |\vec{w}| \quad (417)$$

but we don't know what the constant of proportionality must be like. Clearly this already gives us a scalar and when we double one of the magnitudes, the product doubles, as we would hope for a product, but it doesn't contain any more information than that. What we want is something that has some information also about the relative direction of the vectors. We can define a

scalar product which tells us how similar the two of them are by including the angle in between the two. If the angle in between the two vectors is θ then we can define a product like:

$$\vec{v} \cdot \vec{w} = |\vec{v}| |\vec{w}| \cos \theta \quad (418)$$

Such that when the two vectors are parallel to one another the dot product is just the product of their magnitudes, and when they are orthogonal to one another, the dot product is zero. Now this contains information both about their magnitudes and about their relative directions. This product is greatest when two vectors are pointing in the same direction and reduces to zero when they are at right angles to one another. But it seems like finding the angle between two vectors isn't going to be very easy. Is there an easier way to calculate this? We are going to use a triangle identity to find out how to calculate a dot product in component form. Given the vectors in figure 85

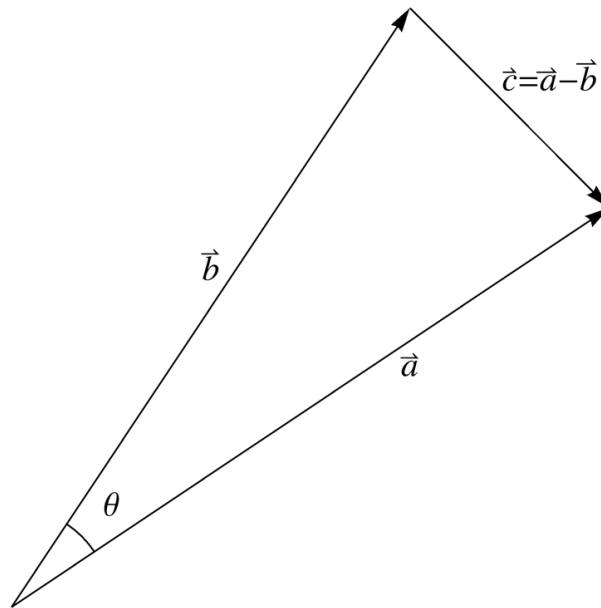


Figure 85: Three vectors forming a triangle with angle θ in between \vec{a} and \vec{b} .

We will use the identity relating the length of the sides of a triangle and one of the angles. The sides of the triangle are of length $|\vec{a}|$, $|\vec{b}|$ and $|\vec{a} - \vec{b}|$, thus we can write that:

$$|\vec{a} - \vec{b}|^2 = |\vec{a}|^2 + |\vec{b}|^2 - 2|\vec{a}| |\vec{b}| \cos \theta \quad (419)$$

If $\vec{a} = \langle a_1, a_2 \rangle$ and $\vec{b} = \langle b_1, b_2 \rangle$ then $\vec{a} - \vec{b} = \langle a_1 - b_1, a_2 - b_2 \rangle$. We know how to calculate the magnitude of a vector in component form, it's just the pythagorean length. We can also see that the term on the right of the equation is precisely the expression we wrote above for the dot product, so let's replace it with the dot product and see what we get:

$$|\vec{a} - \vec{b}|^2 = |\vec{a}|^2 + |\vec{b}|^2 - 2\vec{a} \cdot \vec{b} \quad (420)$$

Rearranging gives:

$$\vec{a} \cdot \vec{b} = -\frac{1}{2} \left(|\vec{a} - \vec{b}|^2 - |\vec{a}|^2 - |\vec{b}|^2 \right) \quad (421)$$

Now plugging in the component forms for the right hand side we get:

$$\vec{a} \cdot \vec{b} = -\frac{1}{2} \left((a_1 - b_1)^2 + (a_2 - b_2)^2 - (a_1^2 + a_2^2) - (b_1^2 + b_2^2) \right) \quad (422)$$

Which can be expanded out and simplifies to:

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 \quad (423)$$

This seems to be suggesting that to calculate the dot product as we defined it above, all we have to do is to multiply together the vectors component by component. Indeed you can show that this is true in any number of dimensions. If we have two n-dimensional vectors, then the dot product of them is:

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^n v_i w_i \quad (424)$$

For instance: $\vec{v} = \langle 4, 3, 2, 1 \rangle$, $\vec{w} = \langle 1, -2, 1, 1 \rangle$ then $\vec{v} \cdot \vec{w} = 4 \times 1 + 3 \times (-2) + 2 \times 1 + 1 \times 1 = 1$. Now we also have a very easy way to find the angle in between two vectors. We know that the dot product of these two vectors is 1 and we also know that the dot product is defined as the product of the magnitudes times the cosine of the angle in between them. This means that:

$$\cos \theta = \frac{\vec{v} \cdot \vec{w}}{|\vec{v}| |\vec{w}|} \quad (425)$$

which in this case gives: $\cos \theta = \frac{1}{\sqrt{30}\sqrt{7}}$.

We can calculate the dot product just as easily using the angle bracket notation as well as the component notation. For instance:

$$\begin{aligned} \langle -1, 7, 4 \rangle \cdot \left\langle 6, 2, -\frac{1}{2} \right\rangle &= -6 + 14 - 2 = 6 \\ (\vec{i} + 2\vec{j} - 3\vec{k}) \cdot (2\vec{j} - \vec{k}) &= 4 + 3 = 7 \end{aligned} \quad (426)$$

We do this knowing that the standard basis vectors are orthogonal to one another and have magnitude one so their dot products with themselves are 1 and with the other basis vectors are 0.

The dot product is a very good way to tell whether two vectors are orthogonal to one another. For instance the fact that:

$$\langle 3, 6 \rangle \cdot \langle 4, -2 \rangle = 12 - 12 = 0 \quad (427)$$

Tells us that these two vectors are perpendicular to one another. Sometimes we are asked to find a unit vector perpendicular to another vector. For instance, if we want to find a unit vector perpendicular to the vector $\langle 4, -3 \rangle$, then we want to find a $\langle a_1, a_2 \rangle$ such that $\sqrt{a_1^2 + a_2^2} = 1$ and that $\langle 4, -3 \rangle \cdot \langle a_1, a_2 \rangle = 0$. This is two equations and two constraints and there are two solutions to this: $\langle -\frac{3}{5}, -\frac{4}{5} \rangle$ and $\langle \frac{3}{5}, \frac{4}{5} \rangle$. Note that these are two unit vectors pointing in opposite directions. These two vectors and the original vector are shown in figure 86.

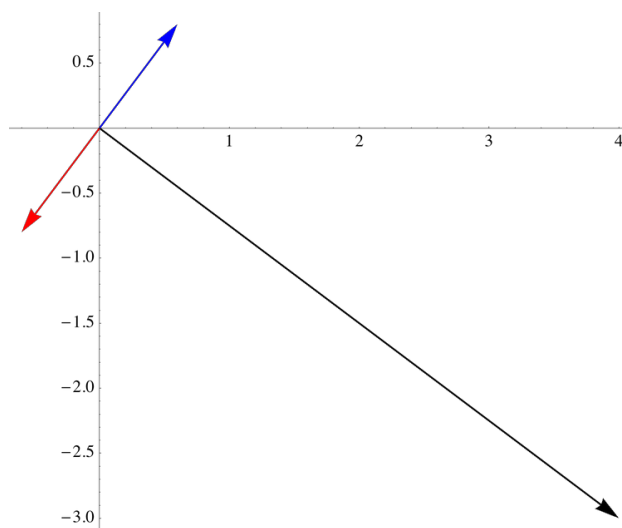


Figure 86: The vector $\langle 4, -3 \rangle$ in black and the two unit vectors orthogonal to it in red and black. Remember that the placement of the vectors is irrelevant, we've just put them here at the origin so that it's easier to see their lengths.

13.2 Scalar and vector projections

Given two vectors, can we ask how much of one vector is pointing in the direction of the other? We can certainly ask how much of the vector $\langle 5, 6 \rangle$

is pointing in the x direction - the answer is just 5. You can think of this as projecting the vector onto the x -axis and asking for its projected length. Similarly we can ask about the projection of a vector into any arbitrary direction. This is illustrated in figure 87. Imagine having a light perpendicular to \vec{b} shining towards it. There is a shadow of the vector \vec{a} cast on the line of \vec{b} . This is the scalar projection of \vec{a} in the direction of \vec{b} , also called the component of \vec{a} in the direction of \vec{b} . When you are looking at this, clearly the size of \vec{b} is unimportant, so you can think of an infinite line stretching in both directions parallel to \vec{b} .

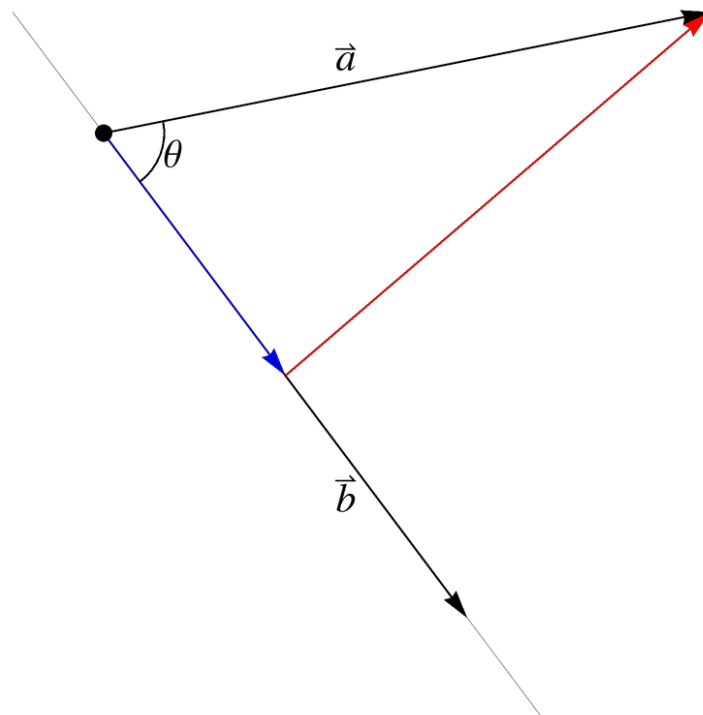


Figure 87: We can form the vector \vec{a} as the addition of the blue vector, and the red vector. The blue vector is the component of \vec{a} which is parallel to \vec{b} and the red vector is the component of \vec{a} which is perpendicular to \vec{b} . Because the length of the vector \vec{b} is unimportant to find the component of \vec{a} in the direction of \vec{b} we can simply think of the direction of \vec{b} as being defined by the infinite line (shown in light grey) as the extension of \vec{b} in both directions.

To calculate this quantity, which we call $comp_{\vec{b}}\vec{a}$ it is clearly just:

$$comp_{\vec{b}}\vec{a} = |\vec{a}| \cos \theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{b}|} \quad (428)$$

Check to make sure that you understand why the last equality is so. We can see that this therefore has nothing to do with the length of \vec{b} as we stated before.

Here we have just asked how much of \vec{a} is pointing in the direction of \vec{b} . We can also define a vector of this length, in the direction of \vec{b} . This is simply the quantity we already have multiplied by a unit vector in the direction of \vec{b} and thus given by:

$$proj_{\vec{b}}\vec{a} = \frac{\vec{a} \cdot \vec{b}}{|\vec{b}|} \frac{\vec{b}}{|\vec{b}|} = \frac{\vec{a} \cdot \vec{b}}{|\vec{b}|^2} \vec{b} \quad (429)$$

Hint: Make sure that you can derive these quantities (the comp and proj) yourself. Don't just remember the formulae but understand where they come from!

13.3 The vector, or cross product

When we took two vectors previously and found a way to multiply them together using the dot product, we ended up with a scalar. However, there is also a way that we can take two vectors and multiply them together to give a vector, but a vector with very specific properties with respect to the first two. What we will define here will be in three dimensions, and, unlike the dot product, does not generalise easily to other dimensions (though it can in fact be extended).

We are going to define the cross product such that it gives a vector which is perpendicular to the two vectors being crossed. This might sound a bit arbitrary but it shows up in a huge number of different situations in physics in particular and can help us to understand the geometric relation between vectors very simply. Clearly if two vectors are parallel to one another in two dimensions, there won't be a unique direction which is perpendicular to both vectors. In fact there will be a whole plane of directions perpendicular to the two vectors. We are going to define the cross product such that its magnitude is proportional to the magnitude of the two vectors being crossed, and the sin of the angle in between them. The direction of the resulting vector will be perpendicular to the two original vectors:

$$\vec{a} \times \vec{b} = (|\vec{a}| |\vec{b}| \sin \theta) \hat{n} \quad (430)$$

where \hat{n} is in the direction perpendicular to \vec{a} and \vec{b} . However, this is not uniquely defined. We can think of two vectors (pointing in opposite directions) which are both parallel to two other vectors and are unit magnitude. This is shown in figure 88 To disambiguate which of the two perpendicular

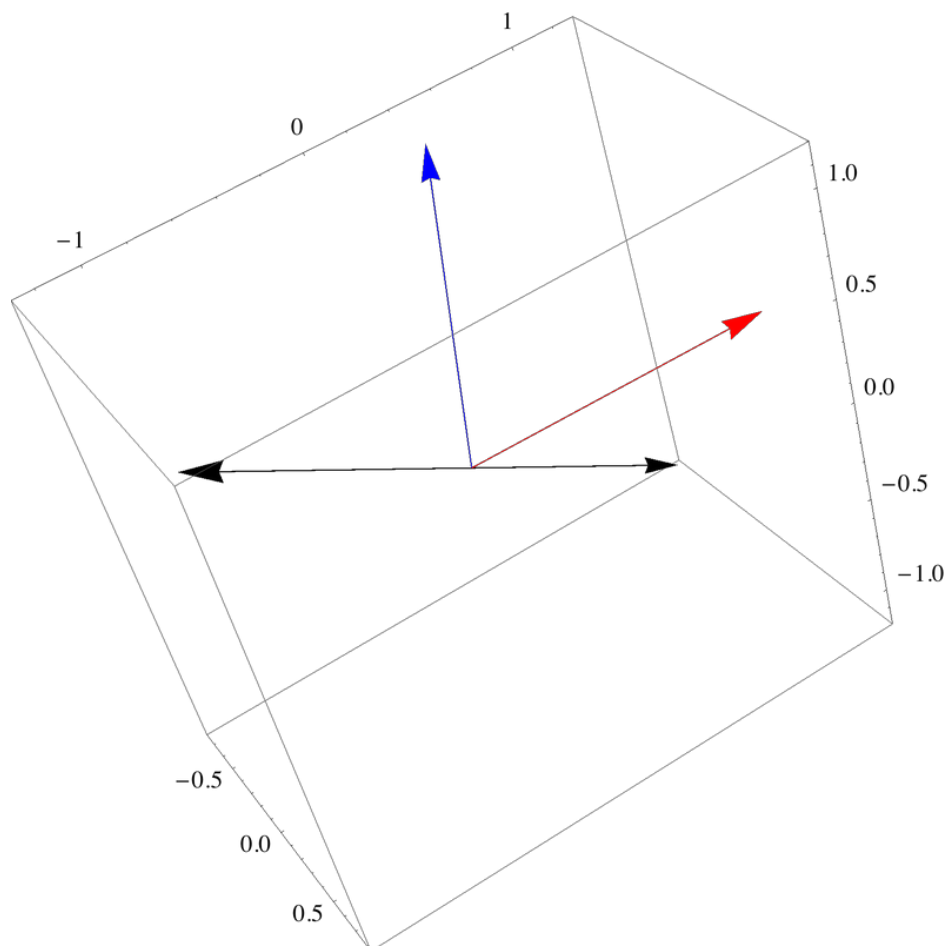


Figure 88: The two black arrows are both perpendicular to the red and blue vectors and have magnitude equal to the product of the magnitude of the two vectors times the sin of the angle in between them. One of these is the cross product. The right hand rule tells us which one.

vectors is the result of the cross product we use the right hand rule. Rather than trying to explain this in words, or in still pictures, I'll simply point you to a video where this is explained clearly.

In order to understand how to take the cross product, it is informative to look at the cross product of the unit basis vectors. Clearly because their

magnitudes are all one and the angle in between them are $\frac{\pi}{2}$, their cross products are going to have unit magnitude, the question is the direction they will be.

For instance, if we want to calculate $\vec{i} \times \vec{j}$ we know that there are two vectors perpendicular to this and of magnitude one. These are \vec{k} and $-\vec{k}$. However, according to the right hand rule we get:

$$\vec{i} \times \vec{j} = \vec{k} \quad (431)$$

Once you have this, all of the other cross products are the cyclic permutations of this expression. That is, you simply shift each vector to the next position in the expression and take the one at the end and put it at the beginning. Shifting each one of these along one, gets you:

$$\vec{k} \times \vec{i} = \vec{j} \quad (432)$$

and shifting it along one more gets:

$$\vec{j} \times \vec{k} = \vec{i} \quad (433)$$

Shifting it along again takes us to the first expression. Swapping around the order of the terms in the cross product changes the result by a sign:

$$\vec{j} \times \vec{i} = -\vec{k} \quad (434)$$

And the same for the other two above. If we take the cross product of a vector with itself then clearly the angle in between the two vectors is zero and so you are left with the zero vector:

$$\vec{i} \times \vec{i} = \vec{0} \quad (435)$$

and similarly for \vec{j} and \vec{k} crossed with themselves. The above facts, along with the distributive property of the cross product, will allow us to calculate the cross product of any vectors when given in component form. We can take a product of two vectors: $\vec{a} = a_1\vec{i} + a_2\vec{j} + a_3\vec{k}$ and $\vec{b} = b_1\vec{i} + b_2\vec{j} + b_3\vec{k}$ for arbitrary components a_i and b_i and find the cross product of the two by expanding everything out:

$$\begin{aligned} \vec{a} \times \vec{b} &= (a_1\vec{i} + a_2\vec{j} + a_3\vec{k}) \times (b_1\vec{i} + b_2\vec{j} + b_3\vec{k}) \\ &= a_1\vec{i} \times b_1\vec{i} + a_1\vec{i} \times b_2\vec{j} + a_1\vec{i} \times b_3\vec{k} + a_2\vec{j} \times b_1\vec{i} \dots etc. \\ &= a_1b_1\vec{0} + a_1b_2\vec{k} - a_1b_3\vec{j} - a_2b_1\vec{k} \dots etc. \\ &= \vec{i}(a_2b_3 - a_3b_2) + \vec{j}(a_3b_1 - a_1b_3) + \vec{k}(a_1b_2 - a_2b_1) \end{aligned} \quad (436)$$

Make sure that you can follow the above and fill in the expressions for "etc.". This looks like a terrible mess but in fact we can come up with a new form of notation that will allow us to remember how to perform this calculation with relative simplicity. First we need to define a determinant of an array.

13.3.1 Determinants

The idea of determinants have been about since around the 3rd century when it first appeared in an ancient Chinese book of Mathematics called The Nine Chapters on the Mathematical Art. It was used originally to define certain properties of systems of linear equations, as we will see later in the section on linear algebra, however for now we will simply use it as a particular way to easily calculate the cross product. Let's take a two by two array of numbers and define the determinant for this.

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad (437)$$

The vertical lines on the left and right are the sign that the we are taking a determinant. For now this is just a definition and we will work with it in what follows. Don't worry too much about where it comes from, but we will see later where it comes from and we will see now why it is useful. The above is for a 2×2 array of numbers. If we have a 3×3 we can write it as follows:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \quad (438)$$

The first term is just the first element in the top row multiplied by the determinant of the bottom two rows, not including the column that a is in. The second term is $-$ the second term in the top row times the determinant of the bottom two rows, not including the column that b is in. The third term is the third element in the top row times the determinant of the bottom two rows, not including the column that c is in. We could then multiply out the 2×2 determinants using the rule in equation 437.

Just to see where the pattern is going, let's look at a 4×4 example. The following defined the determinant of a 4×4 array:

$$\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{vmatrix} = a \begin{vmatrix} f & g & h \\ j & k & l \\ n & o & p \end{vmatrix} - b \begin{vmatrix} e & g & h \\ i & k & l \\ m & o & p \end{vmatrix} + c \begin{vmatrix} e & f & h \\ i & j & l \\ m & n & p \end{vmatrix} - d \begin{vmatrix} e & f & g \\ i & j & k \\ m & n & o \end{vmatrix} \quad (439)$$

Notice that the signs between each term alternate between $+$ and $-$ and we define the determinant recursively. Ok, good, now that we have shown how to calculate a determinant, we can give a very simple expression for the cross

product in terms of the determinant. For our vectors \vec{a} and \vec{b} , we have:

$$\vec{a} \times \vec{b} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \quad (440)$$

Check for yourself that this does indeed give the same expression for the cross product as in equation 436 when you expand out the determinant. What we have now is simply a very simple notational form for calculating the cross product of two vectors. We will see in the section on linear algebra a way to define the determinant of any square matrix.

In fact the cross product has a very geometrical significance. The magnitude of the cross product (ie. $|\vec{a}| |\vec{b}| \sin \theta$) is precisely the area of the parallelogram given by the two ways of adding the vectors \vec{a} and \vec{b} .

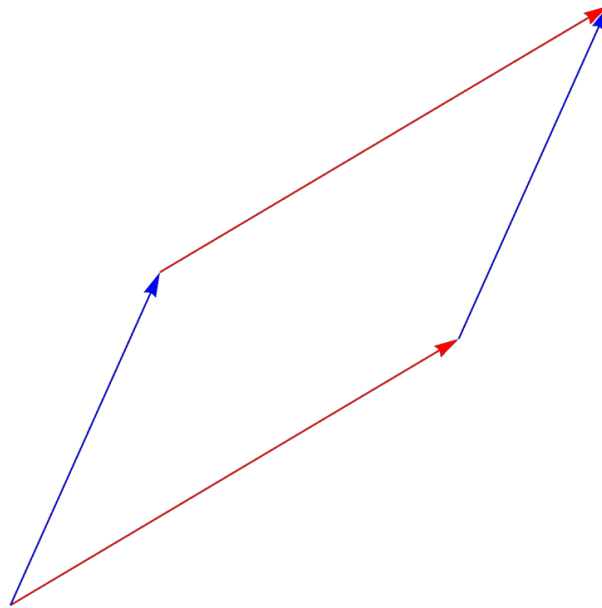


Figure 89: The cross product of the blue and red arrows will give a new vector whose magnitude is equal to the area of the parallelogram between the vectors shown here.

13.4 Properties of the cross product

The fact that the cross product produces a vector which is perpendicular to the two vectors being crossed is extremely useful. In particular it allows us

to find vectors perpendicular to planes. Such a vector is known as a normal vector and can be used to define a plane in 3 dimensional space. Any flat surface in 3 dimensions can be defined by a single point on it and a vector perpendicular to it. To calculate the normal vector we can simply take any 3 points on the plane and define two vectors that go between these three points. As long as the points don't all lie along a straight line we can then take the cross product of the two vectors and the result will be a vector which is perpendicular to the plane. We will see in a bit how this vector can help us to discover relations between points on and off the plane.

Let's imagine a plane that goes through the three points $P(1, 4, 6)$, $Q(-2, 5, 1)$ and $R(1, -1, 1)$. We can form 3 vectors from these three points, but we will only need two.

Let's form the vectors $\vec{PQ} = (-2-1)\vec{i} + (5-4)\vec{j} + (-1-6)\vec{k} = -3\vec{i} + \vec{j} - 7\vec{k}$ and $\vec{PR} = (1-1)\vec{i} + (-1-4)\vec{j} + (1-6)\vec{k} = -5\vec{j} - 5\vec{k}$. We can find a vector normal to these:

$$\vec{PQ} \times \vec{PR} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -3 & 1 & -7 \\ 0 & -5 & -5 \end{vmatrix} = -40\vec{i} - 15\vec{j} + 15\vec{k} \quad (441)$$

How can we check that this vector is perpendicular to the plane? Well, if it is, then it should certainly be perpendicular to any line which lies in the plane, and thus if we take the dot product of any vector that lies in the plane it should give zero. Check for yourselves that when we do the dot product of this vector with the vectors \vec{PQ} and \vec{PR} that it does indeed give zero.

Often we will be most interested not just in a vector which is perpendicular to the plane, but what is known as a unit normal vector, that is a vector of unit magnitude that is perpendicular to the plane. We can thus take the vector we have just found and divide it by its magnitude ($5\sqrt{82}$):

$$\hat{n} = \frac{-40\vec{i} - 15\vec{j} + 15\vec{k}}{5\sqrt{82}} \quad (442)$$

where the hat on the n is sometimes used to denote a unit vector.

In figure 90 we show the vectors described above.

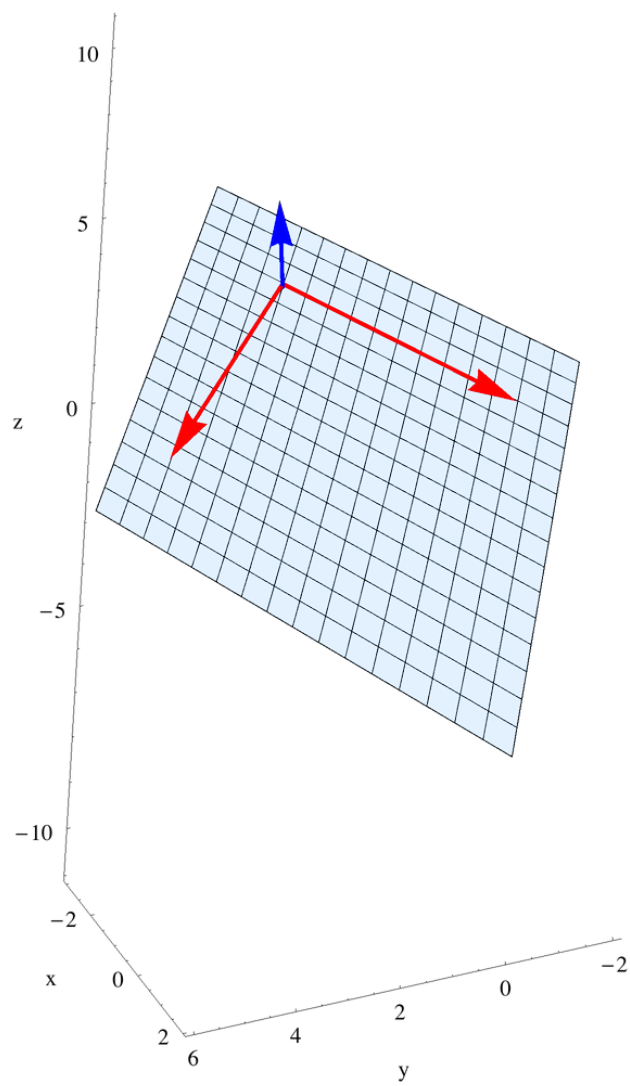


Figure 90: The cross product of the two red arrows which lie in the plane produces the blue arrow (once normalised to one) which lies perpendicular to the plane.

13.5 A multi-technique exercise

We're going to look through an example which includes a fair number of different techniques which we have used in the last few weeks. This will include functions of two variables, partial derivatives, planes, intersections, vector equations for lines, cross and dot products and more.

We will start with a simple function of two variables:

$$f(x, y) = y^2 + x + (x - 1)y \quad (443)$$

We plot this in figure 91 on its own.

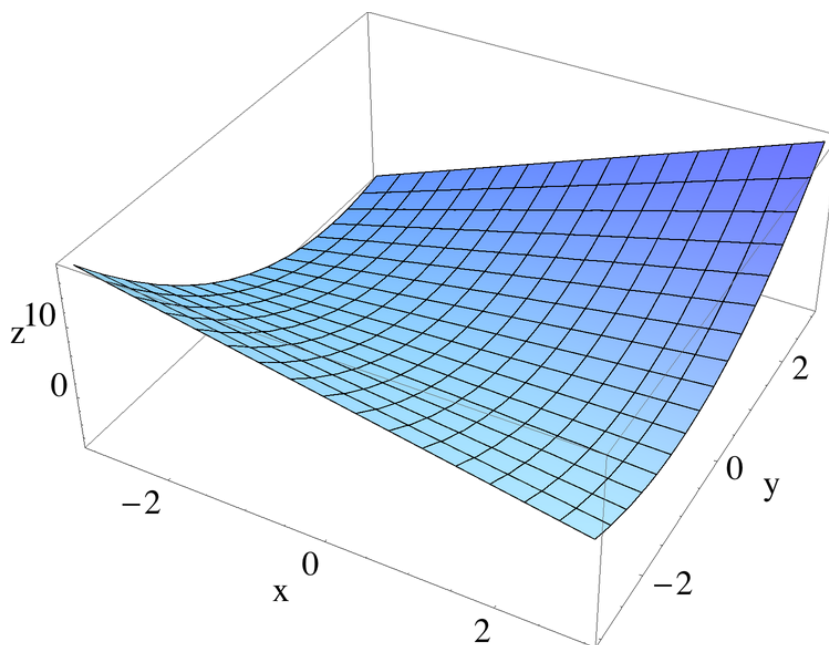


Figure 91: Plot of the function $f(x, y) = y^2 + x + (x - 1)y$.

Now we're going to look at the various traces of this function. The traces (also known as cross sections) are taken by looking at the function at fixed values of x or y or z . Let's start with fixing the value of y . For instance we can ask what the cross-section of the function at $y = -2$ and at $y = 2$ look like. You can imagine intersecting the surface with the surfaces $y = -2$ and $y = 2$. This is shown in figure 92 along with the graph that we get when we look at how the function passes through these two planes.

We can then do the same thing but taking slices in the x direction. This is shown in figure 93

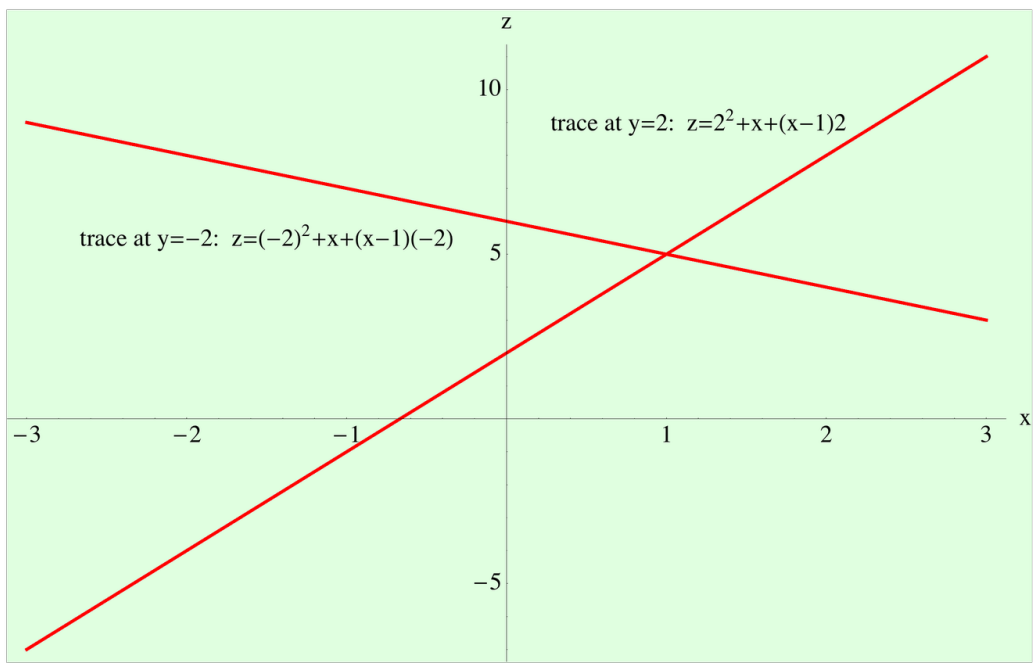
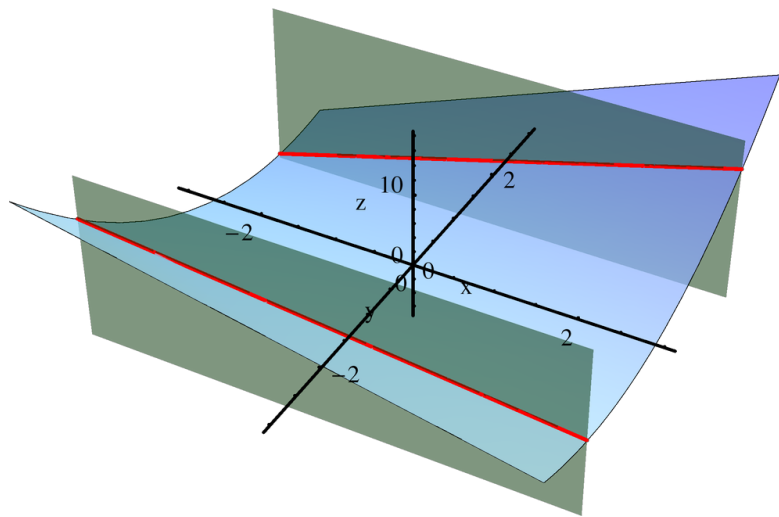


Figure 92: Cross sections of the function at $y = -2$ and $y = 2$.

We can also ask about the level curves of the function. The level curves

are the slices in the z direction - ie. the contours. They are the lines projected into the (x, y) plane which correspond to equal values of z . These are shown in figure 94

OK, so now that we've taken our function and cut it up in various directions to draw out the various traces and level contours, we can start to play some more games with it.

Let's take 3 points on the surface of the function and see what the partial derivatives at these points are. We will take the points $P_0(1, 1, 2)$, $P_1(1, -1, 2)$ and $P_2(-1, 1, -2)$. First check to see that these points are indeed on the surface. These are shown on the surface in figure 95.

Now we're going to ask about the tangent planes to these points. Let's start with $P_0(1, 1, 2)$. We want to find a plane which passes through this point and has the same gradients in the x and y directions as the function itself at that point. The equation for a plane which passes through this point is:

$$z = a(x - 1) + b(y - 1) + 2 \quad (444)$$

where we have still to work out the constants a and b . We will work them out by matching the gradient in the x and y direction with that of the function at that point. Let's first ask about the partial derivatives of the function:

$$\begin{aligned} \partial_x f(x, y) &= 1 + y \\ \partial_y f(x, y) &= 2y + (x - 1) \end{aligned} \quad (445)$$

So, at the point P_0 we get:

$$\begin{aligned} \partial_x f(1, 1) &= 2 \\ \partial_y f(1, 1) &= 2 \end{aligned} \quad (446)$$

So we want the plane to have gradient 2 in the x direction and 2 in the y direction. The equation will thus be:

$$z = 2(x - 1) + 2(y - 1) + 2 \quad (447)$$

Check that this does indeed have the right gradients in the x and y directions. We will call this plane T_0 (the tangent plane through the point P_0). Now let's find the same thing for the tangent planes through P_1 and P_2 . We will call these planes T_1 and T_2 . The equations for these are, for T_1 :

$$z = 0(x - 1) - 2(y + 1) + 2 \quad (448)$$

and for T_2 :

$$z = 2(x + 1) + 0(y - 1) - 2 \quad (449)$$

In figures 96, 97 and 98 we see the 3 tangent planes T_0 , T_1 and T_2 which pass through the three points with the appropriate x and y gradients. Now let's forget about the underlying function for the moment and ask about the 3 tangent planes. Let's first ask about the intersection at T_0 and T_1 . We can guess that they will intersect at a line but let's check. Firstly we will try and find the equation for the line using Gauss reduction and then we will find the equation for the same line as a vector equation.

The two planes and their intersection are shown in figure 99 We have the two equations for T_0 and T_1 and we want to ask about the line along which they intersect. We take the two equations:

$$\begin{aligned} z &= 2(x - 1) + 2(y - 1) + 2 \\ z &= 0(x - 1) - 2(y + 1) + 2 \end{aligned} \quad (450)$$

and we rewrite them as:

$$\begin{aligned} 2x + 2y - z &= 2 \\ -2y - z &= 0 \end{aligned} \quad (451)$$

We can then take these two equations and write this in augmented coefficient matrix form:

$$\left(\begin{array}{ccc|c} 2 & 2 & -1 & 2 \\ 0 & -2 & -1 & 0 \end{array} \right) \quad (452)$$

Let's start to use gauss reduction for this matrix. We start by dividing the top row by 2 and the bottom row by -2 :

$$\left(\begin{array}{ccc|c} 1 & 1 & -\frac{1}{2} & 1 \\ 0 & 1 & \frac{1}{2} & 0 \end{array} \right) \quad (453)$$

We then subtract the bottom row from the top:

$$\left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & \frac{1}{2} & 0 \end{array} \right) \quad (454)$$

We can thus write this as:

$$\begin{aligned} x &= 1 + z \\ y &= -\frac{z}{2} \end{aligned} \quad (455)$$

or alternatively we can write it in parametric form by setting $z = t$:

$$\begin{aligned} x &= 1 + t \\ y &= -\frac{t}{2} \\ z &= t \end{aligned} \quad (456)$$

We can also write this in vector form as:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ -\frac{1}{2} \\ 1 \end{pmatrix} \quad (457)$$

We will now get the same line, but using the vector method. We know that the equation for a line can be given in vector form by:

$$\vec{r} = \vec{r}_0 + t\vec{v} \quad (458)$$

where \vec{r} is a vector which, when we vary t takes us from the origin to all points along the line. \vec{r}_0 is any point on the line and \vec{v} is a vector which is parallel to the line. Let's start by finding a point which is on the line. Remember the equations for the planes, which meet at the line are given by:

$$\begin{aligned} 2x + 2y - z &= 2 \\ -2y - z &= 0 \end{aligned} \quad (459)$$

We want to find any point on this line. Let's see if it passes through the $x = 1$ plane by plugging in $x = 1$:

$$\begin{aligned} 2 + 2y - z &= 2 \\ -2y - z &= 0 \end{aligned} \quad (460)$$

Indeed these two equations can be solved for $y = 0, z = 0$. Thus, the line of intersection of the two planes passes through the point $(1, 0, 0)$. We will then use $\vec{r}_0 = \langle 1, 0, 0 \rangle$. To find \vec{v} we have to find a second point on the plane. Let's see if the line passes through the point $x = 0$ by plugging in $x = 0$ into the equations for the planes:

$$\begin{aligned} 2y - z &= 2 \\ -2y - z &= 0 \end{aligned} \quad (461)$$

This time this is solved for $z = -1, y = \frac{1}{2}$, so it passes through the point $(0, \frac{1}{2}, -1)$. We have two points on the line and thus we can find a vector which is parallel to the line by taking the differences between the points:

$$\vec{v} = \left\langle 0 - 1, \frac{1}{2} - 0, -1 - 0 \right\rangle = \left\langle -1, \frac{1}{2}, -1 \right\rangle \quad (462)$$

If this vector is parallel to the line then clearly so will the negative of this vector (it'll be just pointing in the opposite direction. So let's actually chose:

$$\vec{v} = \left\langle 1, -\frac{1}{2}, 1 \right\rangle \quad (463)$$

So, putting all of this together we have:

$$\vec{r} = \langle 1, 0, 0 \rangle + t \left\langle 1, -\frac{1}{2}, 1 \right\rangle \quad (464)$$

But lo and behold, this is exactly the same as equation 457. The two points used in this construction, as well as the vector (shown by the blue arrow from the origin, giving \vec{r}_0) are shown in figure 100. ok, now we've found the tangent planes and the line corresponding to the intersection of two of them. In fact we've found the equation for this line in two different ways. Now let's look at the intersection of all 3 of them. We have the three equations for the tangent planes given by:

$$\begin{aligned} 2x + 2y - z &= 2 \\ -2y - z &= 0 \\ 2x - z &= 0 \end{aligned} \quad (465)$$

where the last equation is just a rewriting of equation 449. Let's put this into augmented coefficient matrix form to solve the three equations by gauss reduction:

$$\left(\begin{array}{ccc|c} 2 & 2 & -1 & 2 \\ 0 & -2 & -1 & 0 \\ 2 & 0 & -1 & 0 \end{array} \right) \quad (466)$$

We'll start by sending row 3 to row 1 and dividing by 2 and we'll divide row 2 by -2 and row 1 by -1 :

$$\left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ -2 & -2 & 1 & -2 \end{array} \right) \quad (467)$$

Now we'll add 2 lots of row 1 to row 3:

$$\left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & -2 & 0 & -2 \end{array} \right) \quad (468)$$

We notice here that maybe it would have been better to swap rows 2 and 3 before so let's do that and divide by -2 . Let's also multiply row 2 again by 2:

$$\left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 1 & 0 \end{array} \right) \quad (469)$$

Now let's subtract two lots of row 2 from row 3:

$$\left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array} \right) \quad (470)$$

finally let's add $\frac{1}{2}$ of row 3 to row 1:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array} \right) \quad (471)$$

So it seems that the three tangent planes intersect each other at the point $(-1, 1, -2)$. We can make sure that this is true by checking if this point does indeed lie on each plane: Just plug in the numbers to the three equations and make sure that the equations are satisfied. Let's actually calculate this same thing another way. We can write the original 3 equations in the form:

$$\begin{pmatrix} 2 & 2 & -1 \\ 0 & -2 & -1 \\ 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \quad (472)$$

This is just the regular matrix equation form of what we previously wrote as the augmented coefficient matrix. We can solve this equation by finding the inverse:

$$\begin{pmatrix} 2 & 2 & -1 \\ 0 & -2 & -1 \\ 2 & 0 & -1 \end{pmatrix}^{-1} \quad (473)$$

and applying it to both sides of equation 472. We know two ways to find the inverse of a matrix. We can solve the augmented coefficient matrix equation:

$$\left(\begin{array}{ccc|ccc} 2 & 2 & -1 & 1 & 0 & 0 \\ 0 & -2 & -1 & 0 & 1 & 0 \\ 2 & 0 & -1 & 0 & 0 & 1 \end{array} \right) \quad (474)$$

Notice that to get the left hand side into reduced row echelon form we perform exactly the same elementary row operations that we did to solve the augmented coefficient matrix equation in equation 466. Running through these operations again gets us to:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & 1 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right) \quad (475)$$

The left hand side of this is thus the inverse of the original matrix and so we can see that:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & 1 \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix} \quad (476)$$

which, again, is the same solution we came up with before for the point of intersection of the three tangent planes.

Let's calculate the inverse of the matrix using the method of cofactors. We know that the inverse of a matrix is given by:

$$M^{-1} = \frac{Adj(M)}{det(M)} \quad (477)$$

The determinant of the matrix is:

$$\begin{vmatrix} 2 & 2 & -1 \\ 0 & -2 & -1 \\ 2 & 0 & -1 \end{vmatrix} = 2(-2 \times -1) - 2(0 - -2) - 1(0 - -4) = 4 - 4 - 4 = -4 \quad (478)$$

The Adjoint is the transpose of the matrix of cofactors. The cofactor C_{ij} is the signed (ie. multiplied by $(-1)^{i+j}$) determinant of the matrix left over when you delete the i^{th} row and the j^{th} column from the original matrix. We can show that C is given by:

$$\begin{pmatrix} 2 & -2 & 4 \\ 2 & 0 & 4 \\ -4 & 2 & -4 \end{pmatrix} \quad (479)$$

and thus the adjoint is just the transpose of this:

$$\begin{pmatrix} 2 & 2 & -4 \\ -2 & 0 & 2 \\ 4 & 4 & -4 \end{pmatrix} \quad (480)$$

Dividing this by the determinant should give us the inverse which is thus:

$$\frac{\begin{pmatrix} 2 & 2 & -4 \\ -2 & 0 & 2 \\ 4 & 4 & -4 \end{pmatrix}}{-4} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & 1 \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -1 & -1 & 1 \end{pmatrix} \quad (481)$$

Which is precisely the inverse that we found using Gauss reduction.

OK, so we've found the intersection point of the three tangent planes in a number of ways. Let's also look at some more features of the tangent lines. One thing that we can easily calculate is the normal vector to the tangent plane. In fact we can simply read it off from the equation of the planes by looking at the coefficient of the x , y and z terms. For instance, an equation of the form:

$$ax + by + cz + d = 0 \quad (482)$$

will have a normal vector: $\vec{n} = \langle a, b, c \rangle$ but we would like to practice some more of our vector operations, so let's do it in a slightly more long winded way.

Let's take the first tangent plane again:

$$2x + 2y - z = 2 \quad (483)$$

and let's find 3 points on the plane (3 points not in a straight line). Let's find a point at $x = 0, y = 0$. Thus, from the equation for the plane $z = -2$. At $(x, y) = (0, 1), z = 0$ and at $(x, y) = (1, 0), z = 0$ as well. So we have three points $Q_0(0, 0, -2), Q_1(0, 1, 0)$ and $Q_2(1, 0, 0)$. From these three points let's define two vectors $Q_0\vec{Q}_1 = \langle 0, 1, 2 \rangle$ and $Q_0\vec{Q}_2 = \langle 1, 0, 2 \rangle$. Both of these vectors are parallel to the tangent plane and thus if we take the cross product of them we will get a vector which is perpendicular to the tangent plane, ie. a normal vector:

$$\vec{n} = Q_0\vec{Q}_1 \times Q_0\vec{Q}_2 = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 0 & 1 & 2 \\ 1 & 0 & 2 \end{vmatrix} = 2\vec{i} - (-2)\vec{j} + (-1)\vec{k} \quad (484)$$

Note that there is a much easier way of finding a normal vector to a plane, and that is that we can simply read it off from the equation from the plane. For a plane of the form $ax + by + cz = d$ the vector $\langle a, b, c \rangle$ will be a normal vector to the plane, as will any scalar multiple of it.

ok, so we've found a normal vector to one of the planes. We can easily find normal vectors to the other planes and once we have done this we can ask about the angles in between the planes via the dot product of the normal vectors. Normal vectors to the planes T_0, T_1 and T_2 are:

$$\begin{aligned} \vec{n}_0 &= 2\vec{i} + 2\vec{j} - \vec{k} \\ \vec{n}_1 &= -2\vec{j} - \vec{k} \\ \vec{n}_2 &= 2\vec{i} - \vec{k} \end{aligned} \quad (485)$$

Thus, the angles in between the T_0 and T_1 is:

$$\theta = \arccos \frac{\vec{n}_0 \cdot \vec{n}_1}{|\vec{n}_0| |\vec{n}_1|} = \arccos \frac{-4 + 1}{\sqrt{2^2 + 2^2 + 1}\sqrt{2^2 + 1}} = \arccos \frac{-1}{\sqrt{5}} \quad (486)$$

Let's look exactly what this angle corresponds to in terms of the two tangent planes. This is shown in figure 101

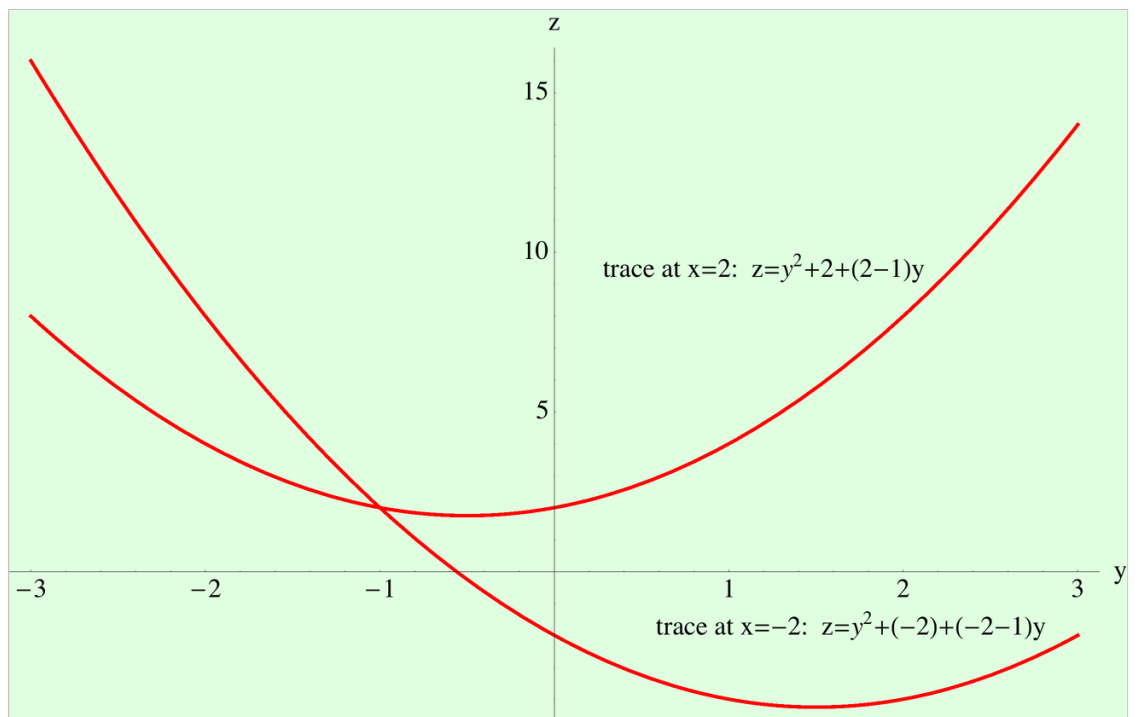
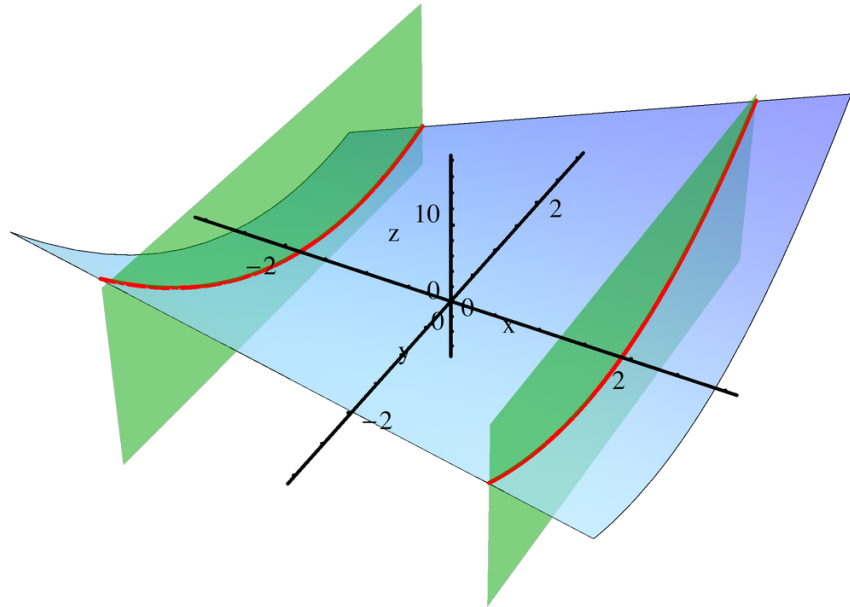


Figure 93: Cross sections of the function at $x = -2$ and $x = 2$.

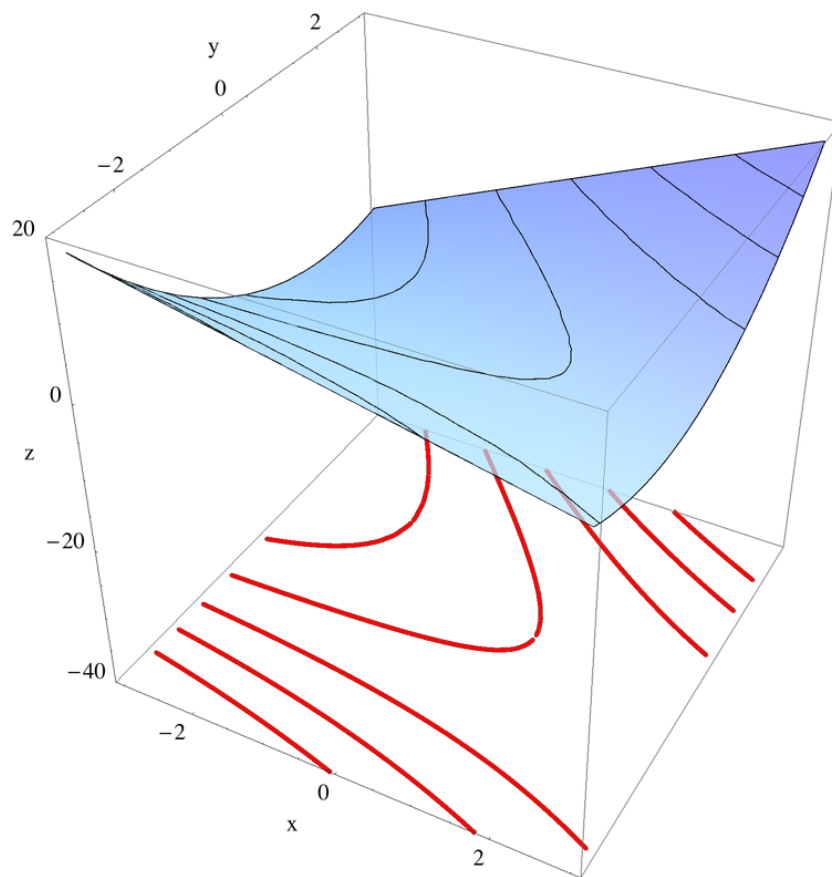


Figure 94: Level curves of the function. These can be shown as contours in the (x, y) plane.

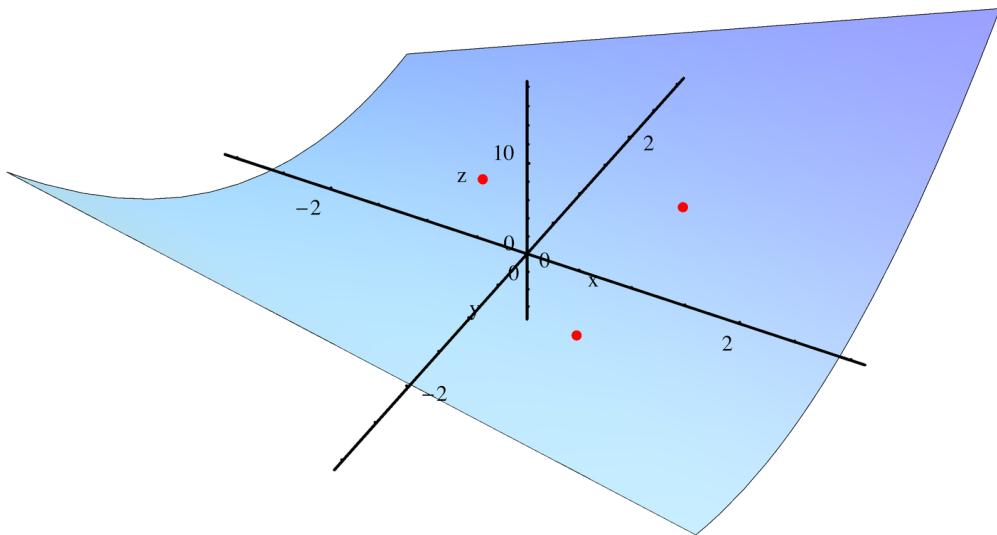


Figure 95: Graph of the function with 3 points drawn on the surface.

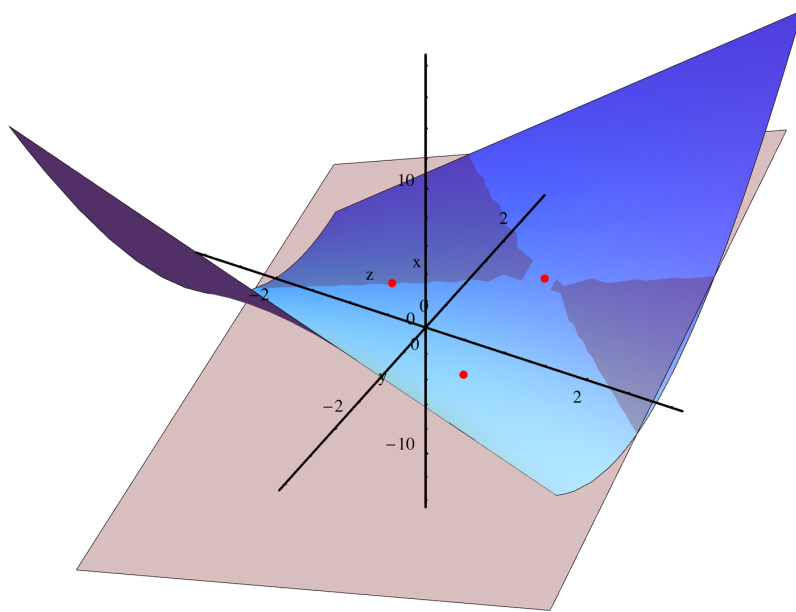


Figure 96: Tangent plane T_0 to the point $P_0(1, 1, 2)$.

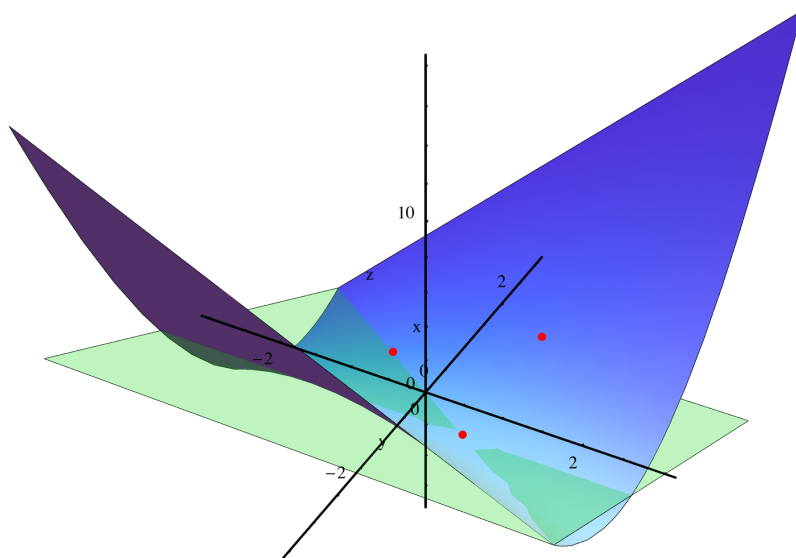


Figure 97: Tangent plane T_1 to the point $P_1(1, -1, 2)$.

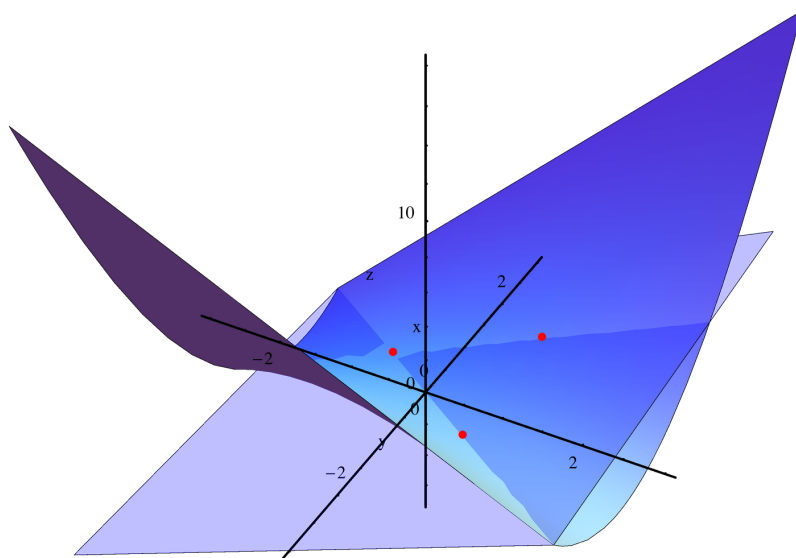


Figure 98: Tangent plane T_2 to the point $P_2(-1, 1, -2)$.

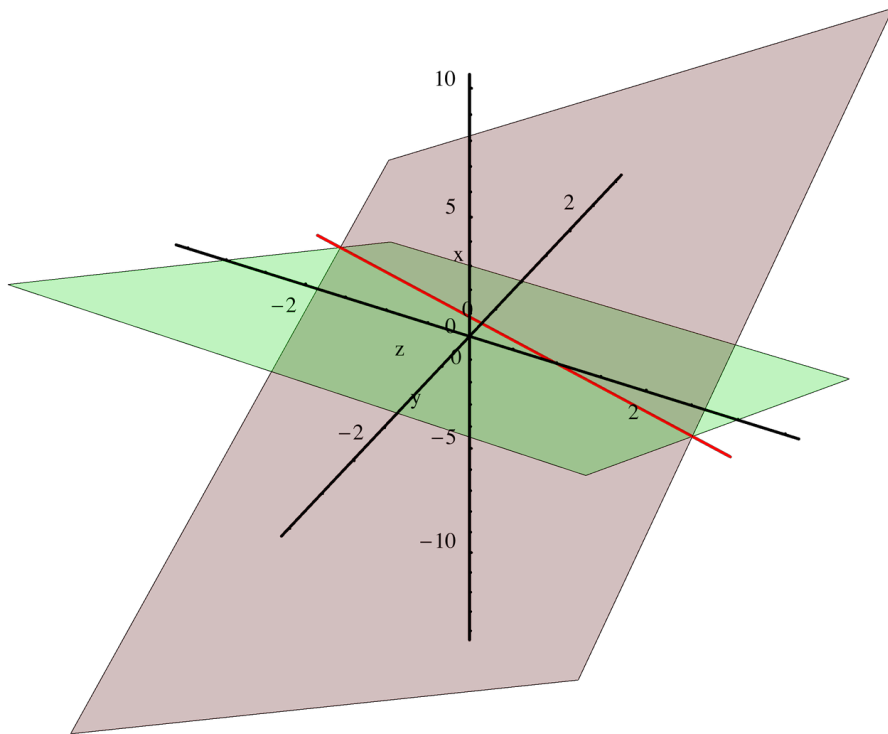


Figure 99: The intersection of T_0 with T_1 and the line of intersection. We will calculate the equation for this line in two different ways.

antialias[]

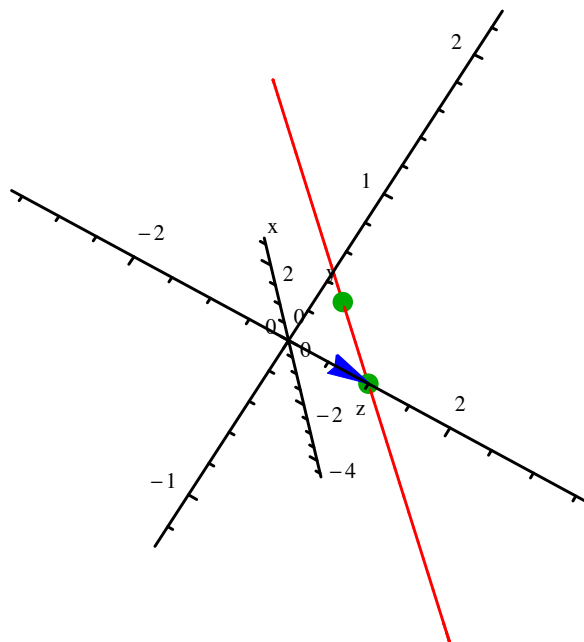


Figure 100: The line given by the intersection of T_0 and T_1 , two points on the line which are used to define the vector parallel to the line (\vec{v}) and the blue arrow giving \vec{r}_0 . Note that without perspective, it's not easy to visualise where this line actually lies in 3 dimensions. It's easier to see it along with the two tangent planes shown in figure 99.

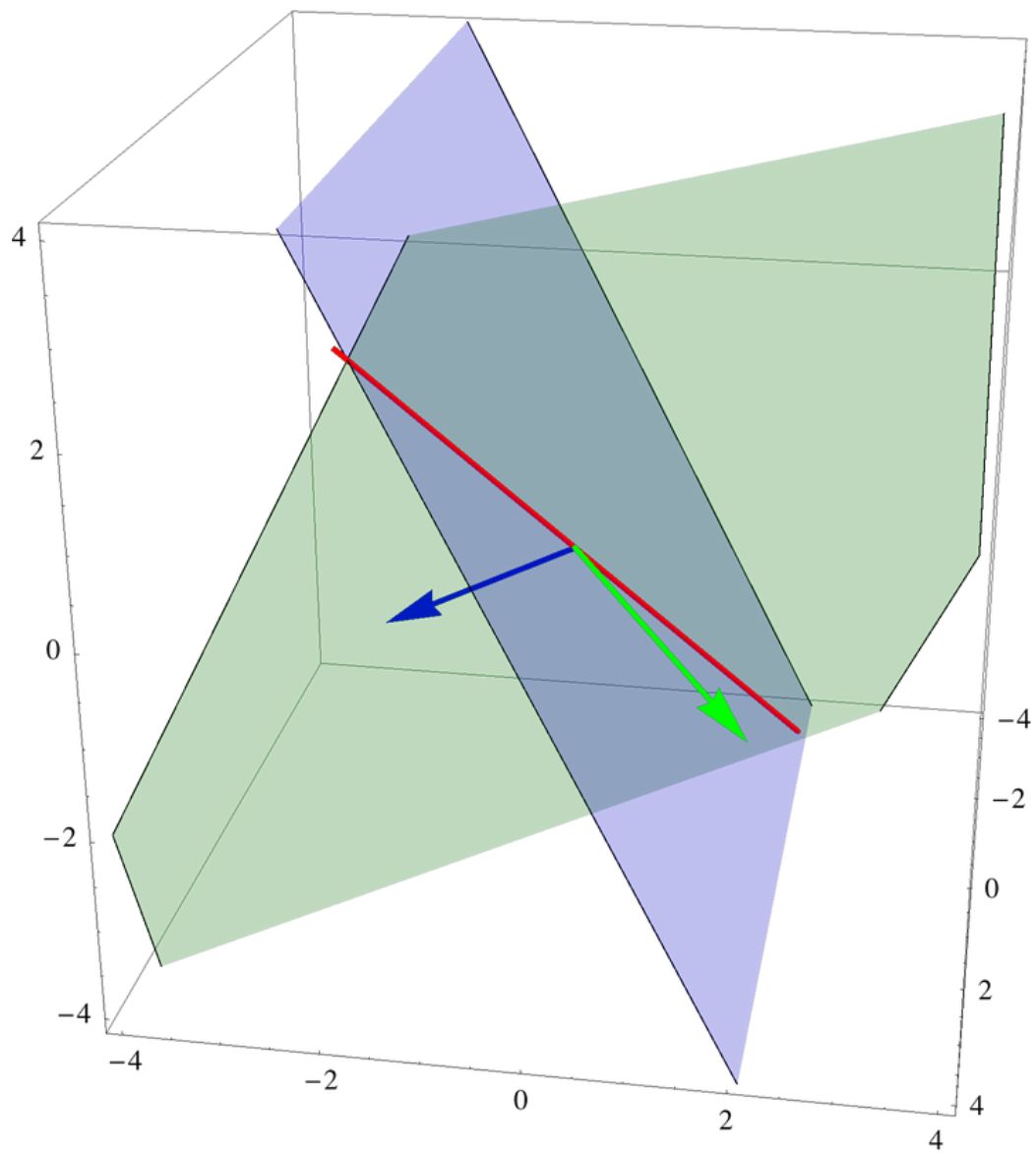


Figure 101: The intersection of T_0 with T_1 and the line of intersection along with two normal vectors to the planes. The angle in between the two normal vectors is calculated using the dot product and gives roughly $\theta = 2 \text{ radians}$.