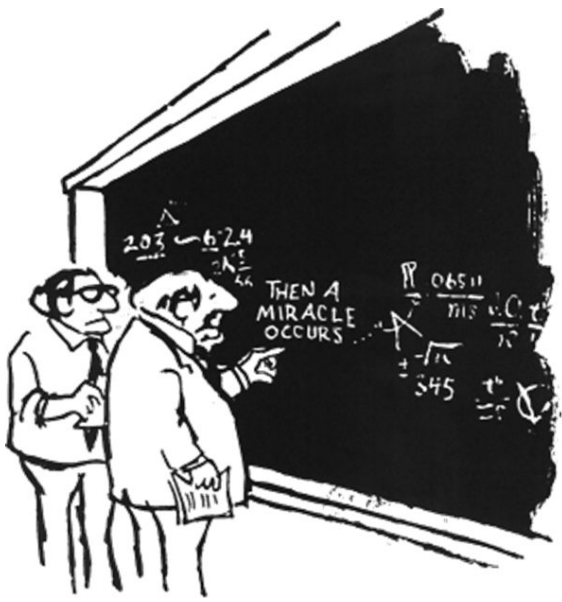
University of Bath
Department of Mechanical Engineering

ME12002 Engineering Mathematics

Dr D A S Rees



``I think you should be more explicit here in step two..."

Contents

0 PRELIMINARIES

0.1 Who am I?

I am Andrew Rees, more formally written as Dr. D. A. S. Rees. I have been stashed away in 4E in the Department of Mechanical Engineering since 1990 and have been happily whiling away my time doing research and also teaching the ever-increasing numbers of students. In 1991 we had approximately 90 students per year for a three year course, whereas now we have close to 320 on a four year course — these are just the Mechanical students. In that time the number of academic staff has only doubled.

My office is 4E 2.54 and my email address is ensdasr@bath.ac.uk, which reflects the old naming convention we had of ensdasr, which is a shorthand for engineering, staff and my initials. Only a few of us old-stagers still have such an email address. An alias is D.A.S.Rees@bath.ac.uk which I prefer, but the Computer Centre doesn't! In your cases your official email addresses consist of your initials and a random number. I have no idea whether we'll get a fourth convention.....

0.2 My background.

What authority do I claim to be able to teach you?

I guess that a 1st class degree in Mathematics from Imperial College, London, in 1980 isn't too shabby. I guess too that the fact that Bath University hasn't yet sacked me for being useless is a note in my favour. Somehow I have managed to get away with it for 32 years....

My Ph.D. was in Applied Mathematics from the University of Bristol in 1986 where I concentrated on convective flows and stability in porous materials. Added to this is an ATCL diploma from the Trinity College of Music, London, in 1977 and membership of the National Youth Orchestra of Wales; so I guess that this means that I do have at least one interest outside of maths and that I also get to see people who aren't colleagues.

I did have a short spell from 1980 to 1982 working for what was then called British Aerospace. Based in Filton, they wanted me to do some flutter calculations, and to devise feedback laws which would stop aircraft wings falling off when they flew faster than their original design speed. This was in those heady days when fly-by-wire technology was being created.

After my doctorate I stayed at Bristol University with the same supervisor doing essentially the same things but doing them better and faster. This was followed by two years' worth of fixed-term lecturing duties in the Maths Department at Exeter University. Then the job in Bath appeared and I just had to say *yes* to the offer of employment because I had a mortgage to service, and a wife and two children to feed.

0.3 Provision of 'content'

Everything that I hall be giving you thi semester may be accessed at my dedicated Engineering Maths website:

https : //people.bath.ac.uk/ensdasr/ME12002.bho/maths1.html

and there is a link to that from the appropriate Moodle page. A second one will be available after Christmas for the second semester. I hardly use Moodle at all because everything swims before my very eyes and, frankly, it is easier to write in basic html for at least I am in control!

The lectures which I recorded for 2020/2021, when we were in a fully remote-learning mode, are also available at the same website. These were done in a conference presentation style, effective but potentially boring — jokes don't work when there is no audience! Note that some of the section numbering will be different — this is because the syllabus has been reordered this year.

There will be 11 main problem sheets for the 10 weeks of lectures. **Please note** that I won't be marking your workings because roughly 320 students multiplied by an estimate of one hour per script adds up to much more than a week! For the great majority of the questions on the problem sheets I will be providing fairly comprehensive solutions, and these will be available at the ME12002 website.

The two hours of tutorials (really they are problems classes) will see about seven or eight of us (me, various postgrads and possibly a postdoc) circulating around the rooms engaging with your mathematical issues.

0.4 Outline of the first semester of ME12002.

The following is a brief breakdown of the content of the unit:

Topic	Number of lectures
Introduction	$\mathbf{0.2}$
Complex numbers	$\mathbf{1.8}$
Differentiation	$\mathbf{3}$
Integration	$\mathbf{3}$
Series	$\mathbf{3}$
ODEs	$\mathbf{5}$
Laplace Transforms	$\mathbf{4}$

As you can see, many of these topics are familiar. It is my experience that different people will have covered different things and, given that all of these topics are deemed to be important, I will need to cover them all. Different people will have met slightly different topics and subtopics before coming to university, and therefore everyone will have already covered a subset of the following content for this semester but not necessarily the same subset. Hence my choice of syllabus. Hopefully, my aim to bring forth some new ideas in various places will mean that there will be a few nuggets of gold to find along with the familiar. On the other hand, the Laplace Transforms section at the end of the semester is likely to be quite new for the vast majority.

0.5 The detailed syllabus.

The following is how I will split the unit content over the 20 lectures.

	Topic	Content
1.	Complex numbers.	Motivation/need for them. Definition. Arithmetic operations. Geometric interpretation. Polar/exponential form.
2.		de Moivre's theorem. Euler's formula and identity. Roots of complex numbers. Further identities.
3.	Differentiation.	Definition and notations. Use of limits. Higher derivatives. Linearity. Product rule. Products of more than two functions.
4.		Chain rule and proof. Nested functions. Advanced cases. Quotient rule.
5.		Critical points. Primary and secondary criteria. Checklist.
6.	Integration.	From sums to integrals. Definite and indefinite integrals. Integration by substitution. f'/f form.
7.		Ratios of polynomials and partial fractions. Top heavy ratios. Repeated factors. Irreducible quadratics.
8.		Integration by parts. Derivation of the Rees method – rules of implementation. Miscellaneous cases.
9.	Series.	Series and sequences. Binomial theorem and Pascal's triangle. Binomial series.
10.		Taylor's series and Maclaurin. Derivation/justification. Two forms of Taylor's series.
11.		Convergence and d'Alembert's test. Examples of numerical and of power series. Radius of convergence. l'Hôpital's rule. Derivations.
12.	ODEs	Classification. Order, linearity, BVP/IVP. Reduction to first order form.
13.		Separation of variables. First order linear ODEs and Integration Factors.
14.		Linear constant-coefficient ODEs. Homogeneous cases.
15.		Linear constant-coefficient ODEs. Inhomogeneous cases 1.
16.		Linear constant-coefficient ODEs. Inhomogeneous cases 2.
17.	Laplace Transforms	Definition. Examples. LT of derivatives. Solutions of some ODEs.
18.		The unit impulse. Solutions of ODEs with an impulse as the forcing function.
19.		The shift theorem in both s and t. The unit step function.
20.		Convolution. Examples. Use in ODE solutions.

The remaining subsections in this introductory chapter cover a variety of preliminary topics. I am not at all sure that these topics are emphasized sufficiently well in pre-university tuition, but I would far prefer to play safe by presenting my ideas to you so that you will know what the standards are that I expect.

0.6 Significant figures.

Here's a good question: which is better, 6 decimal places or 6 significant figures?

The answer depends on how large the number is.

The number $\pi/1000 = 0.003142$ is correct to 6 decimal places but only to 4 significant figures, while the number $100\pi = 314.159265$ is also correct to 6 decimal places but is correct to 9 significant figures.

So it is clear that the number of significant figures is the better or more useful concept because it is independent of the magnitude of the number being considered.

0.6.1 The number of significant figures.

From the practical point of view of computing numbers, the big question is, *How many significant figures do we need?* To answer this we need to understand what happens when we do arithmetic with limited precision.

In engineering, we frequently need only 3 or perhaps 4 significant figures for comparison with experimental work, for that is often the greatest precision with which we can measure. There are, of course, counter-examples to this such as the mass of the electron, the speed of light, gravitational acceleration and the density of water, for each of these have been measured to a very much greater degree of accuracy.

If one were to do a Google search for the value of g, the acceleration due to gravity, then one website gives 9.8m/s^2, another gives 9.81m/s^2 while the 'standard value' is generally defined as 9.80665m/s^2. The first of these has 2SFs, the second has 3SFs, while the third value is apparently correct to 6SFs.

So what is the truth of the matter?

If we take the third as being correct, then the first two are clearly approximations, so all is well....or is it? Although the third has been designated as the standard value, it gives the misleading impression that it is valid everywhere on the earth's surface. Although we know that g decreases as we rise above the Earth's surface and also as we descend under the sea, the standard value is not correct. In fact it varies between 9.779m/s^2 for Mexico City to 9.819m/s^2 for Oslo; these values represent fairly closely the two extreme cases for the earth's land surface when we confine ourselves to cities. Therefore it is pointless using all six significant figures in the standard value unless unless you are at one of the precise places where it applies — Zurich is correct to four significant figures! Perhaps we should only use 9.8m/s^2 and simply note that all our calculations will also have only two SFs of accuracy. And even if we know exactly where we are on the globe, g will also vary with the time of day if we are on open water due to the phase of the tide.

What are the consequences of this variation? One trivial one is that one would weigh 0.4% less in Mexico than in Oslo. Another is that a pendulum in Mexico has a period which is 0.2% longer than an identical one in Oslo. Admittedly these are small values, but at least we know what the possible error is.

However, despite needing only 3 or 4 figures for experimental work, it is not always good to retain only 3 or 4 figures of accuracy during intermediate theoretical calculations. This is because round-off errors can build up catastrophically. In general, I would recommend using 5 or 6 for all calculations, and if the final result is required to 3 significant figures (SFs), one may then round the final answer to the required accuracy. But even then, one must be aware of how accuracy can get eroded.

See `https://units.fandom.com/wiki/Gravity_of_Earth`

Example 0.1 Calculate $\mathbf{0.301 + 0.478 + 1.42 + 18.4 + 101}$ to 3 SFs.

This example illustrates how care must be taken when doing arithmetic. We'll compute the sum using three different 'methods'.

Method (i) Exact arithmetic gives $\mathbf{121.599}$ which is $\mathbf{122}$ to 3SFs. This is the way one would do it using a calculator or pencil and paper.

Method (ii) Adding by beginning with the largest number and rounding each computation before adding the next largest and so on.

$$\begin{aligned} \mathbf{101 + 18.4 = 119.4} &= \mathbf{119} \text{ (3SF)} \\ \mathbf{119 + 1.42 = 120.42} &= \mathbf{120} \text{ (3SF)} \\ \mathbf{120 + 0.478 = 120.478} &= \mathbf{120} \text{ (3SF)} \\ \mathbf{120 + 0.301 = 120.301} &= \mathbf{120} \text{ (3SF)} \end{aligned}$$

Method (iii) Adding by beginning with the smallest number and rounding each computation before adding the next smallest and so on.

$$\begin{aligned} \mathbf{0.301 + 0.478 = 0.779} & \\ \mathbf{0.779 + 1.42 = 2.199} &= \mathbf{2.20} \text{ (3SF)} \\ \mathbf{2.20 + 18.4 = 20.60} &= \mathbf{20.6} \text{ (3SF)} \\ \mathbf{20.6 + 101 = 121.6} &= \mathbf{122} \text{ (3SF)} \end{aligned}$$

Method (i) works well because the precision we are using (namely, exact for pen and paper, 8 to 10 SFs for a calculator, or either 7 or 15 or more for a computer) uses very many more SFs than the precision of the numbers.

Method (ii) is the poorest. By using the largest number first, we are declaring that almost all the information contained after the decimal point in the rest of the data will not contribute to the final result. This is why it provides the least accurate sum. Note also that if we had included another million numbers under $\mathbf{0.5}$ in magnitude, then this would have had no effect on the final answer. So this is a bad way to proceed.

Method (iii) yields the safest method when working with a highly restricted number of SFs.

A further danger point is that the original data that we have added together isn't consistent in terms of the number of decimal places, i.e. we have no idea whether they were already rounded to 3SFs before we started to sum them. If whomsoever gave us the data had rounded the data to 3 SFs, then the largest value, $\mathbf{101}$, could range between $\mathbf{100.500}$ and $\mathbf{101.500}$ in real life and this could alter our final answer. We would have more confidence in the answer if either (a) it is stated that each value is exact or (b) all the data were presented to 3DPs. So the moral of the story is that we must enquire about the provenance of the data before we manipulate it.

The reason I have mentioned the above example is that all calculating devices use a fixed number of SFs, and therefore rounding always happens after each addition, subtraction, exponentiation, multiplication, division, square root, etc... It is rare for this to cause any problems, but it is good to be aware that round-off error can very occasionally yield incorrect results.

However, I often see very severe rounding taking place in computations performed in exams. Even if the method used is correct, the final answer can often be very poor. Always be suspicious of numbers and where they come from.

Example 0.2 Solve the quadratic $x^2 - 2.01x + 1.01 = 0$ giving the answer to 4SFs.

It is straightforward to find the exact answer,

$$\begin{aligned} x &= \frac{2.01 \pm \sqrt{2.01^2 - 4 \times 1.01}}{2} \\ &= \frac{2.01 \pm \sqrt{0.0001}}{2} \\ &= \frac{2.01 \pm 0.01}{2} \\ &= 1.01,\ 1. \end{aligned}$$

Now let us rerun the analysis rounding to 4SFs after each arithmetic operation. First we note that $2.01^2 = 4.0401$ which is equal to 4.040 to 4SFs. Therefore

$$\begin{aligned} x &= \frac{2.01 \pm \sqrt{2.01^2 - 4 \times 1.01}}{2} \\ &= \frac{2.01 \pm \sqrt{4.040 - 4.04}}{2} \\ &= \frac{2.01 \pm 0}{2} \\ &= 1.005,\ 1.005 \end{aligned}$$

In this case, we have an error of 0.5% even with 4SF arithmetic. Although this loss of precision isn't too catastrophic, it has changed the qualitative nature of the solution. We had two different roots when using exact arithmetic, and now they are two identical roots. This has arisen because we subtracted two numbers which are almost equal. It is in these situations where the loss of significant figures is felt most.

An alternative view of the above is to analyse the effect of possible inaccuracies in the coefficients, 2.01 and 1.01. These values could have been obtained from an experiment, and, given that they have been quoted to 3SFs or 2DPs, the likely maximum error in their values is ± 0.005. Let us therefore consider the equation

$$x^2 - (2.01 + \epsilon_1)x + (1.01 + \epsilon_2) = 0.$$

Using exact arithmetic, with the final answers rounded to 3DPs, we get the following possible solutions.

$$\epsilon_1 = 0.005, \quad \epsilon_2 = 0.005: \qquad x = 1.015,\ 1.000,$$

$$\epsilon_1 = 0.005, \quad \epsilon_2 = -0.005: \qquad x = 1.108,\ 0.907,$$

$$\epsilon_1 = -0.005, \quad \epsilon_2 = -0.005: \qquad x = 1.005,\ 1.000,$$

$$\epsilon_1 = -0.005, \quad \epsilon_2 = 0.005: \qquad x = 1.0025 \pm 0.1000i.$$

Note that the last case has complex roots!

So for this example we can get an error of up to 10% in real answers for only a 0.5% change in a coefficient. This is a problem which is very sensitive to the degree of accuracy of the basic data — these are dangerous problems! I emphasize that this is an extreme circumstance, but occasionally these circumstances do arise, and it is important to acquire a good intuition of how accurate one's arithmetic truly is.

0.6.2 Too few significant figures?

This is motivated by the fact that I often see students providing answers with too few significant figures in exams. If the exact answer to a maths problem were to be $1/\sqrt{2}$, then how should we present it? First, I have to come clean and say that I am very happy for the answer to be written as $1/\sqrt{2}$, although some people prefer to have the square root in the numerator: $\sqrt{2}/2$. An alternative notation is $2^{-1/2}$. If this were the final answer, then I would prefer one of these three exact solutions to be written. If someone were to prefer to write it out, then five or six SFs is fine. Hence 0.707107 is generally ok.

Suppose now that we are asked to evaluate $\sin 2\pi$. The great majority of us (assuming that 2π is in radians) would say that the value is zero. The following table shows what happens if 2π is first converted to a number before the sine is computed.

#/SFs	2π	$\sin 2\pi$
2	6.3	0.016814
3	6.28	-0.003185
4	6.283	-0.000185
5	6.2832	0.000015
6	6.28319	0.000005
7	6.283185	-0.000000

So the approximation to 2π using 3 or 4 SFs isn't particularly good given that the maximum value of the sine function is 1 and that the exact value of $\sin 2\pi$ should be zero. It is much better to retain as much accuracy as possible.

0.6.3 Too many significant figures?

The above went into detail about what could happen if we take too few significant figures. But how many is sufficient?

One may quote π to 36 decimal places,

$$\pi \sim 3.141\,592\,653\,589\,793\,238\,462\,643\,383\,279\,502\,884,$$

as a clever party trick, but the last significant digit shown, which is in the 36th decimal place, represents a very incredibly small quantity. Just how small is it? One way of attempting a visualization is to consider 10^{36} sugar crystals where I will assume that each crystal is a cube of side 0.5mm. It doesn't take long to find out that 10^{36} crystals is equivalent to a cube of side $5,000,000$km, which is roughly 12 times the distance of the moon from Earth. If we compare one sugar crystal with this cube, then that is 10^{-36}. Therefore I conclude that 36 significant figures is a waste of effort!

I tend to use roughly 6 SFs for my calculations and I will only increase that if it turns out to be necessary to do so. So keep your error-checking radar on at all times.

0.6.4 How accurate is the usual approximation to the value of π?

Well, we know from school days that $\pi \simeq 22/7$, but how good is this? So we need to compare $22/7 \simeq 3.142857142857...$ with π as given above. To do this we may define the relative error as being,

$$\text{R.E.} = \frac{22/7 - \pi}{\pi},$$

and this comes to 0.00040, i.e. four parts in 10,000, or 0.04 percent, which isn't at all bad.

Is there a better one that isn't too complicated? How about $355/113$? This is 3.1415929204 to 10 decimal places. Here's the comparison:

$$3.1415929204....$$
$$3.14159265358979....$$

The relative error is now about 8.5×10^{-8} which is considerably better.

No doubt you can think of a nice easy way of remembering $355/113$.

Question: How good are the following as approximations to $\sqrt{2}$: $99/70$ and $8119/5741$? I will leave this question to you.

0.6.5 To conclude....

I have just read that lot through again and I reckon that it's somewhat heavy, especially as it is the very first technical material on a Maths unit at university. So perhaps I ought to summarize it in a few recommendations.

1. Check your data. What is its provenance? How accurate is it?
2. Do not do any rounding off until you get to the very final answer.
3. Do not use too few (it's dangerous) or too many (it's pointless) significant figures.
4. Always be on the lookout for potential accuracy issues, e.g. the subtraction of almost equal numbers.

0.7 How to structure a mathematical argument.

The aim of this section is to cover briefly the frequently-neglected subject of the structuring of a mathematical argument and how to write it down.

When one undertakes a piece of mathematics, be it algebra, trigonometry or calculus or a combination of these, then the mathematics should proceed by sequence of logical steps. It is important to be able to write these steps out in a manner which reflects that logic and which allows someone else to be able to follow that logic. A poor example is the following.

Example 0.3 (bad) Evaluate the polynomial $(x + 5)^2 - x + 10$ when $x = 2$.

I have seen the following sequence of steps:

$$(2 + 5)^2 = 49 - 2 = 47 + 10 = 57. \tag{1}$$

The answer is correct, but what do the various equals signs mean in this context? Only the last one is used correctly. This way of doing things is a mathematical version of a 'stream of consciousness' writing.

Apart from the obvious simple substitution, $(2+5)^2 - 2 + 10 = 49 - 2 + 10 = 57$, which is fine, one could be very pedantic and write the problem out this way:

When $x = 2$, then $(x+5)^2 = 49$, and hence

$$\begin{aligned} (x+5)^2 - x + 10 &= 49 - 2 + 10 \\ &= 57. \end{aligned} \tag{2}$$

Example 0.4 (good) Solve the quadratic equation, $x^2 + 2x - 8 = 0$.

We'll do this by completing the square.

$$\begin{array}{lll} & x^2 + 2x - 8 = 0 & \text{Need to add 9 to both sides...} \\ \Rightarrow & x^2 + 2x + 1 = 9 & \text{...since } x^2 + 2x + 1 \text{ is square} \\ \Rightarrow & (x+1)^2 = 9 & \\ \Rightarrow & x + 1 = \pm 3 & \text{taking square roots} \\ \Rightarrow & x = -1 \pm 3 & \\ & \;= -4,\ 2. & \end{array} \tag{3}$$

The 'implies' symbol, $\Rightarrow$, means that the current line follows logically from the preceding line. It must not be used to replace the equals sign. Note that I did not need to use the implies sign on the last line leading to the final solution in Eq. (3), because that was the result of a simple calculation (the same is true for Eq. (2)). Note also the hints on the right, which might aid the reader understand what the logical step was.

So the aim is to write out a mathematical argument the logic of which I (as the marker of your exams) should be able to follow quite easily. This would also mean that you would be able to follow your own work months from now when revising for exams.

One could think of the mathematical argument leading to (3) as being a condensed version of a verbal statement: Given that $x^2 + 2x - 8 = 0$, this implies that it may be rearranged in the form, $x^2 + 2x + 1 = 9$, because the left hand side is now a perfect square and we have a fleeting suspicion that this might be useful to us. Now we take square roots of both sides etc. etc....

0.8 Notations and functions and the avoidance of ambiguity.

0.8.1 Inverse functions.

It is a real pain that inverse functions such as the inverse sine are denoted almost universally by $\sin^{-1}$ these days in the UK — other countries are not afflicted in this way. It happens for other functions too, such as $\cos^{-1}$, $\tan^{-1}$ and $\cosh^{-1}$. There is even a general notation: if $y = f(x)$ then $x = f^{-1}(y)$. The pain is that it looks like a reciprocal.

In the olden days, whenever they were, the inverse sine was written as **arcsin** which isn't at all ambiguous.

If you really wished to write down the reciprocal of $\sin x$, then there are various ways of doing so:

$$\frac{1}{\sin x}, \qquad (\sin x)^{-1} \qquad \text{and} \qquad \text{cosec } x.$$

What about the following four functions; are they the same as one another:

$$(\sin x)^{-2}, \qquad (\sin^{-1} x)^2, \qquad \text{arcsin}^2(x) \qquad \text{and} \qquad \sin^{-2} x?$$

The first is the -2 power of $\sin x$. The second is the square of the inverse sine of x, as is the third. The fourth one is ambiguous. Some would say that it is the same as the first one because the power isn't -1, which might be said to be reserved to denote the inverse function. Others would say that it is the second, although I would hope that almost no-one would fall into that trap. As for me, my view is that, because it is ambiguous, it should never appear; one's mathematical writing *must* never have even the smallest taint of ambiguity.

As for the reciprocal of the inverse sine — I have never seen anyone use this — then the following four will work:

$$(\sin^{-1} x)^{-1}, \qquad \frac{1}{\sin^{-1} x}, \qquad (\text{arcsin}\, x)^{-1}, \qquad \frac{1}{\text{arcsin}\, x}.$$

All of these are correct although the first is an absolute horror.

0.8.2 Double negatives.

Ah, now these have become very popular in the last ten years or so. By this I mean constructions such as,

$$3 - -2 = 5.$$

In one sense this is ok. However, the first minus sign is formally a subtraction operator while the second indicates the negative of 2, so it isn't really a double negative. Just as we saw that -1 as a superscript could mean two different things and that these usages may be confused, here the difficulty/danger is due to where the two minus signs are when written down. I will illustrate that using the following, which is something that I have often witnessed on exam scripts:

$$3 - -2 \quad = \quad 3 - \; - 2 \quad = \quad 3 -- 2 \quad = \quad 3 — 2 \quad = \quad 3 - 2 = 1.$$

The above generally appears on successive lines of an exam script as the more complicated parts of an equation are simplified, but the minus signs just keep getting closer and closer! So although this isn't really an ambiguity, my dislike of the double negative is primarily because this often happens. So I would recommend writing, $3 - (-2) = 5$, as a safe way to play the mathematical game.

Being a bit of a purist (and it's also a Great British sport to complain about irrelevant matters!), I would far prefer the following usage: $3 \times (-2) = -6$, the brackets still signifying the negative number, as opposed to 3×-2 — uuurgh.

0.9 Logarithms.

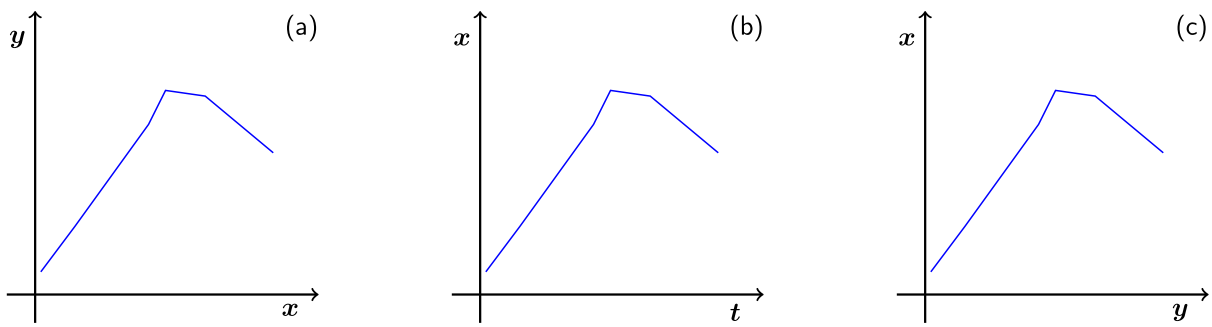
There are three types of logarithm which are used in Engineering. The first is the base-10 logarithm, which is the first one that is taught and which is known as the common logarithm. It is usually written as $y = \log_{10} x$, so if y is increased by 1 (i.e. the addition of 1), then x is multiplied by a factor of 10. Scales such as the Richter Scale behave in the same way: an earthquake which measures 7 on the Richter scale has 10 times the energy of one which measures 6. The amplitude of sound also operates in the same way. I guess that the unit is the *bel* where 10 decibels is equivalent to 1 bel. So a sound of 80 decibels has 100 times the power of one which is rated at 60 decibels.

The natural logarithm, invented by Napier, uses $e = 2.7182818284....$ as its base. This rather strange number arises naturally in very many places and it has many very interesting properties such as the fact that the derivative of the function, e^x, is also e^x. It is generally written as $y = \ln x$, and we simply say that "y is equal to log x" when talking about it, and certainly not in anything involving calculus. There is no need to say, "lun x", because the base-10 version is hardly ever used in Mechanical Engineering.

The base-2 or binary logarithm is written as $y = \log_2 x$. It too is not used often in Mechanical Engineering, but has great use in Computer Science, Information Theory and Music Theory. Given that many of you will also be musicians it is worth dwelling on this for a moment. Violins in an orchestra generally tune their instruments to a note called A which has a frequency of 440Hz. Perhaps this isn't too well-known, but the highest note that the solo violin plays in Dvŏrák's Violin Concerto is also an A, but its frequency is 3520Hz. The ratio of these frequencies is precisely 8 which is also 2^3. Given that $\log_2 8 = 3$, this tells us that this frequency ratio is precisely 3 octaves. Thus the binary logarithm of a frequency ratio tells us the equivalent number of octaves.

[As an aside for the more geeky musicians, a perfect fifth on a stringed instrument corresponds to a frequency ratio of 1.5, whereas the equivalent on an equally-tempered piano, i.e. seven semitones, corresponds to a frequency ratio of $2^{7/12}$ which is 1.498307. Using $\log_2$, the interval for a violin is 0.584963 of an octave, while for the piano it is $7/12 = 0.583333$ (6SFs). This slight difference between these two intervals is why violinists find it difficult to tune to a piano when both the D and the A are sounded. If the violin A is tuned to the piano A then the two D's do not agree, and vice versa. If the violinist tunes the strings perfectly to the piano, then the simultaneous playing of the D and A strings sounds out of tune. Such matters are unresolvable....]

0.10 Naming the axes.



These three graphs look the same. If the first one, labelled (a), were to be written in the form, $y = f(x)$, then the other two are $x = f(t)$ and $x = f(y)$, in turn.

The primary reason for introducing these is the very common but very often inaccurate naming of the vertical axis as the y-axis and the horizontal one as the x-axis. Formally this is correct only for the first one. I have noticed that this can even occur in, say, official government graphs of, to choose an example, the rate of inflation against time, where the rate of inflation is then referred to by the TV reporter as the y-axis and the month by the x-axis. Oops....

In the figure labelled, (b), it seems odd to me that one should say that the y-axis corresponds to the value of x while the x-axis corresponds to the value of t. Even worse, should we say that the y-axis gives the value of x while the x-axis gives the value of y in the figure labelled (c)? Again ambiguity is possible.

One simple resolution is to call them the **vertical and horizontal axes** — no-one will misunderstand you. One could also refer to them as the **ordinate** and the **abscissa**, respectively, which are now rather old-fashioned words that are rarely seen. My recommendation is to use 'vertical' and 'horizontal' because again there is no potential for ambiguity.

The final word. Sometimes one needs 3D graphs. They could take various forms: (i) a projection of of the full 3D curve onto say the (x, y)-plane, (ii) a stereoscopic pair where one needs to go slightly cross-eyed to be able to view its 3D nature, or (iii) an anaglyph form which requires red/blue glasses. So such graphs could be formed of lines (e.g. the trajectory of a particle in 3D space) or of surfaces (e.g. the depiction of the height above sea level of an area of land). The three axes could be (x, y, z) or else (x, y, t) or many others. In all these cases it is best to refer to a chosen axis by the name of the coordinate along that axis. Many examples of stereopairs and anaglyphs may be found using Google.

1 COMPLEX NUMBERS

1.1 Historical Note [For interest only]

As mathematics has developed from early times it has suffered many revolutions in terms of the ideas which are taken as acceptable. Before and up to the time of Pythagoras (approximately 569 BC) it was believed that all quantities could be expressed in terms of whole numbers. What we now call rational fractions were allowed, since they are ratios of whole numbers (e.g. $\frac{3}{4}$), but concepts such as π and $\sqrt{2}$ were unacceptable. Pythagoras himself believed all of this, but this belief sat uncomfortably beside the fact that the diagonal of a square cannot be written in terms of a rational fraction times the side of the square. The technical term for this is that the diagonal is **incommensurate** with its side. Note that the hypotenuse of the right-angled triangle with sides, 3, 4 and 5, is commensurate with the other two sides because the ratios of those sides to the hypotenuse are $\frac{3}{5}$ and $\frac{4}{5}$, i.e. ratios of whole numbers. Neither π nor $\sqrt{2}$ can be written like this.

Returning to the diagonal of a square, Pythagoras's theorem applied to two sides of length 1 which subtend a right angle shows that the length of the hypotenuse is $\sqrt{2}$ — this is well-known. If we wish to prove that this number cannot be expressed as a ratio of whole numbers we may first assume the opposite, namely that $\sqrt{2} = n/m$ where n and m are whole numbers that do not have any common factors. If we square both sides and rearrange slightly we get

$$n^2 = 2m^2. \tag{1.1}$$

As the right hand side is even, this implies that n^2 is even, and, given that only even numbers have squares which are even, this implies that n must be even and so we can write $n = 2p$ where p is a whole number. On substitution into (1.1) we get

$$4p^2 = 2m^2 \qquad \Rightarrow \qquad m^2 = 2p^2. \tag{1.2}$$

In turn this latest expression implies that m is also even. But we have now shown that both n and m have a common factor, 2, which is in contradiction to the original assumption. Therefore $\sqrt{2}$ cannot be expressed as a rational fraction. Consequently it is an example of an **irrational** number.

The scientific community, Pythagoras notwithstanding, eventually accepted irrational numbers since they often arise when solving polynomial equations such as quadratic equations. It has also been shown that numbers like π and e, the base of the natural logarithm, are also irrational, but since they do not arise from solving polynomial equations they are also termed **transcendental**.

Although equations such as

$$x^2 - 6x + 8 = 0 \qquad \text{and} \qquad y^2 - 6y + 7 = 0$$

may be written in the forms

$$(x-3)^2 = 1 \qquad \text{and} \qquad (y-3)^2 = 2$$

and hence have the solutions,

$$x = 2, 4 \qquad \text{and} \qquad y = 3 \pm \sqrt{2},$$

equations such as,

$$z^2 - 6z + 10 = 0, \qquad \text{or} \qquad (z-3)^2 = -1, \tag{1.3}$$

were deemed for a long time not to have a solution. The reason simply was that $z = 3 \pm \sqrt{-1}$ was considered to be nonsensical because there is no number whose square could possibly be equal to -1.

Such a state of affairs lasted for quite some time, and it was only around the time when Cardano (1501-1576) discovered how to solve a cubic equation of the form, $ax^3 + bx^2 + cx + d = 0$, that the resistance to $\sqrt{-1}$ finally faded. The reason for this is that, while a cubic equation may have either one real root or three real roots, Cardan's formula giving the three roots, when they are all real, involves using the square roots of negative numbers; the detailed formula is given in the next section.

The following is a quotation from *Significant Figures* by Ian Stewart which I think is of interest in this context.

> *Ars Magna* [by Cardano] is significant for one other reason. Cardano applied his algebraic methods to find two numbers whose sum is 10 and product is 40, and got the answer $5+\sqrt{-15}$ and $5-\sqrt{-15}$. Since negative numbers have no square roots, he declared this result to be 'as subtle as it is useless'. The formula for cubics also leads to such quantities when all three solutions are real, and in 1572 Rafael Bombelli observed that if you ignore what such expressions mean and just do the sums, you get the correct real solutions. Eventually this line of thinking led to the creation of the system of complex numbers in which -1 has a square root. This extension of the real number system is essential to today's mathematics, physics and engineering.

Thus it was that the symbol i (or, when within the engineering fraternity, j) came to signify the "value" of $\sqrt{-1}$, since it could at least be interpreted as a mathematical trick to obtain results which are nevertheless correct. The symbols i and j denote imaginary numbers, and, together with the real numbers, are termed complex numbers. The solution of Equation (1.3) for z may therefore be written in the form $z = 3 \pm j$.

It is interesting to note throughout this very brief historical sketch the type of words that have been used to denote the new kinds of numbers: irrational, transcendental and imaginary! Perhaps these mark a certain amount of distrust or perhaps awe when compared with nice safe "rational" numbers. Much more recently the term "surreal numbers" has been invented (see wikipedia) and so the old tradition of curious names continues.

1.2 Cardan's formula for the cubic equation [For interest only]

The solution for the **quadratic equation** is well-known. If z satisfies

$$az^2 + bz + c = 0, \tag{1.4}$$

then

$$z_1 = \frac{-b+\sqrt{b^2-4ac}}{2a} \qquad z_2 = \frac{-b-\sqrt{b^2-4ac}}{2a}, \tag{1.5}$$

where, if we set $q = b^2 - 4ac$, we have the three cases

$$\begin{array}{rcc} q > 0 & \Rightarrow & \text{two real roots} \\ q = 0 & \Rightarrow & \text{two equal real roots} \\ q < 0 & \Rightarrow & \text{two complex roots.} \end{array} \tag{1.6}$$

The value, q, is called the discriminant: its value allows us to distinguish between the three potential cases.

The above categorisation holds only when a, b and c are real. But the formulae for z_1 and z_2 remain valid for complex values of a, b and c.

For the **cubic equation**, $z^3 + a_2 z^2 + a_1 z + a_0 = 0$, we let

$$q = \tfrac{1}{3}a_1 - \tfrac{1}{9}a_2^2 \qquad r = \tfrac{1}{6}(a_1 a_2 - 3a_0) - \tfrac{1}{27}a_2^3. \tag{1.7}$$

Again we have three cases:

$$\begin{array}{rcc} q^3 + r^2 > 0 & \Rightarrow & \text{one real root and a pair of complex roots} \\ q^3 + r^2 = 0 & \Rightarrow & \text{all roots are real with at least two equal} \\ q^3 + r^2 < 0 & \Rightarrow & \text{all roots are real.} \end{array} \tag{1.8}$$

To complete the solution we set

$$s_1 = \left[r + \sqrt{(q^3 + r^2)}\right]^{1/3} \qquad s_2 = \left[r - \sqrt{(q^3 + r^2)}\right]^{1/3} \tag{1.9}$$

and then

$$\begin{aligned} z_1 &= (s_1 + s_2) - \tfrac{1}{3}a_2 \\ z_2 &= -\tfrac{1}{2}(s_1 + s_2) - \tfrac{1}{3}a_2 + \tfrac{1}{2}\sqrt{3}(s_1 - s_2)j \\ z_3 &= -\tfrac{1}{2}(s_1 + s_2) - \tfrac{1}{3}a_2 - \tfrac{1}{2}\sqrt{3}(s_1 - s_2)j. \end{aligned} \tag{1.10}$$

There is a much longer formula for the solution of quartic equations (i.e. fourth order polynomials) but there are no general formulae for the solution of quintics or of higher order polynomials. This was proved in 1839 by Abel.

That completes the historical stuff! From now on we shall look at the properties of complex numbers, determine how they may be used and manipulated, and then finally we'll consider a general method for finding the roots of complex numbers (by which we mean the fractional powers of complex numbers such as square roots).

1.3 Manipulation of Complex Numbers

The basic rules for adding, subtracting, multiplying and dividing complex numbers are not too far removed from those involving solely real numbers, except that we will often need to use the result that $j^2 = -1$, and that real and imaginary numbers are usually treated separately.

Addition and Subtraction:
Here we add and subtract in the way one would expect by collecting like terms. For example

$$(1 + 5j) + (2 - 3j) \quad = \quad (1 + 2) + (5 - 3)j \quad = \quad 3 + 2j, \tag{1.11}$$

and

$$(1 + 5j) - (2 - 3j) \quad = \quad (1 - 2) + (5 + 3)j \quad = \quad -1 + 8j. \tag{1.12}$$

So these work in the same way as $(x + 5y) + (2x - 3y) = 3x + 2y$.

Multiplication:
Here we expand the product in the same way as we do for the product,

$$(a + b)(c + d) = ac + bd + ad + bc,$$

except that a little tidying up takes place afterwards using $j^2 = -1$. In general, we have,

$$(a + bj)(c + dj) = ac + bd\,j^2 + adj + bcj = (ac - bd) + (ad + bc)j. \tag{1.13}$$

A numerical example:

$$\begin{aligned} (1 + 5j)(2 + 3j) &= (1 \times 2) + (5j \times 3j) + (1 \times 3j) + (5j \times 2) \\ &= 2 + 15\,j^2 + 3j + 10j \\ &= 2 - 15 + 13j \\ &= -13 + 13j. \end{aligned} \tag{1.14}$$

Note: Integer powers of complex numbers (like squares and cubes and 10th powers etc.) may be evaluated this way but there is a better way which will follow later.

Division:
This is the most complicated operation of the four and it is described best using the concept of the complex conjugate. If $z = x + yj$, then the **complex conjugate** of z is written as $\overline{z}$ and defined as

$$\overline{z} = x - yj. \tag{1.15}$$

Clearly $\overline{3+4j} = 3 - 4j$, $\overline{6} = 6$ and $\overline{5j} = -5j$. One very useful consequence of the complex conjugate is that the product, $z\overline{z}$, is real:

$$\begin{aligned} z\overline{z} &= (x+yj)(x-yj) && \text{by definition} \\ &= x^2 - y^2 j^2 + xyj - xyj && \text{multiplying out} \\ &= x^2 + y^2 && \text{using } j^2 = -1. \end{aligned} \tag{1.16}$$

This provides us with a sneaky way for evaluating the quotient of two complex numbers: we multiply the denominator and the numerator by the complex conjugate of the denominator, and this guarantees that the new denominator is real. Here is an example:

$$\begin{aligned} \frac{1}{3+4j} &= \frac{1}{3+4j} \times \frac{3-4j}{3-4j} \\ &= \frac{3-4j}{(3+4j)(3-4j)} \\ &= \frac{3-4j}{25} \\ &= \frac{3}{25} - \frac{4}{25}j. \end{aligned} \tag{1.17}$$

In general we have

$$\frac{1}{x+yj} = \frac{x-yj}{(x+yj)(x-yj)} = \frac{x-yj}{x^2+y^2} = \frac{x}{x^2+y^2} - \frac{y}{x^2+y^2}j. \tag{1.18}$$

Another numerical example is:

$$\frac{2+j}{1+j} = \frac{(2+j)(1-j)}{(1+j)(1-j)} = \frac{3-j}{2} = \tfrac{3}{2} - \tfrac{1}{2}j.$$

Note: This idea of using a complex conjugate may also help us in an entirely different context, one which involves only real numbers. Here's an example:

$$\begin{aligned} \frac{8}{1+\sqrt{5}} &= \frac{8}{(1+\sqrt{5})} \times \frac{(1-\sqrt{5})}{(1-\sqrt{5})} \\ &= \frac{8(1-\sqrt{5})}{1-5} = 2(\sqrt{5}-1). \end{aligned} \tag{1.19}$$

1.4 Geometrical Interpretation

This will give us some further physical understanding of complex numbers, but will also allow us to do things like finding square roots, tenth roots and $2/5$th powers of complex numbers.

The complex number $z = a + bj$ may be represented as the coordinates of a point in what is called the **complex plane**:

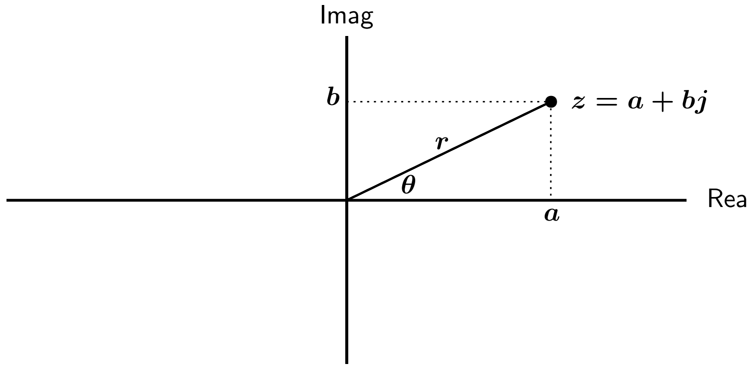


Figure 1.1. Showing the location of the complex number, $z = a + bj$, in the Argand Diagram (or Complex Plane). Also shown are $r = |z| = \sqrt{a^2 + b^2}$ and $\theta = \arg(z) = \tan^{-1}(b/a)$.

The horizontal axis (abscissa) is called the **real axis**, while the vertical axis (ordinate) is called the **imaginary axis**. The whole Figure is also referred to as the **Argand diagram**.

The length of the line joining the origin to z is $\sqrt{a^2 + b^2}$ which is precisely $\sqrt{z\overline{z}}$; this is called the **modulus** of z and is denoted by $|z|$.

The angle that the line joining the origin to z makes with the real axis (measured anti-clockwise) is called the **argument** of z, and is denoted as $\arg z$. If we let $\theta = \arg(z)$ then $\tan\theta = b/a$. Conversely we have $\arg(z) = \tan^{-1}(b/a)$, although it is necessary to specify which of the two possible angles is correct. For example, if $b = a = 1$, then the diagram indicates that $\arg z = \pi/4$, but the inverse tan formula simply says that $\theta = \tan^{-1}(1)$ for which we may have either $\theta = \pi/4$ or $\theta = 5\pi/4$; therefore we either need to state in addition that we require θ to lie in the first quadrant (i.e. $0 \leq \theta \leq \pi/2$), or else quote the correct value in radians. **We note that θ is always measured in the anticlockwise direction.**

The conjugate of z may be located by taking the mirror image of z in the real axis; see Figure 1.2. From this diagram we see that $|z| = |\overline{z}|$ and that $\arg(z) = -\arg(\overline{z})$.

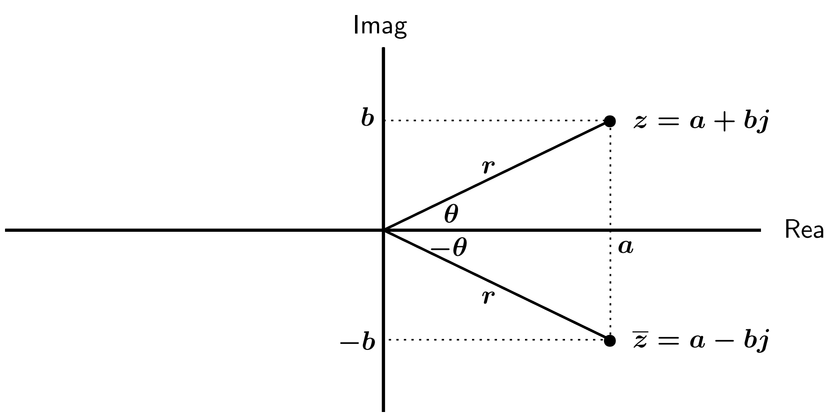


Figure 1.2. Showing the locations of the complex number, $z = a + bj$ and its conjugate, $\overline{z} = a - bj$, in the Complex Plane.

Addition: In the Argand diagram this appears to follow the same rule as the addition of two vectors in that the sum is obtained by completing the rhombus; see Figure 1.3.

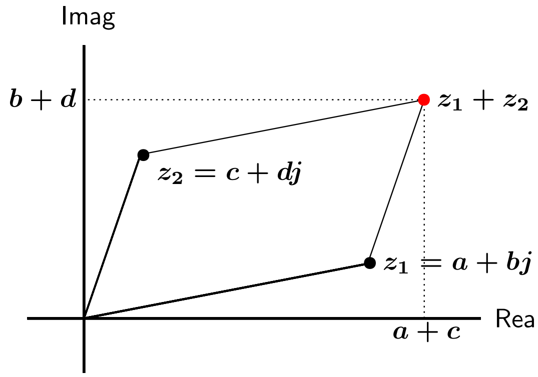


Figure 1.3. Showing how two complex numbers, $z_1 = a+bj$ and $z_2 = c+dj$, may be added by completing the rhombus. Equivalently, we add the real and imaginary parts separately.

Multiplication: For example, we have $(2+4j)\times(1+j) = -2+6j$. On the Complex Plane we have:

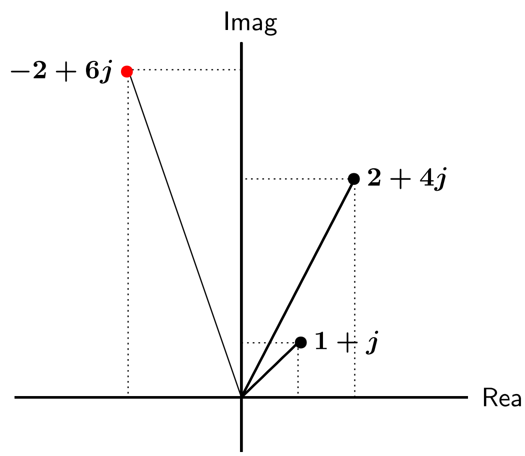


Figure 1.4. Showing the multiplication of the two given complex numbers.

From this Figure it is not immediately obvious what is happening geometrically when we multiply complex numbers. However, if we find the moduli of each of the three numbers:

$$|2+4j| = \sqrt{20}, \qquad |1+j| = \sqrt{2}, \qquad |-2+6j| = \sqrt{40} = \sqrt{2}\times\sqrt{20},$$

then we see that the modulus of the product is equal to the product of the moduli. It is also clear from the Argand diagram that the argument of the product is greater than the (positive) argument of individual numbers forming the product.

We may investigate this in more detail by first considering a partial translation of a complex number into a polar coordinate form. In Fig. 1.1 we can see that

$$a = r\cos\theta \quad \text{and} \quad b = r\sin\theta,$$

or that,

$$a+bj = r(\cos\theta + j\sin\theta). \tag{1.20}$$

Therefore we may write down similar general forms for the following two complex numbers:

$$z_1 = r_1(\cos\theta_1 + j\sin\theta_1) \qquad \text{and} \qquad z_2 = r_2(\cos\theta_2 + j\sin\theta_2). \tag{1.21}$$

Clearly, the moduli of these numbers are r_1 and r_2, respectively, and their arguments are θ_1 and θ_2. Hence

$$\begin{aligned} z_1 z_2 &= r_1 r_2 (\cos\theta_1 + j\sin\theta_1)(\cos\theta_2 + j\sin\theta_2), \\ &= r_1 r_2 \Big[(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + j(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2) \Big], \\ &= r_1 r_2 \Big[\cos(\theta_1+\theta_2) + j\sin(\theta_1+\theta_2) \Big]. \end{aligned} \tag{1.22}$$

where the last line was formed by appealing to some multiple-angle formulae.

This result shows that the **modulus of the product is equal to the *product* of the moduli**, but also that the **argument of the product is equal to the *sum* of the arguments**. This leads us into what is called the polar (or exponential or complex exponential) representation of complex numbers.

1.5 Complex Exponential form

Recalling the expression given in Eq. (1.20), we may write down the following:

$$\begin{aligned} z &= a + bj && \text{by definition} \\ &= (r\cos\theta) + j(r\sin\theta) && \text{using Eq. (1.20)} \\ &= r(\cos\theta + j\sin\theta) \\ &= re^{j\theta}. \end{aligned} \tag{1.23}$$

This last step of rewriting the coefficient of r as a complex exponential needs a little explanation, but this cannot be done satisfactorily without resorting to either Taylor Series, which is a device for writing functions as an infinitely long polynomial, or to the theory of Ordinary Differential Equations; these will be covered later in ME10304 this semester and ME10305 (Mathematics 2) next semester, respectively. You will need to take Eq. (1.23) on trust for now....

Thus complex numbers may be written in either Cartesian form $\boxed{z = a + bj}$ or Polar/Exponential form $\boxed{z = re^{j\theta}}$. [The word, Cartesian, is always spelt with a capital because it is named for Descartes, otherwise known as Cartesius in its Latinized form.] An alternative notation is $\boxed{z = r\angle\theta}$; this is often used in Electrical Engineering contexts, where θ can be allowed to be in degrees. Another occasional notation is $\boxed{r \,\text{cis}\, \theta}$, where cis is a shorthand for $\cos\theta + i\sin\theta$ — this was invented by Hamilton in 1866, but is rarely used now. So the complex exponential form is the generally accepted way.

Warning: When writing $re^{j\theta}$, the angle, θ, **must** always be measured in radians. The same is true for sine and consine. The reason is that derivatives are incorrect when using degrees.

Addition and subtraction are most easily carried out in Cartesian form, while multiplication and division are most easily carried out in Polar form:

$$z_1 z_2 = \Big[r_1 e^{j\theta_1}\Big]\Big[r_2 e^{j\theta_2}\Big] = (r_1 r_2)\, e^{j(\theta_1+\theta_2)},$$

$$z_1/z_2 = \Big[r_1 e^{j\theta_1}\Big]\Big/\Big[r_2 e^{j\theta_2}\Big] = (r_1/r_2) e^{j(\theta_1-\theta_2)}.$$

That said, multiplication and division while in Cartesian form are also pretty straightforward.

The conversion of a complex number from polar/exponential form to Cartesian form is straightforward if one remembers the geometry of the Argand diagram:

$$2e^{j\pi/4} = 2\Big[\cos(\pi/4) + j\sin(\pi/4)\Big] = 2\left[\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\right] = \sqrt{2} + \sqrt{2}j.$$

But the conversion from Cartesian to polar/exponential form requires a little more care due to the fact that the inverse tangent function has more than one solution. For example, $\tan^{-1} 1 = \frac{1}{4}\pi$ or $\frac{5}{4}\pi$. Thus conversion from Cartesian to polar form also requires one to choose the correct inverse tangent.

Example 1.1: Express $-1 + 2j$ in complex exponential form.

We have $-1 + 2j = re^{j\theta}$. Here $r^2 = (-1)^2 + (2)^2 = 5 \Rightarrow r = \sqrt{5}$. The argument satisfies $\tan\theta = (2)/(-1) = -2$, and therefore θ is either -1.107149 (directly from my calculator) or 2.034444 (which is π plus the first value). Given where the complex number lies in the complex plane, i.e. the second quadrant, then $\theta = 2.034444$ or 116.565°.

Note: It is usual to quote angles to lie in the range, $-\pi < \theta \le \pi$ although there may very occasionally be the need to use $0 \le \theta < 2\pi$.

1.6 de Moivre's Theorem

This is a very useful result and is known as a theorem despite the fact its proof takes only two lines (i.e. just one step). It is based on the following well-known property of the powers of exponentials,

$$(a^b)^c = a^{bc}. \tag{1.24}$$

We apply this result to $e^{\theta j}$:

$$\begin{aligned} (e^{\theta j})^n &= e^{n\theta j} \\ \Rightarrow \qquad \Big[\cos\theta + j\sin\theta\Big]^n &= \cos(n\theta) + j\sin(n\theta). \end{aligned} \tag{1.25}$$

From this we may recover the various multiple angle formulae from trigonometry. For example, when $n = 2$ we have

$$\begin{aligned} \cos 2\theta + j\sin 2\theta &= \Big[\cos\theta + j\sin\theta\Big]^2 \\ &= [\cos^2\theta - \sin^2\theta] + j[2\sin\theta\cos\theta]. \end{aligned} \tag{1.26}$$

On equating real and imaginary parts, respectively, we get

$$\cos 2\theta = \cos^2\theta - \sin^2\theta, \qquad \sin 2\theta = 2\sin\theta\cos\theta.$$

When $n = 3$, we have

$$\begin{aligned} \cos 3\theta + j\sin 3\theta &= \Big[\cos\theta + j\sin\theta\Big]^3 \\ &= \cos^3\theta + 3j\cos^2\theta\sin\theta + 3j^2\cos\theta\sin^2\theta + j^3\sin^3\theta \\ &= [\cos^3\theta - 3\cos\theta\sin^2\theta] + j[3\cos^2\theta\sin\theta - \sin^3\theta]. \end{aligned} \tag{1.27}$$

This final line used the fact that $j^3 = j^2 j = -j$, and it may also be tidied up to give

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta, \qquad \sin 3\theta = 3\sin\theta - 4\sin^3\theta.$$

1.7 Roots of complex numbers

It is well-known that there are two square roots of positive numbers and, when that number is 9, then those roots are 3 and -3. But both negative numbers and complex numbers also have two square roots. So let us check out a few examples of square roots and other roots.

Example 1.2. Given that both $j^2 = -1$ and $(-j)^2 = -1$, it is clear that the two square roots of -1 are $\pm j$.

Example 1.3. Likwise, given that $(1+j)^2 = 2j$ and $(-1-j)^2 = 2j,$ then the square roots of $2j$ are $\pm(1+j)$.

Example 1.4. Once more, given that $(2+j)^2 = 3+4j$ and $(-2-j)^2 = 3+4j,$ then the square roots of $3+4j$ are $\pm(2+j)$.

The pattern so far is that each square root is precisely the negative of the other, which is exactly the same as for the square roots of positive numbers. No surprise there! Here's a different case:

Example 1.5. We have,

$$(1+j)^4 = -4, \qquad (-1-j)^4 = -4, \qquad (1-j)^4 = -4 \qquad \text{and} \qquad (-1+j)^4 = -4.$$

Therefore we have shown that -4 has four fourth roots:

$$(-4)^{1/4} = \pm 1 \pm j,$$

where all four possible combinations of plus/minus may be taken.

Example 1.6. One more case with a fourth power:

$$(1+2j)^4 = -7-24j, \qquad (-1-2j)^4 = -7-24j,$$
$$(2-j)^4 = -7-24j \qquad \text{and} \qquad (-2+j)^4 = -7-24j.$$

Therefore the four fourth roots of $(-7-24j)$ are

$$\pm(1+2j) \qquad \text{and} \qquad \pm\,(2-j).$$

Perhaps this still isn't sufficient to see the full pattern yet, although the fact that complex numbers have two square roots and four fourth roots does suggest something. So here is something a little different:

Example 1.7. Consider the following:

$$(-2)^3 = -8, \qquad (1+\sqrt{3}j)^3 = -8 \qquad \text{and} \qquad (1-\sqrt{3}j)^3 = -8.$$

So we have three third roots of -8, namely

$$-2, \quad 1+\sqrt{3}j \quad \text{and} \quad 1-\sqrt{3}j$$

.

Example 1.8. If one were to find the sixth powers of 2, -2, $1+\sqrt{3}j$, $1-\sqrt{3}j$, $-1+\sqrt{3}j$ and $-1-\sqrt{3}j$, then each would yield 64. So 64 has six sixth roots.

So that is a pattern, namely that there will be n n^{th} roots for a complex number. However, that is not the only pattern. The following Figure shows how the roots given within each example above are related to one another.

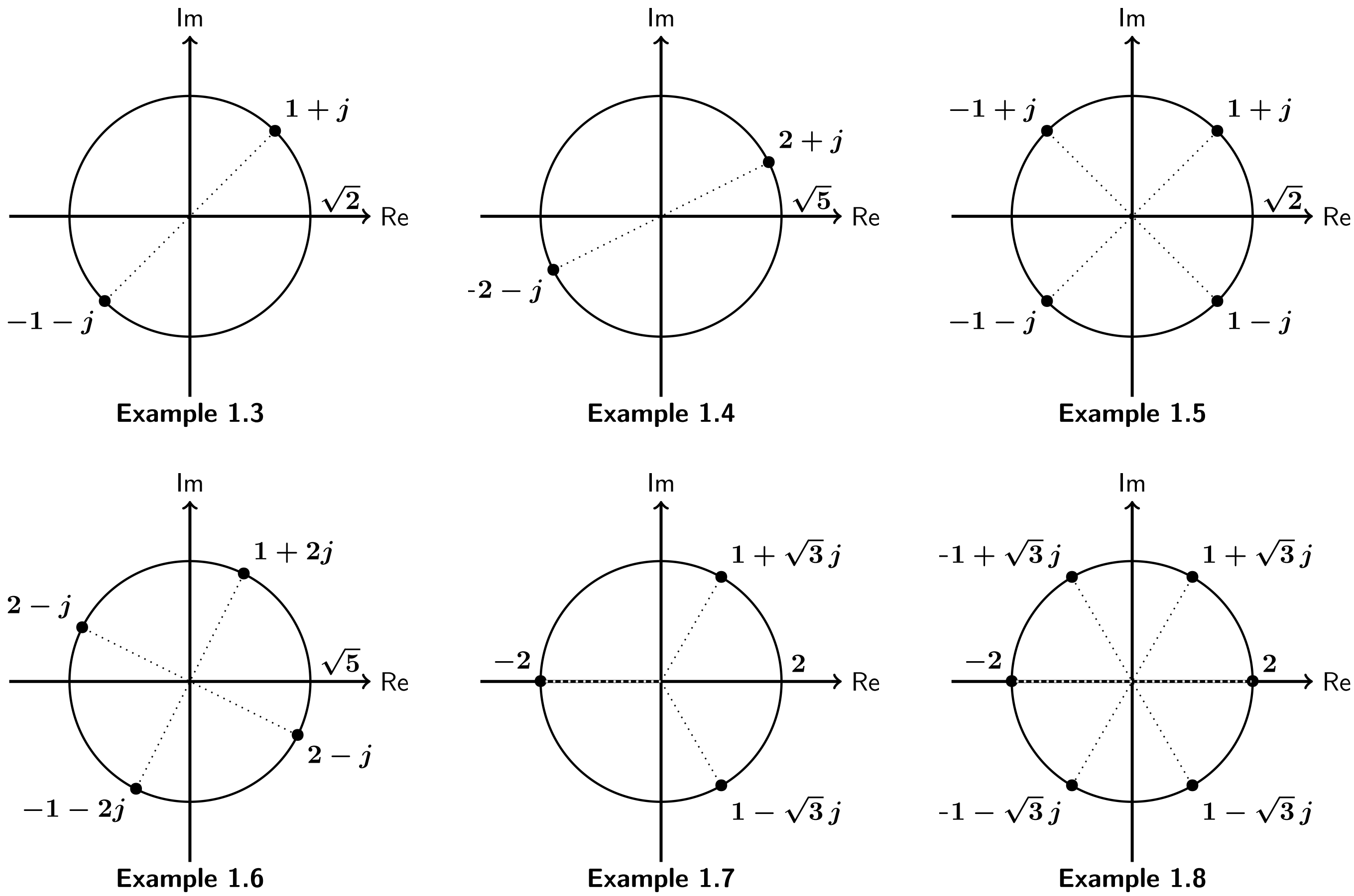


Figure 1.5. Displaying the roots corresponding to Examples 1.3 to 1.8.

The pattern that we need may now be seen in Fig. 1.5 and, for each case, the circle on which all the roots lie is split into equal segments.

The general method. We may motivate the general method using the data in Example 1.5. The four roots are given by

$$\sqrt{2}e^{j\pi/4}, \quad \sqrt{2}e^{3j\pi/4}, \quad \sqrt{2}e^{5j\pi/4} \quad \text{and} \quad \sqrt{2}e^{7j\pi/4}. \tag{1.28}$$

If we raise all of these to the 4th power, then we obtain,

$$4e^{j\pi}, \quad 4e^{3j\pi}, \quad 4e^{5j\pi} \quad \text{and} \quad 4e^{7j\pi}, \tag{1.29}$$

respectively. The arguments of these numbers are π, 3π, 5π and 7π, i.e. they are separated by 2π, and hence each one represents the same point in the Complex Plane. We may write these four values in the alternative form:

$$4e^{(1+2n)\pi j}, \qquad n = 0, 1, 2, 3, \tag{1.30}$$

and therefore the roots listed in Eq. (1.28) may be written as,

$$z_n = \sqrt{2}e^{(1+2n)\pi j/4}, \qquad n = 0, 1, 2, 3. \tag{1.31}$$

Note: In Eq. (1.31) n takes four successive values. Given the way we have motivated the method it is natural that the first value of n will be zero. However, any four successive values of n will be ok.

Note: The general method for finding roots uses the above approach but does it backwards! Three examples follow.

Example 1.9. Find the square roots of j.

We may write j in polar form as either $e^{(\pi/2)j}$ (the obvious one where the argument is $\pi/2$) or as $e^{(5\pi/2)j}$. We may combine these into the form:

$$e^{(\pi/2+2n\pi)j} \quad \text{where} \quad n = 0 \text{ or } n = 1.$$

These two forms for j lie at the same point in the Complex Plane and must indeed do so! However, they will cease to do so when we take the square roots:

$$\sqrt{j} \;=\; e^{(\pi/4)j} \qquad \text{or} \qquad e^{(5\pi/4)j}. \tag{1.32}$$

For this particular example we can write the Cartesian form without the use of a calculator! We have,

$$\begin{aligned} \sqrt{j} \;&=\; \Big(\cos\tfrac{1}{4}\pi + j\sin\tfrac{1}{4}\pi\Big) \qquad \text{or} \qquad \Big(\cos\tfrac{5}{4}\pi + j\sin\tfrac{5}{4}\pi\Big) \\ &=\; \pm\Big(\frac{1+j}{\sqrt{2}}\Big). \end{aligned} \tag{1.33}$$

Example 1.10. Find the fifth roots of $(3+4j)$.

In polar form we have

$$z = 5e^{\theta j},\; 5e^{(\theta+2\pi)j},\; 5e^{(\theta+4\pi)j},\; 5e^{(\theta+6\pi)j},\; 5e^{(\theta+8\pi)j}$$

where we have written five different expressions for z, and where $\tan\theta = \frac{4}{3}$ with θ lying in the first quadrant (i.e. $\theta = 0.927295$). The five fifth roots are

$$z = 5^{1/5}e^{(\theta+2n\pi)j/5} \quad \text{for } n = 0,1,2,3,4.$$

If we were to plot these in the Complex Plane then we would see that the five points are equally distributed around a circle of radius $5^{1/5}$ centred on the origin.

Example 1.11. Find $z = (4+3j)^{2/3}$.

For such a power, it is probably better to find the square first and then take the cube root, although it could be done the other way around. Perhaps I ought to show that they give the same solutions. So we'll square first...

First way: Let

$$z = \Big[(4+3j)^2\Big]^{1/3} = (7+24j)^{1/3}.$$

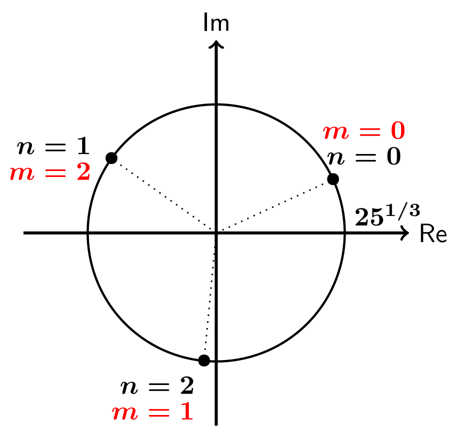
In polar form we have $7+24j = 25e^{(\theta+2n\pi)j}$ for $n = 0,1,2$, where $\tan\theta = \frac{24}{7}$ and θ lies in the first quadrant, we note that the modulus of $(7+24j)$ is 25. Therefore

$$z = 25^{1/3}e^{(\theta+2n\pi)j/3} \quad \text{for } n = 0,1,2 \text{ and where } \theta = 1.287002. \tag{1.34}$$

Second way: Let
$$z = (4+3j)^{2/3}.$$
In polar form we have $4+3j = 5e^{(\phi+2m\pi)j}$ for $m = 0, 1, 2$, where $\tan\phi = \frac{3}{4}$ and where ϕ also lies in the first quadrant. We have used m here instead of n, and ϕ instead of θ to distinguish the two analyses. The modulus of $(4+3j)$ is 5. Therefore

$$z = 5^{2/3} e^{2(\phi+2m\pi)j/3} \quad \text{for } m = 0, 1, 2 \text{ and where } \phi = 0.643501. \tag{1.35}$$

The following figure shows where the roots are as a consequence of these two analyses. The locations of the roots are identical, but manner in which they are counted is different.



Example 1.11

Note: Again I emphasize in the strongest terms that θ MUST be given in radians when using polar form.

1.8 Irrational exponents [For interest only]

Here we are thinking of powers such as $z^{\sqrt{2}}$.

We note that irrational numbers, such as $\sqrt{2}$, may not be expressed as the ratio of two whole numbers, whereas the method we have used above depends on this. Although $\sqrt{2}$ may be approximated by a rational fraction as closely as we might desire, then the more accurate the approximation is the larger is the denominator. For example, the absolute error in the fraction, $\frac{99}{70} = 1.41428571$ (8DP), is less than 10^{-4} and it is the first fraction (as the denominator increases) for which the error is smaller than 10^{-4}. For the errors, 10^{-6} and 10^{-8}, the corresponding fractions are,

$$\sqrt{2} \simeq \tfrac{1393}{985} = 1.41421320 \text{ (8DP)} \qquad \text{and} \qquad \sqrt{2} \simeq \tfrac{19601}{13860} = 1.41421356 \text{ (8DP)}.$$

Therefore increasingly accurate representations of $\sqrt{2}$ result in an increasing number of legitimate roots on a circle in the complex plane: 70, 985 and 13860 in our three examples. But typically the context will dictate if this is sensible, and it generally will not be. Therefore, in such situations we may restrict ourselves to what is called the **principal value** of the complex exponential. By this is meant that, if $z = re^{\theta j}$, then $z^{\sqrt{2}} = r^{\sqrt{2}} e^{\sqrt{2}\theta j}$. However, do be assured that such matters will not arise in engineering applications.

1.9 Relationship with the hyperbolic functions.

The functions $\cos\theta$ and $\sin\theta$ are known as either trigonometric or circular functions, while $\cosh\theta$ and $\sinh\theta$ are the hyperbolic functions. Given that

$$e^{\theta j} = \cos\theta + j\sin\theta \qquad \text{and} \qquad e^{-\theta j} = \cos\theta - j\sin\theta,$$

we may add and subtract these to obtain the relations,

$$\cos\theta = \tfrac{1}{2}\Big[e^{\theta j} + e^{-\theta j}\Big] \qquad \text{and} \qquad \sin\theta = \tfrac{1}{2j}\Big[e^{\theta j} - e^{-\theta j}\Big].$$

These are reminiscent of

$$\cosh\theta = \tfrac{1}{2}\Big[e^{\theta} + e^{-\theta}\Big] \qquad \text{and} \qquad \sinh\theta = \tfrac{1}{2}\Big[e^{\theta} - e^{-\theta}\Big],$$

and explains why there is so much similarity between the circular and hyperbolic functions, particularly when dealing with calculus and differential equations, even though the circular and the hyperbolic functions look so different from one another.

2 DIFFERENTIATION

If we are given a function, $y = y(t)$, where the value of y is a function of t, then we say that y is the **dependent variable** and t is the **independent variable**. This is because the value that y takes *depends* on the value of t. If we interpret y to mean distance and t to mean time, then the dependence of y on t also makes physical sense. The following graph shows a typical function $y(t)$.

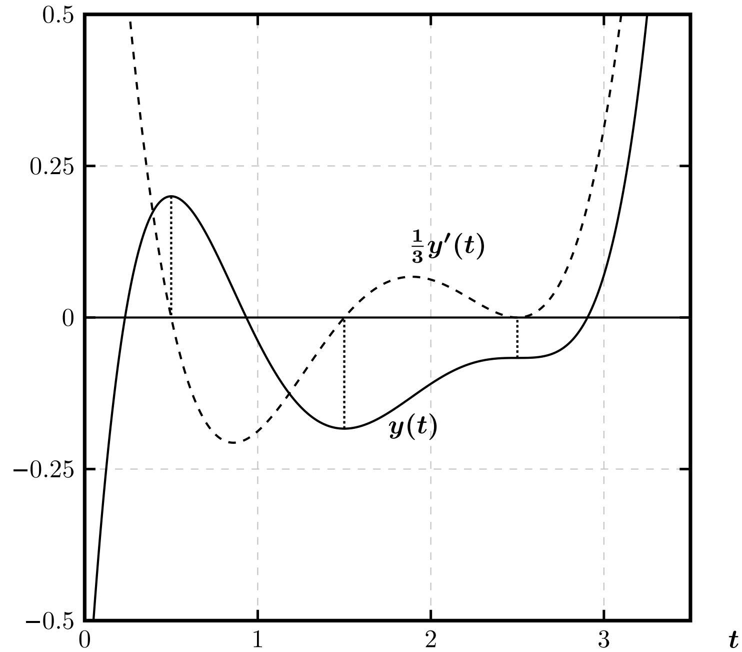


Figure 2.1. Depicting a typical function, $y(t)$ (continuous curve), and its derivative, $y'(t)$ (dashed curve). The dotted lines show where $y(t)$ has a zero slope.

The derivative of $y(t)$ at any point is defined as the slope or gradient of the tangent touching the curve at that point. Places where $y(t)$ has a horizontal tangent are places where the derivative is zero. The gradient of y is also shown in Figure 2.1; note the behaviour of the gradient as compared with the function, especially where the gradient is zero, i.e. where it achieves its maximum or minimum values. The point of inflexion at $t = 2.5$ corresponds to a double zero in $y'(t)$, and therefore $y''(2.5) = 0$ as well.

2.1 First derivative

This derivative is written in various ways:

$$\frac{dy}{dt}, \qquad y'(t), \qquad y', \qquad Dy. \tag{2.1}$$

The first three of these are very common alternatives, while the last occurs only in certain specialised circumstances. The third one is often used as a short-hand form when the identity of the independent variable isn't in question. However, the first notation (that of Leibniz) is reminiscent of one way of defining the derivative mathematically using the notion of limits, as will be illustrated in Figure 2.2, later.

Although we have the common-sense definition that the gradient at a point is the slope of the tangent to the curve at that point, it requires calculus to find that tangent. So it makes sense that we can approximate the gradient of y at $t = t_0$ by the gradient of the straight line joining $y(t_0)$ and $y(t_0 + \delta t)$ since we can do that without calculus; this may be seen in Fig 2.2.

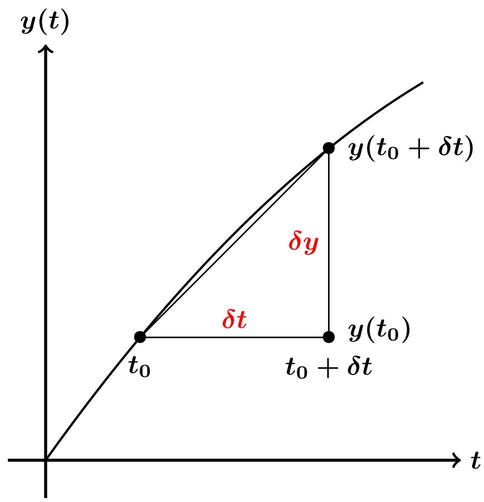


Figure 2.2. A sketch of how a derivative at a point may be approximated by the limit of a slope between two points on a curve.

If the change in y over this interval in time is denoted by δy, then we may define the derivative as

$$\left.\frac{dy}{dt}\right|_{t=t_0} = \lim_{\delta t\to 0}\frac{\delta y}{\delta t}. \tag{2.2}$$

Using the fact that $\delta y = y(t_0+\delta t) - y(t_0)$ we may also write

$$\left.\frac{dy}{dt}\right|_{t_0} = \lim_{\delta t\to 0}\frac{y(t_0+\delta t)-y(t_0)}{\delta t}. \tag{2.3}$$

If we now imagine what happens to the straight line in Fig. 2.2 as δt gets smaller, we should be able to accept that it approximates the curve increasingly well.

Note that the definition of dy/dt given in Eq. (2.3) is not unique — we may also define the derivative in the following way

$$\left.\frac{dy}{dt}\right|_{t_0} = \lim_{\delta t\to 0}\frac{y(t_0+\delta t/2)-y(t_0-\delta t/2)}{\delta t}, \tag{2.4}$$

where the target value of t is halfway between the points where y is evaluated, but the end result is exactly the same. To demonstrate this, we will apply both Eqs. (2.3) and (2.4) to the function $y=t^2$. In the case of Eq. (2.3) we have

$$\begin{aligned}
\frac{dy}{dt} &= \lim_{\delta t\to 0}\left[\frac{(t+\delta t)^2-t^2}{\delta t}\right]\\
&= \lim_{\delta t\to 0}\left[\frac{(t^2+2t\,\delta t+\delta t^2)-t^2}{\delta t}\right]\\
&= \lim_{\delta t\to 0}\left[\frac{2t\,\delta t+\delta t^2}{\delta t}\right]\\
&= \lim_{\delta t\to 0}\Big[2t+\delta t\Big] = 2t.
\end{aligned} \tag{2.5}$$

Using Eq. (2.4) we have,

$$\begin{aligned}\frac{dy}{dt} &= \lim_{\delta t\to 0}\Big[\frac{(t+\delta t/2)^2-(t-\delta t/2)^2}{\delta t}\Big] \\ &= \lim_{\delta t\to 0}\Big[\frac{(t^2+t\,\delta t+\delta t^2/4)-(t^2-t\,\delta t+\delta t^2/4)}{\delta t}\Big] \\ &= \lim_{\delta t\to 0}\Big[\frac{2t\,\delta t}{\delta t}\Big] \\ &= \lim_{\delta t\to 0}\Big[2t\Big]=2t.\end{aligned} \tag{2.6}$$

In the same way we may also show that the derivative of t^n is nt^{n-1} when n is an integer. This will involve the use of the Binomial theorem.

This technique may also be used for other functions, although in the case of $\sin t$, for example, it becomes necessary to use the results,

$$\lim_{\delta t\to 0}\frac{\sin\delta t}{\delta t}=1, \qquad \lim_{\delta t\to 0}\frac{(1-\cos\delta t)}{\delta t}=0. \tag{2.7}$$

These results also require calculus for their proof but we do need to avoid circuar arguments. However, they may at least be verified using a calculator (which must be set to radians); they will also be proved using l'Hôpital's rule, which we will meet later.

2.2 Higher derivatives

There is no reason why we should not take another derivative, which is called the second derivative. If $y(t)$ is again the position of an object at time t, then dy/dt is the **velocity** of that object and the second derivative, which is denoted by,

$$\frac{d^2y}{dt^2}, \qquad y''(t), \qquad y''(t) \qquad \text{or} \qquad D^2y, \tag{2.8}$$

is the **acceleration**. Third derivatives are denoted by

$$\frac{d^3y}{dt^3}, \qquad y'''(t), \qquad y''' \qquad \text{and} \qquad D^3y, \tag{2.9}$$

and fourth derivatives by

$$\frac{d^4y}{dt^4}, \qquad y''''(t), \qquad y'''' \qquad \text{and} \qquad D^4y. \tag{2.10}$$

For yet higher derivatives it is often the case that the 'primed' notation is replaced by a superscript. So 9th derivatives may be written as $y^{(9)}(t)$, as opposed to $y'''''''''(t)$. And occasionally some textbooks use a Roman letter notation: y^i, y^{ii}, y^{iii}, y^{iv}, y^v and so on.

Sometimes third and higher derivatives are of use. In the design of rides such as Nemesis at Alton Towers, the term **jerk** is used to denote y''', the rate of change of acceleration. Further, the fourth derivative, y'''', is called the **jounce**. For a successfully designed ride (i.e. one that is sufficiently exciting, but not too nauseous), the acceleration, jerk and jounce all have to be within certain limits in all three directions.

2.3 Examples of derivatives

It is useful to be able to memorise as many as possible of the following.

$y(t)$	$y'(t)$	$y(t)$	$y'(t)$
c	0	$\sin t$	$\cos t$
t	1	$\sin at$	$a\ \cos at$
t^n	nt^{n-1}	$\cos t$	$-\sin t$
ct^n	nct^{n-1}	$\cos at$	$-a\ \sin at$
e^t	e^t	$\tan t$	$\sec^2 t = 1 + \tan^2 t$
e^{at}	ae^{at}	$\tan at$	$a\ \sec^2 at$
$\ln \lvert t \rvert$	t^{-1}	$\cot at$	$-a\,\mathrm{cosec}\,^2 at$
$\sin^{-1} t$	$\dfrac{1}{\sqrt{1-t^2}}$	$\sinh at$	$a\ \cosh at$
$\tan^{-1} t$	$\dfrac{1}{t^2+1}$	$\cosh at$	$a\ \sinh at$

Some of these will be derived in the following sections.

2.4 Manipulation of derivatives

This is a very short subsection, but the message is extremely important.

It is important to be able to take the derivatives of various functions, even quite complicated functions, with speed and accuracy. In spoken English (or any other language, as far as I know) we may be understood perfectly clearly even when we have expressed ourselves extremely poorly in terms of grammar. But in mathematics it is essential to follow all the rules absolutely precisely in order to avoid errors — all sorts of strange things can happen otherwise. Four of the following five sections give the "rules of the game" for differentiation.

2.5 Linearity

If we are given the two functions $u(t)$ and $v(t)$, then the **first rule of linearity** states that

$$\frac{d(u+v)}{dt} = \frac{du}{dt} + \frac{dv}{dt}. \tag{2.11}$$

If you are already familiar with this result, then it is very likely that you apply it automatically. In many cases this is assumed without even thinking about it, but we need to bear in mind what it means. Essentially it says that you may interchange the order of application of (i) summation and (ii) the taking of derivatives. The left hand side is the derivative of a sum, while the right hand side is the sum of derivatives. This may seem to be a very obvious result, and it can be proved fairly quickly from the limiting definition given in Section 2.1, but similar interchanges of mathematical operations do not always work. For example, the act of walking 100 miles east followed by 100 miles north is not the same as walking 100 miles east followed by 100 north, although there are some isolated locations on Earth where they are the same. In the area of matrices (if you have covered it), if A and B are two matrices, then not only will $AB \neq BA$ in general, but it is also possible that one product exists while the other does not.

The **second rule of linearity** is that

$$\frac{d(ku)}{dt} = k\frac{du}{dt}, \tag{2.12}$$

where k is a constant. Again this may be proved using the limit definition of the derivative.

We may take the example of the function $4t^3 - 2t$ to demonstrate the use of the two linearity rules:

$$\begin{aligned} \frac{d}{dt}\Big[4t^3 - 2t\Big] &= \Big[\frac{d}{dt}(4t^3)\Big] + \Big[\frac{d}{dt}(-2t)\Big] && \text{using (2.11)} \\ &= \Big[4\frac{d}{dt}(t^3)\Big] + \Big[(-2)\frac{d}{dt}(t)\Big] && \text{using (2.12)} \\ &= [4 \times 3t^2] + [(-2) \times 1] = 12t^2 - 2. \end{aligned} \tag{2.13}$$

This was very pedantic indeed and, in practice, we should strive to go immediately from $4t^3 - 2t$ to $12t^2 - 2$ without the intervening lines.

2.6 Product rule

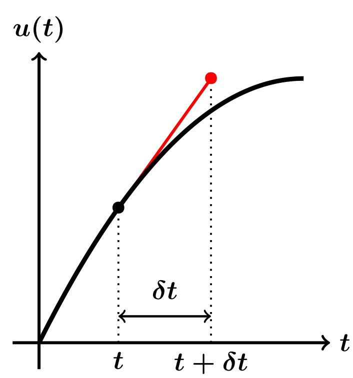
This is a rule which enables us to evaluate the derivative of a product of two functions. It is

$$\frac{d}{dt}(uv) = u\frac{dv}{dt} + v\frac{du}{dt} \qquad \text{or} \qquad (uv)' = uv' + u'v. \tag{2.14}$$

To prove this we first approximate the value of $u(t + \delta t)$ by

$$u(t + \delta t) \;\simeq\; u(t) + \delta t\,\frac{du}{dt} \;=\; u(t) + u'(t)\,\delta t; \tag{2.15}$$

this expression is equivalent to the equation of the straight line through $u(t)$ with the slope of the tangent to the curve at that point, as shown below



Here, the red disk corresponds to the right hand side of Eq. (2.5). We may also write down a similar expression for v. Now if we use the limiting definition of the derivative, see Eq. (2.3), we have the following analysis:

$$\begin{aligned} \frac{d}{dt}(uv) &= \lim_{\delta t \to 0}\Big[\frac{u(t + \delta t)v(t + \delta t) - u(t)v(t)]}{\delta t}\Big] \\ &= \lim_{\delta t \to 0}\Big[\frac{(u + u'\delta t)(v + v'\delta t) - uv}{\delta t}\Big] \\ &= \lim_{\delta t \to 0}\Big[\frac{\delta t(uv' + vu') + \delta t^2\, u'v'}{\delta t}\Big] \\ &= \lim_{\delta t \to 0}\Big[(uv' + vu') + \delta t\, u'v'\Big] \\ &= uv' + vu'. \end{aligned} \tag{2.16}$$

Note that this argument is not completely rigorous, but it is plausible. The lack of rigour takes place between lines 1 and 2 where the functions are replaced by the straight lines. However, the result is correct.

Example 2.1: Find the derivatives of $t^2 \sin t$ and $e^{-4t} \cos 5t$.
Given that the derivative of $\sin t$ is $\cos t$, we may use the product rule to show that

$$\frac{d}{dt}\Big[t^2 \sin t\Big] = t^2\Big[\frac{d}{dt}(\sin t)\Big] + \sin t\Big[\frac{d}{dt}(t^2)\Big] = t^2 \cos t + 2t \sin t. \tag{2.17}$$

Similarly, the derivative of $e^{-4t} \cos 5t$ is

$$\frac{d}{dt}\Big[e^{-4t} \cos 5t\Big] = e^{-4t}(-5 \sin 5t) + (-4)e^{-4t} \cos 5t = -e^{-4t}(5 \sin 5t + 4 \cos 5t). \tag{2.18}$$

The product of three functions.

It is now fairly straightforward to write down the derivative of the product of three or more functions. If we replace v by vw in Eq. (2.14) where $w = w(t)$, then

$$\begin{aligned} \frac{d}{dt}(uvw) &= u\frac{d(vw)}{dt} + (vw)\frac{du}{dt} && \text{using the product rule on } u \text{ and } vw \\ &= u\Big(v\frac{dw}{dt} + w\frac{dv}{dt}\Big) + vw\frac{du}{dt} && \text{using (2.14) on } vw \\ &= uvw' + uv'w + u'vw \\ &= u'vw + uv'w + uvw'. \end{aligned} \tag{2.19}$$

Although an alternative way of writing this final answer is

$$\frac{d}{dt}(uvw) = \Big[\frac{u'}{u} + \frac{v'}{v} + \frac{w'}{w}\Big]uvw, \tag{2.20}$$

I prefer to remember the form given in Eq. (2.19).

Example 2.2: Find the derivative of $t^5 e^{4t} \sin 3t$.

We shall follow the formula given in Eq. (2.19):

$$\begin{aligned} \frac{d}{dt}\Big[t^5 e^{4t} \sin 3t\Big] &= (5t^4)e^{4t} \sin 3t + t^5(4e^{4t}) \sin 3t + t^5 e^{4t}(3 \cos 3t) \\ &= t^4 e^{4t}\Big[(5 + 4t) \sin 3t + 3t \cos 3t\Big]. \end{aligned} \tag{2.21}$$

The product of four functions.

In similar fashion the derivative of the product of four functions, u_1, u_2, u_3 and u_4 is

$$\begin{aligned} \frac{d}{dt}(u_1 u_2 u_3 u_4) &= u_1' u_2 u_3 u_4 + u_1 u_2' u_3 u_4 + u_1 u_2 u_3' u_4 + u_1 u_2 u_3 u_4', \\ &= \Big[\frac{u_1'}{u_1} + \frac{u_2'}{u_2} + \frac{u_3'}{u_3} + \frac{u_4'}{u_4}\Big]u_1 u_2 u_3 u_4. \end{aligned} \tag{2.22}$$

and so on for products of yet more functions. Again, though, I prefer to remember the first line of Eq. (2.22) rather than the second, but both have memorable patterns.

2.7 The chain rule

This rule applies for functions which are functions of other functions of t, two examples of which are $\sin(t^2)$ and $\sin^2 t$. These are the sine of the square of t and the square of the sine of t. In general, if we have $u(t) = u(v(t))$ then the derivative is,

$$\frac{d}{dt}u(v(t)) = \frac{du}{dv}\frac{dv}{dt}. \tag{2.23}$$

This rule may be remembered easily because it *looks* as though the dv may be cancelled on the right hand side of Eq. (2.23). Still, it looks unusual and possibly even implausible, so I have given an outline of the proof of this later in Section (2.8).

Example 2.3: Find the derivatives of $u = \sin t^2$ and of $u = \sin^2 t$.
First we need to decompose/decouple these functions from their basic nested states. For the first one we can say that $u = \sin v$ where $v = t^2$, and therefore

$$\frac{du}{dt} = \frac{du}{dv}\frac{dv}{dt} = [\cos v]\,[2t] = 2t\cos(t^2). \tag{2.24}$$

For the second one we can say that $u = v^2$ where $v = \sin t$. Therefore we have,

$$\frac{du}{dt} = \frac{du}{dv}\frac{dv}{dt} = [2v]\,[\cos t] = 2\sin t\cos t. \tag{2.25}$$

And yes, I know that that last answer may be simplified further! Perhaps more important is that the fact that both answers need to written as functions solely of t since we introduced v solely to help us to do the differentiation, and therefore that final result must not involve v.

Example 2.4: Find the derivative of $u = (t + a)^5$, where a is a constant.

For $u = (t + a)^5$ we let $u = v^5$ where $v = t + a$. The derivative is now

$$\frac{du}{dt} = \frac{du}{dv}\frac{dv}{dt} = [5v^4]\,[1] = 5(t + a)^4. \tag{2.26}$$

That final result is usually memorised.

Example 2.5: Find the derivative of $y = v^{-1}$ where v is a function of t.

To find this derivative we first note that $d(t^{-1})/dt = -t^{-2}$. Now we apply the chain rule to get,

$$\frac{dy}{dt} = \frac{dy}{dv}\frac{dv}{dt} = \Big[-\frac{1}{v^2}\Big] \times \Big[v'\Big] = -\frac{v'}{v^2}. \tag{2.27}$$

Here we are using primes to denote derivatives with respect to t. This result will be used a little later to prove the Quotient Rule.

A triply-nested function

The chain rule may be extended indefinitely. If y is a function of u which is a function of v which is a function of t, i.e.

$$y = y\Big[u\Big(v(t)\Big)\Big],$$

then

$$\frac{dy}{dt} = \frac{dy}{du}\frac{du}{dv}\frac{dv}{dt}. \tag{2.28}$$

Again, this formula may be remembered because of the apparent cancellations hinted at by the colours.

Example 2.6: Find the derivative of $y = e^{\sin t^2}$.

We may decompose this function by setting, $y = e^u$ where $u = \sin v$ and where $v = t^2$. Therefore the derivative is given by

$$\frac{dy}{dt} = \frac{dy}{du}\frac{du}{dv}\frac{dv}{dt} = e^u \times \cos v \times 2t = 2t[\cos(t^2)]\,[e^{\sin(t^2)}]. \tag{2.29}$$

Example 2.7: The Chain Rule may also be used for finding the derivatives of the inverse trigonometrical functions. For example, if $y = \tan^{-1} t$, then $\tan y = t$. The left hand side is of the form of a function of a function, since $y = y(t)$. Using the chain rule, the derivative of $\tan y$ with respect to t is given by

$$\frac{d}{dt}\tan y = \frac{d\tan y}{dy}\frac{dy}{dt} = (1 + \tan^2 y)\frac{dy}{dt}, \tag{2.30}$$

where we have used the result which was stated in the above Table (and which will be proved below) for the derivative of $\tan y$. Hence the time derivative of the equation, $\tan y = t$, is

$$(1 + \tan^2 y)\frac{dy}{dt} = 1 \qquad \Rightarrow \frac{dy}{dt} = \frac{1}{1 + \tan^2 y} = \frac{1}{1 + t^2}. \tag{2.31}$$

Some say that the process used in Eq. (2.31) is called **Implicit Differentiation**, but it is no more that a subset of the chain rule.

Example 2.8: Find the derivative of $y = \ln|t|$.

The derivative of $y = \ln|t|$ may be obtained in like fashion by first taking exponentials of each side, but the detail is a little more complicated because of the presence of the modulus signs.

When $t > 0$ we have

$$y = \ln|t| = \ln t \quad \Rightarrow \quad e^y = t \quad \Rightarrow \quad e^y y' = 1 \quad \Rightarrow \quad y' = e^{-y} = t^{-1}, \tag{2.32}$$

but when $t < 0$ we have,

$$y = \ln|t| = \ln(-t) \quad \Rightarrow \quad e^y = -t \quad \Rightarrow \quad e^y y' = -1 \quad \Rightarrow \quad y' = -e^{-y} = t^{-1}. \tag{2.33}$$

Hence the derivative of $\ln|t|$ is t^{-1}.

2.8 Outline proof of the chain rule [For interest only]

We will let $u = u(v)$ where $v = v(t)$, so that u is ultimately a function of t. We'll write the function as $u = u[v(t)]$ where I've used the differently-shaped brackets here solely for the purposes of clarity/brevity. We will use the limit definition of a derivative as originally given in Eq. (2.3), above. Therefore we have

$$\begin{aligned}
\frac{du}{dt} &= \lim_{\delta t\to 0}\left[\frac{u[v(t+\delta t)]-u[v(t)]}{\delta t}\right] \\
&= \lim_{\delta t\to 0}\left[\frac{u\left[v(t)+\frac{dv}{dt}\delta t+\cdots\right]-u[v(t)]}{\delta t}\right] && v(t+\delta t) \text{ approximated by its tangent} \\
&= \lim_{\delta t\to 0}\left[\frac{u[v(t)+\delta v]-u[v(t)]}{\delta t}\right] && \text{where } \delta v \equiv \frac{dv}{dt}\delta t \text{ for clarity} \\
&= \lim_{\delta t\to 0}\left[\frac{\left(u[v(t)]+\frac{du}{dv}\delta v+\cdots\right)-u[v(t)]}{\delta t}\right] && u(v+\delta v) \text{ approximated by the tangent} \\
&= \lim_{\delta t\to 0}\left[\frac{u[v(t)]+\frac{du}{dv}\frac{dv}{dt}\delta t+\cdots-u[v(t)]}{\delta t}\right] && \text{Using the definition of } \delta v\text{. Red text cancels} \\
&= \frac{du}{dv}\frac{dv}{dt}.
\end{aligned} \tag{2.34}$$

2.9 Quotient rule

This is closely related to the product rule and makes use of the above result, Eq. (2.27), on the derivative of v^{-1}. The rule is derived as follows:

$$\begin{aligned}
\frac{d}{dt}\left(\frac{u}{v}\right) &= \frac{d}{dt}(uv^{-1}) = u\frac{d(v^{-1})}{dt}+v^{-1}\frac{du}{dt} && \text{using the product rule} \\
&= -\frac{uv'}{v^2}+\frac{u'}{v} && \text{using Eq. (2.27)} \\
&= \frac{vu'-uv'}{v^2}.
\end{aligned} \tag{2.35}$$

In the good old(?) days of cathode ray tubes I used to remember this result by stating the first term to be 'VDU' (v delta u or '**V**isual **D**isplay **U**nit' — the old term for a computer terminal or monitor).

Example 2.9: Find the derivative of $y = \tan t$.

This takes the following route:

$$\begin{aligned}\frac{d\tan t}{dt} &= \frac{d}{dt}\Big(\frac{\sin t}{\cos t}\Big)\\ &= \frac{\cos t\,(\sin t)' - \sin t\,(\cos t)'}{\cos^2 t}\\ &= \frac{\cos^2 t - \sin t(-\sin t)}{\cos^2 t} = \frac{\cos^2 t + \sin^2 t}{\cos^2 t}\\ &= \frac{1}{\cos^2 t} \quad \text{or} \quad 1+\tan^2 t \quad \text{or} \quad \sec^2 t.\end{aligned} \tag{2.36}$$

2.10 Critical points

The maximum and minimum values of a function are examples of **critical points** or **extrema** (singular: **extremum**). Points of inflexion* are also critical points. Generally the value of t at which this happens is found by setting the derivative of the function to zero, and solving the resulting equation for t. For example, if $y(t) = t^2 - t$, then $y'(t) = 2t - 1$. Therefore the function has an extremum when $y'(t) = 0$, which, in this case, is when $t = \frac{1}{2}$.

Sometimes it is not clear whether a particular critical point is a local maximum or a local minimum, but it is possible to determine this mathematically in almost all cases by considering the second derivative of the function. Returning to the above function, $y = t^2 - t$, a simple sketch is enough to tell us that this extremum is a minimum, but it is necessary also to have a set of mathematical criteria.

The second derivative at this point is $y''(\frac{1}{2}) = 2$ which is positive; a positive value of the second derivative always corresponds to a minimum (and a negative value corresponds to a maximum). That this is a general result may be seen by considering the typical situation: to the left of a minimum the slope of the function must be negative, whereas the slope is positive to the right of the minimum. Therefore the slope is increasing and hence the second derivative (which is the slope of the slope) is positive.

If the above paragraph has too much verbiage and too little imagery then it may be better to consider Fig. 2.3 on the next page which shows diagrammatically that a minimum in y corresponds to a zero value of y' and a positive value for y''. On the other hand, while a maximum in y also corresponds to a zero value of y', the value of y'' is negative. Hopefully this is clear in Fig. 2.3.

* Some authorities regard an inflexion point as being that value of t at which $y''(t) = 0$ with no condition on $y'(t)$. This means that the curve is S-shaped locally, and changes from being concave to convex or vice-versa. Here, my definition is more restrictive because I require both $y'(t) = 0$ and $y''(t) = 0$ simultaneously; this means that the inflexion point is also a critical point.

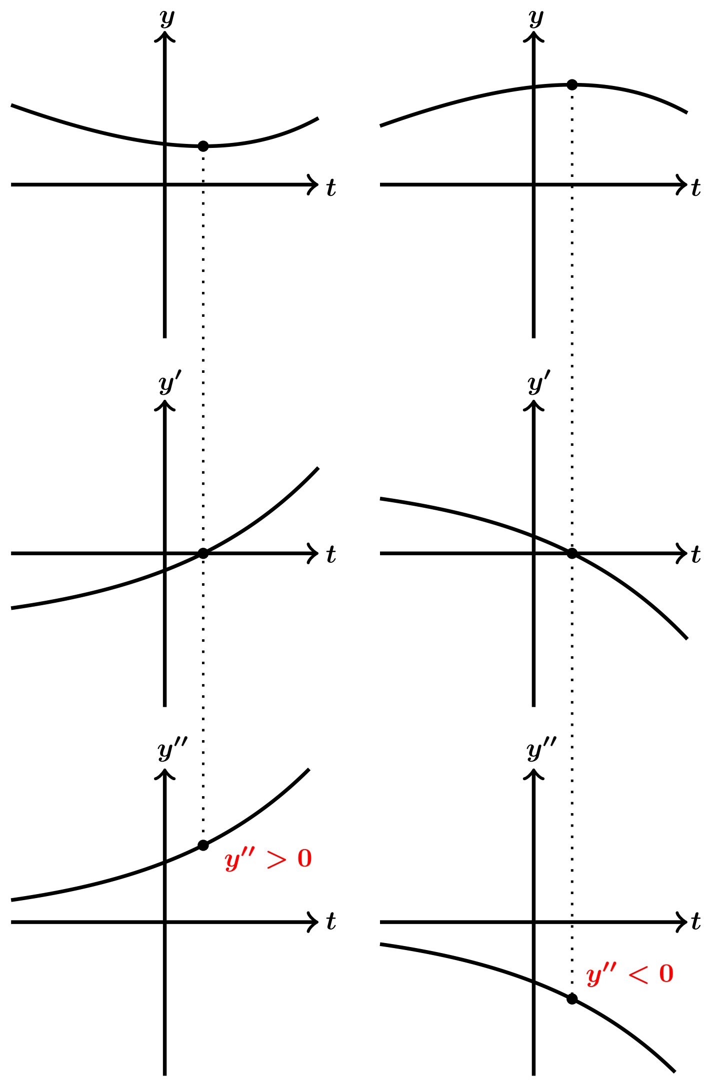


Figure 2.3. Depicting curves for y, y' and y'' (top to bottom) where y shows a minimum (left column) and a maximum (right column). The dotted lines and the black disks indicate the relationship between the behaviours of the y, y' and y'' curves in each case.

Example 2.10: Find the critical points of the function $y = x^3 - 3x$.

Note that we have now changed the identity of the independent variable from t to x. Maxima and minima also occur in space! We have

$$y' = 3x^2 - 3 \qquad \text{which may be factorised:} \quad y' = 3(x-1)(x+1).$$

Hence the critical points are at $x = -1, 1$. These may be classified by evaluating y'' at the critical points. Hence

$$y'' = 6x = \begin{cases} 6 & \text{at } x = 1 \quad \Rightarrow \text{ Minimum} \\ -6 & \text{at } x = -1 \quad \Rightarrow \text{ Maximum} \end{cases} \tag{2.37}$$

Figure 2.4 confirms this analysis.

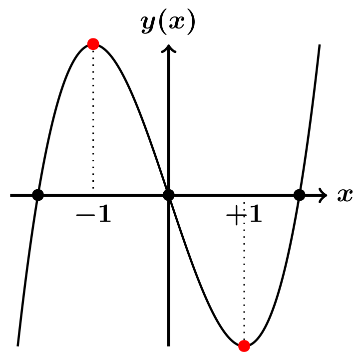


Figure 2.4. Showing $y = x^3 - 3x$ together with its roots (black disks) and the maximum and minimum values (red disks).

While the solution of $y'(x) = 0$ yields where the critical points are, their identification requires the use of secondary criteria. In this case a postive y'' corresponds to a minimum, while a negative value corresponds to a maximum. But what happens if $y'' = 0$ as well? The following example provides an illustration.

Example 2.11: Find the critical points of the function $y = x^3$.

This is a fairly trivial-looking function, but it works well and allows us to see the important details. I could have chosen to use $y = 2(x - 2)^3$ and while the analysis will work, it might prove more difficult to do should I have presented $y(x)$ as a power series, i.e. having multiplied out the cubic.

Clearly $y' = 3x^2$, and the setting of this to zero yields $x = 0$ as the location of the critical point.

If we now evaluate y'' then we obtain, $y''(x) = 6x$. When $x = 0$ then $y''(0) = 0$, and therefore this falls into an intermediate category. We know from applying our curve-sketching ideas (or also the idea of factorisation) that $y = x^3$ has a triple root at $x = 0$ and therefore $x = 0$ corresponds to a point of inflexion. The third derivative is $y'''(x) = 6$ which is positive, and therefore $y'''(0) = 6$ corresponds to a point of inflexion *which is rising*. Figure 2.5 shows this.

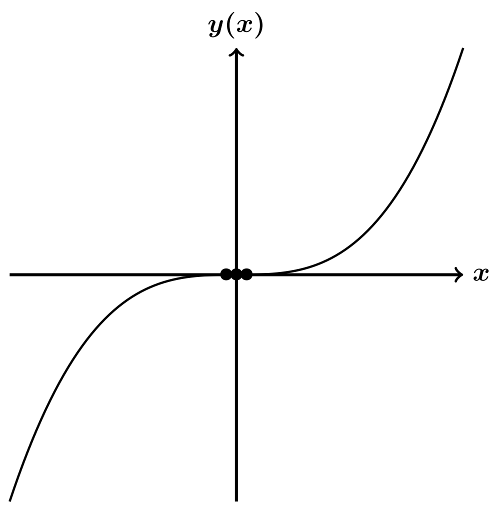


Figure 2.5. Showing $y = x^3$ which has a **rising point of inflexion** at $x = 0$.

If we had had a case where $y'' = 0$ but $y''' < 0$ at the critical point, then we could name it a **descending point of inflexion**. So a point of inflexion has $y' = 0$ and $y'' = 0$ as its primary criteria while the secondary criterion is that $y''' \neq 0$ at the critical point.

But what if $y''' = 0$ as well? Here's an example....

Example 2.12: Find the critical points of the function $y = ax^4$ where a is a **nonzero** constant. We already know from the curve-sketching section that this function has a quadruple root at $x = 0$ and has a quartic minimum when $a > 0$. Our analysis follows:

$$\begin{aligned} y' &= 4ax^3 & &\Rightarrow x = 0 & &\text{is the sole critical point} \\ y'' &= 12ax^2 & &\Rightarrow y''(0) = 0 & &\Rightarrow \text{is not a max/min} \\ y''' &= 24ax & &\Rightarrow y'''(0) = 0 & &\Rightarrow \text{is not an inflexion point} \\ y'''' &= 24a & &\Rightarrow y''''(0) = 24a & &\Rightarrow \text{is a } \textbf{quartic minimum when } a > 0 \\ & & & & &\text{or is a } \textbf{quartic maximum when } a < 0 \end{aligned}$$

The case $a = 1$ is illustrated below.

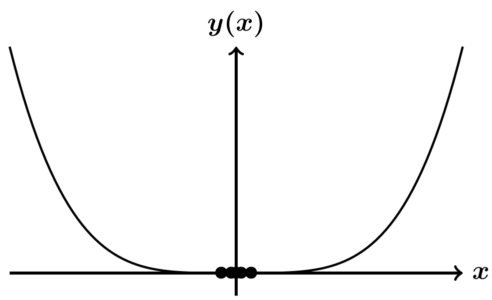


Figure 2.6. Showing $y = x^4$ which has a quartic minimum at $x = 0$.

Example 2.13: Find and classify the critical point/points for $y = x^4 - 8x^3 + 24x^2 - 32x + 16$.

This appears to be a nasty one, and it is, but it isn't a trivial one like Example 2.12 is. The analysis proceeds as follows.

$$\begin{aligned} y' &= 4x^3 - 24x^2 + 48x - 32 & &\Rightarrow x = 2 & &\text{is the sole critical point — you'll need to trust me on this!} \\ y'' &= 12x^2 - 48x + 48 & &\Rightarrow y''(2) = 0 & &\Rightarrow \text{is not a max/min} \\ y''' &= 24x - 48 & &\Rightarrow y'''(2) = 0 & &\Rightarrow \text{is not an inflexion point} \\ y'''' &= 24 & &\Rightarrow y''''(2) = 24 > 0 & &\Rightarrow \text{is a quartic minimum} \end{aligned}$$

Clearly one could now ask what happens if we have $y'' = y''' = y'''' = 0$ at a critical point, but the fifth derivative is nonzero. Well, this will be a quintic inflexion point and it may be a rising or descending one depending on the sign of $y^{(5)}$ at the critical point.

On the next page I have compiled a Table which details how to detect critical points up to and including the septic (!!!) ones. At least you can be assured that the more exotic the critical point is the less likely it is to turn up on an exam paper because of the increased workload which is required — Exercise 2.13 was bad enough!

Table. Primary and secondary criteria for the different types of critical points.

y'	y''	y'''	y''''	$y^{(5)}$	$y^{(6)}$	$y^{(7)}$	
0	+						Minimum
0	−						Maximum
0	0	+					Rising inflexion
0	0	−					Descending inflexion
0	0	0	+				Quartic minimum
0	0	0	−				Quartic maximum
0	0	0	0	+			Rising quintic inflexion
0	0	0	0	−			Descending quintic inflexion
0	0	0	0	0	+		Sextic minimum
0	0	0	0	0	−		Sextic maximum
0	0	0	0	0	0	+	Rising septic inflexion
0	0	0	0	0	0	−	Descending septic inflexion

2.11 Notation

Throughout this part of the unit I have used either t or x as the independent variable. This has been motivated by the fact that very many physical quantities (e.g. speed, acceleration, electrical current, body temperature) are time-dependent or space dependent, and we are often interested in how quickly these quantities change.

However, the are many instances where the independent variable could be something else. At a given pressure, the density of water depends on temperature ($\rho = \rho(T)$). The mean atmospheric pressure depends on the height above the ground ($P = P(h)$). The intensity of an electric field around a point charge depends on the distance from the charge ($E = E(r)$). Therefore we need to be able to take derivatives with respect to any and every possible independent variable. Fortunately, all the rules described above apply, although it may be possible to become a little confused when variables play unusual roles, such as when t is a dependent variable (some people use this to mean temperature). Therefore all the following examples are, in effect, identical — it's just that the names of the variables have changed:

$$y = \sin x^2 \quad \Rightarrow \quad \frac{dy}{dx} = 2x \cos x^2$$

$$x = \sin y^2 \quad \Rightarrow \quad \frac{dx}{dy} = 2y \cos y^2$$

$$t = \sin v^2 \quad \Rightarrow \quad \frac{dt}{dv} = 2v \cos v^2$$

$$\gamma = \sin \alpha^2 \quad \Rightarrow \quad \frac{d\gamma}{d\alpha} = 2\alpha \cos \alpha^2$$

$$\check{\bowtie} = \sin \heartsuit^2 \quad \Rightarrow \quad \frac{d\check{\bowtie}}{d\heartsuit} = 2\heartsuit \cos \heartsuit^2$$

3 INTEGRATION

If we are given a function, $f(x)$, then the area, A, between the function and the x-axis and which lies between $x = a$ and $x = b$ is given by the yellow region below.

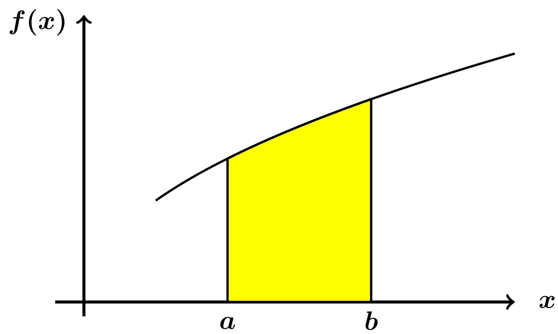


Figure 3.1. Showing the definition of the integral between $x = a$ and $x = b$ as being equivalent to the yellow area under the graph.

It is known as the "area under graph between a and b" or as the "integral of $f(x)$ between a and b". This is denoted by

$$A = \int_a^b f(x)\, dx. \tag{3.1}$$

The notation, $\int$, represents a summation of the areas of small strips, as sketched below:

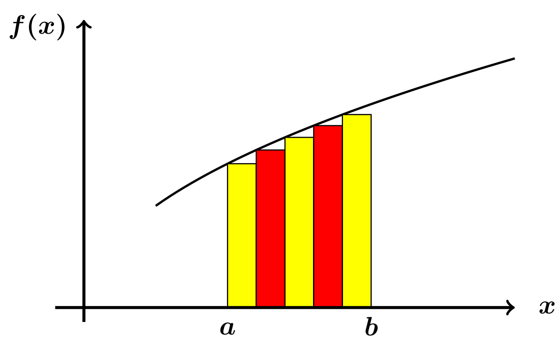


Figure 3.2. Showing how the integral between $x = a$ and $x = b$ may be approximated using 5 strips of equal width.

In this Figure each strip has width $\frac{1}{5}(b - a)$, but if we say more generally that each strip has width, δx, and that the height of each strip is $f(x_i)$, where $x_i = a + i\,\delta x$, then the area under the curve in this general version of Fig. 3.2 may be approximated by

$$A \simeq \sum_{i=0}^{N-1} f(x_i)\,\delta x. \tag{3.2}$$

As with the derivative, the integral may be written in terms of a limiting process where

$$A = \lim_{\delta x \to 0} \sum_{i=0}^{N-1} f(x_i)\,\delta x, \tag{3.3}$$

and where the number of strips is given by $N = (b - a)/\delta x$. So $N \to \infty$ as $\delta x \to 0$. In Fig. 3.3, which shows 20 strips, it is shown how much better Eq. (3.3) approximates the integral.

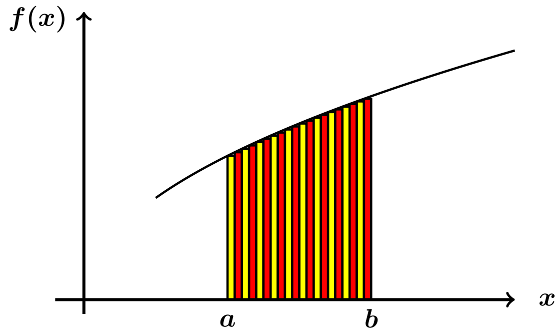


Figure 3.3. Showing the definition of the integral between $x = a$ and $x = b$ as being equivalent to the area under the graph.

The right hand side of Eq. (3.1) is said to be the **integral of $f(x)$ between $x = a$ and $x = b$.** The function $f(x)$ is called the **integrand**, while x is the **variable of integration**.

Note that, since the value of this integral is the constant value, A, then the symbol we use for the variable of integration is irrelevant. Hence

$$\int_a^b f(x)\,dx = \int_a^b f(y)\,dy. \tag{3.4}$$

For this reason we may call x a **dummy variable** because the final result is independent of its identity.

3.1 Relationship with the derivative [For interest and background]

Formally, integration and differentiation are mutually inverse operations. At first glance this seems like madness, for what does the area under a graph have to do with the slope of a graph?

This question may be answered by first defining $I(x)$ according to,

$$I(x) = \int_a^x f(y)\,dy. \tag{3.5}$$

Note the different viewpoint here: this is an integral, or area under the graph, between $y = a$ and $y = x$. We assume that a is a constant. The upper limit, x, is allowed to vary and therefore the value of the integral depends on the value of x. So $I(x)$ is a function of x.

We will now attempt to find the derivative of $I(x)$ using the limiting definition of the derivative which was given in Eq. (3) in the Differentiation notes. Therefore the derivative of $I(x)$ is defined according to,

$$\frac{dI}{dx} = \lim_{\delta x \to 0} \left[\frac{I(x + \delta x) - I(x)}{\delta x} \right], \tag{3.6}$$

or,

$$\frac{dI}{dx} = \lim_{\delta x \to 0} \frac{1}{\delta x} \left[\int_a^{x+\delta x} f(y)\,dy - \int_a^x f(y)\,dy \right]. \tag{3.7}$$

If we now consider our original definition of an integral as an area under a graph, then this expression reduces to the integral of a narrow strip (at least when δx is small):

$$\frac{dI}{dx} = \lim_{\delta x \to 0}\Big[\frac{1}{\delta x}\int_x^{x+\delta x} f(y)\,dy\Big]. \tag{3.8}$$

This transition from Eq. (3.7) to Eq. (3.8) is illustrated in Fig. 3.4:

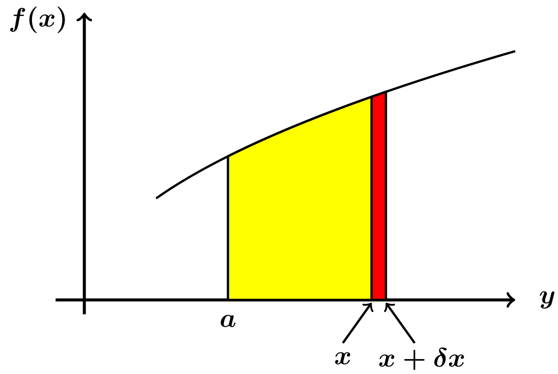


Figure 3.4. Illustrating the process behind finding the derivative of the area defined by Eq. (3.5).

Here $\int_a^x f(y)\,dx$ corresponds to the yellow region, $\int_a^{x+\delta x} f(y)\,dy$ to both the red and yellow regions, and $\int_x^{x+\delta x} f(y)\,dy$ to the red strip. When δx is very small, the thin red strip red has an area which is well-approximated by a trapezium of width δx, and of sides $f(x)$ and $f(x+\delta x)$. Therefore the integral in Eq. (3.8) is, to a very good degree of approximation equal to,

$$\int_x^{x+\delta x} f(y)\,dy \simeq \frac{f(x+\delta x)+f(x)}{2}\delta x. \tag{3.9}$$

Therefore we have

$$\frac{dI}{dx} = \lim_{\delta x \to 0}\Big[\frac{1}{\delta x}\Big(\frac{f(x+\delta x)+f(x)}{2}\delta x\Big)\Big] = f(x). \tag{3.10}$$

Note 1: Should $I(x) = c + \int_a^x f(y)\,dy$ where c is an arbitrary constant, then an identical analysis also yields Eq. (3.10) simply because the derivative of c is zero.

It is this result that tells us that integration and the taking of derivatives are mutually inverse operations. Therefore the art of integration is the determination of a function which, when differentiated, yields the original function. So the integral of $3x^2$ is $x^3 + c$ because the derivative of $x^3 + c$ is $3x^2$. In mathematical terms we write,

$$\int 3x^2\,dx = x^3 + c \tag{3.11}$$

where an integral which is written without limits is generally called an **indefinite integral** and is sometimes called the **antiderivative**. On the other hand, an integral with limits is called a **definite integral**.

Table of some common indefinite integrals

$$\begin{aligned}
\int \cos ax\, dx &= \frac{\sin ax}{a} + c, \\
\int \sin ax\, dx &= -\frac{\cos ax}{a} + c, \\
\int x^n\, dx &= \frac{x^{n+1}}{n+1} + c \qquad (n \neq 1), \\
\int \frac{1}{x}\, dx &= \ln|x| + c, \\
\int e^{ax}\, dx &= \frac{e^{ax}}{a} + c.
\end{aligned}$$

Note 2: Protocol for the evaluation of a definite integral. If we take the expression,

$$I(x) = \int_a^x f(y)\, dy + c, \tag{3.12}$$

then the respective setting of $x = a$ and of $x = b$ yields the following,

$$\begin{aligned}
I(a) &= \int_a^a f(y)\, dy + c = c \qquad \text{identical limits} \Rightarrow \text{zero area} \\
I(b) &= \int_a^b f(y)\, dy + c
\end{aligned} \tag{3.13}$$

By eliminating c we obtain,

$$\int_a^b f(y)\, dy = I(b) - I(a), \tag{3.14}$$

which will, no doubt, be familiar. And this result tells us immediately that swapping the limits is effectively the same as multiplication by -1:

$$\int_b^a f(y)\, dy = I(a) - I(b) = -\int_a^b f(y)\, dy. \tag{3.15}$$

Example 3.1. Given that the indefinite integral of $3x^2$ is $x^3 + c$ (see Eq. (3.11)) find the definite integral between $x = 2$ and $x = 5$.

This is,

$$\int_2^5 3x^2\, dx = \left[x^3 + c\right]_2^5 = (5^3 + c) - (2^3 + c) = 117. \tag{3.16}$$

This example also tells us that we don't need to use the arbitrary constant when evaluating definite integrals, i.e. we are, for once, allowed to be lazy and not bother with the arbitrary constant.

3.2 Linearity Rules

Integration obeys the same linearity rules as differentiation:

$$\int \Big[f(x) + g(x) \Big] \, dx = \int f(x) \, dx + \int g(x) \, dx, \tag{3.17}$$

$$\int k \, f(x) \, dx = k \int f(x) \, dx \qquad \text{where } k \text{ is a constant.} \tag{3.18}$$

Therefore we may interchange the order of summation and integration and the order of multiplication by a constant and integration. Here's an example.

Example 3.2.

$$\begin{aligned} \int (x^2 - 2x + 1) \, dx &= \int x^2 \, dx + \int (-2)x \, dx + \int 1 \, dx && \text{using (3.17)} \\ &= \int x^2 \, dx + (-2) \int x \, dx + \int 1 \, dx && \text{using (3.18)} \\ &= \tfrac{1}{3}x^3 - x^2 + x + c. \end{aligned}$$

This was long-winded and pedantic; in practice one should aim to go directly to the answer when the integrand consists of functions like those in the above Table.

3.3 Integration by substitution

Very few functions that we need to integrate take the precise form given in the above examples and exercises. Although differentiation may always be done even if it is lengthy, it is often the case that it can be quite a puzzle to find out how to do some integrations. This is what makes the topic more tricky than differentiation. Indeed, there are some integrals that cannot be done, an example being that of e^{-x^2}.

We will illustrate the "Integration by substitution" method by means of examples.

Example 3.3. Evaluate $I = \displaystyle\int \cos(\omega x) \, dx$.

We know how to integrate $\cos x$, but we will pretend not to know how deal with $\cos \omega x$, and therefore our aim will be to reduce the integrand to the form we know. This aim is achieved using the substitution, $\xi = \omega x$. We may now alter the integration by means of the following equivalent of the chain rule:

$$\begin{aligned} I = \int \cos(\omega x) \, dx &= \int \cos \xi \, \frac{dx}{d\xi} \, d\xi \\ &= \int \frac{\cos \xi}{\omega} \, d\xi \\ &= \frac{1}{\omega} \int \cos \xi \, d\xi && \text{taking the constant outside the integral} \quad (3.19) \\ &= \frac{\sin \xi}{\omega} + c && \text{on integration} \\ &= \frac{\sin \omega x}{\omega} + c && \text{returning to the original variable.} \end{aligned}$$

The equivalent of the chain rule is involved in the step with the red font. Generally, though, most people do not follow the above method even though it looks more rigorous and reasonable than the method which is usually taught and which I will describe now.

Given that $\xi = \omega x$ we may deduce that $d\xi = \omega\, dx$ by differentiation of both sides — essentially this says if we change x to $x + \delta x$, where δx is small, then we expect ξ to change too. Hence, $\xi + \delta\xi = \omega\,(x + \delta x)$, and hence $\delta\xi = \omega\,\delta x$.

We now have all the information required to change variable from x to ξ. Therefore the above integration-by-substitution analysis becomes the following.

$$\begin{aligned} I &= \int \cos\omega x\, dx \\ &= \int \cos\xi\, \frac{d\xi}{\omega} \qquad \text{the substitution, all in one go.} \left(dx = \frac{d\xi}{\omega}\right) \\ &= \frac{1}{\omega}\int \cos\xi\, d\xi \\ &= \frac{\sin\xi}{\omega} + c \\ &= \frac{\sin(\omega x)}{\omega} + c. \qquad \text{returning from } \xi \text{ to } x \end{aligned} \tag{3.20}$$

Example 3.4. Evaluate $I = \displaystyle\int 2x\sin(x^2)\, dx$.

We know how to evaluate the integral of $\sin x$, but $\sin(x^2)$ is not one of the standard integrals. However, we may make this part of the integrand integrable by using the substitution $y = x^2$. From this we have $dy = 2x\, dx$ where the $2x$ is the derivative of x^2; this also follows from the small-increment analysis: $y + \delta y = (x + \delta x)^2 = x^2 + 2x\delta x + \ldots$.

Hence the integrand immediately transforms to

$$I = \int \underbrace{\sin(x^2)}_{\sin y}\ \underbrace{2x\, dx}_{dy} = \int \sin(y)\, dy = -\cos(y) + c = -\cos(x^2) + c. \tag{3.21}$$

Example 3.5. Find $I = \displaystyle\int \sin^2(t)\cos(t)\, dt$.

As $\cos t$ is the differential of $\sin t$, and there is only one $\cos t$, we shall let $z = \sin t$ and hence $dz = \cos t\, dt$. The integral becomes,

$$I = \int \underbrace{\sin^2 t}_{z^2}\ \underbrace{\cos t\, dt}_{dz} = \int z^2\, dz = \tfrac{1}{3}z^3 + c = \tfrac{1}{3}\sin^3 t + c. \tag{3.22}$$

Of course, in all of these cases one may check the final answer by differentiating it.

Example 3.5 (reprise). Find $I = \displaystyle\int_0^{\pi/2} \sin^2(t)\cos(t)\, dt$.

This rerun of Example 3.5 will show us how important it is to consider carefully the limits of a definite integral when using the substitution method. I'll pretend that Example 3.5 hasn't happened$\cdots$.

We have the following pieces of information which will need to be used:

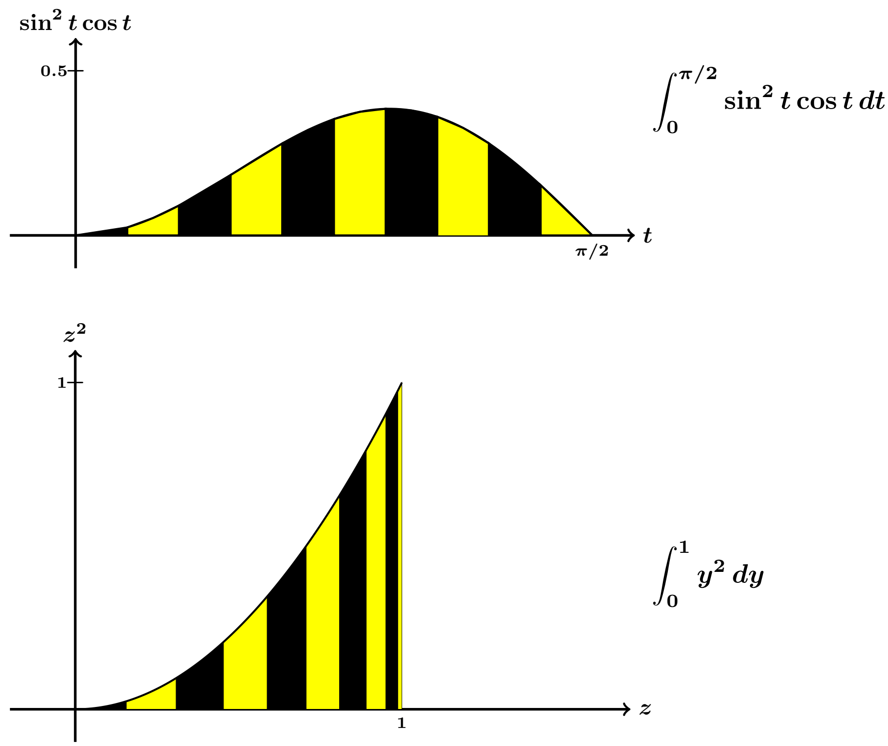
$$\boxed{z = \sin t \;\Rightarrow\; dz = \cos t\, dt.} \qquad \boxed{t = 0 \;\Rightarrow\; z = 0.} \qquad \boxed{t = \tfrac{1}{2}\pi \;\Rightarrow\; z = 1.}$$

Now we may transform the given integral immediately from being a t-integral to being a z-integral:

$$I = \int_0^{\pi/2} \sin^2 t \, \cos t \, dt \; = \; \int_0^1 z^2 \, dz \; = \; \Big[\tfrac{1}{3}z^3 + c\Big]_0^1 \; = \; \tfrac{1}{3}. \tag{3.23}$$

Of course, this final numerical value could have been obtained easily by using the final answer in Eq. (3.22) and evaluating that between $t = 0$ and $t = \frac{1}{2}\pi$.

Although the integrands and the limits in two integrals in Eq. (3.22) are so very different, the fact is that the overall value of the integral is conserved. We may illustrate this fairly well in the following figure.



Showing the manner in which areas are conserved when using Integration-by-Substitution, as given by the reprise of Example 3.5.

Each of the integrals shown above is split into ten intervals. In the t-integral the strips have equal widths and the edges of these strips are located at $t = (n/10) \times (\pi/2) = n\pi/20$ for $n = 0, 1, \cdots, 9, 10$. This yields ten separate regions that are illustrated as alternating black and yellow strips. One the other hand the edges of the strips in the z-integral correspond to where $z = \sin(n\pi/20)$, which is why they are of unequal widths. However, each of the ten strips for the z-integral has an equal area to the corresponding strip in the t-integral. This looks very feasible for the first two or three strips (working from the left) but it remains true even for the very last strips on the right where the two tenth strips have very different aspect ratios.

Example 3.6. Find the value of $I = \int_0^{\sqrt{\pi}} 2x \sin(x^2)\, dx$.

The integrand is another reprise, but of Example 3.4, above. Using the notation of Example 3.4, when $x = 0$ then $y = 0$, and when $x = \sqrt{\pi}$ then $y = \pi$. Therefore the present integral is

$$\begin{aligned} I &= \int_0^{\sqrt{\pi}} 2x \sin(x^2)\, dx \\ &= \int_0^{\pi} \sin y\, dy && \text{the } x \text{ to } y \text{ substitution, including the limits} \\ &= \Big[-\cos y\Big]_0^{\pi} = 2. \end{aligned} \tag{3.24}$$

Note: It is essential that such integrals use only one dummy variable at any time — there must never be a "half-way house" containing both x and y simultaneously — if that were to happen, then which are the correct limits, those belonging to x or those for y? The transformation from one dummy variable to the other must take place instantaneously and then the maths is completely unambiguous.

An alternative route which avoids transforming the limits of integration is as follows. First find the indefinite integral using substitution, then write the answer in terms of the original variable, and finally apply the limits. So here we would have the following:

$$\begin{aligned} \text{Let } I &= \int 2x \sin(x^2)\, dx \\ &= \int \sin y\, dy && \text{the } x \text{ to } y \text{ substitution} \\ &= \Big[-\cos y\Big] \\ &= -\cos(x^2) && \text{back to } x. \text{ No need for the arbitrary constant} \end{aligned} \tag{3.25}$$

$$\begin{aligned} \text{Hence } & \int_0^{\sqrt{\pi}} 2x \sin(x^2)\, dx \\ &= \Big[-\cos(x^2)\Big]_0^{\sqrt{\pi}} = 2. \end{aligned}$$

Example 3.7. Find $I = \int \frac{x}{\sqrt{1 - x^2}}\, dx$.

This integral may be obtained using either the substitution, $y = x^2$ or $y = 1 - x^2$, but I will use a third substitution, $x = \sin\theta$. This one is motivated by the fact that the denominator of the integrand will transform to $\cos\theta$, while the dx part will also involve a $\cos\theta\, d\theta$, and there will be the potential for cancellation. Given our chosen transformation, we have

$$x = \sin\theta \qquad \Rightarrow \qquad dx = \cos\theta\, d\theta,$$

and hence

$$\begin{aligned}
I &= \int \frac{x}{\sqrt{1-x^2}}\,dx \;=\; \int \frac{\sin\theta}{\sqrt{1-\sin^2\theta}}\;\cos\theta\,d\theta && \text{the substitution}\\
&= \int \frac{\sin\theta\,\cos\theta}{\cos\theta}\,d\theta && \text{tidying up}\\
&= \int \sin\theta\,d\theta && \text{some cancellation}\\
&= -\cos\theta + c \;=\; -\sqrt{1-x^2}+c.
\end{aligned} \tag{3.26}$$

In the last line of this evaluation of I we chose $\cos\theta = +\sqrt{1-\sin^2\theta}$ rather than $-\sqrt{1-\sin^2\theta}$. The correctness of this choice may (and perhaps should) be tested by differentiating the final answer.

If we were to consider the very similar integral, $I = \int \frac{x}{\sqrt{a^2-x^2}}\,dx$, then we could use $x = a\sin\theta.$

Example 3.8. Find the integral of $\int \frac{1}{x^2+a^2}\,dx$.

This is a classic integral. The secret here is try find a substitution which, like the previous one, employs a trigonometric identity to simplify the integral. This one will be based on the result that $\sec^2\theta = 1+\tan^2\theta$. Given the a^2 in the integral, we shall use the substitution,

$$x = a\tan\theta \qquad\Rightarrow\qquad dx = a\sec^2\theta\,d\theta. \tag{3.27}$$

Hence,

$$\int \frac{1}{x^2+a^2}\,dx \;=\; \int \frac{a\sec^2\theta}{a^2(1+\tan^2\theta)}\,d\theta \;=\; \int \frac{a\sec^2\theta}{a^2\sec^2\theta}\,d\theta \;=\; \int \frac{1}{a}\,d\theta \;=\; \frac{1}{a}\theta + c \;=\; \frac{1}{a}\tan^{-1}\frac{x}{a}. \tag{3.28}$$

Example 3.9. $I = \int \frac{1}{x^2+2x+2}\,dx$.

Given that the denominator may be rewritten as $(x+1)^2+1$ by the process of completing the square, we see that the integral is in the same category as the previous one. So we set,

$$x+1 = \tan\theta \qquad\Rightarrow\qquad dx = \sec^2\theta\,d\theta,$$

and hence,

$$\begin{aligned}
I &= \int \frac{1}{x^2+2x+2}\,dx\\
&= \int \frac{1}{(x+1)^2+1}\,dx && \text{using the completing of the square}\\
&= \int \frac{\sec^2\theta}{\sec^2\theta}\,d\theta && \text{using } x+1=\tan\theta\\
&= \theta + c\\
&= \tan^{-1}(x+1)+c. && \text{returning to } x.
\end{aligned} \tag{3.29}$$

Example 3.10. $I = \displaystyle\int \frac{1}{x^2+4x+13}\,dx.$

Another on the same theme. The denominator may be rewritten as $(x+2)^2+9$ and therefore we need the substitution,

$$x+2 = 3\tan\theta \qquad \Rightarrow \qquad dx = 3\sec^2\theta\,d\theta,$$

and hence,

$$\begin{aligned}
I &= \int \frac{1}{x^2+4x+13}\,dx \\
&= \int \frac{1}{(x+2)^2+3^2}\,dx && \text{using the completing of the square} \\
&= \int \frac{3\sec^2\theta}{9\sec^2\theta}\,d\theta && \text{using } x+2=3\tan\theta \\
&= \tfrac{1}{3}\theta + c \\
&= \tfrac{1}{3}\tan^{-1}\Big(\tfrac{1}{3}(x+2)\Big) + c && \text{returning to } x.
\end{aligned} \tag{3.30}$$

Example 3.11. Evaluate $I = \displaystyle\int \frac{f'(x)}{f(x)}\,dx.$

Here $f(x)$ is an unknown function of x. This particular example of an integral is very important indeed and the final result is frequently used as a *method* in and of itself. The evaluation of I follows from the substitution, $y = f(x)$, and hence $dy = f'(x)\,dx$. The integral transforms to

$$I = \int \frac{1}{y}\,dy = \ln|y| + c = \ln|f(x)| + c. \tag{3.31}$$

This general integral illustrates what is often called the f dashed over f form. So if it is recognised that the numerator is the derivative of the denominator, then this result may be applied immediately. The following example uses this.

Example 3.12. Determine $I = \displaystyle\int \frac{2x+2}{x^2+2x+2}\,dx.$

This has the same denominator as the one in Example 3.9, but we note first that the integrand has the precise form, f'/f where $f = x^2+2x+2$. Hence,

$$I = \ln|x^2+2x+2| + c. \tag{3.32}$$

Example 3.13. It is rare that an integrand has the precise f'/f form but it may be close. This example illustrates such a case: Determine $I = \displaystyle\int \frac{x}{x^2+2x+2}\,dx.$

The numerator, x, may be coerced to yield an integral which involves cases that we have already seen in Examples 3.9 and 3.11. Here we go....

$$\begin{aligned} I &= \int \frac{x}{x^2+2x+2}\,dx \\ &= \tfrac{1}{2}\int \frac{2x+2-2}{x^2+2x+2}\,dx && \text{red is the } f'/f \text{ case, black is Ex. 3.9} \\ &= \tfrac{1}{2}\int \frac{2x+2}{x^2+2x+2}\,dx - \int \frac{1}{x^2+2x+2}\,dx && \text{to get the two known forms} \\ &= \tfrac{1}{2}\ln|x^2+2x+2| - \tan^{-1}(x+1) + c. \end{aligned} \tag{3.33}$$

Example 3.14. Find $\int \cot x\,dx$.

Given that $\cot x = \cos x/\sin x$ then the integrand is already in f'/f form. Hence the integral is $I = \ln|\sin x| + c$.

3.4 Integration using Partial Fractions

The technique of partial fractions may be used to integrate functions which are the ratio of two polynomials such as $1/(1-x^2)$ and $(x-1)/(x^3+1)$. It begins with the factorisation of the denominator of the fraction, and the original function then gets split into smaller and more easily integrable parts.

Example 3.15. Determine $I = \int \frac{1}{1-x^2}\,dx$.

We proceed by first reducing the integrand to more manageable bits (i.e. the more easily integrable bits).

$$\text{Let } \frac{1}{1-x^2} = \frac{1}{(1-x)(1+x)} = \frac{A}{1-x} + \frac{B}{1+x}. \tag{3.34}$$

The expression involving the constants A and B is the essential core idea of the method of Partial Fractions. We now need to find A and B, and I prefer to start it by multiplying by the original denominator to get

$$1 = (1+x)A + (1-x)B. \tag{3.35}$$

We may now follow either of two ways, although the first way I present becomes very cumbersome when applied to cases with three or more unknown constants.

<u>**Method 1**</u>: We equate coefficients of 1 and of x separately...

$$\text{Constant terms: } 1 = A + B$$
$$\text{Coefficients of } x\text{: } 0 = A - B$$

This pair of simultaneous equations is solved easily to give $A = B = \frac{1}{2}$.

<u>**Method 2**</u>: choose appropriate values of x ...

$$x = 1 \quad \Rightarrow \quad 1 = 2A$$
$$x = -1 \quad \Rightarrow \quad 1 = 2B.$$

Clearly the two methods yield the same valus for A and B, but when we choose values of x which are roots of the original denominator ($(x-1)(x+1)$, in this case), then the equations for the unknown constants decouple, and the answer is obtained much more quickly.

To complete this problem we have

$$\begin{aligned}
\int \frac{1}{1-x^2}\,dx &= \int \Big[\frac{\frac{1}{2}}{1-x} + \frac{\frac{1}{2}}{1+x}\Big]\,dx \\
&= \int \Big[-\frac{\frac{1}{2}}{x-1} + \frac{\frac{1}{2}}{x+1}\Big]\,dx \qquad \text{I prefer the first integral this way — it's safer} \\
&= \tfrac{1}{2}\Big[\ln|x+1| - \ln|x-1|\Big] + c \\
&= \tfrac{1}{2}\ln\Big|\frac{x+1}{x-1}\Big| + c.
\end{aligned} \tag{3.36}$$

Example 3.16. $I = \displaystyle\int \frac{2}{x^3+3x^2+2x}\,dx.$

The denominator of the integrand factorises to

$$x^3 + 3x^2 + 2x = x(x+1)(x+2),$$

and hence we set

$$\frac{2}{x^3+3x^2+2x} = \frac{2}{x(x+1)(x+2)} = \frac{A}{x} + \frac{B}{x+1} + \frac{C}{x+2},$$

where A, B and C are to be found. On multiplying by the original denominator we obtain,

$$2 = A(x+1)(x+2) + Bx(x+2) + Cx(x+1). \tag{3.37}$$

We will use the three values, $x = 0$, -1 and -2, to evaluate the constants since these are the roots of the linear factors of the original denominator. Substituting these in turn into Eq. (3.37) yields,

$$2 = 2A \qquad 2 = -B \qquad 2 = 2C,$$

in turn and hence $A = C = 1$ and $B = -2$. The integral now becomes,

$$\begin{aligned}
I &= \int \Big[\frac{1}{x} - \frac{2}{x+1} + \frac{1}{x+2}\Big]\,dx \\
&= \ln|x| - 2\ln|x+1| + \ln|x+2| + c \\
&= \ln\Big|\frac{x(x+2)}{(x+1)^2}\Big| + c \\
&= \ln\Big|\frac{x^2+2x}{x^2+2x+1}\Big| + c.
\end{aligned} \tag{3.38}$$

Example 3.17. Evaluate $I = \displaystyle\int \frac{t-1}{t^3+t^2+t+1}\,dt.$

This example illustrates what happens when one cannot find real linear factors of the denominator. First note that the denominator may be factorised to get

$$t^3+t^2+t+1 = (t+1)(t^2+1),$$

where (t^2+1) does not have real linear factors and is said to be an **irreducible quadratic**. The modification to the standard method of using partial fractions is to use a linear numerator (rather than a constant) when dealing with a quadratic denominator which has complex roots. Mathematically we set

$$\frac{t-1}{t^3+t^2+t+1} = \frac{t-1}{(t+1)(t^2+1)} = \frac{A}{t+1} + \frac{Bt+C}{t^2+1}. \tag{3.39}$$

Multiplying by (t^3+t^2+t+1) we obtain,

$$t-1 = A(t^2+1) + (Bt+C)(t+1).$$

We may find A, B and C using a mixture of the methods used in Example 3.15. It's a bit of an art spotting the quickest/safest way through.

$$\begin{aligned}
\text{Let } t=-1: &\qquad 2A=-2 \quad\Rightarrow\quad A=-1 \\
\text{Coefficient of } t^2: &\qquad A+B=0 \quad\Rightarrow\quad B=1 \\
\text{Let } t=0: &\qquad A+C=-1 \quad\Rightarrow\quad C=0.
\end{aligned} \tag{3.40}$$

Using this information we are now in a position to find the original integral, which again uses the $\int f'/f$ result in the middle, somewhere — I'll leave you to spot it!

$$\begin{aligned}
I &= \int \frac{t-1}{t^3+t^2+t+1}\,dt \\
&= \int \Big[\frac{t}{t^2+1} - \frac{1}{t+1}\Big]\,dt \\
&= \tfrac{1}{2}\int \frac{2t}{t^2+1}\,dt - \int \frac{1}{t+1}\,dt \\
&= \tfrac{1}{2}\ln|t^2+1| - \ln|t+1| + c \\
&= \ln\Big|\frac{(t^2+1)^{1/2}}{t+1}\Big| + c.
\end{aligned} \tag{3.41}$$

Note: The final page of this Chapter gives a brief checklist of the "rules of the game" when applying the method of partial fractions.

Example 3.17 (reprise). Evaluate $I = \displaystyle\int \frac{t-1}{t^3+t^2+t+1}\,dt.$

I'll repeat Example 3.17 solely to demonstrate why one needs to use a linear function, $Bt+C$, in the numerator of the quotient, $(Bt+C)/(t^2+1)$, in Eq. (3.39). The quadratic, t^2+1, is irreducible but it is so only when one is seeking real linear factors. Given that $t^2+1=(t-j)(t+j)$, we could proceed with the analysis using these instead. So we have,

$$\frac{t-1}{t^3+t^2+t+1} = \frac{t-1}{(t+1)(t^2+1)} = \frac{t-1}{(t+1)(t-j)(t+j)} = \frac{A}{t+1}+\frac{B}{t-j}+\frac{C}{t+j}. \quad (3.42)$$

Multiplying up in the usual way we have,

$$t-1 = A(t-j)(t+j)+B(t+1)(t+j)+C(t+1)(t-j). \quad (3.43)$$

Now we evaluate the unknown coefficients:

$$\begin{aligned}
\text{Let } t=-1: &\quad -2 = A(-1-j)(-1+j) = 2A \quad \Rightarrow \quad A=-1\\
\text{Let } t=j: &\quad -1+j = 2j(1+j)B = 2B(j-1) \quad \Rightarrow \quad B=\tfrac{1}{2}\\
\text{Let } t=-j: &\quad -1-j = -2j(1-j)C = -2(j+1) \quad \Rightarrow \quad C=\tfrac{1}{2}.
\end{aligned} \quad (3.44)$$

Hence we obtain,

$$\begin{aligned}
\frac{t-1}{t^3+t^2+t+1} &= \frac{-1}{t+1}+\frac{1/2}{t-j}+\frac{1/2}{t+j}\\
&= \frac{-1}{t+1}+\frac{\frac{1}{2}(t+j)+\frac{1}{2}(t-j)}{t^2+1}\\
&= \frac{-1}{t+1}+\frac{t}{t^2+1},
\end{aligned} \quad (3.45)$$

as in Example 3.17.

So clearly the strategy of using a linear function as the numerator when the denominator is an irreducible quadratic works for this particular case, but does it work in general? Yes, it does, and while I won't supply a general proof of this, the following analysis comes close.

Suppose we have the following two complex linear factors of an irreducible quadratic: $t=a\pm bj$. Hence this general case may be written down as,

$$\frac{\text{polynomial}}{(\text{polynomial})\times(t-a-bj)(t-a+bj)} = \frac{A}{t-a-bj}+\frac{B}{t-a+bj}+\text{other terms}, \quad (3.46)$$

where neither of the two polynomials contain either of these complex factors. Given that the left hand side of Eq. (3.46) is real then the right hand side must also be real. On the right hand side the linear factors are complex conjugates of one another, and this implies that A and B must also be complex conjugates of one another. So we may therefore let $A=C+Dj$ and $B=C-Dj$.

Hence,

$$\begin{aligned}
\frac{A}{t-a-bj}+\frac{B}{t-a+bj} &= \frac{C+Dj}{t-a-bj}+\frac{C-Dj}{t-a+bj} \\
&= \frac{(C+Dj)(t-a+bj)+(C-Dj)(t-a-bj)}{(t-a-bj)(t-a+bj)} \\
&= \frac{2C(t-a)-2Db}{(t+a)^2+b^2} \\
&= \frac{\mathcal{C}t+\mathcal{D}}{(t+a)^2+b^2},
\end{aligned} \tag{3.47}$$

where the as-yet unknown real constants, $\mathcal{C}$ and $\mathcal{D}$, are given in terms of the earlier unknown constants, C and D, and these may be traced back to the original A and B. So the use of the form,

$$\frac{A}{t-a-bj}+\frac{B}{t-a+bj},$$

for the two complex linear factors is equivalent to the linear/quadratic form,

$$\frac{\mathcal{C}t+\mathcal{D}}{(t+a)^2+b^2}, \tag{3.48}$$

which was used without explanation in Eq. (3.39). In practice it is usually much quicker to use the linear/quadratic form given in (3.48).

3.5 Integration of products

In many important applications, such as Fourier series and integral transforms both of which you will meet in ME10305 Mathematics 2 next semester, it is often necessary to integrate a product of two functions. I find that the classic method for doing this, which is called **integration by parts**, is very time-consuming and susceptible to more errors than one would expect of a technique which, after all, is often taught in school and subsequently used frequently in undergraduate work. The fault, I believe, lies not in the ability of those who teach but in the formula itself which, as far as I am aware, is used almost universally.

So there we are, outrageously I have trashed something that almost everyone uses! Maybe I had better demonstrate a better way in the next few pages.

The standard derivation uses the well–known result for the derivative of the product of two functions, $u(x)$ and $v(x)$:

$$\frac{d(uv)}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}. \tag{3.49}$$

This is rearranged slightly and integrated to obtain the result,

$$\int u\frac{dv}{dx}\,dx = uv - \int v\frac{du}{dx}\,dx, \tag{3.50}$$

or the other way around,

$$\int v\frac{du}{dx}\,dx = uv - \int u\frac{dv}{dx}\,dx. \tag{3.51}$$

As already mentioned, the application of this expression should be straightforward enough — in Eq. (3.50) one chooses which function is u, which is dv/dx, then find both v and du/dx, substitute into Eq. (3.50) and then see what happens next. Here's an example.

Example 3.18. Integrate $\int x\,e^{ax}\,dx$ which involves the product of x and e^{ax}.

The overall aim of the formula in Eq. (3.50) is to replace the integral on the left with a right hand one which may be solved or at least is easier to solve. In the light of this, if we were to identify dv/dx with x then the resulting integral would involve $\frac{1}{2}x^2$ which yields a more difficult product to integrate. On the other hand, the identification of v with x means that the resulting integral involves 1 instead.

Therefore a typical way in which this is written down is like this:

$$\text{Let} \quad \left\{ \begin{array}{l} \dfrac{du}{dx} = e^{ax} \\[2ex] v = x \end{array} \right. \qquad \Longrightarrow \qquad \left\{ \begin{array}{l} u = e^{ax}/a \\[2ex] \dfrac{dv}{dx} = 1. \end{array} \right. \tag{3.52}$$

Substitution into Eq. (3.50) gives,

$$\int x\,e^{ax}\,dx \;=\; \frac{x\,e^{ax}}{a} - \int \frac{1 \times e^{ax}}{a}\,dx. \tag{3.53}$$

This right hand side integral becomes e^{ax}/a^2, and hence the final result is,

$$\int x\,e^{ax}\,dx \;=\; \frac{x\,e^{ax}}{a} - \frac{e^{ax}}{a^2} \;=\; \frac{1}{a^2}(ax-1)e^{ax}, \tag{3.54}$$

after a little tidying-up. Of course there should be an arbitrary constant in there too.

So what is the problem with this? I reckon that this example is fine, but if we were to be asked to integrate $x^2\,e^{ax}$ then the selection and writing out of information corresponding to Eq. (3.52) would need to be done twice as well the equivalent of Eq. (3.53) twice. This will be getting on for half a page of writing.

If, in the other hand, I were to ask for the integral of x^5e^{ax}, then this represents five successive integrations by parts, and correct choices would need to be made for each one, together with more than a page of writing which would be difficult to check. I regard that as unsatisfactory.

There has to be a better way....

....and there is. For integrals such as $x^5 e^{ax}$ there is a way to write the answer down in one line, but we'll need to do a little derivation first. Sorry about that...

Referring to Equation (3.50), it seems to me rather unnatural to start off with the product of two functions, one of which is a *derivative*. Since we are faced with a product of two functions, it would be better for both functions to be represented explicitly in the integrand on the left hand side of the equation. Therefore let us set

$$u(x) = f(x) \qquad \text{and} \qquad \frac{dv}{dx} = g(x)$$

in Eq. (3.50). We will need to use the superscript notation mentioned in the Differentiation chapter as a shorthand, for convenience. Therefore we let,

$$f^{(1)} = \frac{df}{dx}, \qquad f^{(2)} = \frac{d^2 f}{dx^2},$$

and

$$g^{(-1)} = \int g\,dx, \qquad g^{(-2)} = \int g^{(-1)}\,dx,$$

and so on. So positive superscripts denote the number of differentiations, while negative superscripts denote how many times an indefinite integral has been taken.

Equation (3.50) now becomes,

$$\int fg\,dx = fg^{(-1)} - \int f^{(1)} g^{(-1)}\,dx. \tag{3.55}$$

As before, we choose one function to be differentiated (called f here) and one to be integrated (called g here).

Thus far we have only achieved a change in notation. The reason for that change becomes clearer when we apply this latest formula recursively. In other words, we will apply the way the formula works to the right hand side integral. This will give,

$$\int f^{(1)} g^{(-1)}\,dx = f^{(1)} g^{(-2)} - \int f^{(2)} g^{(-2)}\,dx. \tag{3.56}$$

Nice! And now we may substitute Eq. (3.56) into Eq. (3.55) to obtain,

$$\int fg\,dx = fg^{(-1)} - f^{(1)} g^{(-2)} + \int f^{(2)} g^{(-2)}\,dx, \tag{3.57}$$

and this represents two integrations by parts. One may continue this process as many times as one wishes. So the formulae for five and six integrations by parts are,

$$\int fg\,dx = fg^{(-1)} - f^{(1)} g^{(-2)} + f^{(2)} g^{(-3)} - f^{(3)} g^{(-4)} + f^{(4)} g^{(-5)} - \int f^{(5)} g^{(-5)}\,dx. \tag{3.58}$$

$$\int fg\,dx = fg^{(-1)} - f^{(1)} g^{(-2)} + f^{(2)} g^{(-3)} - f^{(3)} g^{(-4)} + f^{(4)} g^{(-5)} - f^{(5)} g^{(-6)} + \int f^{(6)} g^{(-6)}\,dx. \tag{3.59}$$

So do I wish you to remember all of these new formulae? No, not at all. Definitely not. Rather, I would prefer everyone to follow a set of rules that will work every time and for however many integrations-by-parts that are needed. Here they are:

1. The first product on the right hand side, $fg^{(-1)}$, has the g-term integrated but the f-term has been left alone. So we start by integrating, which feels natural given that the overall aim is to find an integral!

2. Each successive product which appears is related to the previous product by having the term called f being differentiated and the term called g being integrated.

3. The signs alternate between the terms on the right hand side of (3.59) including the one for the final integral.

4. The final integral (should it be required) uses a final differentiation of the preceding f term.

Do check all of these statements against the formula given in Eq. (3.59). Use of these rules will make Integration by Parts quite astonishingly fast and easily checkable afterwards.

The following two examples won't need to use the final integral because the power of x will eventually become zero after a sufficient number of differentiations.

Example 3.19. Find the integral, $\int x^2 e^{-ax}\,dx$.

We will choose to differentiate the x^2 terms and to integrate the exponential. We obtain the following,

$$\int \underbrace{x^2}_{D}\;\underbrace{e^{-ax}}_{I}\,dx = \underbrace{\left[x^2\right]}_{D_0}\underbrace{\left[\frac{e^{-ax}}{-a}\right]}_{I_1} - \underbrace{\left[2x\right]}_{D_1}\underbrace{\left[\frac{e^{-ax}}{a^2}\right]}_{I_2} + \underbrace{\left[2\right]}_{D_2}\underbrace{\left[\frac{e^{-ax}}{-a^3}\right]}_{I_3} - \underbrace{\left[0\right]}_{D_3}\underbrace{\left[\frac{e^{-ax}}{a^4}\right]}_{I_4} + c. \qquad (3.60)$$

Here I have placed red annotations under the different terms to indicate how many differentiations or integrations have taken place — if needed, these may be checked against Eq. (3.59). The final product, $D_3 I_4$, is zero, and therefore there is no need to continue because every further term will also be zero. And I have included the arbitrary constant.

The beauty of the solution in Eq. (3.60) is that one may hop through all the D terms to check that all the differentiations are correct, and through all the I terms to check the integrations.

The final point to make is that all of the D and I terms are sitting cosily inside their own pair of brackets. This means that there is no need to try to tidy up any minus signs during the writing down of a solution like Eq. (3.60) because that will be done on a second line. So the first line, Eq. (3.60), is the integration while the second line, which I am just about to do, is the tidying up. So here's the tidied-up version:

$$\int x^2\,e^{-ax}\,dx = -\frac{(a^2x^2 + 2ax + 2)}{a^3}e^{-ax} + c. \qquad (3.61)$$

Example 3.20. Find the integral, $\int x^5 e^{ax}\,dx$.

This one would normally be the stuff of nightmares, but now we can take it in our stride.... So it's the same as the last Example where we differentiate the power.

$$\begin{aligned}\int \underbrace{x^5}_{D}\;\underbrace{e^{ax}}_{I}\,dx &= \underbrace{\left[x^5\right]}_{D_0}\underbrace{\left[\frac{e^{ax}}{a}\right]}_{I_1} - \underbrace{\left[5x^4\right]}_{D_1}\underbrace{\left[\frac{e^{ax}}{a^2}\right]}_{I_2} + \underbrace{\left[20x^3\right]}_{D_2}\underbrace{\left[\frac{e^{ax}}{a^3}\right]}_{I_3} - \underbrace{\left[60x^2\right]}_{D_3}\underbrace{\left[\frac{e^{ax}}{a^4}\right]}_{I_4} \\ &\quad + \underbrace{\left[120x\right]}_{D_4}\underbrace{\left[\frac{e^{ax}}{a^5}\right]}_{I_5} - \underbrace{\left[120\right]}_{D_5}\underbrace{\left[\frac{e^{ax}}{a^6}\right]}_{I_6} + c \qquad (3.62) \\ &= \frac{(a^5x^5 - 5a^4x^4 + 20a^3x^3 - 60a^2a^2 + 120ax - 120)}{a^6}e^{ax} + c.\end{aligned}$$

Example 3.21. Find the integral, $\int x^2 \cos ax\, dx$.

We have,

$$\int \underbrace{x^2}_{D}\ \underbrace{\cos ax}_{I}\, dx = \underbrace{\left[x^2\right]}_{D_0}\underbrace{\left[\frac{\sin ax}{a}\right]}_{I_1} - \underbrace{\left[2x\right]}_{D_1}\underbrace{\left[\frac{-\cos ax}{a^2}\right]}_{I_2} + \underbrace{\left[2\right]}_{D_2}\underbrace{\left[\frac{-\sin ax}{a^3}\right]}_{I_3}, \tag{3.63}$$

which may, of course, be tidied up a bit and the arbitrary constant included.

Example 3.22. Find $\int e^{ax} \sin bx\, dx$.

This one is a little different because the integration by parts process doesn't terminate. Given that exponentials yield exponentials when differentiated or integrated, and the sines are reproduced every two differentiations or integrations, then (i) it doesn't matter which of the two functions is differentiated and (ii) after two integrations one should obtain an integral which is proportional to the original one.

For this I will choose to integrate the sine, but it worth you attempting the other way to see if a correct solution will be obtained. Hence,

$$\text{Let } I = \int \underbrace{e^{ax}}_{D}\ \underbrace{\sin bx}_{I}\, dx = \underbrace{\left[e^{ax}\right]}_{D_0}\underbrace{\left[\frac{-\cos bx}{b}\right]}_{I_1} - \underbrace{\left[ae^{ax}\right]}_{D_1}\underbrace{\left[\frac{-\sin bx}{b^2}\right]}_{I_2} + \int \underbrace{\left[a^2e^{ax}\right]}_{D_2}\underbrace{\left[\frac{-\sin bx}{b^2}\right]}_{I_2},$$

$$= \frac{e^{ax}}{b^2}\Big[-b\cos bx + a\sin bx\Big] - \frac{a^2}{b^2}I. \tag{3.64}$$

The terms involving I may be brought together on the left to obtain,

$$\left[1 + \frac{a^2}{b^2}\right]I = \frac{e^{ax}}{b^2}\Big[a\sin bx - b\cos bx\Big] \tag{3.65}$$

and therefore we obtain,

$$I = \frac{e^{ax}}{a^2+b^2}\Big[a\sin bx - b\cos bx\Big]. \tag{3.66}$$

Example 3.23. Find $\int \ln|x|\, dx$.

Another classic, and although it doesn't look like a product we can make it so: $\int 1 \times \ln x\, dx$. Once you know the trick, you know the trick. We will differentiate the $\ln x$ because we can! We will only integrate by parts once because it isn't entirely obvious what will happen after that, so then we'll take stock at that point.

$$\begin{aligned}\int \underbrace{1}_{I} \times \underbrace{\ln|x|}_{D}\, dx &= \underbrace{\left[x\right]}_{I_1}\underbrace{\left[\ln|x|\right]}_{D_0} - \int \underbrace{\left[x\right]}_{I_1}\underbrace{\left[\frac{1}{x}\right]}_{D_1} dx && \text{one integration by parts} \\ &= x\ln x - \int 1\, dx && \text{tidied up the integral} \\ &= x\ln x - x + c. \end{aligned} \tag{3.67}$$

Note: An additional rule — if in doubt, do it once!

3.6 Mean and Root Mean Square

It is frequently necessary to evaluate the mean value of a function and its root mean square (or RMS). The mean of a function, $f(x)$, over the interval $a \le x \le b$, is defined as

$$\mu = \text{mean of } f(x) = \frac{1}{b-a}\int_a^b f(x)\,dx. \tag{3.68}$$

This definition may be derived easily by insisting that a rectangle of length $b - a$ and with a height which is the mean of $f(x)$ has the same area as the graph of $f(x)$ in the same range. This is illustrated in Fig. 3.5.

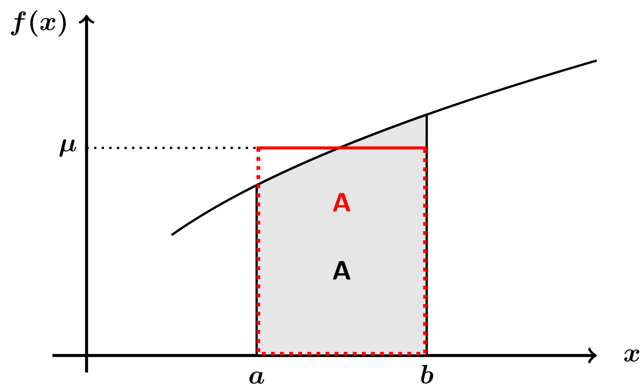


Figure 3.5. Demonstrating the mean value of the function, $f(x)$, in the range $a \le x \le b$. The areas of the grey region and of the red rectangle need to be the same.

The RMS is precisely what it says it is, the square root of the mean of the square of the function. Therefore we have the following definition,

$$\text{RMS of } f(x) = \left[\frac{1}{b-a}\int_a^b [f(x)]^2\,dx\right]^{1/2}. \tag{3.69}$$

Example 3.24. What are the mean and RMS of the function $f(x) = x$ in the interval $1 \le x \le 3$?

They are,

$$\mu = \frac{1}{3-1}\int_1^3 x\,dx = 2, \qquad \text{RMS} = \left[\frac{1}{3-1}\int_1^3 x^2\,dx\right]^{1/2} = \sqrt{\frac{13}{3}}. \tag{3.70}$$

Appendix — a checklist for Partial Fractions

This is a checklist of the various complications which can arise when using partial fractions to simplify the ratio of two polynomials, and it shows what procedure to use in each case.

I will assume that the degree of the polynomial in the denominator is less than that of the numerator, i.e. that the ratio is "bottom heavy". An example would be a quadratic divided by a cubic. If the ratio is not bottom heavy (e.g. a quintic divided by a cubic), then it is always possible to make it so by subtracting out suitable polynomials.

If all the factors of the denominator are different, then let

$$\frac{?}{(x+a)(x+b)(x+c)\cdots} = \frac{A}{x+a} + \frac{B}{x+b} + \frac{C}{x+c} + \cdots .$$

If one factor is repeated, then let

$$\frac{?}{(x+a)^2(x+b)(x+c)\cdots} = \frac{A_1}{x+a} + \frac{A_2}{(x+a)^2} + \frac{B}{x+b} + \frac{C}{x+c} + \cdots .$$

A more complicated example: let

$$\frac{?}{(x+a)^3(x+b)^2(x+c)(x+d)\cdots} = \frac{A_1}{x+a} + \frac{A_2}{(x+a)^2} + \frac{A_3}{(x+a)^3} + \frac{B_1}{x+b} + \frac{B_2}{(x+b)^2} + \frac{C}{x+c} + \frac{D}{x+d} + \cdots .$$

If one of the factors is an irreducible quadratic then let

$$\frac{?}{(x^2+ax+b)(x+c)(x+d)\cdots} = \frac{Ax+B}{x^2+ax+b} + \frac{C}{x+c} + \frac{D}{x+d} + \cdots .$$

A repeated quadratic factor:

$$\frac{?}{(x^2+ax+b)^2(x+c)(x+d)\cdots} = \frac{Ax+B}{x^2+ax+b} + \frac{Cx+D}{(x^2+ax+b)^2} + \frac{E}{x+c} + \frac{F}{x+d} + \cdots .$$

Note that it is always possible to reduce (what is generally known as) an irreducible quadratic factor to two complex linear factors, and then one may proceed in the usual way for linear factors. It might take a little while!

An example of a top-heavy case:

$$\begin{aligned} \frac{x^3}{(x+a)(x+b)} &= \frac{x(x+a)(x+b) + A(x+a)(x+b) + Bx + C}{(x+a)(x+b)} \\ &= x + A + \frac{Bx+C}{(x+a)(x+b)} \quad = \quad x + A + \frac{D}{x+a} + \frac{E}{x+b}, \end{aligned}$$

where $a \neq b$ here, and where A, D and E are to be found.

4 SERIES

4.1 Definition of a sequence

A sequence in an ordered list of numbers which may be either random or generated by experimental readings or else generated according to a given rule. It may be finite or infinite in length. Examples include the following,

$$1, \quad 2, \quad 4, \quad 8, \quad 16, \quad 32, \quad 64, \quad 128, \quad \cdots \qquad (i)$$

$$1, \quad \tfrac{1}{2}, \quad \tfrac{1}{4}, \quad \tfrac{1}{8}, \quad \tfrac{1}{16}, \quad \tfrac{1}{32}, \quad \tfrac{1}{64}, \quad \tfrac{1}{128}, \quad \cdots \qquad (ii)$$

$$2, \quad 4, \quad 6, \quad 8, \quad 10, \quad 12, \quad 14, \quad 16, \quad \cdots \qquad (iii)$$

$$1, \quad -1, \quad 1, \quad -1, \quad 1, \quad -1, \quad 1, \quad -1, \quad \cdots \qquad (iv)$$

$$1, \quad -\tfrac{1}{2}, \quad \tfrac{1}{4}, \quad -\tfrac{1}{8}, \quad \tfrac{1}{16}, \quad -\tfrac{1}{32}, \quad \tfrac{1}{64}, \quad -\tfrac{1}{128}, \quad \cdots \qquad (v)$$

$$1, \quad 4, \quad 1, \quad 4, \quad 2, \quad 1, \quad 3, \quad 5, \quad \cdots \qquad (vi)$$

$$19, \quad 13, \quad 15, \quad 6, \quad 19, \quad 17, \quad 1, \quad 3, \quad \cdots \qquad (vii)$$

Can you identify the rule which governs each member of the different sequences?

Generally we use a subscript notation for sequences and series: x_1, x_2, x_3 and so on. Or, should there be a good reason to do, we could begin with a zero subscript: x_0, x_1, x_2 and so on. Therefore we write the general term (or the n^{th} term) simply as x_n.

Sequence (i) is formed by doubling the preceding term. We could write it in the form,

$$x_n = 2^{n-1} \quad \text{for} \quad n = 1, 2, 3, \cdots, \tag{4.1}$$

or in the alternative way,

$$x_n = 2^n \quad \text{for} \quad n = 0, 1, 2, \cdots . \tag{4.2}$$

or even,

$$x_n = 2^{n-163} \quad \text{for} \quad n = 163, 164, 165, \cdots, \cdots . \tag{4.3}$$

Tastes differ, but I prefer the second one because it is a little tidier. The third is correct but just plain silly.

Sequence (ii). The terms here are the reciprocals of the respective terms in Sequence (i). Therefore,

$$x_n = 2^{-(n-1)} \quad \text{for} \quad n = 1, 2, 3, \cdots \qquad \text{or} \qquad x_n = 2^{-n} \quad \text{for} \quad n = 0, 1, 2, \cdots . \tag{4.4}$$

Sequence (iii). Clearly these are twice the value of the positive integers. Hence

$$x_n = 2n, \quad \text{for} \quad n = 1, 2, 3, \cdots . \tag{4.5}$$

Sequence (iv). Alternating signs. This feature happens very frequently in sequences and series and it is worth knowing about them now, right at the outset.

$$x_n = (-1)^n, \quad \text{for} \quad n = 0, 1, 2, \cdots \qquad \text{or} \qquad x_n = (-1)^{n+1}, \quad \text{for} \quad n = 1, 2, 3, \cdots . \tag{4.6}$$

Note that an alternative form for $(-1)^n$ is $\cos \pi n$. Strange but true. This cosine version will eventually crop up naturally in Fourier Series in ME10305 Mathematics 2 and we usually convert it back into the $(-1)^n$ form.

Sequence (v) is a combination of (ii) and (iv). Hence,

$$x_n = \frac{(-1)^n}{2^n} \quad \text{for} \quad n = 0, 1, 2, \cdots, \qquad \text{or} \qquad x_n = \frac{(-1)^{n-1}}{2^{n-1}} \quad \text{for} \quad n = 1, 2, 3, \cdots. \tag{4.7}$$

In this case I prefer the first version because it looks tidier.

Sequence (vi) is a bit of a trick question. These are the successive digits of $\sqrt{2}$, and although this knowledge is sufficient to determine, say, the 100^{th} significant figure, there isn't a simple formula to tell us what that is without finding all the previous ones first. The following analysis is for interest only, but it is possible to write down some sort of expression; it's a bit of a rocky ride to get there and has no relevance to the unit but I hope that some may enjoy the experience.

First I will introduce the floor function. This is the greatest integer less than or equal to a real number. For positive numbers it is more easily defined as what you get when the decimal places are ignored! So $\mathbf{floor}\,(2.4) = 2$ and $\mathbf{floor}\,(\pi) = 3$. Given that $\sqrt{2} = 1.414\,213\,562$ to 9SFs we can see that $\mathbf{floor}\,(\sqrt{2}) = 1$ and $\mathbf{floor}\,(100\sqrt{2}) = 141$. To be honest I hadn't come across this analysis or even the idea that it is possible before typing these notes, so what follows marks out my voyage of discovery. Therefore a little bit of playing first:

$$\begin{aligned}
x_1 = 1 &= \mathbf{floor}\left[\sqrt{2}\right] \\
x_2 = 4 &= \mathbf{floor}\left[14.14213562... - 10\right] &&= \mathbf{floor}\left[10\sqrt{2} - 10x_1\right] \\
x_3 = 1 &= \mathbf{floor}\left[141.4214562... - 140\right] &&= \mathbf{floor}\left[100\sqrt{2} - 100x_1 - 10x_2\right] \\
x_4 = 4 &= \mathbf{floor}\left[1414.214562... - 1410\right] &&= \mathbf{floor}\left[1000\sqrt{2} - 1000x_1 - 100x_2 - 10x_3\right] \\
x_5 = 2 &= \mathbf{floor}\left[14142.14562... - 14140\right] &&= \mathbf{floor}\left[10000\sqrt{2} - 10000x_1 - 1000x_2 - 100x_3 - 10x_4\right]
\end{aligned} \tag{4.8}$$

This gives the idea of the pattern. The general case, then, is

$$x_n = \mathbf{floor}\left[10^{n-1}\sqrt{2} - 10^{n-1}x_1 - 10^{n-2}x_2 - \cdots - 10x_{n-1}\right]. \tag{4.9}$$

and this may be written more compactly using the summation notation:

$$x_n = \mathbf{floor}\left[10^{n-1}\sqrt{2} - \sum_{i=1}^{n-1} 10^{n-i}x_i\right]. \tag{4.10}$$

Worthy, perhaps interesting, but not at all useful! Just to say that I don't expect this of you in the exams (!!!) but these ideas may prove useful at some point. Possibly.

Sequence (vii). This one is random. The numbers came from the first eight rolls of an icosahedral die where the faces are numbered from 1 up to 20, as opposed to from 1 up to 6 for a standard cubical die.

4.2 Definition of a series

However, the present topic is about series. A series (singular) may be defined as the sum of the terms in a sequence, and series (plural) may also be either finite or infinite in length. The following contrasts sequences and series.

$$\text{Sequence:} \quad x_1,\, x_2,\, x_3,\, x_4,\, \cdots x_n,$$
$$\text{Series:} \quad x_1 + x_2 + x_3 + x_4 + \cdots + x_n = \sum_{i=1}^{n} x_i = S_n. \tag{4.11}$$

So we may create a series by adding the terms in a sequence, but it is possible too to make a sequence from a series. The value, S_n, which is the sum of the first n terms of the series and which is called the n^{th} **partial sum**), can form a sequence: S_1, S_2, S_3 and so on. In more detail, this sequence is,

$$S_1 = x_1, \qquad S_2 = x_1 + x_2, \qquad S_3 = x_1 + x_2 + x_3, \qquad S_4 = x_1 + x_2 + x_3 + x_4, \cdots\cdots. \tag{4.12}$$

In many cases we are usually interested in determining the sum of the series and, when it is of infinite length, whether that sum converges to a finite limit or not. The convergence of an infinite series is not a straightforward concept but there are two main ways of judging whether or not a series converges and these will be covered below. As a taster, it is quite clear that the series formed by adding the terms in sequence (i), above, will diverge and become infinite. On the other hand, the series corresponding to sequence (ii) is a geometric series and converges to 2.

In some cases it is possible to find ways of summing series analytically, or else of proving divergence, but these methods are often quite specific to the type of series being considered. A bit of a pain, really.

We will be interested here in creating series which represent functions; this will involve both Binomial expansions and Taylor's Series. When these power series exist we also need to determine their convergence properties, by which we mean, for what range of values of x do the series converge?

Ideas drawn from all of this will be used to answer the age-old question of determining the values of $0/0$ and ∞/∞. The general non-committal (but completely correct) answer is that it all depends.....

4.3 The binomial expansion, theorem and series

First it is essential to state that $(x + y)$ is a **binomial** because there are two quantities involved. Although much less well-known, the following is a **trinomial**: $(x+y+z)$. No doubt one can guess what $(a+b+c+d)$ should be called!

We may find an alternative form for $(x + y)^2$ easily by multiplying out the brackets in $(x + y)(x + y)$; it is $x^2 + 2xy + y^2$. This fact may be visualised by means of the area of the following square.

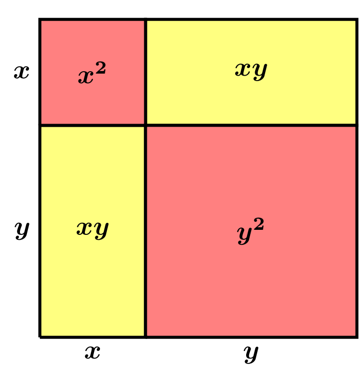


Figure 4.1. Illustrating the binomial expansion of $(x + y)^2$ using areas within a square of side, $(x + y)$.

Further multiplication by an $(x+y)$ factor yields

$$(x+y)^3 = (x+y)(x^2+2xy+y^2) = x^3+3x^2y+3xy^2+y^3.$$

This again may be visualised by means of the volume of a suitable cube.

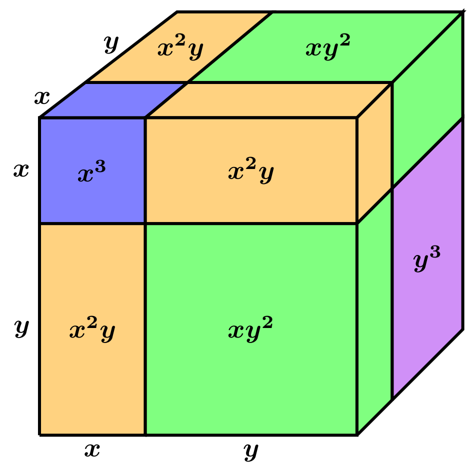


Figure 4.2. Illustrating the binomial expansion of $(x+y)^3$ using cuboids.

Finding higher powers is tedious, and we run out of dimensions with which to visualise them. Fortunately Pascal's triangle allows us to calculate fairly quickly the next few sets of coefficients. It is

$$\begin{array}{l ccccccccccccc}
n \\
0: & & & & & & & 1 \\
1: & & & & & & 1 & & 1 \\
2: & & & & & 1 & & 2 & & 1 \\
3: & & & & 1 & & 3 & & 3 & & 1 \\
4: & & & 1 & & \color{red}{4} & & \color{red}{6} & & 4 & & 1 \\
5: & & 1 & & 5 & & \color{red}{10} & & 10 & & 5 & & 1 \\
6: & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}$$

Figure 4.3. Pascal's triangle, which is one method for evaluating the Binomial coefficients. The red coefficient on row 5 is the sum of the two in red on row 4; this is a general property of the pattern.

So if we take the line 5 then the corresponding row of coefficients is equivalent to

$$(x+y)^5 = x^5+5x^4y+10x^3y^2+10x^2y^3+5xy^4+y^5. \tag{4.13}$$

The coloured data in Pascal's triangle illustrates the well-known property that any entry is the sum of the two just above. That this is so is a simple consequence of multiplying by $(x+y)$:

$$\left(x^4+\boxed{4x^3y}+\boxed{6x^2y^2}+4xy^3+y^4\right)\left(\boxed{x}+\boxed{y}\right) = x^5+5x^4y+\boxed{10x^3y^2}+10x^2y^3+5xy^4+y^5. \tag{4.14}$$

Here we add the product of the two red terms and the product of the blue terms on the left to give the purple term on the right, and this demonstrates the addition-of-previous-coefficients property of the Triangle.

Example 4.1. The binomial expansion may be used as a means of computing quick approximations by hand. For example, if we set $x = 1$ and $y = 0.01$ in (4.13) we get

$$\begin{aligned}(1.01)^5 &= 1^5 + [5 \times 1^4 \times 0.01] + [10 \times 1^3 \times 0.01^2] + \cdots \\ &= 1 + 0.05 + 0.001 + \cdots \\ &\simeq 1.051.\end{aligned} \tag{4.15}$$

Note firstly that one always has to be careful about how large the first neglected term is. In this case it will be 10×0.01^3, i.e. it is roughly 10^{-5} in magnitude and therefore negligible if we intend to keep to 3DPs. Secondly, this method works well only if $x \gg y$ (or the other way around, obviously). It would not be quite as efficient to evaluate 1.1^5 by this method since all the terms would be required.

Example 4.3. A second example is to find a 4DP approximation to 2.004^5.

$$\begin{aligned}2.004^5 &= 2^5(1.002)^5 && \text{to get a leading "1"} \\ &= 2^5(1+0.002)^5 && \text{now in Binomial form} \\ &= 2^5\Big[1 + [5 \times 0.002] + [10 \times 0.002^2] + \cdots\Big] && \text{three terms should be enough?} \\ &= 2^5\Big[1 + 0.01 + 0.000004 + \cdots\Big] \\ &= 2^5 \times 1.01004 = 32.3214 && \text{4DPs.}\end{aligned} \tag{4.16}$$

Here we factored out the 2 in order to make life a little easier — such expansions are easier and safer with a leading 1. Those who have used a calculator to check the above answer will note that the last decimal place is incorrect and that the number should be 32.3213 to four decimal places. This indicates how much care should be taken with such expansions, and I should have taken one more term in the above calculation.

4.4 Explicit formula for the Binomial coefficients

The binomial coefficients are of particular use when the binomial is raised to a large power, but it is clearly impractical to use Pascal's triangle for, say, a 100th power! Fortunately there is an alternative way.

$$(x+y)^n = \binom{n}{0} x^n + \binom{n}{1} x^{n-1}y + \binom{n}{2} x^{n-2}y^2 + \cdots + \binom{n}{n-1} xy^{n-1} + \binom{n}{n} y^n, \tag{4.17}$$

where the Binomial coefficients are defined as,

$$\binom{n}{i} = \frac{n!}{i!(n-i)!} \qquad = \qquad {}^nC_i, \tag{4.18}$$

and where $0!$ is defined as being equal to 1.

If you think that that is a bit of a cheat, then if we were to define $n!$ using the following formula,

$$n! = \int_0^\infty x^n e^{-x}\, dx,$$

for integer values of n, it is straightforward to show that $0! = 1$ by setting $n = 0$. This definition also provides for strange quantities such as $(\frac{1}{2})!$.

Note: the symmetry about the vertical midline of Pascal's triangle is reflected in its coefficients:

$$\binom{n}{i} = \frac{n!}{i!(n-i)!} = \binom{n}{n-i}. \tag{4.19}$$

So let us check if Eqs. (4.17) and (4.18) work. We'll use $n = 5$:

$$\begin{aligned}
\binom{5}{0} &= \frac{5!}{0!\,5!} &= 1, \\
\binom{5}{1} &= \frac{5!}{1!\,4!} &= 5, \\
\binom{5}{2} &= \frac{5!}{2!\,3!} &= 10, \\
\binom{5}{3} &= \frac{5!}{3!\,2!} &= 10, \\
\binom{5}{4} &= \frac{5!}{4!\,1!} &= 5, \\
\binom{5}{5} &= \frac{5!}{5!\,0!} &= 1.
\end{aligned} \tag{4.20}$$

That matches up nicely with Eq. (4.13).

There is a second check that we can do (which is beyond the examinable part of this unit, but it is quick). Referring to the red digits in Fig. 4.3 and the summation property of the coefficients which was demonstrated in Eq. (4.14), we may try to derive the general case. So,

$$\begin{aligned}
& \binom{n}{m} + \binom{n}{m+1} && \text{neighbouring coeffiicents on row } n \\
=\; & \frac{n!}{m!\,(n-m)!} + \frac{n!}{(m+1)!\,(n-m-1)!} && \text{by definition} \\
=\; & \frac{n!}{(m+1)!\,(n-m)!} \times [(m+1) + (n-m)] && \text{need your wits on this one!} \\
=\; & \frac{n! \times (n+1)}{(m+1)!\,(n-m)!} && \\
=\; & \frac{(n+1)!}{(m+1)!\,[(n+1)-(m+1)]!} && \text{sneaky treatment of the bottom-right term} \\
=\; & \binom{n+1}{m+1}. &&
\end{aligned} \tag{4.21}$$

That was a bit of a tour-de-force, but that very last step could have been predicted from the values of n and m corresponding to the red values in Fig. 4.3.

4.5 From the Binomial expansion to the Binomial series.

We may now maniplate this new explicit definition of the binomial coefficients, Eq. (4.18), in order to rewrite them in a differerent form. On taking the general case we have,

$$\begin{aligned}
\binom{n}{0} &= \binom{n}{n} = \frac{n!}{0!\,n!} = 1,\\
\binom{n}{1} &= \binom{n}{n-1} = \frac{n!}{1!\,(n-1)!} = n,\\
\binom{n}{2} &= \binom{n}{n-2} = \frac{n!}{2!\,(n-2)!} = \frac{n(n-1)}{2},\\
\binom{n}{3} &= \binom{n}{n-3} = \frac{n!}{3!\,(n-3)!} = \frac{n(n-1)(n-2)}{3!},
\end{aligned} \tag{4.22}$$

and so on following that pattern. Therefore Eq. (4.17) with $y = 1$ begins as

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \frac{n(n-1)(n-2)(n-3)}{4!}x^4 + \cdots. \tag{4.23}$$

Again this has a pattern that is useful to memorise.

Surprisingly, this latest formula, Eq. (4.23), is also valid for both positive fractional values of n and for negative values — this will be proved using Taylor's series which will follow later.

Example 4.4. Find the Binomial Series for $(1+x)^{-1}$.

We simply apply Eq. (4.23) with $n = -1$. Hence

$$\begin{aligned}
(1+x)^{-1} &= 1 + (-1)x + \frac{(-1)(-2)}{2}x^2 + \frac{(-1)(-2)(-3)}{3!}x^3 + \cdots\\
&= 1 - x + x^2 - x^3 + \cdots = \sum_{n=0}^{\infty}(-1)^n x^n.
\end{aligned} \tag{4.24}$$

We went as far as x^3 on the first line because this was far enough to observe the pattern which was then simplified (though cancellations) on the next line, and then eventually expressed in summation form. The acquisition of a general term is the usual final objective of problems like this.

Note that it is possible to use the series in Eq. (4.24) to write down others. For example, if we were to replace all instances of "x" by "$-x$" then we would obtain,

$$(1-x)^{-1} = 1 + x + x^2 + x^3 + \cdots = \sum_{n=0}^{\infty} x^n. \tag{4.25}$$

We could also replace all instances of "x" in Eq. (4.24) by "x^2" to obtain,

$$(1+x^2)^{-1} = 1 - x^2 + x^4 - x^6 + \cdots = \sum_{n=0}^{\infty}(-1)^n x^{2n}. \tag{4.26}$$

Honestly, that isn't cheating!

Example 4.5. Find the Binomial Series for $(1-x)^{-2}$.

Again we apply Eq. (4.23) with $n=-2$ where x is replaced by $-x$. Hence

$$\begin{aligned}(1-x)^{-2} &= 1+(-2)(-x)+\frac{(-2)(-3)}{2}(-x)^2+\frac{(-2)(-3)(-4)}{3!}(-x)^3+\cdots \\ &= 1+2x+3x^2+4x^3+\cdots \\ &= \sum_{n=0}^{\infty}(n+1)x^n.\end{aligned} \tag{4.27}$$

Example 4.6 Find the Binomial series for $(1+x)^{-1/2}$. This one is much trickier.

We apply Eq. (4.23) with $n=-1/2$. Hence

$$(1+x)^{-1/2}=1+(-\tfrac{1}{2})x+\frac{(-\frac{1}{2})(-\frac{3}{2})}{2!}x^2+\frac{(-\frac{1}{2})(-\frac{3}{2})(-\frac{5}{2})}{3!}x^3+\cdots. \tag{4.28}$$

This looks like a mess, and while it is indeed true that it is a mess it can be tidied up after some work. So we'll have a look at the coefficient of x^4 as an exemplar. This is the Binomial series equivalent of white water rafting; we have,

$$\begin{aligned}\text{Coefficient of } x^4 &= \frac{(-1)^4\,1\times3\times5\times7}{2^4\,4!} && \text{leave the } (-1)^4 \text{ as it is} \\ &= \frac{(-1)^4\,1\times2\times3\times4\times5\times6\times7\times8}{2^4\,4!\,\times2\times4\times6\times8} && \text{now we have a factorial in the numerator} \\ &= \frac{(-1)^4 8!}{2^4\,4!\times2^4\,4!} && \text{note carefully what happens with the red terms.}\end{aligned} \tag{4.29}$$

The form taken by this coefficient tells us that the general coefficient is $\dfrac{(-1)^n\,(2n)!}{2^{2n}\,n!\,n!}$, and hence

$$(1+x)^{-1/2}=\sum_{n=0}^{\infty}\frac{(-1)^n\,(2n)!}{2^{2n}\,n!\,n!}x^n. \tag{4.30}$$

Note that it is possible to integrate or differentiate these power series in order to derive others although the constant of integration needs to be calculated when integrating. Two very useful examples of such an integration follow.

Example 4.7. Find a power series representation of $y = \ln(1+x)$ using Binomial series.

In its present form we cannot apply Eq. (4.23), but we will be able to after differentiation:

$$\begin{aligned} y &= \ln(1+x) \\ \Rightarrow \frac{dy}{dx} &= (1+x)^{-1} \\ &= 1 - x + x^2 - x^3 + x^4 \cdots && \text{using the previous result in Eq. (4.24)} \\ \Rightarrow y &= c + x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{4}x^4 + \tfrac{1}{5}x^5 \cdots && c \text{ is an arbitrary constant} \\ &= x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{4}x^4 + \tfrac{1}{5}x^5 \cdots && y(0) = 0, \text{ hence } c = 0 \\ &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}x^n}{n}. \end{aligned} \tag{4.31}$$

Example 4.8. Find the power series representation of $\sin^{-1}(x)$.

The following is how one may write the inverse sine function in integral form.

$$\sin^{-1} x = \int_0^x \frac{dt}{\sqrt{1-t^2}}. \tag{4.32}$$

Give our experience with Example 4.7, we shall begin by finding the Binomial series of $(1-t^2)^{-1/2}$. Rather than undertaking that expansion from scratch, we may use Eq. (4.30) and replace all instances of 'x' by '$-t^2$'. This gives us,

$$(1-t^2)^{-1/2} = \sum_{n=0}^{\infty} \frac{(2n)!}{2^{2n}\, n!\, n!} t^{2n}. \tag{4.33}$$

We may now integrate each term in turn and evaluate these between $t = 0$ and $t = x$. Hence,

$$\sin^{-1} x = \int_0^x (1-t^2)^{-1/2}\, dt = \sum_{n=0}^{\infty} \frac{(2n)!}{(2n+1)2^{2n}\, n!\, n!} x^{2n+1}. \tag{4.34}$$

Hence,

$$\begin{aligned} \sin^{-1} x &= \sum_{n=0}^{\infty} \frac{(2n)!}{(2n+1)2^{2n}\, n!\, n!} x^{2n+1} \\ &= x + \frac{2!}{3 \times 2^2 \times 1! \times 1!} x^3 + \frac{4!}{5 \times 2^4 \times 2! \times 2!} x^5 + \frac{6!}{7 \times 2^6 \times 3! \times 3!} x^7 + \cdots, \end{aligned} \tag{4.35}$$

where the first terms have been written out.

Note: One has to be a little careful with $(2n)!$ — this is **not** the same as $2 \times (n!)$. So if $n = 5$, then $(2n)! = 10! = 3628800$ while $2(n!) = 2 \times 5! = 240$. This is another example where it is possible to write something down ambiguously, or to make an incorrect assumption about what a formula means. Consider the following three expressions:

$$(2n)! \qquad 2n! \qquad 2(n!).$$

Of these the first and the third are written correctly, and the brackets have been used (and hopefully will be interpreted) correctly. The problem with the second on is that it could be interpreted as either the first of the third, and therefore my recommendation is that it should never be used. One could fit the act of taking factorials into the BODMAS/PEMDAS/BEDMAS/BIDMAS rules by saying that it has a higher priority than multiplication and division. In that case $2n!$ ought to be the double of $n!$, but I still think that it is open to misinterpretation and therefore it should be avoided.

4.6 Power series, Maclaurin's series and Taylor's series

Any series of the type

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \cdots \tag{4.36}$$

is called a **power series** simply because it consists of powers of x. An example of such a series is

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}, \tag{4.37}$$

but how is such an expression obtained, particularly since we cannot use the Binomial expansion?

Let us first assume that it is possible to express an arbitrary function, $f(x)$, in a power series form with positive integer powers, and therefore we set

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \cdots . \tag{4.38}$$

We may find the value of a_0 by substituting $x = 0$ into Eq. (4.38). This yields $\boxed{a_0 = f(0)}$. Now we shall differentiate Eq. (4.38) and substitute $x = 0$ again. We get,

$$f'(x) = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \cdots , \qquad \Rightarrow \boxed{a_1 = f'(0)}. \tag{4.39}$$

Now, two terms is rarely a good basis for extrapolation — yes, a_2 will not be equal to $f''(0)$! (By the way, that was an exclamation mark, not a factorial!). So we'll differentiate Eq. (4.39),

$$f''(x) = 2a_2 + (3\times 2)a_3 x + (4\times 3)a_4 x^2 + (5\times 4)a_5 x^3 \cdots , \qquad \Rightarrow \boxed{a_2 = \tfrac{1}{2} f''(0)}. \tag{4.40}$$

Actually, three terms are also not a good basis for extrapolation. The denominators of the numerical coefficients for a_0, a_1 and a_2 are, $1,\ 1,\ 2$, which could represent various things including the Fibonacci sequence. So we need to go a little further. Actually, we'll do it twice more, and we get, in turn:

$$f'''(x) = (3\times 2)a_3 + (4\times 3\times 2)a_4 x + (5\times 4\times 3)a_5 x^2 + \cdots \qquad \Rightarrow \qquad \boxed{a_3 = f'''(0)/(3\times 2)}, \tag{4.41}$$

and

$$f''''(x) = (4\times 3\times 2)a_4 + (5\times 4\times 3\times 2)a_5 x + \cdots \qquad \Rightarrow \qquad \boxed{a_4 = f''''(0)/(4\times 3\times 2)}. \tag{4.42}$$

So these denominators are factorials. We have just derived what is usually called **Maclaurin's series**:

$$\begin{aligned} f(x) &= f(0) + f'(0)\,x + \frac{f''(0)\,x^2}{2!} + \frac{f'''(0)\,x^3}{3!} + \frac{f''''(0)\,x^4}{4!} + \cdots \\ &= \sum_{n=0}^{\infty} \left.\frac{d^n f}{dx^n}\right|_{x=0} \frac{x^n}{n!}. \end{aligned} \tag{4.43}$$

The Maclaurin expansion focuses on $x = 0$ as the point 'about which' the function is expanded. More generally we may expand about any other point, and therefore we could also have

$$\begin{aligned} f(x) &= f(a) + f'(a)\,(x-a) + \frac{f''(a)\,(x-a)^2}{2!} + \frac{f'''(a)\,(x-a)^3}{3!} + \frac{f''''(a)\,(x-a)^4}{4!} + \cdots \\ &= \sum_{n=0}^{\infty} \left.\frac{d^n f}{dx^n}\right|_{x=a} \frac{(x-a)^n}{n!}. \end{aligned} \tag{4.44}$$

Equation (4.44) may also be derived in exactly the same way as for Eq. (4.43), namely by successive differentiation followed by the substitution of $x = a$. This more general expression is called **Taylor's series about $x = a$**. Thus Maclaurin's series is simply Taylor's series about $x = 0$.

Note: that while Eq. (4.44) is the one which we shall use for this unit, the University's Formula Book has it written down slightly differently; there it is

$$f(x) \;=\; f(x_0) + \delta f'(x_0) + \frac{\delta^2}{2!} f''(x_0) + \frac{\delta^3}{3!} f'''(x_0) \cdots + \frac{\delta^n}{n!} f^{(n)}(x_0) + R_{n+1}. \tag{4.45}$$

In this expression $f^{(n)}$ is the n^{th} derivative, $\delta = x - x_0$ and the value, R_{n+1} is the error due to using only a finite series. Clearly x_0 plays the same role that a does in the above analysis, but we will not be considering the error term.

Example 4.9. Find the Taylor's series of e^{bx} about $x = 0$.

An alternative way of phrasing this example is: Expand e^{bx} about $x = 0$.

Essentially all we need to do is find successive derivatives of the function, let $x = 0$ in those derivatives and then assemble the Taylor's series using Eq. (4.43). I prefer to use a Table to do this because it is tidy and may be checked easily. So we have,

n	$f^{(n)}$	$f^{(n)}(0)$
0	e^{bx}	1
1	be^{bx}	b
2	b^2e^{bx}	b^2
3	b^3e^{bx}	b^3

Equation (4.43) yields

$$e^{bx} = 1 + bx + \frac{b^2x^2}{2!} + \frac{b^3x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{b^n x^n}{n!}. \tag{4.46}$$

Once more this type of result may be manipulated. By the simple act of replacing b by $-b$ we obtain,

$$e^{-bx} = 1 - bx + \frac{b^2x^2}{2!} - \frac{b^3x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n b^n x^n}{n!}. \tag{4.47}$$

This may look like a sleazy manoeuvre but it is correct and rigorous. A more acceptable way might be to replace b by c ini (4.46) and afterwards to use the substitution, $c = -b$.

Example 4.10. Find the Taylor's series of $\cos bx$ about $x = 0$.

Again using a Table:

n	$f^{(n)}$	$f^{(n)}(0)$
0	$\cos bx$	1
1	$-b\sin bx$	0
2	$-b^2\cos bx$	$-b^2$
3	$b^3\sin bx$	0
4	$b^4\cos bx$	b^4
5	$-b^5\sin bx$	0

I took a few more terms to be assured of the pattern of coefficients. The Taylor's series is

$$\cos bx = 1 - \frac{b^2x^2}{2!} + \frac{b^4x^4}{4!}\cdots = \sum_{n=0}^{\infty}\frac{(-1)^n b^{2n}x^{2n}}{(2n)!}. \tag{4.48}$$

Note carefully the uses of n and $2n$ in the final summation and how this enables us to get the alternating signs. A similar analysis yields,

$$\sin bx = bx - \frac{b^3x^3}{3!} + \frac{b^5x^5}{5!}\cdots = \sum_{n=0}^{\infty}\frac{(-1)^n b^{2n+1}x^{2n+1}}{(2n+1)!}. \tag{4.49}$$

Sorry, but I want to play briefly with these last two results. If all the instances of b in Eq. (4.46) were to be replaced by bj then we have,

$$e^{jbx} = 1 + jbx + \frac{j^2b^2x^2}{2!} + \frac{j^3b^3x^3}{3!} + \frac{j^4b^4x^4}{4!} + \frac{j^5b^5x^5}{5!} + \cdots = \sum_{n=0}^{\infty}\frac{j^nb^nx^n}{n!}. \tag{4.50}$$

This may be rearranged to give,

$$\begin{aligned} e^{jbx} &= \left[1 - \frac{b^2}{2!}x^2 + \frac{b^4}{4!}x^4 - \frac{b^6}{6!}x^6 + \cdots\right] + j\left[bx - \frac{b^3}{3!}x^3 + \frac{b^5}{5!}x^5 - \frac{b^7}{7!}x^7 + \cdots\right] \\ &= \sum_{n=0}^{\infty}\frac{(-1)^n b^{2n}x^{2n}}{(2n)!} + j\sum_{n=0}^{\infty}\frac{(-1)^n b^{2n+1}x^{2n+1}}{(2n+1)!} \\ &= \cos bx + j\sin bx \qquad \text{using Eqs. (4.48) and (4.49).} \end{aligned} \tag{4.51}$$

So this is the proof of $e^{jbx} = \cos bx + j\sin bx$ which was promised earlier in the Complex Numbers chapter.

Example 4.11. Find the Taylor's series of $(1+x)^{-1}$ about $x=0$ and about $x=1$.

Here's the Table of data for the two cases:

n	$f^{(n)}$	$f^{(n)}(0)$	$f^{(n)}(1)$
0	$(1+x)^{-1}$	1	$1/2$
1	$-(1+x)^{-2}$	-1	$-1/2^2$
2	$2(1+x)^{-3}$	2	$2/2^3$
3	$-3!\,(1+x)^{-4}$	$-3!$	$-3!/2^4$
4	$4!\,(1+x)^{-5}$	$4!$	$4!/2^5$
5	$-5!\,(1+x)^{-6}$	$-5!$	$-5!/2^6$

From these data and Eq. (4.44) we obtain the desired Taylor's series:

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - x^5 + \cdots = \sum_{n=0}^{\infty} (-1)^n x^n, \tag{4.52}$$

and

$$\frac{1}{1+x} = \frac{1}{2}\Big[1 - \frac{x-1}{2} + \frac{(x-1)^2}{2^2} - \frac{(x-1)^3}{2^3} + \frac{(x-1)^4}{2^4} - \frac{(x-1)^5}{2^5} + \cdots\Big] = \frac{1}{2}\sum_{n=0}^{\infty} \frac{(-1)^n (x-1)^n}{2^n}. \tag{4.53}$$

Example 4.12 Expand the integral $y(x) = \int_0^x e^{-t^2}\,dt$ about $x=0$ and hence evaluate $y(0.2)$.

This integral is very closely related to the Normal Distribution but the indefinite integral of e^{-x^2} cannot be written in terms of familiar functions, which is a bit awkward. Bizarrely, it is possible to show that $y(x) \to \sqrt{\pi}/2$ as $x \to \infty$, as we found in the final integration problem sheet.

We begin by differentiating the integral; we get

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

an exponential function. Now, I would not advise trying to apply the Taylor's series formula directly. Rather it is better to note that we already have a Taylor's series representation of e^{-x} in Eq. (4.46) and therefore we may simply replace all occurrences of x by x^2 to find that,

$$y'(x) = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} \cdots . \tag{4.54}$$

Integration of this (noting that $y(0)=0$ so that we can evaluate the constant of integration) yields,

$$y(x) = x - \frac{x^3}{3} + \frac{x^5}{5.2!} - \frac{x^7}{7.3!} + \frac{x^9}{9.4!} \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)\times n!}. \tag{4.55}$$

Three terms of this series gives $y(0.2) = 0.197365$. The fourth term has magnitude 3×10^{-7} and therefore just these three terms are enough for six decimal places of accuracy since four terms yields no change.

4.7 The convergence of power series

Following on from Example 4.12 a natural question is whether we can use any value of x in Eq. (4.55) or any other series. It is this which forms the topic of the present subsection, but we will study a couple of series from earlier in this chapter in order to illustrate that different power series may have different convergence properties. For convenience Eqs. (4.52), (4.49) (with $b = 1$) and (4.46) (also with $b = 1$) are quoted here:

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - x^5 + \cdots \,, \tag{4.56}$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \tag{4.57}$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \tag{4.58}$$

4.7.1 Numerical demonstration of convergence/divergence

The series in Eq. (4.56) is a geometric series and therefore we could have found its sum, the left hand side, using those ideas. When we choose $x = 0.1$ then it feels certain that the right hand side will converge because the successive terms are decreasing rapidly. Even more obvious is the choice, $x = -0.1$, for then the partial sums become, 1, 1.1, 1.11, 1.111 and on to the value $10/9$. On the other hand, if we were to choose $x = 2$ then the series looks like,

$$1 - 2 + 2^2 - 2^3 + 2^4 - 2^5 + \cdots$$

and there is no hope of this converging to a finite value. Although the signs alternate, the magnitudes of the **partial sums** increase: 1, -1, 3, -5, 11,-21, 43.... After a bit of work I determined that this sequence satisfies $[(-1)^2 2^{n+1} + 1]/3$ — big deal! Our conclusion is that convergence for the series in Eq. (4.56) is conditional: if x is small enough then it converges, but it will diverge otherwise. The boundary between those behaviours is likely to be $x = 1$ for then the series is

$$1 - 1 + 1 - 1 + 1 \cdots \,,$$

and the partials sums alternate: 1, 0, 1, 0, and so on.

A similar playing around with numbers for Eqs. (4.57) and (4.58) using a calculator takes time. The fact that the denominators are factorials gives a hint that these series might well be convergent. If one were to choose $x = 0.1$ in Eq. (4.58), then only a few terms are required to obtain, say, 6DP accuracy. This may be seen in the following Table (where u_n represents the successive terms, and S_n the partial sums):

Showing successive terms and partial sums for Eq. (4.58) when $x = 0.1$

n	u_n	S_n
0	1.000000	1.000000
1	0.100000	1.100000
2	0.005000	1.105000
3	0.000167	1.105167
4	0.000004	1.105171
5	4×10^{-7}	1.105171

That final answer is correct to 6DPs.

On the other hand, if we choose $x = 10$ then we will have the following.

Showing successive terms and partial sums for Eq. (4.58) when $x = 10$

n	u_n	S_n
0	1.000000	1.000000
1	10.000000	11.000000
2	50.000000	61.000000
3	166.666667	227.666667
4	416.666667	644.333333
5	833.333333	1477.666667
6	1388.888889	2866.555556
7	1984.126984	4850.682540
8	2480.158730	7330.841270
9	2755.731922	10086.573192
10	2755.731922	12842.305115
11	2505.210839	15347.515953
12	2087.675699	17435.191652
13	1605.904384	19041.096036
14	1147.074560	20188.170595
15	764.716373	20952.886969
16	477.947733	21430.834702
17	281.145725	21711.980427
18	156.192070	21868.172497
19	82.206352	21950.378849
20	41.103176	21991.482026
21	19.572941	22011.054967
22	8.896791	22019.951758
23	3.868170	22023.819928
24	1.611738	22025.431666
25	0.644695	22026.076361
26	0.247960	22026.324321
27	0.091837	22026.416157
28	0.032799	22026.448956
29	0.011310	22026.460266
30	0.003770	22026.464036
31	0.001216	22026.465252
32	0.000380	22026.465632
33	0.000115	22026.465748
34	0.000034	22026.465781
35	0.000010	22026.465791
36	0.000003	22026.465794
37	0.000001	22026.465795
38	0.000000	22026.465795
39	0.000000	22026.465795
40	0.000000	22026.465795

My apologies for the length of this Table but there are some important points which arise. The first is that, even for a value of x that is as large as 10, the series converges. The second point is that the magnitude of successive terms in the series decrease once $n > 10$. Given the form of the series, successive terms will decrease when $x = 100$ is taken only when $n > 100$, and therefore this suggests that the series will converge for any value of x. The third point is illustrated by compiling a similar Table for $x = -10$:

Showing successive terms and partial sums when $x = -10$

n	u_n	S_n
0	1.000000	1.000000
1	−10.000000	−9.000000
2	50.000000	41.000000
3	−166.666667	−125.666667
4	416.666667	291.000000
5	−833.333333	−542.333333
6	1388.888889	846.555556
7	−1984.126984	−1137.571429
8	2480.158730	1342.587302
9	−2755.731922	−1413.144621
10	2755.731922	1342.587302
11	−2505.210839	−1162.623537
12	2087.675699	925.052162
13	−1605.904384	−680.852222
14	1147.074560	466.222338
15	−764.716373	−298.494035
16	477.947733	179.453698
17	−281.145725	−101.692027
18	156.192070	54.500042
19	−82.206352	−27.706310
20	41.103176	13.396866
21	−19.572941	−6.176075
22	8.896791	2.720716
23	−3.868170	−1.147454
24	1.611738	0.464284
25	−0.644695	−0.180411
26	0.247960	0.067548
27	−0.091837	−0.024289
28	0.032799	0.008510
29	−0.011310	−0.002800
30	0.003770	0.000970
31	−0.001216	−0.000246
32	0.000380	0.000134
33	−0.000115	0.000019
34	0.000034	0.000053
35	−0.000010	0.000043
36	0.000003	0.000046
37	−0.000001	0.000045
38	0.000000	0.000045
39	−0.000000	0.000045
40	0.000000	0.000045

The terms in the u_n column here have the same magnitude as their counterparts when $x = 10$ but they alternate in sign. Despite the cancellations due to having both positive and negative values with a large magnitude, the final converged value of this series is $e^{-10} \simeq 0.00004540$, which is the same value that I get on my calculator.

For a more extreme case, $x = -40$, the 'converged' value of the series turns out to be about -3.166 which is clearly wrong — exponentials can't be negative. So what has happened? This has arisen because the 13 significant figures of accuracy in my computer calculations are no longer sufficient to cope with the very great amount of cancellation that arises when adding the terms in the series. For $x = -40$ my calculator gives 4.248×10^{-18} which I have checked to be correct via other means, and therefore I conclude that my calculator doesn't use Taylor's series! Even here, one must always be aware of significant figures and whether accuracy has been compromised because of the necessary use of finite precision arithmetic.

4.7.2 Graphical representation of convergence/divergence

Figure 4.3 shows the variation of $\sin x$ and of the first 11 partial sums of the Taylor's series (the third partial sum, for example, being the first three terms, $x - x^3/3! + x^5/5!$):

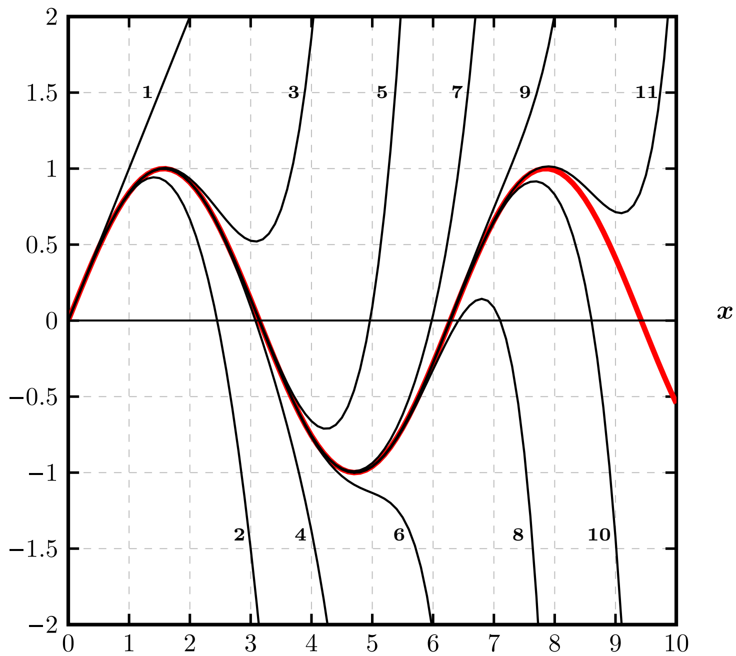


Figure 4.3. Demonstrating the convergence properties of the Taylor's series for $\sin x$. Showing the function itself (red) and the first 11 partial sums.

The first partial sum, x, forms the tangent to $\sin x$ at $x = 0$, and this figure demonstrates that just this one term is quite a good approximation up to, perhaps $x = 0.5$.

The second partial sum, $x - x^3/3!$, matches the sine wave over a larger range of values of x, and the graph suggests that that series does well when $x < 1$.

As further terms in the Taylor's series are included, the sequence of curves gradually continues to envelop the sine wave, and it appears that the ninth partial sum provides quite an accurate estimate for its first period. There is no suggestion or hint here that the Taylor's series will diverge for some larger value of x, although we have already seen that there will be a limitation on the usefulness of the series for sufficiently large values of x due to the use of fixed precision arithmetic.

So our tentative conclusion is that convergence happens irrespective of the value of x.

Figure 4.4 shows the variation of $(1+x)^{-1}$ and of the first 10 partial sums of the Taylor's series

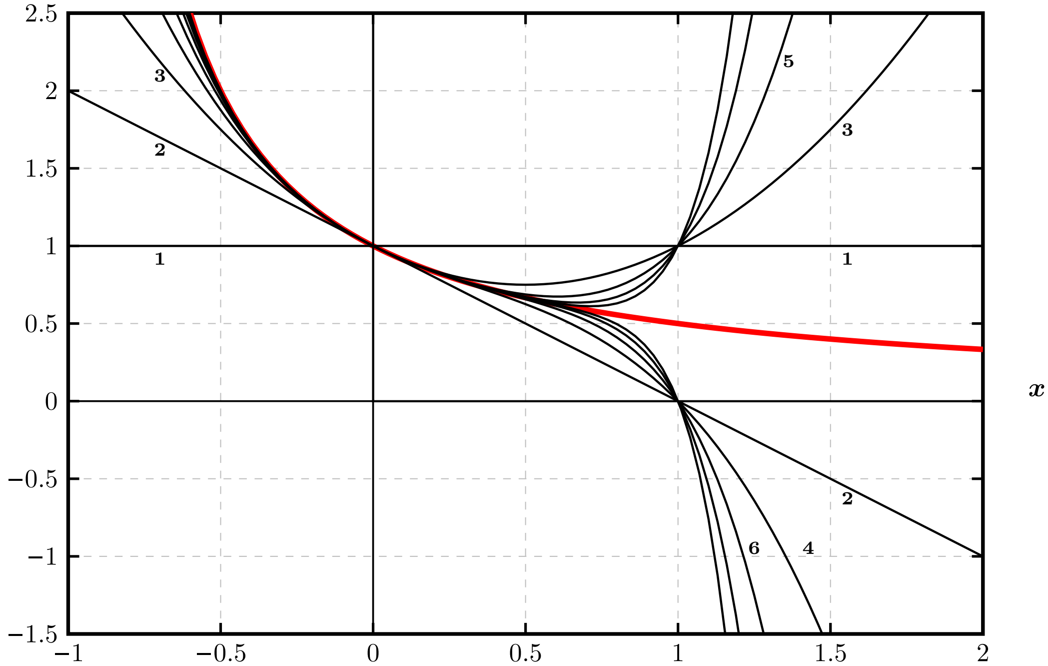


Figure 4.4. Demonstrating the convergence properties of the Taylor's series for $(1+x)^{-1}$. Showing the function itself (red) and the first 10 partial sums.

Once more, when x lies within a certain range, the partial sums approach the original curve. For positive values of x this certainly happens when $x < 1$ (strictly less than 1). Once $x > 1$ be it ever so slightly, then successive partial sums diverge away from the original function.

When $x < 0$ the convergence towards $(1+x)^{-1}$ is monotonic. It is more difficult to see what happens when x is large and negative, but $(1+x)^{-1}$ tends to infinity as $x \to -1$ and this presents a difficulty. When $x < -1$ the successive terms in the Taylor's series get larger in magnitude and therefore convergence fails to happen.

Our tentative conclusion from this numerical experiment is that we have convergence when $-1 < x < 1$.

4.8 d'Alembert's ratio test

If the series under consideration has the form

$$S = \sum_{n=1}^{\infty} u_n = u_1 + u_2 + u_3 + \cdots\,, \tag{4.59}$$

then **d'Alembert's ratio test** states the following:

$$\begin{array}{|lcl|}\hline \lim_{n\to\infty}\left|\dfrac{u_{n+1}}{u_n}\right| < 1 & \Rightarrow & \text{absolute convergence} \\ \lim_{n\to\infty}\left|\dfrac{u_{n+1}}{u_n}\right| > 1 & \Rightarrow & \text{divergence} \\ \lim_{n\to\infty}\left|\dfrac{u_{n+1}}{u_n}\right| = 1 & \Rightarrow & \text{inconclusive.} \\ \hline \end{array} \tag{4.60}$$

Note: that this test relies on the long-term behaviour of the coefficients, and therefore the starting point (viz. $n = 1$ in Eq. (4.59)) is of no consequence.

d'Alembert's test is an excellent method for finding the range of values of x for power series, but when the series consists solely of numerical data inconclusive conclusions apply to a depressingly wide range of useful series. You'll see what I mean....

4.8.1 Convergence of numerical series

Example 4.13. The Maclaurin's series for e^1 is $e^1 = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots$. Determine if it converges to a value.

The first step always is to determine an expression for the general term. In this case the general term is the reciprocal of successive factorials. Hence $u_n = 1/n!$ and $u_{n+1} = 1/(n+1)!$. The second step is to evaluate $|u_{n+1}/u_n|$:

$$\left|\frac{u_{n+1}}{u_n}\right| = \left|\frac{n!}{(n+1)!}\right| = \left|\frac{n!}{n! \times (n+1)}\right| = \frac{1}{n+1}. \tag{4.61}$$

Hence,

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\frac{1}{n+1} = 0. \tag{4.62}$$

Given that $0 < 1$, this series converges.

Example 4.14. Use d'Alembert's test to check for the convergence of the geometric series, $r + r^2 + r^3 + \cdots$.

Here u_n may be taken to be r^n. Hence $u_{n+1} = r^{n+1}$ and so $u_{n+1}/u_n = r$. Therefore we have,

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = |r|. \tag{4.63}$$

For convergence we require $|r| < 1$.

Example 4.15. Consider the sum of the reciprocals of the positive integers,

$$\sum_{n=1}^{\infty}\frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots.$$

Here $u_n = 1/n$ and therefore

$$\left|\frac{u_{n+1}}{u_n}\right| = \left|\frac{n}{n+1}\right| = \frac{1}{1+\frac{1}{n}} \longrightarrow 1 \text{ as } n \longrightarrow \infty. \tag{4.64}$$

Therefore the test is inconclusive because the limit is precisely 1.

However, it is possible to show that the series is divergent by considering the series in the following way.

$$\begin{aligned} & 1 + \tfrac{1}{2} + \left(\tfrac{1}{3} + \tfrac{1}{4}\right) + \left(\tfrac{1}{5} + \tfrac{1}{6} + \tfrac{1}{7} + \tfrac{1}{8}\right) + \left(\tfrac{1}{9} + \cdots \tfrac{1}{16}\right) + \cdots \\ > \; & 1 + \tfrac{1}{2} + \left(\tfrac{1}{4} + \tfrac{1}{4}\right) + \left(\tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8}\right) + \left(\tfrac{1}{16} + \cdots \tfrac{1}{16}\right) + \cdots \\ = \; & 1 + \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \cdots \end{aligned} \tag{4.65}$$

Clearly this process continues indefinitely and if we continue to N such sets of brackets then we will accumulate N instances of $\frac{1}{2}$, and therefore

$$S_{2^N} = \sum_{n=1}^{2^N} \frac{1}{n} > 1 + \tfrac{1}{2}N,$$

and hence the series diverges.

This is a trick that can only very rarely be deployed, but it is interesting.

Having settled the convergence properties of this example, we may return to Example 4.7 which yields the following when $x = 1$ is substituted,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \cdots .$$

This differs from the present example by having alternating signs, yet Example 4.7 shows that this series adds to $\ln 2$ and hence it is convergent.

These two series demonstrate the idea of **conditional convergence**. The respective magnitudes of the terms in the two series are identical but the pattern of signs differ, so the convergence properties then depend on that pattern of signs.

Example 4.16. Consider the sum of the reciprocals of the squares of the integers,

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots .$$

Here $u_n = 1/n^2$ and hence

$$\left|\frac{u_{n+1}}{u_n}\right| = \left|\frac{n^2}{(n+1)^2}\right| = \frac{1}{(1+\frac{1}{n})^2} \longrightarrow 1 \text{ as } n \longrightarrow \infty. \tag{4.66}$$

Again we have an inconclusive conclusion from d'Alembert! This is a little depressing for these coefficients decay more rapidly than those in Example 4.15. This series does in fact converge, but it requires us to use Fourier Series methods (to be done in ME10305 Mathematics 2) to show that the sum is $\pi^2/6$. The equivalent sum of the reciprocals of the fourth powers of the positive integers is $\pi^4/90$, although d'Alembert's test remains inconclusive even for this series.

4.9 Application of d'Alembert's test to power series

Now we take the power series,

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 \cdots , \tag{4.67}$$

and apply d'Alembert's ratio test to it using $u_n = a_n x^n$ as the general term. Therefore we obtain convergence when

$$\lim_{n\to\infty} \left|\frac{a_{n+1}x^{n+1}}{a_n x^n}\right| < 1. \tag{4.68}$$

Although this may be simplified and rearranged to get the convergence criterion,

$$|x| < \lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right|. \tag{4.69}$$

I really don't like to teach this form because the subscripts here (n on top, $n+1$ below) are the opposite way around from those in the original ratio test given in Eq. (4.60) and this could be confusing. My preference is to employ the formula in Eq. (4.60) as it is and then to follow whatever the maths tells us. I will do this in all of the examples below.

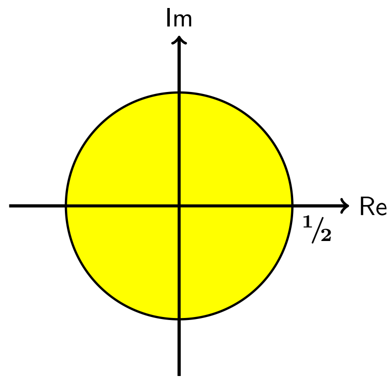
Example 4.17. Find the values of x for which the following power series converges.

$$1 + 2x + 4x^2 + 8x^3 + 16x^4 + \cdots .$$

The general term may be written as $u_n = 2^n x^n$ and therefore we have

$$\left|\frac{u_{n+1}}{u_n}\right| = \left|\frac{2^{n+1}x^{n+1}}{2^n x^n}\right| = 2|x|. \tag{4.70}$$

Therefore we have convergence when $2|x| < 1$, i.e. when $|x| < \frac{1}{2}$. Perhaps this may be a surprise, but this expression is also valid for complex values of x and therefore we may sketch the region of convergence in the complex plane as follows:



The yellow region is the region of convergence while the radius of the black perimeter is called the **radius of convergence** of the series; in this case we say that the radius of convergence is $\frac{1}{2}$. Thus values of x which are larger than $\frac{1}{2}$ in magnitude will cause the series to diverge.

Example 4.18. Find the values of x for which the Taylor series given in Eq. (4.53) converges.

The series is

$$\frac{1}{2}\sum_{n=0}^{\infty}\frac{(-1)^n(x-1)^n}{2^n}$$

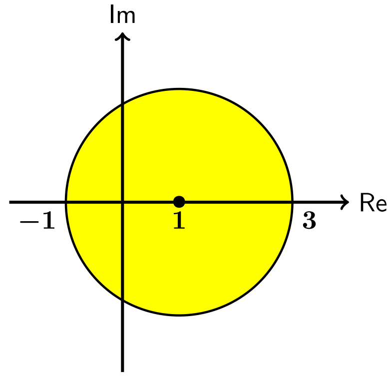
and therefore the general term is

$$u_n = \frac{(-1)^n(x-1)^n}{2^{n+1}}.$$

Hence,

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\left|\frac{(-1)^{n+1}(x-1)^{n+1}}{2^{n+2}} \times \frac{2^{n+1}}{(-1)^n(x-1)^n}\right| = \lim_{n\to\infty}\left|\frac{(x-1)}{2}\right| = \tfrac{1}{2}|x-1|. \tag{4.71}$$

This needs to be less than 1 for convergence, which means that $|x-1| < 2$. So the radius of convergence is 2, as illustrated in the following Figure.



Should we wish to confine ourselves to real values of x, this means that convergence will then take place only when $-1 < x < 3$, noting the strict inequalities in this expression.

Example 4.19. Find the radius of convergence of the Taylor's series for e^x given in Eq. (4.58).

The series is $e^x = 1 + x + x^2/2! + x^3/3! + x^4/4! + \cdots$. Hence the general term is $u_n = x^n/n!$, and therefore d'Alembert's test gives,

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\left|\frac{x^{n+1}\,n!}{(n+1)!\,x^n}\right| = \lim_{n\to\infty}\left|\frac{x}{(n+1)}\right| = 0. \qquad (4.72)$$

While we saw this sort of result in Example 4.13, in the present context the limit is zero for all possible values of x, and therefore this function has **an infinite radius of convergence**. This confirms our deduction from numerical evidence in Section 4.7.1.

Example 4.20. Find the radius of convergence of the Taylor's series for $\sin x$ which is given in Eq. (4.57).

The series is $\sin x = x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \dfrac{x^7}{7!} + \cdots$ and therefore the general term is $u_n = \dfrac{(-1)^n x^{2n+1}}{(2n+1)!}$.

The correct expression for u_{n+1} is $\dfrac{(-1)^{n+1}x^{2n+3}}{(2n+3)!}$. **Note** that the replacement of n by $n+1$ in $2n+1$ yields $2n+3$ rather than $2n+2$, a favourite error of so many undergraduates in an exam!

Hence

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\left|\frac{(-1)^{n+1}x^{2n+3}}{(2n+3)!}\times\frac{(2n+1)!}{(-1)^n x^{2n+1}}\right| = \lim_{n\to\infty}\left|\frac{x^2}{(2n+3)(2n+2)}\right| = 0. \qquad (4.73)$$

So again we have an infinite radius of convergence, just as we suspected from the graphical evidence in Fig. 4.3.

Finally, I'll leave it to you to show that the series, $1 + (1!)x + (2!)x^2 + (3!)x^3 + \cdots$, has a zero radius of convergence. By this is meant that the series always diverges except for when $x = 0$.

4.10 l'Hôpital's rule

4.10.1 A graphical motivation

This rule allows us to evaluate the ratio of two functions, $f(x)$ and $g(x)$, as $x \to a$ when both $f(a) = 0$ and $g(a) = 0$. The situation is well-illustrated by Figure 4.5.

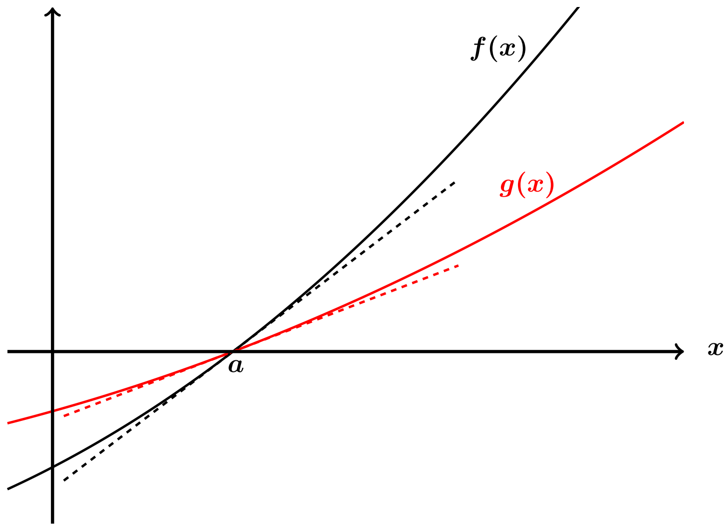


Figure 4.5. Displaying $f(x)$ and $g(x)$ both of which are zero at $x = a$, together with their tangents (dotted lines) at $x = a$.

What is zero-divided-by-zero? More precisely in this context, we wish to find the value of $f(x)/g(x)$ when $x = a$ and this diagram tells us the basic answer: it depends on how quickly the two functions move towards zero as x moves towards a. Figure 4.5 shows two generic functions, $f(x)$ and $g(x)$, and both of them are zero at $x = a$. Also shown as dashed lines are the tangents to each curve at $x = a$. One could say that as $x = a$ is approached then each curve more closely resembles its tangent and when $x = a$ they are identical. So this produces what I call the graphical motivation for finding the value of the ratio of $f(x)$ and $g(x)$ at $x = a$.

The tangents are given, respectively, by

$$f(x) \simeq (x - a)f'(a) \qquad g(x) \simeq (x - a)g'(a). \tag{4.74}$$

These may be written down immediately when it is noted (i) tangents are linear functions, (ii) these specific tangents must pass through $x = a$ and hence $(x - a)$ must be a factor, and (iii) their slopes are the slopes of the functions at $x = a$.

Therefore the ratio of $f(x)$ and $g(x)$ may be written as,

$$\frac{f(x)}{g(x)} = \frac{(x - a)f'(a)}{(x - a)g'(a)} = \frac{f'(a)}{g'(a)}. \tag{4.75}$$

Given that $f(x)$ isn't quite equal to $(x - a)f'(a)$ unless $x = a$, it is better to write this down as

$$\boxed{\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}.} \tag{4.76}$$

OK, if you are someone like me then your reaction is going to be that this seems plausible, but you would rather have a more rigorous proof than this. Well, we can bring Taylor's series into the fray and let Taylor provide a better proof.

4.10.2 Proof of l'Hôpital's rule using Taylor's series

Equations (4.74) use what is, in reality, the first nonzero term of the Taylor's series of the two functions. We need to use a few more just to check that all is well. So let us write out this ratio again and work on it.

$$\begin{aligned}
\frac{f(x)}{g(x)} &= \frac{\cancel{f(a)} + (x-a)f'(a) + \frac{1}{2}(x-a)^2 f''(a) + \frac{1}{3!}(x-a)^3 f'''(a) + \cdots}{\cancel{g(a)} + (x-a)g'(a) + \frac{1}{2}(x-a)^2 g''(a) + \frac{1}{3!}(x-a)^3 g'''(a) + \cdots} && \text{by definition} \\
&= \frac{f'(a) + \frac{1}{2}(x-a)f''(a) + \frac{1}{3!}(x-a)^2 f'''(a) + \cdots}{g'(a) + \frac{1}{2}(x-a)g''(a) + \frac{1}{3!}(x-a)^2 g'''(a) + \cdots} && \text{cancelling } (x-a) \\
&\longrightarrow \frac{f'(a)}{g'(a)} \text{ as } x \to a.
\end{aligned} \tag{4.77}$$

The conclusion here is that the terms we neglected in Eq. (4.74) do not contribute to the final answer, but it was necessary to check this for our peace of mind. So Eq. (4.76) has been retrieved using a better analysis.

Example 4.21 Find $\lim_{x\to 0} \frac{\sin 2x}{x}$.

Both the numerator and the denominator are zero at $x = 0$ and therefore we apply l'Hôpital's rule.

$$\lim_{x\to 0} \frac{\sin 2x}{x} = \frac{2\cos 2x}{1}\Big|_{x=0} = 2.$$

Example 4.22 Find $\lim_{x\to \pi} \frac{\sin x}{x-\pi}$.

The numerator and the denominator are zero at $x = \pi$ and therefore,

$$\lim_{x\to \pi} \frac{\sin x}{x-\pi} = \frac{\cos x}{1}\Big|_{x=\pi} = -1.$$

Example 4.23 Find $\lim_{x\to 0} \frac{\sin^2 x}{x}$.

We have

$$\lim_{x\to 0} \frac{\sin^2 x}{x} = \frac{2\sin x \cos x}{1}\Big|_{x=0} = 0.$$

This example shows that zero is a legitimate final answer.

Example 4.24 Find $\lim_{x\to 0} \frac{\sin x}{x^2}$.

We have

$$\lim_{x\to 0} \frac{\sin x}{x^2} = \frac{\cos x}{2x}\Big|_{x=0} = \infty.$$

So this example shows that it is legitimate for an infinite solution to be obtained. One can say that x^2 tends to zero much faster than $\sin x$ does.

4.10.3 Nested applications of l'Hôpital's rule

This follows from the question: what happens to the derivation of l'Hôpital's rule if, in addition to $f(a) = g(a) = 0$, we also have $f'(a) = g'(a) = 0$? These conditions tell us that both $f(x)$ and $g(x)$ have a double zero at $x = a$ and therefore we are comparing two **parabolae**.

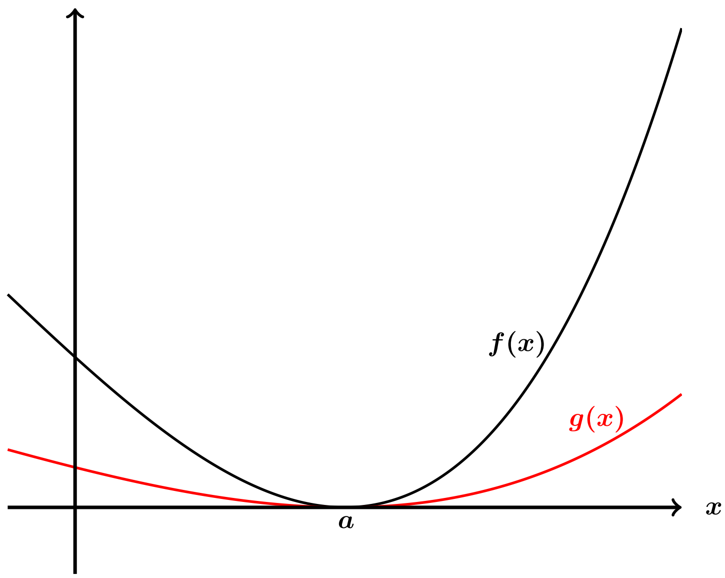


Figure 4.6. Displaying $f(x)$ and $g(x)$ both of which are zero and have a zero first derivative at the point $x = a$.

In this case Eq. (4.77) becomes,

$$\begin{aligned}\frac{f(x)}{g(x)} &= \frac{\cancel{f(a)} + \cancel{(x-a)f'(a)} + \frac{1}{2}(x-a)^2 f''(a) + \frac{1}{3!}(x-a)^3 f'''(a) + \cdots}{\cancel{g(a)} + \cancel{(x-a)g'(a)} + \frac{1}{2}(x-a)^2 g''(a) + \frac{1}{3!}(x-a)^3 g'''(a) + \cdots} && \text{by definition} \\ &= \frac{\frac{1}{2}f''(a) + \frac{1}{3!}(x-a)f'''(a) + \cdots}{\frac{1}{2}g''(a) + \frac{1}{3!}(x-a)g'''(a) + \cdots} && \text{cancelling } (x-a)^2 \\ &\longrightarrow \frac{f''(a)}{g''(a)} \text{ as } x \to a. \end{aligned} \tag{4.78}$$

So when the functions and their derivatives are all zero at $x = a$, l'Hôpital's rule becomes,

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f''(a)}{g''(a)}. \tag{4.79}$$

Although we won't prove this, if these second derivatives are also zero then we need to compute the ratio of the third derivatives, and so on. So it is my view that the best way to write down l'Hôpital's rule is the following way:

$$\boxed{\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.} \tag{4.80}$$

There is a subtle difference between this and Eq. (4.76) in that the right hand side also uses the limiting process. Thus if we have applied the rule to a zero-divide-zero ratio and have obtained another one, then we merely apply the same formula to the new case. In essence this is a recursive formula which can continue to be applied to successive ratios until the ratio has a definitive limit (i.e. zero, nonzero or infinite, as opposed to being indeterminate).

Example 4.25 Find $\lim_{x\to 0} \dfrac{1-\cos ax}{x^2}$.

$$\begin{aligned}
\lim_{x\to 0} \frac{1-\cos ax}{x^2} &\overset{l'H}{=\!=} \lim_{x\to 0} \frac{a \sin ax}{2x} && \text{This too is a } 0/0 \text{ case, hence l'H again} \\
&\overset{l'H}{=\!=} \lim_{x\to 0} \frac{a^2 \cos ax}{2} && \text{This isn't a } 0/0 \text{ case} \\
&= \tfrac{1}{2}a^2.
\end{aligned}$$

Note that I have introduced my own notation here: $\overset{l'H}{=\!=}$. You won't find it in textbooks, but I am using it simply to indicate that I have used l'Hôpital's rule as opposed to merely simplifying what is inside the limit.

Example 4.26 Find $\lim_{x\to 0} \dfrac{\sin x - x}{x^3}$.

$$\begin{aligned}
\lim_{x\to 0} \frac{\sin x - x}{x^3} &\overset{l'H}{=\!=} \lim_{x\to 0} \frac{\cos x - 1}{3x^2} && \text{This too is a } 0/0 \text{ case, hence l'H again} \\
&\overset{l'H}{=\!=} \lim_{x\to 0} \frac{-\sin x}{6x} && \text{This too is a } 0/0 \text{ case, hence l'H again} \\
&\overset{l'H}{=\!=} \lim_{x\to 0} \frac{-\cos x}{6} = \frac{1}{6} && \text{This isn't a } 0/0 \text{ case}
\end{aligned}$$

4.10.4 A few more exotic examples

Example 4.27 Find $\lim_{x\to\infty} \dfrac{e^{ax}}{x}$.

This example demonstrates two different variations on the standard l'Hôpital's rule theme. First we are considering $x \to \infty$ and second we are considering two functions that become infinite in that limit. In this case l'Hôpital's rule compares the speed at which the two functions grow in magnitude. We have

$$\lim_{x\to\infty} \frac{e^{ax}}{x} \overset{l'H}{=\!=} \left.\frac{ae^{ax}}{1}\right|_{x\to\infty} = \infty.$$

Example 4.28 [A more difficult example.] Find $\lim_{x\to 0} x^x$.

This doesn't look like a l'Hôpital's rule question because there is no ratio! However, if we let $y = x^x$, then

$$\ln y = x \ln x = \frac{\ln x}{1/x}.$$

This is an $x \to 0$ problem where the ratio is an infinity-over-infinity case. Now we may apply l'Hôpital's rule:

$$\lim_{x\to 0} \ln y \overset{l'H}{=\!=} \lim_{x\to 0} \frac{\ln x}{1/x} \qquad \text{This is an } \infty/\infty \text{ case, hence use l'H}$$

$$\overset{l'H}{=\!=} \lim_{x\to 0} \frac{1/x}{-1/x^2} \qquad \text{Also } \infty/\infty \text{ but we can simplify the quotient}$$

$$= \lim_{x\to 0}(-x) \qquad \text{I haven't used l'Hôpital's rule here, just normal algebra}$$

$$= 0.$$

Since $\ln y \to 0$ then $y \to 1$. Therefore the $x \to 0$ limit of x^x is 1.

Example 4.29 [Also a more difficult example.] Find $\lim_{n\to\infty} \Big(1 + \dfrac{1}{n}\Big)^n$.

This problem is associated with compound interest. When $n = 1$ the factor $(1 + \frac{1}{n})^n$ may be said to be equivalent to the adding of 100% interest to an investment, I, after a year. The yield is that the investment doubles to $2I$.

When $n = 2$, then 50% interest is applied twice per year, and the return is now $(1 + \frac{1}{2})^2 I = 2.25I$.

When $n = 4$, then 25% interest is added quarterly, and the return is $(1 + \frac{1}{4})^4 I = 2.441406I$.

If monthly, then the return will be $(1 + \frac{1}{12})^{12} I = 2.613035I$, and if it is daily then the return will be $(1 + \frac{1}{365})^{365} I = 2.714567I$. The ultimate question is: what is your return if interest is added continuously at the appropriate infinitesimal rate?

Again this is not a ratio, but again we may take logs. So if $y = (1 + \frac{1}{n})^n$ then

$$\ln y = n \ln\Big(1 + \frac{1}{n}\Big) = \frac{\ln\Big(1 + \dfrac{1}{n}\Big)}{\dfrac{1}{n}}.$$

This looks very messy, but with a little sleight of hand we may transform a large–n problem into a small–x problem by setting $x = 1/n$. Therefore the problem becomes,

$$\lim_{x\to 0} \frac{\ln(1+x)}{x} \overset{l'H}{=\!=} \lim_{x\to 0} \frac{1/(1+x)}{1} = 1.$$

This, then, is the value of $\ln y$ and therefore the $n \to \infty$ limit of y is e. The return on the investment will be eI or $2.718282I$ (6DPs). You won't get any more than that!

5 ORDINARY DIFFERENTIAL EQUATIONS

Ordinary differential equations are equations which contain differentials, i.e. derivatives! I cannot imagine that anyone would find that fact surprising. Therefore this section of the unit is devoted to solving equations which contain differentials. These equations are called ordinary differential equations (ODEs), rather than partial differential equations (PDEs) which contain partial derivatives (which will arise in semester 2 of this unit).

Ordinary differential equations (ODEs) use dependent and independent variables. If y is a function of t, then y is called the **dependent** variable (since its value *depends* on the value of t) and t is the **independent** variable. The function is written in the form $y(t)$, or often just as y if the context implies that it is a function of t. Likewise, should y be a function of x, distance, then we could write it as $y(x)$, or often just as y if the context is clear.

5.1 Motivation

But before we begin all of this we need to answer the question:

What is the point of solving ODEs?

The briefest answer that I can think of is that virtually everything in science (i.e. anything within the known universe with the possible exception of the interior of a black hole) satisfies an equation or a set of equations of some sort. The topic of differential equations is an extremely vast and diverse part of science and engineering. The design of buildings, bridges, aircraft and power supplies, and the modelling and description of natural phenomena such as tides, cloud patterns, the buckling of rock layers, planetary orbits, our wretched UK weather, the aerodynamics of a cricket ball, and blood flow in the human body may all be modelled by differential equations.

One of the most straightforward examples is the equation describing the motion of a mass on a spring. If it is assumed that the force which is required to maintain an extension of length x of the spring from the equilibrium position is kx, then a mass at the end of such a spring which has that extension suffers a force of $-kx$. As Newton's law states that the rate of change of momentum is equal to the applied force, we immediately obtain the differential equation,

$$\underbrace{\frac{d}{dt}\left(m\frac{dx}{dt}\right)}_{\text{rate of change of momentum}} = \underbrace{-\,kx}_{\text{restoring force}}. \tag{5.1}$$

As the mass is usually constant for a mass/spring system, we obtain

$$m\frac{d^2x}{dt^2} = -kx. \tag{5.2}$$

In this (ODEs) part of the unit we will find out how to solve equations like this in order to determine exactly how this mass moves.

Note: that we will often abbreviate the second derivative in (5.2) as x'' for convenience.

5.2 Classification

Just as different types of algebraic equation (e.g. linear, quadratic, cubic, transcendental) have their different methods of solution, so do ordinary differential equations. Therefore it is necessary at the outset to classify the equation at hand in order to select which method is suitable for its solution. So our first chunk of technical stuff won't cover solutions at all.

We will classify ODEs in three different ways:

- IVP or BVP (i.e. Initial value problem or Boundary Value Problem),
- Linear or Nonlinear,
- The order of the equation.

Whilst we might have at best a diaphanous or wraith-like sense of what these mean, there remains the question,

What is the point of classifying ODEs?

The answer is that the methods which are needed to solve ODEs will depend very much on what the classification is. This is true not only for the analytical methods we shall be dealing with, but also for numerical/computational methods which will taught next year.

5.2.1 IVP or BVP?

An equation containing only a first derivative, such as

$$\frac{dy}{dt} + 3y = 0, \tag{5.3}$$

is incomplete because it requires an extra condition for it to be solved. Therefore a **boundary condition** such as,

$$y(0) = 1, \tag{5.4}$$

could be provided and then it is possible to provide a full solution, as we shall see later.

For first order equations like this, ones which have a first derivative, such a boundary condition is also called an **initial condition** and that the equation together with the initial condition is called an **Initial Value Problem** (IVP). These can also be called evolution equations because the solution then evolves in time from the initial condition.

That this is so may be made clearer (hopefully) by the following thought experiment. Suppose that we are at the given initial time (say $t = 0$) and y is equal to the given initial value (say $y = 1$). The substitution of these values of t and y into the ODE then gives us the value of y', which is the initial slope of y. If we now use that slope to predict what value y takes a very small time interval later, then that will give us a new y-value at the new time. So we have travelled a very small way along the tangent to the curve. Having the values of y and t at this slightly later time means that we can compute the new (but slightly different) value of y' and progress a little further. Thus we can repeat this procedure for as long as we wish, and it will trace out at least an approximation to the exact solution.

Actually, I have just described what is called Euler's method, a well-known numerical scheme for solving ODEs. It isn't particularly accurate — there are much much better ones available than this — but its accuracy improves as the time steps are reduced in size. Nevertheless, this shows that the solution evolves from the **initial condition** and that the problem being solved (ODE plus initial condition) forms an IVP.

If instead we have the following ODE with a second derivative (which represents an undamped mass/spring system) together with two initial conditions,

$$\frac{d^2y}{dt^2} + 3y = 0 \qquad \text{subject to} \qquad y(0) = 1 \text{ and } y'(0) = 0, \tag{5.5}$$

then this too forms an IVP because both of the boundary conditions are given at the same point in time. An alternative thought experiment, a physical one in this instance, may be used to justify this classification as an IVP. The initial conditions correspond to a unit displacement and a zero velocity. These are precisely equivalent to what we might do in practice, namely to extend the spring to a given extension and to release it from rest at $t = 0$. Then what happens next is that the mass begins to move in the negative y-direction and the displacement from the initial state begins to decrease. So these two conditions are sufficient to determine the future evolution of the system. The system is an IVP.

In general, then, when all the Boundary Conditions are Initial Conditions (i.e. the values of y and a sufficient number of derivatives are given at the same point in time), then this will always be an IVP.

As an example of a BVP we may use the equation given in Eq. (5.5) but with the alternative boundary conditions,

$$y(0) = 1 \text{ and } y(1) = 0. \tag{5.6}$$

The most important feature of these conditions is that they are given at **two different points in time.** Formally this is called a **two-point boundary value problem** although it is exceptionally rare for BVPs to involve boundary conditions which involve three or more points. So one can get away with calling it a **boundary value problem**. Physically, we now have a mass/spring system which is subject to a unit initial displacement but with an unknown initial velocity. However, that velocity must be the correct one to yield a zero displacement when $t = 1$. Mathematically, this doesn't provide too much of a difficulty when compared with an equivalent IVP, but numerically it requires quite different techniques to find the solution: the initial velocity must be iterated upon in order to find the correct one which satisfies the condition at $t = 1$. IVPs do not need such iterations. So the choice of the method depends on the classification.

As a final example consider the following coupled pair of equations:

$$\frac{d^2x}{dt^2} + 2x - y = 0, \qquad \frac{d^2y}{dt^2} + 2y - x = 0. \tag{5.7}$$

These represent the motion of two masses in an undamped mass/spring system. If we had to solve this subject to the conditions:

$$x(0) = 1,\ x'(0) = 0,\ y(0) = 0,\ y'(0) = 1, \tag{5.8}$$

then the ODE and its boundary conditions form an IVP. In this case one mass (x) has a unit displacement with a zero velocity at $t = 0$, while the other has a zero displacement and a unit velocity which are also at $t = 0$.

On the other hand if we had to solve the same equation subject to the boundary conditions,

$$x(0) = 1,\ x'(0) = 0,\ y(1) = 0,\ y'(0) = 1, \tag{5.9}$$

or

$$x(0) = 1,\ x'(5) = 0,\ y(5) = 0,\ y'(0) = 5, \tag{5.10}$$

then both of these cases would form examples of BVPs.

5.2.2 Linearity and Nonlinearity

I offer the following definition of a linear equation which is precise, although it may need to be read quite a few times before it makes sense!

An equation or system of coupled equations is linear when all the dependent variables and their derivatives are multiplied either by constants or by functions of the independent variable, otherwise the equation or system is nonlinear.

If we confine ourselves to a single ODE for $y(t)$ with as many derivatives appearing as we wish, then the most complicated **linear equation** will take the form,

$$f_n(t)\frac{d^n y}{dt^n} + f_{n-1}(t)\frac{d^{n-1}y}{dt^{n-1}} + \cdots + f_2(t)\frac{d^2 y}{dt^2} + f_1(t)\frac{dy}{dt} + f_0(t)y = F(t), \tag{5.11}$$

where $f_i(t)$ $(i = 0, 1, 2, \cdots, n)$ and $F(t)$ are given functions of t at worst. At the opposite extreme, the coefficients of y and its derivatives could be constants:

$$a_n\frac{d^n y}{dt^n} + a_{n-1}\frac{d^{n-1}y}{dt^{n-1}} + \cdots + a_2\frac{d^2 y}{dt^2} + a_1\frac{dy}{dt} + a_0 y = F(t), \tag{5.12}$$

and we will be solving equations like these later. The extension of this definition of linearity to systems of equations is straightforward: all of the ODEs making up the system must be linear. An example is Eq. (5.7).

Example 5.1: $\dfrac{dy}{dt} + ay = 0$ is linear. This is so because both y and y' are multiplied by constants.

Example 5.2: $\dfrac{d^2y}{dt^2} + ty = e^{-t}$ is is also linear. This is so because both y is multiplied by a function of t and y' is multiplied by a constant. The function of t on the right hand is irrelevant.

Example 5.3: $\dfrac{d^2A}{dx^2} + A - A^3 = 0$ is nonlinear because of the A^3, a power of the dependent variable.

Example 5.4: $\dfrac{d^3y}{dt^3} + \underbrace{y\dfrac{dy}{dt}}_{\text{nonlinear}} + \underbrace{t^5 y}_{\text{linear}} = 0$ is nonlinear. We have the product of two dependent variables.

Example 5.5: $\dfrac{d^2\theta}{dt^2} + \dfrac{g}{L}\sin\theta = 0$ is nonlinear because $\sin\theta \simeq \theta - \frac{1}{3!}\theta^3 + \cdots$, using Taylor's series.

Example 5.6: $y'' + ty - z = 0, \quad z''' + 2z - yz = 0$. The first equation is linear but the second is nonlinear due to the presence of yz, a product of dependent varables. Overall, this is a nonlinear system because the nonlinearity in the second equation effectively contaminates the full system.

Note: There is one big difference between linear and nonlinear systems. If there are two different solutions of a linear equation or system of equations, then the sum (or even a weighted sum) will also be a solution. However, this does not happen for nonlinear equations/systems. As an example, let both $y_1(t)$ and $y_2(t)$ satisfy the equation $y'' + Ky = 0$, and we shall test if $y_1 + y_2$ also satisfies it. So if we let $y = y_1 + y_2$ in the ODE, then

$$\begin{aligned} y'' + Ky &= (y_1'' + y_2'') + K(y_1 + y_2) \\ &= \cancel{(y_1'' + Ky_1)} + \cancel{(y_2'' + Ky_2)} \\ &= 0. \end{aligned} \tag{5.13}$$

If, on the other hand, we were to test this idea out with the equation, $y'' + y^2 = 0$, then we would obtain,

$$
\begin{aligned}
y'' + y^2 &= (y_1'' + y_2'') + (y_1 + y_2)^2 \\
&= \cancel{(y_1'' + y_1^2)} + \cancel{(y_2'' + y_2^2)} + 2y_1 y_2 \\
&= 2y_1 y_2 \\
&\neq 0 \quad \text{in general.}
\end{aligned}
\tag{5.14}
$$

Therefore a sum of two solutions of a nonlinear ODE doesn't necessarily satisfy the same ODE.

Note: The fact that we can add two solutions of a linear ODE to obtain another solution will be used later when we solve Linear Constant-Coefficient ODEs.

5.2.3 The order of ODEs and systems of ODEs

The **order** of an equation is the order of the highest derivative appearing in that equation. Thus

$$\frac{dy}{dt} = -ay \quad \text{is of 1st order,} \tag{5.15}$$

while

$$\frac{d^3y}{dt^3} + \left(\frac{dy}{dt}\right)^{10} + y = 0 \qquad \text{is of 3rd order.} \tag{5.16}$$

In Eq. (5.16) one must not be put off by the 10th power appearing on the 1st derivative term; the 10th power does not affect the fact that the highest derivative is a 3rd derivative.

In many physical systems such as complex circuits or coupled mass/spring configurations more thn just than one differential equation is the rule, rather than the exception. For such systems **the order of the system is equal to the sum of the orders of the individual equations**. For example, the system,

$$
\begin{aligned}
m\frac{d^2x}{dt^2} + k(2x - y) = 0 \\
m\frac{d^2y}{dt^2} + k(2y - x) = 0,
\end{aligned}
\tag{5.17}
$$

is of 4th order because it is composed of two 2nd order equations: $2 + 2 = 4$. If we have a system which is composed of one 1st order equation, two 3rd order equations and one 5th order equation, then the whole system is of 12th order simply because $1 + 3 + 3 + 5 = 12$.

This seems straightforward enough, but very occasionally there might be slight confusion. In the following system,

$$
\begin{aligned}
\frac{dx}{dt} + x - \frac{d^2y}{dt^2} = t \\
\frac{d^3y}{dt^3} + 2y - x = 0,
\end{aligned}
\tag{5.18}
$$

the first equation (for x) is of 1st order while the second equation (for y) is of 3rd order. Hence the system is of 4th order. The presence of y'' in the first equation doesn't make that equation to be of 2nd order because this is the equation for x. The troublesome y'' is of lower order (a 2nd derivative) than the y''' (a third derivative) in the second equation. This illustrates the fact that, for physical systems, each equation in a system will be associated/identified with a unique dependent variable. However, I can assure you that it is very rare that such a potential conundrum will arise in real life.

5.2.4 Reduction to First order Form

The idea of *the order of a system of equations* may also be seen when such a system is reduced to what is called **first order form**.

For example, if we take Eq. (5.16) above, then we may transform it into first order form by first defining three new independent variables, y_1, y_2 and y_3, each of which are functions of t, according to

$$y_1 = y, \qquad y_2 = \frac{dy}{dt}, \qquad y_3 = \frac{d^2y}{dt^2}. \tag{5.19}$$

We now form 1st order equations for each of these three variables, either by using their definitions in Eq. (5.19) or by using the original Eq. (5.16) suitably translated into the new notation. Thus we obtain,

$$\begin{aligned} \frac{dy_1}{dt} &= y_2 \\ \frac{dy_2}{dt} &= y_3 \\ \frac{dy_3}{dt} &= -y_1 - y_2^{10}, \end{aligned} \tag{5.20}$$

and therefore we have three 1st order equations replacing the original one 3rd order equation: $3 \times 1 = 1 \times 3$. Note carefully that the first two of these equations come directly from the definitions in Eq. (5.19) while the third equation is essentially a translation of Eq. (5.16) into the new notation. So we have achieved the primary aim, namely first order equations for *each* of the three new dependent variables.

We may also reduce the system given in Eq. (5.17) to first order form. As both of these equations are of 2nd order, we need to use two new variables for each of the original dependent variables. We define

$$z_1 = x, \qquad z_2 = \frac{dx}{dt}, \qquad z_3 = y, \qquad z_4 = \frac{dy}{dt}, \tag{5.21}$$

and therefore the system (5.17) becomes,

$$\begin{aligned} \frac{dz_1}{dt} &= z_2 \\ m\frac{dz_2}{dt} &= -k(2z_1 - z_3) \\ \frac{dz_3}{dt} &= z_4 \\ m\frac{dz_4}{dt} &= -k(2z_3 - z_1). \end{aligned} \tag{5.22}$$

Thus a 4th order system which was composed of two 2nd order equations has been reduced to four 1st order equations: $2 \times 2 = 4 \times 1$.

Note: This technique of reduction to first order form is especially useful when solving ODEs using numerical methods; this topic will be covered next year. The subscripts on z in Eq. (5.21), for example, represent array indices in Matlab or Fortran.

5.2.5 Full classifications and reduction to first order form

In this subsubsection we'll undertake a full classification on an ODE and on a system of ODEs and reduce each together with their boundary conditions to first order form.

Example 5.7: Consider the fifth order equation,

$$y''''' + t^2 y'''(1 - y') + yy'' = t, \tag{5.23}$$

subject to the boundary conditions,

$$y(0) = y'(0) = 0, \quad y''(0) = 1, \quad y'(1) = 1, \; y'''(1) = 0. \tag{5.24}$$

This may be seen to be a **5th order nonlinear BVP**. Given that it is of 5th order, we need to define five new variables to replace the original y. I like to tackle this simply by creating a Table of data, as follows.

Variable	Equation	BC($t = 0$)	BC($t = 1$)
$y_1 = y$	$y_1' = y_2$	$y_1(0) = 0$	•
$y_2 = y'$	$y_2' = y_3$	$y_2(0) = 0$	$y_2(1) = 1$
$y_3 = y''$	$y_3' = y_4$	$y_3(0) = 1$	•
$y_4 = y'''$	$y_4' = y_5$	•	$y_4(1) = 0$
$y_5 = y''''$	$y_5' = t - y_1 y_3 - t^2 y_4(1 - y_2)$	•	•

(5.25)

The bullet symbols have been used solely to indicate those initial and final conditions which haven't been specified. Thus we can see that there are no initial conditions for y_4 and y_5, and therefore these would need to be found so that the conditions for y_2 and y_4 at $t = 1$ are satisfied.

Example 5.8: Consider the following system of ODEs,

$$\begin{aligned} &\psi''' + 3\psi\psi'' - 2(\psi')^2 + \theta + N\phi = 0, \\ &\theta'' + 3a\,\psi\theta' = 0, \\ &\phi'' + 3b\,\psi\phi' = 0, \end{aligned} \tag{5.26}$$

where a, b and N are constants. These equations arise in the free convective boundary layer flow due to heat and solute being supplied into a fluid. The primes denote derivatives with respect to distance, x. The boundary conditions are that ψ and ψ' are zero at $x = 0$ while both θ and ϕ are equal to 1. As $x \to \infty$, ψ', θ and ϕ all tend to zero.

This is a seventh order system $(3 + 2 + 2)$ and so we need seven variables to replace the present dependent variables, ψ, θ and ϕ. Three will be needed for ψ and two each for θ and ϕ. Therefore we get the following:

Variable	Equation	BC($x = 0$)	BC($x \to \infty$)
$v_1 = \psi$	$v_1' = v_2$	$v_1 = 0$	•
$v_2 = \psi'$	$v_2' = v_3$	$v_2 = 0$	$v_2 \to 0$
$v_3 = \psi''$	$v_3' = -3v_1 v_3 + 2v_2^2 - v_4 - N v_6$	•	•
$v_4 = \theta$	$v_4' = v_5$	$v_4 = 1$	$v_4 \to 0$
$v_5 = \theta'$	$v_5' = -3a v_1 v_5$	•	•
$v_6 = \phi$	$v_6' = v_7$	$v_6 = 1$	$v_6 \to 0$
$v_7 = \phi'$	$v_7' = -3b v_1 v_7$	•	•

(5.27)

So this is a **7th order nonlinear boundary value problem**. The initial conditions for v_3, v_5 and v_7 are unknown and will need to be iterated upon in a numerical scheme to enable the the three boundary conditions as $x \to \infty$ to be satisfied — again, this numerical aspect is something that you won't need to worry about until next year.

Finally we return the apparently awkward case given in (5.18) which is repeated here:

$$\begin{aligned} &\frac{dx}{dt} + x - \frac{d^2y}{dt^2} = t \\ &\frac{d^3y}{dt^3} + 2y - x = 0, \end{aligned} \tag{5.28}$$

The reduction to first order form (without boundary conditions, for none were given) is as follows:

Variable	Equation
$v_1 = x$	$v_1' = t - v_1 + v_4$
$v_2 = y$	$v_2' = v_3$
$v_3 = y'$	$v_3' = v_4$
$v_4 = y''$	$v_4' = v_1 - 2v_2.$

(5.29)

While one might have thought that a serious problem will ensue by having a second y-derivative in the equations for x, this reduction to first order form has derivatives only on the left hand sides of the transformed equations.

5.3 A very simple 1st order ODE

The most straightforward case of an ODE is when

$$\frac{dy}{dt} = f(t) \tag{5.30}$$

and the solution is obtained by integrating both sides of the equation with respect to t to get

$$y = \int f(t)\,dt + c \tag{5.31}$$

where c is the constant of integration and the integral is an indefinite integral. The constant of integration is obtained by having an initial condition whereby the value of y at a given value of t is known. Therefore if we were solving

$$\frac{dy}{dt} = f(t) \qquad \text{subject to} \quad y(1) = A, \tag{5.32}$$

then the solution may be written in the form,

$$y = \int_1^t f(\alpha)\,d\alpha + A, \tag{5.33}$$

where α is a dummy variable of integration. The choice of the lower limit means that the value of the integral is zero when $t = 1$, and therefore the initial condition is seen to be satisfied.

Example 5.9: If we have the equation

$$\frac{dy}{dt} = t^2 \qquad \text{subject to } y(1) = 3, \tag{5.34}$$

then the solution is

$$y = \int_1^t \alpha^2\,d\alpha + 3 \quad = \quad \Big[\tfrac{1}{3}\alpha^3\Big]_1^t + 3 \quad = \quad \Big[\tfrac{1}{3}t^3 - \tfrac{1}{3}(1)^3\Big] + 3 \quad = \quad \tfrac{1}{3}t^3 + \tfrac{8}{3}. \tag{5.35}$$

An alternative (but more familiar) method is to use the indefinite integral, as in Eq. (5.31), and then to substitute the initial condition to obtain c. Thus

$$y = \tfrac{1}{3}t^3 + c \qquad \text{but} \qquad y(1) = 3 \implies 3 = \tfrac{1}{3}(1)^3 + c \qquad \implies \qquad c = \tfrac{8}{3} \qquad \Rightarrow \qquad y = \tfrac{1}{3}t^3 + \tfrac{8}{3}. \tag{5.36}$$

Generally people tend to use the latter method as it is less complicated.

Note: Really it's a bit of a cheat to call this an ODE when it is really an integration exercise in disguise.

5.4 Separation of variables

This technique applies to 1st order ODEs which take the form,

$$\frac{dy}{dt} = f(y)g(t), \tag{5.37}$$

where f and g are given functions. The general solution is obtained by **separating the variables,** i.e. by taking all y-dependent terms to one side and all t-dependent terms to the other, and then integrating. Thus the general solution follows in this manner:

$$\begin{aligned} & \frac{dy}{dt} = f(y)g(t) \\ \Longrightarrow \quad & \frac{dy}{f(y)} = g(t)\,dt \\ \Longrightarrow \quad & \int \frac{dy}{f(y)} = \int g(t)\,dt. \end{aligned} \tag{5.38}$$

This splitting of the dy/dt into a separate dy and dt is meant to be thought of in a "$\lim_{\delta t \to 0}$" sense, such as we saw many times in Maths 1.

So this method of solving a **variables-separable** equation reduces down to the finding of two integrals. The final solution is them obtained when an initial condition is applied. We shall apply this to a few examples.

Example 5.10: Solve the ODE, $\dfrac{dy}{dt} = 2\sqrt{y}\,\cos t$ subject to $y(0) = 0$.

Clearly this is a variables-separable ODE and therefore we may proceed as above:

$$
\begin{aligned}
& & \frac{dy}{dt} &= 2\sqrt{y}\,\cos t \\
&\Longrightarrow & \frac{dy}{\sqrt{y}} &= 2\cos t\,dt \\
&\Longrightarrow & \int \frac{dy}{\sqrt{y}} &= \int 2\cos t\,dt \\
&\Longrightarrow & 2\sqrt{y} &= 2\sin t + c & & c \text{ is arbitrary} \\
&\Longrightarrow & \sqrt{y} &= \sin t + \tfrac{1}{2}c \\
&\Longrightarrow & y &= (\sin t + \tfrac{1}{2}c)^2 & & \text{Now apply the Initial Condition} \ldots \\
&\Longrightarrow & y &= \sin^2 t & & y(0) = 0 \;\Rightarrow\; c = 0
\end{aligned}
\tag{5.39}
$$

Note: This is the typical way that a separation-of-variables problem proceeds. The final solution is as difficult to obtain as the integrals are which make it up. In the next two examples we will consider slightly different cases where the respective presence of a square root and of a logarithm will require some careful treatment.

Example 5.11: Solve the ODE, $y' = 3t^2/y$ subject to $y(1) = 2$.

We have

$$
\begin{aligned}
\frac{dy}{dt} = \frac{3t^2}{y} \quad &\Longrightarrow \quad \int y\,dy = \int 3t^2\,dt \\
&\Longrightarrow \quad \tfrac{1}{2}y^2 = t^3 + c \\
&\Longrightarrow \quad y = \pm\sqrt{2t^3 + 2c}.
\end{aligned}
\tag{5.40}
$$

Application of the initial condition shows that $2 = \pm\sqrt{2 + 2c}$. The only way that this can be solved is for $c = 1$ to be chosen *and* the positive sign to be taken. The final solution is

$$
y = \sqrt{2t^3 + 2}. \tag{5.41}
$$

In our analysis we have evaluated the arbitrary constant right at the end. It is possible to do so at an earlier point. So if we have got as far as $\frac{1}{2}y^2 = t^3 + c$ and applied $y(1) = 2$, then we will obtain, $c = 1$, perhaps not surprisingly. Hence $y^2 = 2t^3 + 2$. Now we need to take the square root: $y = \pm\sqrt{2t^3 + 2}$, and then appeal a second time to the initial condition in order to confirm that we need the positive square root. This way of doing things does feel a little strange but it is entirely consistent with the first way.

Example 5.12: Solve the ODE, $y' = 2ty$, subject to $y(0) = -2$.

On separating the variables we get

$$\int \frac{dy}{y} = \int 2t\,dt, \tag{5.42}$$

and hence

$$\ln|y| = t^2 + c. \tag{5.43}$$

If we apply the Initial Condition at this point, then we obtain $c = \ln 2$. Setting $e^{\text{LHS}} = e^{\text{RHS}}$ gives

$$|y| = e^{t^2+\ln 2} \qquad \Longrightarrow \qquad |y| = 2e^{t^2} \qquad \Longrightarrow \qquad y = \pm 2e^{t^2}, \tag{5.44}$$

and again we need to invoke the Initial Condition a second time to confirm that we need the negative option. Hence the solution is $y = -2e^{t^2}$.

That was fine and in many ways the outcome (viz. the choosing of the correct sign) was determined in the same way as in Ex. 5.10. However, many people prefer to find y explicitly in terms of both t and the arbitrary constant before applying the initial condition, so lets do that. Equation (5.43) yields,

$$|y| = e^{t^2+c} = e^c e^{t^2}. \tag{5.45}$$

Now the application of $y(0) = -2$ gives $e^c = 2$, and hence $|y| = 2e^{t^2}$. Strictly, we can now remove the modulus signs and account for that with the introduction of a $\pm$ on the right hand side, and then we can choose the correct sign for the final answer, as before. But I have often seen the modulus signs just disappear when in the heat of the battle in the exam and then the initial condition gives $e^c = -2$, and thereby disaster follows!

My preferred safe route through a problem like this would be to use the following as the next line after Eq. (5.45):

$$y = e^{t^2+c} = Ae^{t^2} \tag{5.46}$$

where A is still arbitrary for now, but it does allow us to obtain the correct sign for y without any choice being made.

OK, this analysis has been far too wordy and disjointed, and therefore the best thing to do is for me to run this all again from scratch as a full analysis.

$$\begin{aligned}
& \frac{dy}{dt} &&= 2ty \\
\Longrightarrow\quad & \int \frac{dy}{y} &&= \int 2t\,dt \\
\Longrightarrow\quad & \ln|y| &&= t^2 + c \\
\Longrightarrow\quad & |y| &&= e^{t^2+c} = e^c\,e^{t^2} \\
\Longrightarrow\quad & y &&= Ae^{t^2} && \text{replacement of the arbitrary constant} \\
\Longrightarrow\quad & y &&= -2e^{t^2} && \text{using } A = -2 \text{ from the initial conditions.}
\end{aligned} \tag{5.47}$$

I hope this is also a safe route for you.

5.5 1st order linear equations

These equations fall into the general form,

$$\frac{dy}{dt} + P(t)y = Q(t), \tag{5.48}$$

where $P(t)$ and $Q(t)$ are given functions of t. There is a general method for solving these equations, but before it is presented let us consider the following example.

Example 5.13: Solve the equation,

$$\frac{dy}{dt} + \frac{2}{t}y = 5t^2. \tag{5.49}$$

Believe it or not, this equation is simplified slightly by multiplying both sides by t^2. On doing this we get

$$t^2\frac{dy}{dt} + 2ty = 5t^4. \tag{5.50}$$

Although this latest equation still doesn't look simple, the left hand side is the exact derivative of t^2y since $d(t^2y)/dt = t^2y' + 2ty$. Hence this equation may be rewritten in the form,

$$\frac{d}{dt}\left(t^2y\right) = 5t^4. \tag{5.51}$$

We may now integrate both sides to obtain

$$t^2y = t^5 + c \qquad \text{and hence} \qquad y = t^3 + ct^{-2} \tag{5.52}$$

is the solution, where c is an arbitrary constant.

Note: Once more, the ODE has been solved using an integration, but only after the ODE had been modified by multiplication by a function that I appeared to pluck out of thin air! So the above method is straightforward enough except for the reason why we chose to multiply throughout by t^2. Our progress was facilitated by choosing the correct function to make the left hand side equal to an exact differential. The question is: **How do we find that function?** The following subsection is a derivation of the formula for finding that function — this is for interest only, but it does explain why the weird formula works.

5.5.1 The Integrating Factor

If we return to the general equation given in Eq. (5.48), reproduced here:

$$\frac{dy}{dt} + P(t)y = Q(t),$$

then let the function we require be $F(t)$. After multiplication by F, we get,

$$Fy' + FPy = FQ. \tag{5.53}$$

Now we insist that the left hand side is an exact derivative. If it is, then it must be the derivative of Fy, given the presence of Fy' in Eq. (5.53). As the differential of Fy is $Fy' + F'y$, this means that Eq. (5.53) must also be written precisely as,

$$Fy' + F'y = FQ. \tag{5.54}$$

So the terms in red in Eqs. (5.53) and (5.54) must be identical and hence we must have,

$$FP = F'. \tag{5.55}$$

Equation (5.55) may be rearranged to get

$$\frac{1}{F}\frac{dF}{dt} = P(t), \tag{5.56}$$

which is of variables-separable type although this may not be too obvious. Therefore

$$\int \frac{dF}{F} = \int P(t)\,dt \quad \Longrightarrow \quad \ln|F| = c + \int P(t)\,dt \quad \Longrightarrow \quad F = Ae^{\int P(t)\,dt}. \tag{5.57}$$

We have introduced the constant of integration, c, as usual, and then set $A = e^c$ as we did in Eq. (5.46).

Note: we always neglect to use A. This is because we shall be multiplying the original equation by F, and therefore A will multiply both sides of the resulting equation and may be cancelled. So in practice we always use the following formula for F, which is referred to as the **Integrating Factor**,

$$\boxed{F = e^{\int P(t)\,dt}.} \tag{5.58}$$

For the example ODE given above in Eq. (5.49), the Integrating Factor is

$$F = e^{\int (2/t)\,dt} = e^{2\ln t} = e^{\ln t^2} = t^2, \tag{5.59}$$

which is indeed the function by which we multiplied.

This method also yields a formula for the general solution,

$$y = \left[c + \int FQ\,dt\right]/F \qquad \text{where} \qquad F = e^{\int P(t)\,dt}, \tag{5.60}$$

but I would very definitely recommend remembering only the formula for F, and then proceeding as in the following examples.

Example 5.14: Solve the equation, $y' + \dfrac{3y}{t} = \dfrac{2}{t^2}$ subject to $y(1) = 2$.

The coefficient of y is $3/t$, and therefore the Integrating Factor is

$$e^{\int (3/t)\,dt} \;=\; e^{3\ln|t|} \;=\; e^{\ln|t^3|} = t^3. \tag{5.61}$$

Note: that this is one of the rare occasions when one doesn't need to worry about the modulus signs in a logarithm! So let us multiply the original ODE by t^3. We get

$$t^3y' + 3t^2y = 2t. \tag{5.62}$$

Our theory guarantees that the left hand side is an exact derivative of something. In this case it is the derivative of t^3y — check this using the product rule. So Eq. (5.62) becomes,

$$\begin{aligned}
(t^3y)' &= 2t \\
\Longrightarrow \qquad t^3y &= t^2 + c \qquad \text{upon integrating} \\
\Longrightarrow \qquad y &= \frac{1}{t} + \frac{c}{t^3}.
\end{aligned} \tag{5.63}$$

The application of the initial condition, $y(1) = 2$, yields $c = 1$ and therefore the final solution is,

$$y = \frac{1}{t} + \frac{1}{t^3}. \tag{5.64}$$

Example 5.15: Solve the equation, $y' + 2xy = 4xe^{-x^2}$, subject to $y(0) = 1$.

OK, the right hand side looks horrible, but let us just get on with the task of finding the Integrating Factor and leave any potential trouble until later. Also, note that this equation has x as the independent variable, not t, but this makes no difference at all to what we do. Given that the coefficient of y is $2x$ the Integrating Factor is,

$$e^{\int 2x\,dx} = e^{x^2}. \tag{5.65}$$

On multiplying the given ODE by the Integrating Factor we get,

$$e^{x^2}y' + 2xe^{x^2}y = 4x. \tag{5.66}$$

So the right hand side has simplified nicely but the left hand side has now become a mess. However, we are guaranteed that the left hand side is an exact derivative. So the rest of the analysis proceeds as follows.

$$\begin{aligned} e^{x^2}y' + 2xe^{x^2}y &= 4x \\ \Longrightarrow \qquad \left(e^{x^2}y\right)' &= 4x \\ \Longrightarrow \qquad e^{x^2}y &= 2x^2 + c \qquad\qquad \text{on integrating} \\ \Longrightarrow \qquad y &= \left[2x^2 + c\right]e^{-x^2}. \end{aligned} \tag{5.67}$$

Application of the Initial Conditon, $y(0) = 1$, yields $c = 1$. Hence the required solution is,

$$y = \left[2x^2 + 1\right]e^{-x^2}. \tag{5.68}$$

Example 5.16: Solve the equation, $(\cot t)\,y' + y = \cot t \cos t$, subject to $y(0) = 1$.

The very very first thing to be done is to note that the coefficient of y' must be 1 so that we may apply the theory we derived earlier. Clearly we need to divide the full equation by $\cot t$. The equation becomes,

$$y' + (\tan t)\,y = \cos t. \tag{5.69}$$

The integrating factor is,

$$F = e^{\int \tan t\,dt} = e^{-\int(-\sin t/\cos t)\,dt} = e^{-\ln\cos t} = 1/\cos t. \tag{5.70}$$

Note how the integrand was manipulated to get it into an f'/f form, which necessitated the use of two minus signs. Multiplication by the Integrating Factor yields,

$$(\sec t)\,y' + (\tan t \sec t)\,y = 1. \tag{5.71}$$

As fearsome as this looks, again we can reduce our collective blood pressures by noting that the left hand side is, by design, an exact derivative. Therefore our analysis takes the following route:

$$\begin{aligned} (\sec t)\,y' + (\tan t\sec t)\,y &= 1 \\ \Longrightarrow \qquad \Big[(\sec t)\,y\Big]' &= 1 \\ \Longrightarrow \qquad (\sec t)\,y &= t + c \qquad\qquad \text{on integration} \\ \Longrightarrow \qquad y &= (t + c)\cos t. \end{aligned} \tag{5.72}$$

Application of the initial condition, $y(0) = 1$, gives $c = 1$ and hence the final solution is,

$$y = (t + 1)\cos t. \tag{5.73}$$

Example 5.17: Solve the equation, $(\cot t)\, y' - y = \cot t \cos t$, subject to $y(0) = 1$.

This is the same as the previous example except for the replacement of a plus by a minus. Clearly we need to attain a unit coefficient for the y' once more and therefore division by $\cot t$ yields,

$$y' - (\tan t)\, y = \cos t. \tag{5.74}$$

The integrating factor is,

$$F = e^{\int -\tan t\, dt} = e^{\int (-\sin t/\cos t)\, dt} = e^{\ln\cos t} = \cos t. \tag{5.75}$$

Multiplication by the Integrating Factor yields,

$$(\cos t)\, y' - (\sin t)\, y \;=\; \cos^2 t \;=\; \tfrac{1}{2}(1 + \cos 2t), \tag{5.76}$$

using a multiple angle formula. The rest of the analysis now follows:

$$\begin{aligned}
(\cos t)\, y' - (\sin t)\, y &= \tfrac{1}{2}(1 + \cos 2t) \\
\Longrightarrow \quad \Big[(\cos t)\; y\Big]' &= \tfrac{1}{2}(1 + \cos 2t) \\
\Longrightarrow \quad (\cos t)\, y &= \tfrac{1}{2}t + \tfrac{1}{4}\sin 2t + c && \text{upon integrating} \\
&= \tfrac{1}{2}(t + \sin t \cos t) + c && \text{multiple angles again} \\
\Longrightarrow \quad y &= \tfrac{1}{2}(t\sec t + \sin t) + c\sec t.
\end{aligned} \tag{5.77}$$

Now we apply the initial condition, $y(0) = 1$, to yield $c = 1$. Hence the final solution is,

$$y = \tfrac{1}{2}\Big[(t + 2)\sec t + \sin t\Big]. \tag{5.78}$$

So the change of one sign between Examples 5.16 and 5.17 results in two very different solutions.

5.6 Solution of homogeneous linear, constant coefficient ODEs

The most general n^{th} order ODE of this type may be written in the form

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \cdots a_1 \frac{dy}{dt} + a_0 y = F(t), \tag{5.79}$$

where

- $F(t)$ is a given real function,
- all the a_i coefficients are real and
- a_n is nonzero (since otherwise it would not be an n^{th} order ODE!).

Given that $a_n \neq 0$, we could divide Eq. (5.79) by a_n to yield an n^{th} order ODE where $d^n y/dt^n$ has a unit coefficient.

Such equations (and indeed coupled systems of these equations) form the core of ODE theory and the modelling of physical systems. Mass/spring systems, electrical circuits, hydraulic circuits, vibrating structures of many different kinds including bridges, buildings and aircraft, may all be modelled by linear constant-coefficient ODEs. Therefore they assume a huge importance in Engineering. We will be revisiting ODE theory occasionally in later topics in this unit: Laplace Transforms, Fourier Series and Matrices. By the end of the unit you should have acquired a great facility in the various methods of their solution.

At the outset it should be noted that these ODEs may be split into two closely-related families depending on the function, $F(t)$, which is on the right hand side of Eq. (5.79). Here follows some terminology.

- When $F(t) = 0$ then Eq. (5.79) is called **homogeneous**.
- When the equation is homogeneous, nonzero solutions can only arise if at least one initial condition is nonzero.
- When $F(t) \neq 0$ then Eq. (5.79) is called **inhomogeneous**.
- The term, $F(t)$, is called the **forcing term** or the **inhomogeneous term**.
- When the equation is inhomogeneous, then the solutions will be nonzero even if all of the initial conditions are zero.

Some of these ideas will be unpacked later in the unit

My intention is to develop a unified approach to solving these equations through a sequence (not series, hah!) of examples with suitable comments to show how the exposition is evolving. We will begin with homogeneous ODEs and then extend our expertise to inhomogeneous ODEs afterwards.

Example 5.1: Solve the ODE $y' + ay = 0$.

Obviously $y = 0$ is a solution, but that's boring and there is a nonzero solution to be found.

Now this is the simplest possible 1st order ODE. It is linear and has no forcing term. Indeed, the theory from §5.5 could be used because it is of first order linear form. Actually, it is also of variables-separable form, and so the theory of §5.4 could also be applied. However, I wish to develop an approach which will work with all linear constant-coefficient equations whatever their order.

First, I'll rearrange the equation:

$$y' = -ay. \tag{5.80}$$

An equation is always the expression of a balance of some kind, but this equation also says that, whatever shape y has, then y' must have precisely the same shape. In other words, y and y' must be identical functions apart, perhaps, from their amplitudes. The only function which, when differentiated, yields exactly the same function is an exponential. Sines and cosines don't do this after one differentiation, and polynomials don't do it either. The question then is, which exponential? We may work it out by substituting $y = Ae^{\lambda t}$ into Eq. (5.80), where λ is to be found and where A is arbitrary. We get,

$$A\lambda e^{\lambda t} = -aAe^{\lambda t} \qquad \Longrightarrow \qquad A(\lambda + a)e^{\lambda t} = 0. \tag{5.81}$$

Given that the exponential cannot be zero and that $A = 0$ leaves us with only the boring zero solution, this means that,

$$\lambda + a = 0. \tag{5.82}$$

So $\lambda = -a$ and therefore the solution is,

$$y = Ae^{-at}, \tag{5.83}$$

where A is an arbitrary constant. If I had provided an initial condition then A could be found.

For example, if we had $y(0) = 2$ then (5.83) yields $A = 2$ and therefore the final solution is $y = 2e^{-at}$. Alternatively, if we were to have $y(1) = 2$ as the initial condition, then (5.83) yields $A = 2e^{a}$ and the final solution could be written in the form, $y = 2e^{-a(t-1)}$ — do check that out!

More generally, if $y(b) = c$ then $y = ce^{-a(t-b)}$.

As an exercise solve this example equation again, but using the separation-of-variables and first-order-linear methods in turn. This is worth doing just to see how these methods cope with the equation.

Example 5.2: Solve $y'' + 3y' + 2y = 0$.

Given our experience with Example 5.1 we shall use the same substitution. This is likely to work because *all* derivatives of $e^{\lambda t}$ are proportional to $e^{\lambda t}$. Therefore we obtain,

$$\begin{aligned}
& \lambda^2 e^{\lambda t} + 3\lambda e^{\lambda t} + 2e^{\lambda t} = 0 & \\
\Longrightarrow\quad & (\lambda^2 + 3\lambda + 2)e^{\lambda t} = 0 & \\
\Longrightarrow\quad & \underbrace{\lambda^2 + 3\lambda + 2 = 0}_{\text{Auxiliary Equation}} & \text{since } e^{\lambda t} \text{ cannot be zero} \qquad (5.84)\\
\Longrightarrow\quad & (\lambda + 1)(\lambda + 2) = 0 & \text{by factorisation}\\
\Longrightarrow\quad & \lambda = -1, -2. &
\end{aligned}$$

Note that the equation labelled, **Auxiliary Equation**, is also known as the **Indicial Equation** or even as the **Characteristic Equation**. All three terms are used by many people, so it's worth knowing their names (the terms, that is, not the people).

So we have two possible choices for λ, but which one should we choose? The answer is both, and we do so simultaneously. Therefore the general solution to the ODE is

$$y = Ae^{-t} + Be^{-2t}, \qquad (5.85)$$

where A and B are arbitrary constants. If we wished to do so then we may show that Eq. (5.85) satisfies the original ODE by substitution into the ODE.

In practice we would be provided with two boundary/initial conditions to satisfy in order to find A and B. An example of an Initial Value Problem might be:

$$\begin{cases} y(0) = 0 \\ y'(0) = 1 \end{cases} \Longrightarrow \begin{cases} A + B = 0 \\ -A - 2B = 1 \end{cases} \Longrightarrow \begin{cases} A = 1 \\ B = -1 \end{cases} \Longrightarrow y = e^{-t} - e^{-2t}. \qquad (5.86)$$

An example of a Boundary Value Problem is

$$\begin{cases} y(0) = 1 \\ y(1) = 0 \end{cases} \Longrightarrow \begin{cases} A + B = 1 \\ Ae^{-1} + Be^{-2} = 0 \end{cases} \Longrightarrow \begin{cases} A = 1/(1-e) \\ B = -e/(1-e) \end{cases} \Longrightarrow y = \frac{e^{-t} - e^{1-2t}}{1-e}. \qquad (5.87)$$

The final solution did need a few lines of manipulation to get it into such a compact form.

Note that it is quite easy to write down the auxiliary equation because the coefficients of powers of λ correspond to the coefficients of the respective derivatives of y:

$$y'' + 3y' + 2y = 0 \quad \Longrightarrow \quad \lambda^2 + 3\lambda + 2 = 0.$$

This also works the other way:

$$3\lambda^4 - 2\lambda^2 + \lambda + 6 = 0 \quad \Longrightarrow \quad 3y'''' - 2y'' + y' + 6y = 0.$$

Therefore an homogeneous ODE has a unique auxiliary equation and vice versa; knowledge of one means that we have knowledge of the other.

Example 5.3: Solve the equation, $y''' - 2y'' - y' + 2y = 0$.

Proceeding a little more quickly, the substitution of $y = e^{\lambda t}$ yields the auxiliary equation,

$$\begin{aligned} \lambda^3 - 2\lambda^2 - \lambda + 2 &= (\lambda - 2)(\lambda^2 - 1) \\ &= (\lambda - 2)(\lambda + 1)(\lambda - 1) = 0, \end{aligned} \tag{5.88}$$

for which the roots are $\lambda = 2, 1, -1$. Hence the solution is

$$y = Ae^{2t} + Be^{t} + Ce^{-t},$$

where we have the three arbitrary constants, A, B and C.

Note: we could introduce further examples of this type which are of 4th order, 5th order and so on, and where the auxiliary equation has different roots, all real and all nonzero, but nothing new arises. So the above three examples illustrate the general/standard case. The rest of this section is devoted to exceptions to this general case. Clearly it will become increasingly difficult to find all the λ-values as the degree of the polynomial for λ increases. Given that an n^{th} order ODE yields an n^{th} order polynomial for λ and hence there will be n aribtrary constants in the geneal solution, the application of boundary conditions also becomes more difficult as the degree of the polynomial increases. For example, in the case of a 5th order ODE there will be five boundary conditions to satisfy and hence five algebraic equations to solve for the five unknown constants. Nasty.

Example 5.4: Solve the equation, $y'' + 2y' = 0$.

The auxiliary equation is $\lambda^2 + 2\lambda = 0$. So $\lambda(\lambda + 2) = 0$ which means that $\lambda = 0, -2$. Although one of these λ-values is zero, we may proceed as usual:

$$y \;=\; Ae^{0t} + Be^{-2t} \;=\; A + Be^{-2t}. \tag{5.89}$$

So the presence of the root, $\lambda = 0$, means that the corresponding solution is that y is equal to a constant.

Example 5.5: Solve the ODE, $y'' + 9y = 0$.

The auxiliary equation is $\lambda^2 + 9 = 0$ from which we obtain,

$$\lambda^2 = -9 \qquad \Longrightarrow \qquad \lambda = \pm 3j. \tag{5.90}$$

The auxiliary equation has purely imaginary roots. Surely this is a problem? No, it isn't, for we may again proceed as usual:

$$y = Ae^{3jt} + Be^{-3jt}. \tag{5.91}$$

Given our theory to date, this is the obvious way to write down the solution. However, we have real equations and we expect to have real solutions rather than complex ones. However, we have already met complex exponentials and so we may play around a little with these:

$$\begin{aligned} y &= Ae^{3jt} + Be^{-3jt} \\ &= A(\cos 3t + j \sin 3t) + B(\cos 3t - j \sin 3t) \\ &= (A + B)\cos 3t + (Aj - Bj)\sin 3t \\ &= C \cos 3t + D \sin 3t, \end{aligned} \tag{5.92}$$

where $C = A + B$ and $D = (A - B)j$. When first seen, the last line in Eq. (5.92) looks like sleight-of-hand. Indeed, it will be automatically assumed by almost everyone that A and B are real, possibly because they have been so in all of the Examples before this one. It will also be assumed that C and D are real, but this is clearly inconsistent with the definitions of C and D in terms of A and B.

If A and B are both real, then C is real but D is purely imaginary. On the other hand, if A and B are complex conjugates of one another (which isn't crazy because the complex exponentials, e^{3jt} and e^{-3jt}, are complex conjugates of one another), then C and D are real. Specifically, we may take $A = C - Dj$ and $B = C + Dj$ and everything then the final solution in Eq. (5.92) is real.

Note 1: There are two main ways of writing down the solution for the present equation. They are

$$y = Ae^{3jt} + Be^{-3jt} \qquad \text{and} \qquad y = C\cos 3t + D\sin 3t.$$

The one which is chosen will almost always be the one involving sinusoids, although there are some circumstances when the one involving complex exponentials is better.

Note 2: If we were to solve this equation with the *real* initial conditions, $y(0) = 1$ and $y'(0) = 0$, then the final solution will be real. This happens even if the complex exponential form of the general solution is used. Let us check this out. On taking the complex exponential form, $y = Ae^{3jt} + Be^{-3jt}$, we have

$$\begin{cases} y(0) = 1 \\ y'(0) = 0 \end{cases} \Longrightarrow \begin{cases} A + B = 1 \\ 3j(A - B) = 0 \end{cases} \Longrightarrow \begin{cases} A = 1/2 \\ B = 1/2 \end{cases} \Longrightarrow y = \tfrac{1}{2}\left[e^{3jt} + e^{-3jt}\right] = \cos 3t. \tag{5.93}$$

On taking the sinusoidal form, $y = C\cos 3t + D\sin 3t$, we have

$$\begin{cases} y(0) = 1 \\ y'(0) = 0 \end{cases} \Longrightarrow \begin{cases} C = 1 \\ 3D = 0 \end{cases} \Longrightarrow y = \cos 3t. \tag{5.94}$$

So both forms of solution will yield the correct real answer, but the one involving sinusoids is a little quicker.

Example 5.6: Solve the ODE, $y'' + 4y' + 13y = 0$.

The auxiliary equation is,

$$\begin{aligned} & \lambda^2 + 4\lambda + 13 = 0 \\ \Longrightarrow\quad & (\lambda + 2)^2 + 9 = 0 \qquad \text{completing the square} \\ \Longrightarrow\quad & (\lambda + 2)^2 = -9 \\ \Longrightarrow\quad & \lambda + 2 = \pm 3j \\ \Longrightarrow\quad & \lambda = -2 \pm 3j. \end{aligned} \tag{5.95}$$

We have a pair of values for λ which are a complex conjugate pair. These are fully complex, unlike those in Example 5.5 which were purely imaginary. Given that we have set $y = e^{\lambda t}$, the solution may be written as follows:

$$\begin{aligned} & y = Ae^{(-2+3j)t} + Be^{(-2-3j)t} \\ \Longrightarrow\quad & y = e^{-2t}\left[Ae^{3jt} + Be^{-3jt}\right] \\ \text{or}\quad & y = e^{-2t}\left[C\cos 3t + D\sin 3t\right] \qquad \text{c.f. Ex. 5.5.} \end{aligned} \tag{5.96}$$

Again, we have two forms of the solution, one involving complex exponentials and the other involving sinusoids. In general, should $\lambda = a \pm bj$ then the general solution would be,

$$y = e^{at}\Big[Ae^{bjt} + Be^{-bjt}\Big] \quad \text{or} \quad y = e^{at}\Big[C\cos bt + D\sin bt\Big]. \tag{5.97}$$

Solutions of this form are always associated with damped vibrating systems such as structures and stringed musical instruments. Hence a will always be negative for these applications.

Example 5.7: Solve the equation, $y'' + 4y' + 4y = 0$.

This is the first example of a really special case, one where the auxiliary equation has a repeated root. The auxiliary equation is

$$\begin{aligned} & \lambda^2 + 4\lambda + 4 = 0 \\ \implies \quad & (\lambda + 2)^2 = 0 \\ \implies \quad & \lambda = -2,\ -2. \end{aligned} \tag{5.98}$$

We have a repeated value for λ because the equation for λ has a repeated root. But how do we treat such cases?

Clearly we cannot use just $y = Ae^{-2t}$ because a second order ODE requires two boundary/initial conditions, but we have only one constant. So that guess is no good. However, it is clear that e^{-2t} must play some sort of role because $\lambda = -2, -2$. Perhaps the best thing would be to factor out the e^{-2t} dependence by means of the following substitution.

$$\begin{aligned} \text{Let} \quad & y = e^{-2t} z(t) & \\ \implies \quad & y' = e^{-2t}[z' - 2z] & \text{(product rule and tidying up)} \\ \implies \quad & y'' = e^{-2t}[z'' - 4z' + 4z] & \text{(product rule and more tidying up).} \end{aligned} \tag{5.99}$$

Substitution of this into the ODE gives,

$$\underbrace{e^{-2t}[z'' - 4z' + 4z]}_{y''} + \underbrace{4e^{-2t}[z' - 2z]}_{4y'} + \underbrace{4e^{-2t}[z]}_{4y} = 0. \tag{5.100}$$

All but one of these terms then cancel leaving just,

$$e^{-2t} z'' = 0 \qquad \implies \qquad z'' = 0. \tag{5.101}$$

Two successive integrations using the appropriate constants of integration yield $z = A + Bt$, from which we obtain the final solution,

$$y = (A + Bt)e^{-2t}. \tag{5.102}$$

Note 1: For future reference we are going to interpret the form of this solution in the following way. The constant term, A, corresponds to the first of the two repeated λ-values, while the Bt corresponds to the second λ-value. The extreme usefulness of this interpretation will be seen later in further examples.

Note 2: Although we have demonstrated what happens when $\lambda = -2, -2$, the same form of solution applies for other pairs of repeated λ-values. As a further example, the ODE $y'' + 10y' + 25y = 0$ has the auxiliary equation, $\lambda^2 + 10\lambda + 25 = 0$, for which $\lambda = -5, -5$ is the solution. Hence $y = (A + Bt)e^{-5t}$ is the solution of the ODE.

Note 3: In this example I used the substitution, $y = e^{-2t}z(t)$, solely to determine what solution corresponds to a repeated value of λ. I won't expect this to be done in an exam unless it is asked for specifically. Thus I would expect one to go immediately from the values of λ in Eq.(5.98) to the solution in Eq. (5.102).

Example 5.8: Solve the ODE, $y''' + 3y'' + 3y' + y = 0$.

The auxiliary equation for this ODE is,

$$\lambda^3 + 3\lambda^2 + 3\lambda + 1 = 0 \quad \Longrightarrow \quad (\lambda + 1)^3 = 0 \quad \Longrightarrow \quad \lambda = -1, -1, -1. \tag{5.103}$$

Thus the auxiliary equation has a triple root and we have three instances of $\lambda = -1$. The solution for y in this case is,

$$y = (A + Bt + Ct^2)e^{-t}, \tag{5.104}$$

where the A-term corresponds to the first instance of $\lambda = -1$, the Bt-term to the second instance and the Ct^2-term to the third instance.

Note 1: The correctness of Eq.(5.104) may be confirmed using the substitution, $y = e^{-t}z(t)$, which then yields the ODE, $z''' = 0$, for which its solution is $z = A + Bt + Ct^2$. Do check this.

Note 2: It may now be conjectured that an n-times repeated value of λ, will correspond to a polynomial of order $n - 1$, i.e. it has n terms. This is indeed correct. For example, should we have $(\lambda - 4)^6 = 0$ as the auxiliary equation, then the solution of equivalent ODE would be

$$(A + Bt + Ct^2 + Dt^3 + Et^4 + Ft^5)e^{4t}. \tag{5.105}$$

Example 5.9: Solve the ODE, $y'''' + 4y''' + 4y'' = 0$.

The auxiliary equation in this case is,

$$\lambda^4 + 4\lambda^3 + 4\lambda^2 = 0 \quad \Longrightarrow \quad \lambda^2(\lambda + 2)^2 = 0 \quad \Longrightarrow \quad \lambda = 0, 0, -2, -2. \tag{5.106}$$

This auxiliary equation has two pairs of repeated roots. The solution of the ODE is,

$$y = \underbrace{(A + Bt)e^{0t}}_{\lambda = 0,0} + \underbrace{(C + Dt)e^{-2t}}_{\lambda = -2,-2}, \tag{5.107}$$

where the distinct values of λ, (i.e. 0 and -2) have been treated separately but each has been in the same way as before for twice-repeated roots. In addition, the $\lambda = 0$ components of the solution have been written in the form, e^{0t}, merely to show that $\lambda = 0$ has been used in the solution. Of course, it looks better to write Eq. (5.107) in the form,

$$y = A + Bt + (C + Dt)e^{-2t}. \tag{5.108}$$

Note 1: The main over-riding message from this section so far is that different values of λ (namely, zero, nonzero real and complex) may be treated in exactly the same way via exponentials. In the case of $\lambda = 0$ the corresponding function of t is a constant, and in the case of complex conjugate pairs, the complex exponentials may be replaced by a real exponential multiplied by the appropriate cosine and sine.

Note 2: A secondary but nevertheless important message is that any multiplicities in the values of λ transform into polynomial coefficients of the exponential, and that different values of λ act independently in this regard. Given that I am aware that this Note is quite tricky to understand, its message may be gleaned from the following crazy example.

Example 5.10: If the auxiliary equation for an ODE is $(\lambda+3)^4(\lambda+1)(\lambda+1+2j)^2(\lambda+1-2j)^2\lambda^3=0$, then what is the solution of the equivalent ODE?

The roots of the auxiliary equation are,

$$\lambda = -3, -3, -3, -3, \qquad -1, \qquad -1\pm 2j, -1\pm 2j, \qquad 0, 0, 0, \tag{5.109}$$

and therefore the solution of the equivalent 12$^{\mathrm{th}}$ ODE is,

$$y = \underbrace{(A+Bt+Ct^2+Dt^3)e^{-3t}}_{\lambda=-3,-3,-3,-3} + \underbrace{Ee^{-t}}_{\lambda=-1} + \underbrace{\Big[(F+Gt)\cos 2t + (H+It)\sin 2t\Big]e^{-t}}_{\lambda=-1\pm 2j,-1\pm 2j} + \underbrace{(J+Kt+Lt^2)}_{\lambda=0,0,0}. \tag{5.110}$$

For each different value of λ, namely -3, -1, $-1\pm 2j$ and 0, note how their multiplicities are reflected in the solution, especially for the complex conjugate pair.

Should you be interested the ODE is

$$y^{(12)}+17y^{(11)}+132y^{(10)}+628y^{(9)}+2026y^{(8)}+4750y^{(7)}+7860y^{(6)}+8964y^{(5)}+6345y''''+2025y'''=0, \tag{5.111}$$

and you may be relieved to know that I constructed the ODE from the factorised auxiliary equation, rather than the other way around!

5.6.1 A checklist of examples.

Some examples of how the solution of an ODE is related to the roots of the auxiliary equation and their multiplicity:

Roots	Solution
$2,\ -2$	$Ae^{2t}+Be^{-2t}$
$2,\ 3$	$Ae^{2t}+Be^{3t}$
$2,\ 3,\ 4,\ 5,\ 10$	$Ae^{2t}+Be^{3t}+Ce^{4t}+De^{5t}+Ee^{10t}$
$2,\ 2$	$(A+Bt)e^{2t}$
$2,\ 2,\ 4$	$(A+Bt)e^{2t}+Ce^{4t}$
$2,\ 2,\ 4,\ 4$	$(A+Bt)e^{2t}+(C+Dt)e^{4t}$
$2,\ 2,\ 2,\ 2,\ 4,\ 4,\ 5$	$(A+Bt+Ct^2+Dt^3)e^{2t}+(E+Ft)e^{4t}+Ge^{5t}$
$0,\ 2$	$A+Be^{2t}$
$0,\ 0,\ 0,\ 2$	$(A+Bt+Ct^2)+De^{2t}$
$\pm 2j$	$A\cos 2t+B\sin 2t$
$\pm 2j,\ \pm 5j,\ 2$	$A\cos 2t+B\sin 2t+C\cos 5t+D\sin 5t+Ee^{2t}$
$\pm 2j,\ \pm 2j$	$(A+Bt)\cos 2t+(C+Dt)\sin 2t$
$2\pm 3j$	$e^{2t}(A\cos 3t+B\sin 3t)$
$2\pm 3j,\ 2\pm 4j$	$e^{2t}\left(A\cos 3t+B\sin 3t+C\cos 4t+D\sin 4t\right)$
$2\pm 3j,\ 4\pm 5j$	$e^{2t}\left(A\cos 3t+B\sin 3t\right)+e^{4t}\left(C\cos 5t+D\sin 5t\right)$
$2\pm 3j,\ 2\pm 3j$	$e^{2t}\left[(A+Bt)\cos 3t+(C+Dt)\sin 3t\right]$
$2\pm 3j,\ 2\pm 3j,\ 2\pm 3j$	$e^{2t}\left[(A+Bt+Ct^2)\cos 3t+(D+Et+Ft^2)\sin 3t)\right]$
$2\pm 3j,\ 2\pm 3j,\ 2\pm 3j,\ 2\pm 3j$	$e^{2t}\left[(A+Bt+Ct^2+Dt^3)\cos 3t+(E+Ft+Gt^2+Ht^3)\sin 3t)\right]$

5.6.2 A general note on the effect of the substitution $y = e^{ct}z(t)$

For background information only.

In Example 5.7, above, we considered the ODE, $y'' + 4y' + 4y = 0$, for which $\lambda = -2, -2$. Of course we now know how to write down the solution immediately; it is $y = (A + Bt)e^{-2t}$. This solution was found by using the sustitution, $y = e^{-2t}z(t)$, which yielded $z'' = 0$. The solution for z was found by integrating twice, but if we try to solve it using $z = e^{\sigma t}$ (note that I have used σ here, not λ), then the auxiliary equation for the z-ODE is now $\sigma^2 = 0$. The roots of this auxiliary equation are $\sigma = 0, 0$, and therefore the repeated $\sigma = 0$ yields the solution, $z = A + Bt$.

So we have $\lambda = -2, -2$ as the roots of the auxiliary equation for the y-ODE, and this transforms into $\sigma = 0, 0$ as the roots of the auxiliary equation for the z-ODE when using the substitution, $y = e^{-2t}z(t)$. So the act of stripping out a e^{-2t} factor changes the values of the roots of the auxiliary equation by 2. **This is no accident**, and the purpose of this subsection is to generalise this observation. First, a specific case and then we generalise.

Example 5.11: Solve the ODE, $y'' - 7y' + 12y = 0$. Then use the substitution, $y = e^{t}z(t)$, and solve the resulting ODE for z.

Letting $y = e^{\lambda t}$ in the ODE for y gives $\lambda^2 - 7\lambda + 12 = 0$. Factorisation gives $(\lambda - 3)(\lambda - 4) = 0$, and hence $\lambda = 3, 4$. The solution for y is, therefore,

$$y = Ae^{3t} + Be^{4t}. \tag{5.112}$$

The substitution, $y = e^{t}z(t)$, gives $z'' - 5z' + 6z = 0$ as the ODE for z. The auxiliary equation which follows the substitution $z = e^{\sigma t}$ is $\sigma^2 - 5\sigma + 6 = 0$. Hence $(\sigma - 2)(\sigma - 3) = 0$ which gives $\sigma = 2, 3$, and the solution for z is,

$$z = Ae^{2t} + Be^{3t}. \tag{5.113}$$

So we had $\lambda = 3, 4$ from the equation for y, while the substitution, $y = e^{t}z(t)$, means that $\sigma = 2, 3$ from the equation for z. So the act of factoring out e^{t} from y reduces the value of each of the roots of the auxiliary equation by 1.

Example 5.12: Solve the ODE, $y'' - (a + b)y' + aby = 0$. Then use the substitution, $y = e^{ct}z(t)$, and solve the resulting ODE for z.

I will merely summarize the results of the same process — it is worth checking this if you have time.

The auxiliary equation for y gives $\lambda = a, b$, and hence the solution,

$$y = Ae^{at} + Be^{bt}. \tag{5.114}$$

After substituting $y = e^{ct}z(t)$ into the ODE for y we obtain the ODE,

$$z'' - (a + b - 2c)z' + (a - c)(b - c)z = 0. \tag{5.115}$$

The substitution, $z = e^{\sigma t}$, yields the auxiliary equation the roots for which are, $\sigma = a - c, b - c$. Hence the solution for z is,

$$z = Ae^{(a-c)t} + Be^{(b-c)t}. \tag{5.116}$$

To summarise:

$$y = Ae^{at} + Be^{bt} \qquad \xrightarrow{y=e^{ct}z(t)} \qquad z = Ae^{(a-c)t} + Be^{(b-c)t}. \tag{5.117}$$

So the factoring out of e^{ct} decreases the roots of the auxiliary equation by c. This will also be true for a linear constant-coefficient ODE of any order.

5.7 Solution of inhomogeneous linear, constant coefficient ODEs

Now we turn to the solution of inhomogeneous equations such as Eq. (5.79) for which $F(t)$, the inhomogeneous or forcing term, is nonzero. This ODE is repeated here for convenience:

$$a_n \frac{d^n y}{dt^n} + a_{n-1}\frac{d^{n-1}y}{dt^{n-1}} + \cdots a_1 \frac{dy}{dt} + a_0 y = F(t). \tag{5.118}$$

Analytical progress is assured for those cases where $F(t)$ takes exponential, sinusoidal or polynomial form. Other cases typically require numerical solution.

When solving inhomogeneous ODEs the resulting solution is composed of two parts, the **Complementary Function** and the **Particular Integral**. The Complementary Function (CF) is the full solution of the corresponding homogeneous equation while the Particular Integral (PI) is any solution of the full equation (but it is generally only that part which is intimately associated with the presence of $F(t)$). It is probably best to illustrate the roles of these two components of the full solution with a simple example.

Example 5.13: Solve the ODE, $y' + y = e^{2t}$.
This is solved in two parts:

- (i) Solve $y' + y = 0$. This yields the Complementary Function. This needs to be undertaken first — reasons to follow later.
- (ii) Solve the full equation, $y' + y = e^{2t}$, but focussing solely on the consequence of having a nonzero right hand side. This yields the Particular Integral.

The Complementary Function is the solution of $y' + y = 0$. This follows the ideas of §5.6, so the auxiliary equation is $\lambda + 1 = 0$. Therefore $\lambda = -1$ and hence,

$$y_{\mathrm{cf}} = Ae^{-t}. \tag{5.119}$$

For the Particular Integral we need to solve,

$$y' + y = e^{2t}, \tag{5.120}$$

but we have to concentrate on the contribution of e^{2t} to the solution. First we need to be reminded that, since differentials of e^{at} are proportional to e^{at}, then it makes some sense to let $y_{\mathrm{pi}} = Be^{2t}$ in order to find the value of B. Substitution into Eq. (5.120) yields,

$$\underbrace{2Be^{2t}}_{y'} + \underbrace{Be^{2t}}_{y} = \underbrace{e^{2t}}_{e^{2t}} \qquad \Longrightarrow \qquad 3B = 1 \qquad \Longrightarrow \qquad B = \tfrac{1}{3}. \tag{5.121}$$

So the Particular Integral is,

$$y_{\mathrm{pi}} = \tfrac{1}{3}e^{2t}. \tag{5.122}$$

Given that the type of ODEs which we are solving are linear, we may add together both the Complementary Function and the Particular Integral to obtain the General Solution:

$$y = y_{\mathrm{cf}} + y_{\mathrm{pi}} = Ae^{-t} + \tfrac{1}{3}e^{2t}. \tag{5.123}$$

Note: At this stage of the analysis the Complementary Function always has arbitrary constants as coefficients whereas the Particular Integral doesn't.

If, in addition, we had been given the initial condition, $y(0) = 1$, then it is straightforward to show that $A = \frac{2}{3}$. Hence the final solution would then be,

$$y = \tfrac{2}{3}e^{-t} + \tfrac{1}{3}e^{2t}. \tag{5.124}$$

Example 5.14: Solve the equation, $y'' + 3y' + 2y = e^{at}$. The value, a is an unspecified constant.

The Complementary Function is found by solving $y''+3y'+2y = 0$. The auxiliary equation is $\lambda^2+3\lambda+2 = 0$, and the roots of this are $\lambda = -1, -2$. Hence the CF is,

$$y_{\rm cf} = Ae^{-t} + Be^{-2t}. \tag{5.125}$$

The Particular Integral is found by solving the full ODE using the substitution, $y_{\rm pi} = Ce^{at}$. Hence,

$$Ce^{at}\Big[a^2 + 3a + 2\Big] = e^{at} \qquad \Longrightarrow \qquad C = \frac{1}{a^2 + 3a + 2} = \frac{1}{(a+1)(a+2)}, \tag{5.126}$$

and therefore the PI is,

$$y_{\rm pi} = \frac{e^{at}}{(a+1)(a+2)} \tag{5.127}$$

The general solution is

$$y = y_{\rm cf} + y_{\rm pi} = Ae^{-t} + Be^{-2t} + \frac{e^{at}}{(a+1)(a+2)}. \tag{5.128}$$

Let us consider a few different values of a.

$$\begin{aligned}
a = -3 \quad &\Longrightarrow \quad y'' + 3y' + 2y = e^{-3t} &&\Longrightarrow \quad y = Ae^{-t} + Be^{-2t} + \tfrac{1}{2}e^{-3t} \\
a = 10 \quad &\Longrightarrow \quad y'' + 3y' + 2y = e^{10t} &&\Longrightarrow \quad y = Ae^{-t} + Be^{-2t} + \tfrac{1}{132}e^{10t} \\
a = 0 \quad &\Longrightarrow \quad y'' + 3y' + 2y = 1 &&\Longrightarrow \quad y = Ae^{-t} + Be^{-2t} + \tfrac{1}{2} \\
a = -1.01 \quad &\Longrightarrow \quad y'' + 3y' + 2y = e^{-1.01\,t} &&\Longrightarrow \quad y = Ae^{-t} + Be^{-2t} - \tfrac{10000}{99}e^{-1.01t} \\
a = -0.99 \quad &\Longrightarrow \quad y'' + 3y' + 2y = e^{-0.99t} &&\Longrightarrow \quad y = Ae^{-t} + Be^{-2t} + \tfrac{10000}{101}e^{-0.99t}.
\end{aligned} \tag{5.129}$$

In these cases, and in almost every other case, the solution given in Eq. (5.128) works and is correct. However, there are difficulties when $a = -1$ or $a = -2$ for then the denominator of $y_{\rm pi}$ is zero and the coefficient of e^{at} becomes infinite. We may see the approach to the difficulties which arise when $a = -1$ by observing the numerical coefficient of the PIs in Eq. (5.129) when $a = -1.01$ and $a = -0.99$. These exceptional values of a are the roots of the auxiliary equation which we used to find $y_{\rm cf}$ — this is not an accident.

So when $a = -2$ then the forcing term is e^{-2t}, but this function is identical to one of the components of $y_{\rm cf}$ in (5.125). If we were to choose to think about all of this in terms of λ-values (as in $e^{\lambda t}$), then $y_{\rm cf}$ corresponds to $\lambda = -1, -2$ while the forcing term may be regarded as being *equivalent* to $\lambda = -2$, a second instance of this λ-value. When we encountered this type of situation in Example 5.7, a second instance of $\lambda = -2$ there was shown to lead to a solution of the form, te^{-2t}. The same is true here, but we'll consider a different example first to demonstrate this more convincingly before returning to the present example.

Example 5.15: Solve the equation, $y' + 2y = e^{-2t}$.

For this ODE the auxiliary equation for the Complementary Function is $\lambda + 2 = 0$ and hence $\lambda = -2$. The inhomogeneous term is e^{-2t} which may be said to be equivalent to a second instance of $\lambda = -2$. Therefore this ODE is a prototype of Example 5.14. We will solve this using two different methods.

Method 1: Being a first order linear equation we may solve this ODE using an Integrating Factor. This factor is $e^{\int 2\,dt}$ which is e^{2t}. So we shall multiply the ODE by e^{2t} and eventually obtain the final solution:

$$\begin{aligned}
& & y' + 2y &= e^{-2t} & &\text{the original ODE}\\
&\Longrightarrow & e^{2t}\Big[y' + 2y\Big] &= 1 & &\text{multiplied by } e^{2t}\\
&\Longrightarrow & \Big[e^{2t}\,y\Big]' &= 1 & &\text{LHS is an exact derivative}\\
&\Longrightarrow & e^{2t}\,y &= t + A & &\text{integrating}\\
&\Longrightarrow & y &= \underbrace{Ae^{-2t}}_{\text{CF}} + \underbrace{te^{-2t}}_{\text{PI}} & &
\end{aligned} \tag{5.130}$$

Although this method doesn't use the terms, Complementary Function and Particular Integral, I have indicated which terms are which in terms of the language of the present section. We see that the PI does turn out to be proportional to te^{-2t}, as we guessed it might be.

Given that all the available λ-values are equal to one another, we could also use the substitution, $y = e^{-2t}z(t)$, to obtain, $z' = 1$. This yields $z = t + A$, and hence we obtain the above solution for y.

Method 2: Now let us rerun this Example using a CF/PI approach. So we are solving,

$$\underbrace{y' + 2y}_{\lambda=-2} = \underbrace{e^{-2t}}_{\lambda=-2}, \tag{5.131}$$

where I have indicated the λ-values associated with the left hand side (i.e. the roots of the auxiliary equation) and the right hand side (i.e. the coefficient of t in the exponent).

Given that the CF has $\lambda = -2$ as the root of its auxiliary equation, we may state immediately that $y_{\text{cf}} = Ae^{-2t}$. Given that the equivalent value of λ from the forcing term is $\lambda = -2$, a second instance, then we need to let $y_{\text{pi}} = Bte^{-2t}$ and then find the value of B by substituting it into the full ODE:

$$\begin{aligned}
& & \underbrace{Be^{-2t}(1-2t)}_{y_{\text{pi}}'} + \underbrace{2Bte^{-2t}}_{2y_{\text{pi}}} &= e^{-2t} & &\\
&\Longrightarrow & Be^{-2t} &= e^{-2t} & &\text{all the terms involving } te^{-2t} \text{ cancel}\\
&\Longrightarrow & B &= 1. & &
\end{aligned} \tag{5.132}$$

Hence $y_{\text{pi}} = te^{-2t}$ and therefore we recover the solution given in Eq. (5.130).

Note: This latest example suggests that the repetition of a λ-value is treated in exactly the same way as in the last section, namely that increasing powers of t appear depending on how many repetitions there are. Clearly we haven't got a general proof of this, but the following few examples will show this in action.

Example 5.16: Solve the equation, $y'' + 3y' + 2y = e^{-2t}$. This is the $a = -2$ instance of Example 1.14.

Guided by Example 5.15, we'll write out the ODE and classify it according its λ-values. We have,

$$\underbrace{y'' + 3y' + 2y}_{\lambda=-1,-2} = \underbrace{e^{-2t}}_{\lambda=-2}, \tag{5.133}$$

The Complementary Function is $y_{\text{cf}} = Ae^{-t} + Be^{-2t}$, as found in Example 5.14. With regard to the Particular Integral, the substitution we need is now determined by the fact that the λ-value corresponding to the forcing term is the second instance of $\lambda = -2$, and therefore we have to set $y_{\mathrm{pi}} = Cte^{-2t}$. We get,

$$\begin{aligned} & y'' + 3y' + 2y = e^{-2t} \\ \Longrightarrow \quad & Ce^{-2t}\Big[\underbrace{-4 + \cancel{4t}}_{y''} + \underbrace{3(1 - \cancel{2t})}_{3y'} + \underbrace{\cancel{2t}}_{2y} \Big] = e^{-2t} \\ \Longrightarrow \quad & -Ce^{-2t} = e^{-2t} \quad \Longrightarrow \quad C = -1. \end{aligned} \tag{5.134}$$

Hence $y_{\mathrm{pi}} = -te^{-2t}$. The general solution is,

$$y = y_{\mathrm{cf}} + y_{\mathrm{pi}} = Ae^{-t} + Be^{-2t} - te^{-2t}. \tag{5.135}$$

If we had chosen to solve $y'' + 3y' + 2y = e^{-t}$, then the forcing term represents a second instance of $\lambda = -1$ and therefore we would need to use $y_{\mathrm{pi}} = Cte^{-t}$. In this case it is worth checking that $C = 1$ is the correct value. The general solution in this case is,

$$y = y_{\mathrm{cf}} + y_{\mathrm{pi}} = Ae^{-t} + Be^{-2t} + te^{-t}. \tag{5.136}$$

As a further twist on this problem, if we had wished to solve $y'' + 3y' + 2y = ae^{-t} + be^{-2t}$ then each of the forcing terms should be considered separately, and then the general solution may be found to be,

$$y = y_{\mathrm{cf}} + y_{\mathrm{pi}} = Ae^{-t} + Be^{-2t} + ate^{-t} - bte^{-2t}. \tag{5.137}$$

Example 5.17: Solve the ODE, $y'' + 3y' + 2y = te^{-2t}$.

The left hand side of this ODE is the same as for Examples 5.14 and 5.16, and therefore the Complementary Function is the same. We note too that the λ-values forming the roots of the auxiliary equation are $\lambda = -1, -2$. But what should we make of the present forcing term?

In Example 5.7 we interpreted two instances of $\lambda = -2$ as being equivalent to e^{-2t} (for the first $\lambda = -2$) and te^{-2t} (for the second $\lambda = -2$). In the present Example we shall do the same, but the equivalence will be taken in the opposite direction. Therefore we shall adopt the point of view that the presence of te^{-2t} as the forcing term is equivalent to having $\lambda = -2, -2$. To illustrate this we may write the ODE with suitable labelling:

$$\underbrace{y'' + 3y' + 2y}_{\lambda=-1,-2} = \underbrace{te^{-2t}}_{\lambda=-2,-2}. \tag{5.138}$$

As before the Complementary Function is given by $y_{\rm cf} = Ae^{-t} + Be^{-2t}$. Given that the forcing term is now to be regarded as the second and third instances of $\lambda = -2$, we need to set $y_{\rm pi} = (Ct + Dt^2)e^{-2t}$. My belief is that, if this is understood, then nothing else in this topic holds any fears apart from the length of the algebra.

I will omit much of the algebra for this example, but eventually we get to,

$$\begin{aligned} y'' + 3y' &+2y \\ &= \Big[\Big(3C + 2D - 4C\Big) + t\Big(\cancel{2C} + 6D - \cancel{6C} - 4D + \cancel{4C} - 4D\Big) + t^2\Big(\cancel{2D} - \cancel{6D} + \cancel{4D}\Big)\Big]\, e^{-2t} \\ &= (2D - C - 2Dt)e^{-2t} \quad = \quad te^{-2t}. \end{aligned} \tag{5.139}$$

It is much to be recommended that the analysis leading to Eq. (5.139) is checked. So we see that all the t^2 terms cancel, and those involving C for the t-coefficents have also cancelled. This serves as a check that one's algebra is correct! This final right hand side should now be equal to the original forcing term, te^{-2t}. Hence $2D = -1$ (matching the coefficients of t) and $2D - C = 0$ (matching the constants). Hence $D = -\frac{1}{2}$ and $C = -1$. The Particular Integral is $y_{\rm pi} = (-t + \frac{1}{2}t^2)e^{-2t}$, and therefore the general solution is,

$$y = y_{\rm cf} + y_{\rm cf} = Ae^{-t} + Be^{-2t} + (-t - \tfrac{1}{2}t^2)e^{-2t}. \tag{5.140}$$

Example 5.18: Solve the ODE, $y''' + 5y'' + 8y' + 4y = te^{-2t}$.

I have contrived this example so that the auxiliary equation takes the form, $(\lambda+1)(\lambda+2)^2 = 0$, so that we have $\lambda = -1, -2, -2$. From this I reconstructed the left hand side of the ODE that is seen above, and then included the te^{-2t} on the right hand side. With labels, the ODE is,

$$\underbrace{y''' + 5y'' + 8y' + 4y}_{\lambda=-1,-2,-2} = \underbrace{te^{-2t}}_{\lambda=-2,-2}. \tag{5.141}$$

So we have two instances of $\lambda = -2$ in the auxiliary equation and the equivalent of another two from the forcing term. Hopefully it is now not too surprising that we shall take,

$$y_{\rm cf} = Ae^{-t} + (B + Ct)e^{-2t} \qquad \text{and} \qquad y_{\rm pi} = (Dt^2 + Et^3)e^{-2t}. \tag{5.142}$$

Here the values, A, B and C are arbitrary, while D and E may be found by substitution into the full ODE. These values turn out to be, $D = -\frac{1}{2}$ and $E = -\frac{1}{6}$, and therefore the full solution is,

$$y = y_{\rm cf} + y_{\rm pi} = Ae^{-t} + (B + Ct)e^{-2t} + (-\tfrac{1}{2}t^2 - \tfrac{1}{6}t^3)e^{-2t}. \tag{5.143}$$

The amount of algebra that is required to find the Particular Integral is quite large, but there is an easier route for this ODE. Given that there are so many instances of $\lambda = -2$, we may factor an e^{-2t} out using $y = e^{-2t}z(t)$ and this yields,

$$z''' - z'' = t. \tag{5.144}$$

Later we will find out how to deal with powers of t on the right hand side. The solution for z is quicker to obtain than the one for y. We will return to this ODE later as Example 5.26.

Example 5.19: Find the solutions of $y''' + 3y'' + 3y' + y = t^5 e^{-t}$.

Yes, this is a seriously extreme example, but the CF is $y_{\mathrm{cf}} = (A+Bt+Ct^2)e^{-t}$, given that $\lambda = -1, -1, -1$ from the auxiliary equation (see Example 5.8). The constants, A, B and C are arbitrary.

Given the t^5 multiplying the e^{-t} on the right hand side of the ODE, we have the equivalent of six further repetitions of $\lambda = -1$. So we may label the ODE as follows,

$$\underbrace{y''' + 3y'' + 3y' + y}_{\lambda = -1, -1, -1} \quad = \quad \underbrace{t^5 e^{-t}}_{\lambda = -1 \text{ six times.}} \tag{5.145}$$

Note that e^{-t} is equivalent to one instance of $\lambda = -1$, te^{-t} to two, t^2e^{-t} to three and so on. Therefore we need to substitute

$$y_{\mathrm{pi}} = (Dt^3 + Et^4 + Ft^5 + Gt^6 + Ht^7 + Jt^8)e^{-t} \tag{5.146}$$

in order to find the Particular Integral.

For this very extreme case, and upon noting that the only value that λ takes is -1, then we may solve the whole problem in one go by substituting $y = z(t)e^{-t}$. This shifts all the λ values from -1 to 0; see Example 5.12. The ODE transforms to

$$\begin{aligned} & z''' = t^5 \\ \Longrightarrow \quad & z = A + Bt + Ct^2 + \tfrac{1}{336}t^8 \qquad \text{using three integrations} \qquad (5.147) \\ \Longrightarrow \quad & y = (A + Bt + Ct^2 + \tfrac{1}{336}t^8)e^{-t}. \end{aligned}$$

Therefore the constants introduced in Eq. (5.146) are

$$D = E = F = G = H = 0, \quad J = \tfrac{1}{336}. \tag{5.148}$$

A summary. OK, we need to pause briefly here to take stock of what has been achieved with all of these Examples. In §5.6 and in the present section so far I have attempted to introduce a unified way of determining from the ODE what forms are taken by the Complementary Function and by the Particular Integral. These are intimately associated with the λ-values. Hopefully, it has become clear that the form the Particular Integral takes depends on what has happened with the Complementary Function. This is why I have been focussed so strongly on the values of λ, as we have defined them here, and on their multiplicity. This is also why the Complementary Function *must* be found first.

As an attempt to describe this *unified way* one could say that repeated values of λ involve the use of increasing powers of t multiplying the associated function, $e^{\lambda t}$, the power increasing by 1 for every subsequent repetition. This happens for both the Complementary Function and the Particular Integral, but the counting must start with the auxiliary equation for the Complementary Function.

Here is a Table of examples of "what to do when" when faced with an ODE where all the λ-values are real quantities. I will take the value, $\lambda = 2$, to be the one which tends to be repeated, although the very last instance in the Table is slightly different. As always with such ODEs, the constants in the Complementary Function are arbitrary, whereas those of the Particular Integral need to be found.

ODE	λ(CF)	λ(PI)	CF	PI
$y' - 3y = e^{2t}$	3	2	Ae^{3t}	Be^{2t}
$y' - 2y = e^{3t}$	2	3	Ae^{2t}	Be^{3t}
$y' - 2y = e^{2t}$	2	2	Ae^{2t}	Bte^{2t}
$y' - 2y = te^{2t}$	2	$2, 2$	Ae^{2t}	$(Bt + Ct^2)e^{2t}$
$y' - 2y = t^2e^{2t}$	2	$2, 2, 2$	Ae^{2t}	$(Bt + Ct^2 + Dt^3)e^{2t}$
$y'' - 4y' + 3y = e^{2t}$	$1,\ 3$	2	$Ae^t + Be^{3t}$	Ce^{2t}
$y'' - 3y' + 2y = e^{3t}$	$1,\ 2$	3	$Ae^t + Be^{2t}$	Ce^{3t}
$y'' - 3y' + 2y = e^{2t}$	$1,\ 2$	2	$Ae^t + Be^{2t}$	Cte^{2t}
$y'' - 3y' + 2y = te^{2t}$	$1,\ 2$	$2,\ 2$	$Ae^t + Be^{2t}$	$(Ct + Dt^2)e^{2t}$
$y'' - 3y' + 2y = t^2e^{2t}$	$1,\ 2$	$2,\ 2$	$Ae^t + Be^{2t}$	$(Ct + Dt^2 + Et^3)e^{2t}$
$y'' - 4y' + 4y = e^{3t}$	$2,\ 2$	3	$(A + Bt)e^{2t}$	Ce^{3t}
$y'' - 4y' + 4y = e^{2t}$	$2,\ 2$	2	$(A + Bt)e^{2t}$	Ct^2e^{2t}
$y'' - 4y' + 4y = te^{2t}$	$2,\ 2$	$2, 2$	$(A + Bt)e^{2t}$	$(Ct^2 + Dt^3)e^{2t}$
$y'' - 4y' + 4y = t^2e^{2t}$	$2,\ 2$	$2, 2, 2$	$(A + Bt)e^{2t}$	$(Ct^2 + Dt^3 + Et^4)e^{2t}$
$y''' - 6y'' + 12y' - 8y = e^{3t}$	$2,\ 2,\ 2$	3	$(A + Bt + Ct^2)e^{2t}$	De^{3t}
$y''' - 6y'' + 12y' - 8y = e^{2t}$	$2,\ 2,\ 2$	2	$(A + Bt + Ct^2)e^{2t}$	Dt^3e^{2t}
$y''' - 6y'' + 12y' - 8y = te^{2t}$	$2,\ 2,\ 2$	$2, 2$	$(A + Bt + Ct^2)e^{2t}$	$(Dt^3 + Et^4)e^{2t}$
$y''' - 6y'' + 12y' - 8y = t^2e^{2t}$	$2,\ 2,\ 2$	$2, 2, 2$	$(A + Bt + Ct^2)e^{2t}$	$(Dt^3 + Et^4 + Ft^5)e^{2t}$
$y''' - 3y' - 2y = te^{2t}$	$-1,\ -1,\ 2$	$2, 2$	$(A + Bt)e^{-t} + Ce^{2t}$	$(Dt + Et^2)e^{2t}$

The next few Examples will consider how to deal with sinusoidal forcing terms.

Example 5.20: Solve the ODE, $y'' + 3y' + 2y = \cos bt$.

With regard to the Complementary Function things are straightforward. The auxiliary equation is $\lambda^2+3\lambda+2 = 0$ and hence $\lambda = -1, -2$. So the Complementary Function is $y_{\rm cf} = Ae^{-t} + Be^{-2t}$.

For the Particular Integral, one *could* argue as follows. If we were to set $y_{\rm pi}$ to be proportional to $\cos bt$, then the y' term in the ODE yields a sine and it means that our initial substitution is incorrect. So we cannot use just a cosine as the substitution. Therefore we need to use $y_{\rm pi} = C\cos bt + D\sin bt$ as the substitution. While this form for $y_{\rm pi}$ is indeed correct for this Example, in general we need to make sure that there are no repeated λ-values. We know from ME10304, Maths 1, that

$$\cos bt = \tfrac{1}{2}(e^{bjt} + e^{-bjt}). \tag{5.149}$$

Hence the right hand side of the ODE is equivalent to $\lambda = \pm bj$. Restating the ODE with labelling we have,

$$\underbrace{y'' + 3y' + 2y}_{\lambda=-1,-2} = \underbrace{\cos bt}_{\lambda=\pm bj}, \tag{5.150}$$

and hence there are no repeated factors.

We shall solve this ODE in two different ways. The first is to use the substitution, $y_{\rm pi} = C\cos bt + D\sin bt$, while the second will replace $\cos bt$ with e^{bjt} (which means that we have added an imaginary component to the forcing term) and then we take the real part of the resulting complex Particular Integral.

Method 1: The substitution of $y_{\rm pi} = C\cos bt + D\sin bt$ into Eq. (5.150) yields,

$$\underbrace{\Big[-b^2C\cos bt - b^2D\sin bt\Big]}_{y''} + \underbrace{3\Big[-bC\sin bt + bD\cos bt\Big]}_{3y'} + \underbrace{2\Big[C\cos bt + D\sin bt\Big]}_{2y} = \cos bt. \tag{5.151}$$

Now we collect like terms:

$$\begin{aligned} \cos bt \text{ terms:}\quad & (2-b^2)C + 3bD = 1, \\ \sin bt \text{ terms:}\quad & -3bC + (2-b^2)D = 0. \end{aligned} \tag{5.152}$$

This pair of simultaneous equations may be solved in the usual way to obtain,

$$C = \frac{2-b^2}{b^4+5b^2+4}, \qquad D = \frac{3b}{b^4+5b^2+4}, \tag{5.153}$$

and hence the PI is
$$y_{\rm pi} = \frac{(2-b^2)\cos bt + 3b\sin bt}{b^4+5b^2+4}. \tag{5.154}$$

An alternative way of solving Eqs. (5.152) is to recast them in matrix/vector form:

$$\begin{pmatrix} (2-b^2) & 3b \\ -3b & (2-b^2) \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{5.155}$$

Using the formula for the inverse matrix (see later in these notes) we have,

$$\begin{pmatrix} C \\ D \end{pmatrix} = \frac{1}{b^4+5b^2+4} \begin{pmatrix} (2-b^2) & -3b \\ 3b & (2-b^2) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{b^4+5b^2+4} \begin{pmatrix} (2-b^2) \\ 3b \end{pmatrix}, \tag{5.156}$$

which is identical to Eq. (5.153).

Note: For those who haven't studied matrices yet, this will all make sense by the end of the semester!

Method 2. This proceeds by replacing the forcing term, $\cos bt$, by e^{bjt}. So we have added the imaginary term, $j\sin bt$, to the right hand side. This subterfuge makes it much easier to find the Particular Integral, at least in terms of e^{bjt}, although we will then need to find the real part of this version of the Particular Integral to obtain the one that is needed. So we shall solve,

$$y'' + 3y' + 2y = e^{bjt} \tag{5.157}$$

using the substitution, $y_{\rm pi} = Ce^{bjt}$. This yields,

$$Ce^{bjt}\Big[-b^2 + 3bj + 2\Big] = e^{bjt} \qquad \Longrightarrow \qquad C = \frac{1}{2 - b^2 + 3bj}. \tag{5.158}$$

Hence the Particular Integral is,

$$y_{\rm pi} = \frac{e^{bjt}}{2 - b^2 + 3bj}. \tag{5.159}$$

The real part of this expression is the solution of $y'' + 3y' + 2y = \cos bt$, while the imaginary part is the solution of $y'' + 3y' + 2y = \sin bt$. We need the real part here, so

$$\frac{e^{bjt}}{2 - b^2 + 3bj} = \frac{\cos bt + j\sin bt}{2 - b^2 + 3bj} \qquad \text{expanding the complex exponential}$$

$$= \frac{(\cos bt + j\sin bt)(2 - b^2 - 3bj)}{(2 - b^2 + 3bj)(2 - b^2 - 3bj)} \qquad \text{using complex conjugates}$$

$$= \frac{(2 - b^2)\cos bt + 3b\sin bt + j(-3b\cos bt + (2 - b^2)\sin bt)}{(2 - b^2)^2 + 9b^2} \qquad \text{multiplying out}$$

$$= \left[\frac{(2 - b^2)\cos bt + 3b\sin bt}{b^4 + 5b^2 + 4}\right] + j\left[\frac{-3b\cos bt + (2 - b^2)\sin bt}{b^4 + 5b^2 + 4}\right]. \tag{5.160}$$

The real part of this final expression is what we were aiming for, although the imaginary part is an added bonus, namely the solution corresponding to having $\sin bt$ as the forcing term.

Note: This rather lengthy Example shows that you have a choice of methods for solving equations with sinusoidal forcing terms. One requires the solution of simultaneous equations (a task that is made a little quicker by using the matrix/vector variant), while other dives off into complex numbers. It is worth practicing both ways a few times to see which you find to be quicker and more reliable.

Note: There are some interesting comments that may be made about the physical meaning of this example. First, this is an over-damped system because the Complementary Function consists solely of decaying exponentials. This behaviour is what happens with many types of door restraints, such as those in 4 East! So the amplitude would naturally rather than to decay in an oscillatory manner.

This damped system is being perturbed by $\cos bt$ and the PI is of most interest because it is what remains when the transient, the CF, has decayed. When the perturbations have a very low frequency then $b \ll 1$. We may therefore write the PI in the form,

$$y_{\rm pi} = \frac{(2 - b^2)\cos bt + 3b\sin bt}{b^4 + 5b^2 + 4} \simeq \tfrac{1}{2}\cos bt,$$

where the greyed-out terms are almost negligible compared with the terms that have remained black. Physically, the velocity and acceleration are negligible and we obtain an in-phase response.

Very fast perturbations correspond to $b \gg 1$. Hence,

$$y_{\rm pi} = \frac{(2 - b^2)\cos bt + 3b\sin bt}{b^4 + 5b^2 + 4} \simeq -\frac{\cos bt}{b^2}.$$

This response corresponds to y'' dominating the left hand side of the ODE (i.e. y and y' are negligible). We have obtained a very small amplitude and an out-of-phase response.

Both of these responses may be confirmed using a mass attached to the end of a long elastic string/band. When the string is moved up and down with a very low frequency then the mass follows passively, but when it is jiggled up and down at a high frequency then the mass hardly moves at all but the movement that it does make is out of phase with your hand.

The next three examples will bring us slowly to a point where we will be able to deal with repeated complex values of λ in the context of inhomogeneous ODEs.

Example 5.21: Solve the ODE, $y'' + 9y = \cos bt$, where b is an unspecified constant.

This equation represents the effect of a periodic forcing on an undamped mass/spring system. The auxiliary equation for the Complementary Function is $\lambda^2 + 9 = 0$ and hence $\lambda = \pm 3j$. The ODE, with labelling, is,

$$\underbrace{y'' + 9y}_{\lambda = \pm 3j} = \underbrace{\cos bt}_{\lambda = \pm bj}, \tag{5.161}$$

and therefore we do not have any repeated values of λ unless $b = 3$. In this example we shall assume, therefore, that $b \neq 3$, and then we'll treat the special case, $b = 3$, in the next example.

Given that $\lambda = \pm 3j$, the Complementary Function is

$$y_{\text{cf}} = A\cos 3t + B\sin 3t. \tag{5.162}$$

The arguments used in Example 5.20 about how to choose the form of the Particular Integral also apply here, and therefore we could let $y_{\text{pi}} = C\cos bt + D\sin bt$. In the present case, the ODE doesn't have a first derivative, and therefore we could let $y_{\text{pi}} = C\cos bt$ without any problem. This is because the equations for C and D decouple. Nevertheless, we shall use the full substitution in order to illustrate these comments. Equation (5.161) yields,

$$(-b^2 + 9)C\cos bt + (-b^2 + 9)D\sin bt = \cos bt, \tag{5.163}$$

and therefore we obtain,

$$\begin{array}{ll} \cos bt \text{ terms:} & (-b^2+9)C = 1 \\ \sin bt \text{ terms:} & (-b^2+9)D = 0. \end{array} \tag{5.164}$$

So the equations for C and D have decoupled, and therefore $C = 1/(9 - b^2)$ and $D = 0$. The Particular Integral is,

$$y_{\text{pi}} = \frac{\cos bt}{9 - b^2}, \tag{5.165}$$

and the general solution is,

$$y = y_{\text{cf}} + y_{\text{pi}} = A\cos 3t + B\sin 3t + \frac{\cos bt}{9 - b^2}. \tag{5.166}$$

This solution is valid when $b \neq 3$ but the amplitude of the Particular Integral becomes infinite as $b \to 3$. An alternative solution is required when $b = 3$, and this is considered in the next Example.

Example 5.22: Solve the ODE, $y'' + 9y = \cos 3t$.

The auxiliary equation has roots, $\lambda = \pm 3j$, and the forcing term may be regarded as being equivalent to $\lambda = \pm 3j$. Therefore both $\lambda = 3j$ and $\lambda = -3j$ are repeated. Knowing how repeated values of λ are dealt with when λ is real, this means that we could write,

$$y_{\text{cf}} = Ae^{3jt} + Be^{-3jt}, \qquad y_{\text{pi}} = t\Big[Ce^{3jt} + De^{-3jt}\Big]. \tag{5.167}$$

Using and extending the results of Example 5.5 means that we may rewrite this solution in the form,

$$y_{\text{cf}} = A\cos 3t + B\sin 3t, \qquad y_{\text{pi}} = t\Big[C\cos 3t + D\sin 3t\Big], \tag{5.168}$$

where the values of A, B, C and D in Eq. (5.167) are not the same as the ones in Eq. (5.168). The general solution is

$$y = y_{\text{cf}} + y_{\text{pi}} = A\cos 3t + B\sin 3t + t\Big[C\cos 3t + D\sin 3t\Big], \tag{5.169}$$

and substitution of this into the ODE yields $C = 0$ and $D = \frac{1}{6}$. Hence the solution is,

$$y = y_{\text{cf}} + y_{\text{pi}} = A\cos 3t + B\sin 3t + \tfrac{1}{6}t\sin 3t. \tag{5.170}$$

Note: that the Particular Integral here has an amplitude which grows linearly in time. This is quite typical of undamped systems which are perturbed at one of its resonant frequencies. Generally, structures tend to be at least lightly damped and therefore this growth eventually attentuates leaving behind what could still be a rather large response but at least it doesn't continue to grow unboundedly.

Example 5.23: Solve the ODE, $y'' + 6y' + 25y = te^{-3t}\cos 4t$.

This is a rather strange example due to the form of the forcing term, but first we need to consider the roots of the auxiliary equation before making sense of that forcing term. The auxiliary equation is,

$$\lambda^2 + 6\lambda + 25 = 0 \qquad \Longrightarrow \qquad (\lambda+3)^2 + 16 = 0 \qquad \Longrightarrow \qquad \lambda = -3 \pm 4j, \tag{5.171}$$

and hence

$$y_{\text{cf}} = e^{-3t}\Big[A\cos 4t + B\sin 4t\Big] \tag{5.172}$$

is the Complementary Function. Now we see that the forcing term in the ODE is equivalent to a second and a third instance of $\lambda = -3 \pm 4j$. Repeating the ODE with labels, we have,

$$\underbrace{y'' + 6y' + 25y}_{\lambda = -3 \pm 4j} \qquad = \qquad \underbrace{te^{-3t}\cos 4t.}_{\lambda = -3 \pm 4j,\, -3 \pm 4j} \tag{5.173}$$

So the Particular Integral takes the form,

$$y_{\text{pi}} = e^{-3t}\Big[(Ct + Dt^2)\cos 4t + (Et + Ft^2)\sin 4t\Big]. \tag{5.174}$$

I have to admit that I haven't determined what values C, D, E and F take, but it is possible to find them if you have a day or two free.

Finally, we turn to having polynomials as forcing functions. Although we haven't yet considered these explicitly, the process of determining the Particular Integral is no different from when we have nonzero real values of λ. All one has to do is to keep in mind that there is a ghostly e^{0t} present even if it is not written down. I will offer three examples of this.

Example 5.24: Solve $y' + 3y = 6$.

If we label this ODE with the appropriate values of λ, then we have

$$\underbrace{y' + 3y}_{\lambda=-3} \quad = \quad \underbrace{6}_{\lambda=0} \quad . \tag{5.175}$$

These values of λ are different and hence we may write $y_{\rm cf} = Ae^{-3t}$ and, to find the Particular Integral, we let $y_{\rm pi} = Be^{0t} = B$. Substitution into the ODE yields $B = 2$, and hence the general solution is,

$$y = y_{\rm cf} + y_{\rm pi} = Ae^{-3t} + 2. \tag{5.176}$$

A simple initial condition such as, $y(0) = 1$, yields $A = -1$, and hence the final solution is $y = 2 - e^{-3t}$.

Note that we have already seen the solution of a second order ODE with a constant forcing term in the $a = 0$ cases in Eq. (5.129).

Example 5.25: Solve $y' + 3y = t^2$.

In this example the forcing term is equivalent to three instances of $\lambda = 0$, i.e. we may label the equation as follows:

$$\underbrace{y' + 3y}_{\lambda=-3} \quad = \quad \underbrace{t^2}_{\lambda=0,0,0} \quad . \tag{5.177}$$

The Complementary Function is the same as for Example 5.24, but in view of the triple instance of $\lambda = 0$, we need to let $y_{\rm pi} = B + Ct + Dt^2$. After substitution, we obtain, $D = \frac{1}{3}$, $C = -\frac{2}{9}$ and $B = \frac{2}{27}$, in turn. Hence the general solution is,

$$y = y_{\rm cf} + y_{\rm pi} = Ae^{-3t} + \tfrac{2}{27} - \tfrac{2}{9}t + \tfrac{1}{3}t^2. \tag{5.178}$$

Example 5.26: Solve $z''' - z'' = t$.

This was the equation we obtained at the end of Example 5.18 and now we have the tools to solve it. Let us rewrite the ODE with labelling:

$$\underbrace{z''' - z''}_{\lambda=1,0,0} \quad = \quad \underbrace{t}_{\lambda=0,0} \quad , \tag{5.179}$$

and hence we will write,

$$z = z_{\rm cf} + z_{\rm pi} = \underbrace{Ae^t + B + Ct}_{\text{CF}} + \underbrace{Dt^2 + Et^3}_{\text{PI}}. \tag{5.180}$$

Substitution into the ODE yields, $D = -\frac{1}{2}$ and $E = -\frac{1}{6}$, and hence the general solution is,

$$z = z_{\rm cf} + z_{\rm pi} = Ae^t + B + Ct - \tfrac{1}{2}t^2 - \tfrac{1}{6}t^3. \tag{5.181}$$

This final solution is consistent with the one obtained in Example 5.18.

Final remarks

Throughout the whole of these two sections on homogeneous and inhomogeneous ODEs I have striven to emphasize a common approach to finding the forms for both the Complementary Function and the Particular Integral. This approach involves the identification of all of the λ-values which are associated with the CF and the PI and also their multiplicity. Imaginary or complex values are usually interpreted in terms of the appropriate sines and cosines, while zero values involve powers of t beginning with the zeroth power, i.e. 1, but otherwise everything works within this uniform approach.

Therefore I hope that it is possible in all situations to identify both y_{cf} and y_{pi}, with the latter being decided upon *after* the former. Thereafter it is a matter of determining the unknown coefficients which are associated with the PI, and this algebraic tedium usually takes more than half of the time needed for the full solution. Should initial and/or boundary conditions be given, then yet more tedium is involved in computing the formerly arbitrary constants in the Complementary Function.

Post Script

Those who have studied ODEs before will notice the omission of Bernoulli's equation and of equidimensional equations. These appear in the problem sheets.

A large omission here is the solution of systems of linear constant-coefficient ODEs. Well, the ODE notes so far cover five lectures' worth of content and I think we need to give it a rest for a moment. Although we will touch briefly on the solution of ODEs and systems of ODEs in the Laplace Transform section, we'll do a much more general analysis in the Matrices section after Christmas using eigenvalues and eigenvectors.

6 LAPLACE TRANSFORMS

6.1 Motivation

This is the first topic this academic year where I am pretty sure that only a small handful at most will have met it before and so some motivation is needed. So what are Laplace Transforms used for? Here's a short list:

1. to solve linear constant-coefficient ODEs;
2. to "translate" electrical/hydraulic circuits into a form which is the *equivalent* of an ODE or system of ODEs;
3. to solve feedback sytems;
4. to solve some of the Partial Differential Equations which arise in fluid machanics and heat transfer.

In this unit we will definitely be covering item (1) because of the present emphasis on ODEs, and we will also be developing the background theory behind item (2).

Items (2) and (3) form part of the Systems and Control unit next year.

Item (4) isn't covered using Laplace Transforms in this unit, but it can be done.

The chief objectives in these four lectures are to acquire some facility in the use the Laplace Transform itself, to acquire some of the terminology that is used in Control Theory (not that we will be doing Control Theory) and to use Laplace Transforms to solve some ODEs. You'll also meet some weird animals in the Laplace Transform zoo, namely the unit impulse and the unit step function.

6.2 What is the Laplace Transform?

The Laplace Transform is an integral which changes a function of time, $y(t)$, say, into $Y(s)$, a function of s. The variable, s, is known as the **Laplace Transform variable.** It may be described informally as being equivalent to a time derivative with a hint of initial condition! Sounds mad, but you'll see why later.

If one has a constant-coefficient ODE for $y(t)$, then the first step in its solution via Laplace Transforms is to **take the Laplace Transform** of that equation. This process results in an algebraic system for $Y(s)$, the transform of the dependent variable, which is then solved. Finally the process of taking the Inverse Laplace Transform takes place; this results in a function of time which, somewhat magically, is the desired solution of the original equation. Diagrammatically, we could summarise this process as follows:

ODE for $y(t)$ $\longrightarrow$ algebraic equation for $Y(s)$ $\longrightarrow$ solve for $Y(s)$ $\longrightarrow$ solution for $y(t)$.

The final step, namely the recovery of $y(t)$ from $Y(s)$, is known as **taking the Inverse Laplace Transform**.

If you were studying this topic in a Mathematics Department, then taking of the inverse Laplace Transform would require the use of a contour integral in the complex plane called the Bromwich Contour Integral. This certainly does need to be done for quite complicated types of $Y(s)$, but these are ones which we won't meet, fortunately. Instead, we will be creating a toolchest of results that may be used to invert the sorts of $Y(s)$ which arise when solving linear constant-coefficient ODEs. So no worries then.

The Laplace Transform of the function $y(t)$ is defined in the following way,

$$\boxed{\mathcal{L}[y(t)] \quad = \quad \int_0^\infty y(t)e^{-st}\,dt \quad = \quad Y(s).} \tag{6.1}$$

Note that the symbol, $\mathcal{L}[\ \ ]$, merely means **'Laplace Transform of'** whatever happens to be in the square bracket, and this is how it is stated when talking about it. Thus the Laplace Transform process changes the function $y(t)$ into a new function $Y(s)$.

Once one has $Y(s)$, the corresponding $y(t)$ is found using,

$$\mathcal{L}^{-1}[Y(s)] \quad = \quad y(t). \tag{6.2}$$

Here the symbol, $\mathcal{L}^{-1}[\ \ ]$, means the **'Inverse Laplace Transform of'** whatever happens to be in the square bracket. At this stage I will say no more, but this concept is much easier than you might fear it will be.

6.3 Some examples of Laplace Transforms

Example 6.1: Find $\mathcal{L}[1]$.

This may be done simply by writing down the definition of the Laplace Transform and performing the integration:

$$\mathcal{L}[1] = \int_0^\infty 1\, e^{-st}\, dt = \Big[-\frac{e^{-st}}{s}\Big]_0^\infty = \frac{1}{s}. \tag{6.3}$$

In this integration it is important to note that the answer is correct only when $s > 0$. If s had been negative then the integral would be infinite. Thus $s > 0$ is an existence condition for the transform. All Laplace Transforms must have a range of values of s for which the integral exists. That being said, it is not often that one needs to consider this aspect of Laplace Transforms for the functions which we will be transforming.

Example 6.2: Find $\mathcal{L}[e^{-at}]$.

Again, by definition, we have

$$\mathcal{L}[e^{-at}] = \int_0^\infty e^{-at}\, e^{-st}\, dt = \int_0^\infty e^{-(s+a)t}\, dt = \frac{1}{s+a}. \tag{6.4}$$

This Laplace Transform exists when $s + a > 0$, i.e. when $s > -a$. When $a = 0$ we recover the result of Example 6.1. Other examples are when $a = 1$ and $a = -2$ we get

$$\mathcal{L}[e^{-t}] = \frac{1}{s+1} \qquad \text{and} \qquad \mathcal{L}[e^{2t}] = \frac{1}{s-2}. \tag{6.5}$$

Hopefully it is really easy to see that $\mathcal{L}[Ae^{-at}] = A/(s+a)$ so that multiplication of a function by a constant is equivalent to multiplication of its Laplace Transform by the same constant. This will be used freely with no further comment.

Example 6.3: Find $\mathcal{L}[\cos at]$.

By definition we have,

$$\mathcal{L}[\cos at] = \int_0^\infty \cos at \, e^{-st} \, dt = \frac{s}{s^2 + a^2}. \quad (6.6)$$

This integral could have been performed using either integration by parts in the usual way, or by changing $\cos at$ into e^{ajt} and then taking the real part of the final answer. In addition we have,

$$\mathcal{L}[\sin at] = \frac{a}{s^2 + a^2}. \quad (6.7)$$

The easiest proof of the results in (6.6) and (6.7) follows:

$$\begin{aligned} \mathcal{L}[\cos at + j\sin at] & \\ &= \mathcal{L}[e^{ajt}] \\ &= \int_0^\infty e^{ajt} \, e^{-st} \, dt \\ &= \int_0^\infty e^{-(s-aj)t} \, dt && \text{combining the exponentials} \\ & && \text{(careful with the } aj \text{ term)} \quad (6.8) \\ &= \frac{1}{s - aj} && \text{yes, this works for a complex exponent} \\ &= \frac{s + aj}{s^2 + a^2} && \text{using the complex conjugate} \\ &= \left(\frac{s}{s^2 + a^2}\right) + j\left(\frac{a}{s^2 + a^2}\right) \end{aligned}$$

As can be seen, the **real** part and the **imaginary** part have been colour-coded.

Example 6.4: Find $\mathcal{L}[t]$.

Using one integration by parts we obtain,

$$\mathcal{L}[t] = \int_0^\infty t \, e^{-st} \, dt = \frac{1}{s^2}. \quad (6.9)$$

This is worth checking. The Laplace Transforms for higher powers of t are given by,

$$\begin{aligned} \mathcal{L}[t^2] &= \int_0^\infty t^2 \, e^{-st} \, dt = \frac{2}{s^3}, \\ \mathcal{L}[t^3] &= \int_0^\infty t^3 \, e^{-st} \, dt = \frac{6}{s^4} = \frac{3!}{s^4}, \quad (6.10) \\ \mathcal{L}[t^n] &= \int_0^\infty t^n \, e^{-st} \, dt = \frac{n!}{s^{n+1}}. \end{aligned}$$

All of these were covered in ME10304 Mathematics 1.

Example 6.5: Find $\mathcal{L}[te^{-at}]$.

Application of the definition of the Laplace Transform yields,

$$\mathcal{L}[te^{-at}] = \int_0^\infty t\,e^{-at}\,e^{-st}\,dt = \int_0^\infty t\,e^{-(s+a)t}\,dt = \frac{1}{(s+a)^2}. \tag{6.11}$$

Again, this integral may be found using integration by parts. However, it is interesting (and quite consequential) to note that the role which is played by $\boxed{s}$ in the integration in Eq. (6.9) is identical to the role played by $\boxed{s+a}$ in Eq. (6.11): both s and $s+a$ are constants.

This is a foretaste of what we be calling the s-shift theorem a little later.

6.4 Laplace Transforms of derivatives

Although we have derived many useful Laplace Transform results, we aren't yet in a position to solve some linear constant-coefficient ODEs. Therefore we need to find the Laplace Transforms of some derivatives in order to enable us to do this.

Example 6.6: Find $\mathcal{L}[y'(t)]$.

We start with the statement that $\mathcal{L}[y(t)] = Y(s)$. Using the definition of the Laplace Transform we have,

$$\begin{aligned}
\mathcal{L}[y'] &= \int_0^\infty \underbrace{y'}_{I}\ \underbrace{e^{-st}}_{D}\,dt \\
&= \underbrace{\Big[y\Big]}_{I_1}\underbrace{\Big[e^{-st}\Big]_0^\infty}_{D_0} - \int_0^\infty \underbrace{\Big[y\Big]}_{I_1}\underbrace{\Big[-se^{-st}\Big]}_{D_1}\,dt && \text{one integration by parts} \\
&= -y(0) + s\int_0^\infty y\,e^{-st}\,dt && \text{where } ye^{-st}\to 0 \text{ as } t\to\infty \\
&= sY - y(0).
\end{aligned} \tag{6.12}$$

In other words, $\mathcal{L}[y] = Y \implies \mathcal{L}[y'] = sY - y(0)$.

Note 1: We have assumed that $ye^{-st} \to 0$ as $t \to \infty$ in the above analysis. Even if y were a growing exponential, then there is always a value of s above which this limit is satisfied.

Note 2: If one were trying to work out if there is a physical meaning for s, then one might say that multiplication by s is essentially equivalent to a time derivative but there is an additional contribution from the initial condition for y at $t = 0$ — the hint of initial condition! This fits well with solving ODEs that are Initial Value Problems.

With two and three integrations by parts, one may also find that,

$$\mathcal{L}[y''] = s^2Y - y'(0) - sy(0), \tag{6.13}$$

and

$$\mathcal{L}[y'''] = s^3Y - y''(0) - sy'(0) - s^2y(0), \tag{6.14}$$

and so on. I will leave these as exercises for you to do, but these results now make it very clear that Laplace Transforms will be well-suited to solve Initial Value Problems where all the initial conditions are at $t = 0$.

An alternative way of deriving Eq. (6.13) is to apply the result of Eq. (6.12) recursively. If we define $v = y'$ then

$$\begin{aligned} \mathcal{L}[v'] &= s\mathcal{L}[v] - v(0) && \text{using Eq. (6.12) on } v \\ \implies \quad \mathcal{L}[y''] &= s\mathcal{L}[y'] - y'(0) && \text{on translating back to } y \\ \implies \quad \mathcal{L}[y''] &= s\Big(s\mathcal{L}[y] - y(0)\Big) - y'(0) && \text{using Eq. (6.12) on } y \\ &= s^2 Y - y'(0) - sy(0). \end{aligned} \tag{6.15}$$

A similar trick works for $\mathcal{L}[y''']$ and higher derivatives.

In addition, given that integration is the inverse process to differentiation, it might not come as a surprise that,

$$\mathcal{L}\left[\int_0^t y(\tau)\,d\tau\right] = \frac{Y(s)}{s}. \tag{6.16}$$

In this result the variable τ is a dummy variable of integration. Proof of this will require integration by parts again, but the integral we see in Eq. (6.16) will need to be differentiated where use is made of the result,

$$\frac{d}{dt}\Big[\int_0^t y(\tau)\,d\tau\Big] = y(t). \tag{6.17}$$

This will form a question on a problem sheet.

6.5 Solutions of ODEs

We will now consider three examples of ODEs which will be solved using Laplace Transforms. Use will be made of some of the above results in order to illustrate the process of applying the inverse Laplace Transform.

Example 6.7: Solve the ODE, $y' + 2y = e^{-t}$, subject to the initial condition, $y(0) = 0$.

In the language I used in the ODEs section, the Complementary Function is characterised by $\lambda = -2$ while the forcing term is equivalent to $\lambda = -1$. From that point of view the writing down of $y_{\mathbf{cf}}$ and $y_{\mathbf{pi}}$ is straightforward, but let us see what happens with Laplace Transforms.

We will solve this equation by applying Laplace Transforms to each term in turn. We already know the following:

$$\begin{aligned} \mathcal{L}[y'] &= -y(0) + sY(s) && \text{(see Example 6.6)} \\ \mathcal{L}[2y] &= 2Y \\ \mathcal{L}[e^{-t}] &= \frac{1}{s+1}. && \text{(see Example 6.2)} \end{aligned} \tag{6.18}$$

Therefore we may write the following,

$$\begin{aligned}
& y' + 2y &=& \quad e^{-t} & \\
\Longrightarrow \quad & sY - y(0) + 2Y &=& \quad \frac{1}{s+1} & \\
\Longrightarrow \quad & (s+2)Y &=& \quad \frac{1}{s+1} & \text{since } y(0) = 0 \\
\Longrightarrow \quad & Y &=& \quad \frac{1}{(s+1)(s+2)} & \\
\Longrightarrow \quad & Y &=& \quad \frac{1}{(s+1)} - \frac{1}{(s+2)} & \text{using partial fractions} \\
\Longrightarrow \quad & y &=& \quad e^{-t} - e^{-2t} & \text{taking the inverse LT}
\end{aligned} \tag{6.19}$$

That final step used Eq. (6.4), namely $\mathcal{L}[e^{-at}] = 1/(s+a)$, with $a = 1$ and $a = 2$.

Example 6.8: Solve the equation, $y' + y = e^{-t}$, subject to the initial condition, $y(0) = c$, where c is a known, but unspecified, constant.

This equation is almost identical to that of Example 6.7, but the right hand side forcing term is now proportional to the Complementary Function of the ODE. In terms of λ-values, the auxiliary equation yields $\lambda = -1$ and the forcing term is also equivalent to $\lambda = -1$. So it will be of interest to see how Laplace Transforms cope with this special case with a repeated λ-value.

On taking Laplace Transforms of each term in the ODE we obtain,

$$sY - c + Y = \frac{1}{s+1}, \tag{6.20}$$

and therefore

$$Y = \frac{1}{(s+1)^2} + \frac{c}{s+1}. \tag{6.21}$$

The first term may be inverted using Example 6.5 with $a = 1$, and hence

$$y = te^{-t} + ce^{-t}. \tag{6.22}$$

In this solution the first term is the Particular Integral, while the second is the Complementary Function. Clearly the Laplace Transform takes this situation in its stride with no additional difficulties, but we need previously-derived results such as the one given by Example 6.5 as part of the toolchest of results to draw upon. But let us now consider a second order ODE.

Example 6.9: Solve the equation $y'' + 4y = 5e^{-t}$ subject to $y(0) = 0$ and $y'(0) = -1$.

On using the formula for $\mathcal{L}[y'']$ which is given in Eq. (6.15), the ODE transforms into

$$\underbrace{s^2Y - y'(0) - sy(0)}_{y''} + \underbrace{4Y}_{4y} = \underbrace{\frac{5}{s+1}}_{5e^{-t}}, \tag{6.23}$$

which may be rearranged into the form,

$$(s^2 + 4)Y + 1 = \frac{5}{s+1}. \tag{6.24}$$

Therefore Y is given by,

$$Y = \frac{5}{(s+1)(s^2+4)} - \frac{1}{s^2+4}. \tag{6.25}$$

We may now proceed either by expanding the first fraction using the method of partial fractions, or else combining the two fractions together to get $(4-s)/[(s+1)(s^2+4)]$, and then using the method of partial fractions. Either way, we obtain,

$$Y = \frac{1}{s+1} - \frac{s}{s^2+4}. \tag{6.26}$$

Results already derived (Example 6.2 with $a = 1$ and Example 6.3 with $a = 2$) are now sufficient to invert this expression; we get

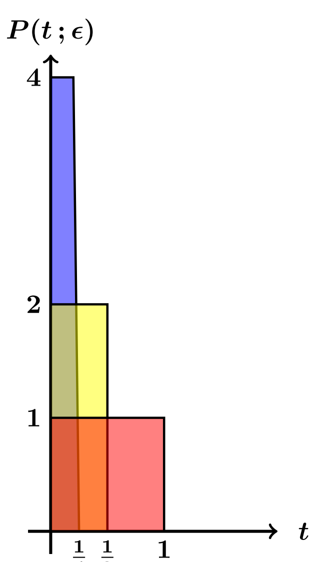
$$y = e^{-t} - \cos 2t. \tag{6.27}$$

6.6 The unit impulse

6.6.1 The definition

This is one of two unusual functions which are of great utility in Laplace Transforms. The other is the unit step function which is considered a little later.

We define $P(t\,;\epsilon)$ to be the **unit pulse of duration,** ϵ, beginning at $t = 0$. It has a **unit area**, and therefore Fig 6.1 shows three specific examples, while Eq. (6.28) gives the mathematical definition.



$$P(t\,;\epsilon) = \begin{cases} 1/\epsilon & (0 < t < \epsilon) \\ 0 & (\epsilon < t). \end{cases} \tag{6.28}$$

Figure 6.1. Showing $P(t\,;\epsilon)$, for $\epsilon = 1$, $\frac{1}{2}$, and $\frac{1}{4}$.

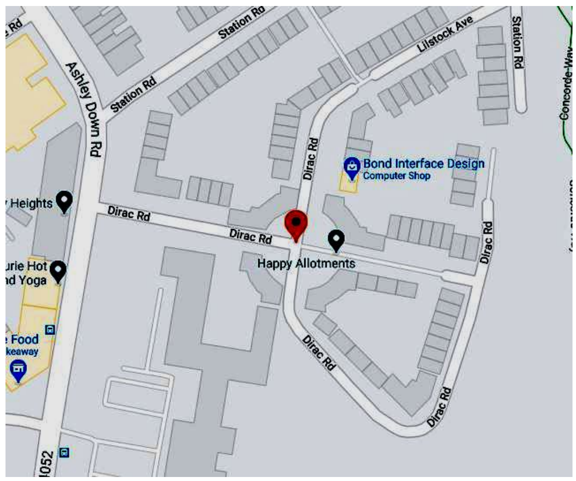
The **unit impulse** is what is obtained when ϵ becomes infinitesimally small. In this limit, $\epsilon \to 0$, the unit pulse now has an infinite strength over an interval of length zero, but the total area remains equal to 1 by definition.

The unit impulse is also known as the **delta function** or, in physics contexts, as the **Dirac delta function**. It is denoted by $\delta(t)$ and may be defined formally as,

$$\delta(t) = \lim_{\epsilon \to 0} P(t\,;\epsilon). \tag{6.29}$$

The unit impulse is used to as an idealised model of an impact, and has an important role in many areas of physics, mathematics and different branches of engineering,

Given the extremely tall and very narrow form of the Dirac delta function, it is surprising how poorly the Bristol-born Dirac has been commemorated by the Bristol and South Gloucestershire Councils:



Courtesy of Google Maps

6.6.2 The Laplace Transform of the unit impulse

We now need to find the Laplace Transform of the delta function, and we shall do this by first taking the Laplace Transform of the unit pulse, and then taking the $\epsilon \to 0$ limit. The Laplace Transform of the unit pulse of duration, ϵ, is

$$\begin{aligned}\mathcal{L}[P(t\,;\epsilon)] &= \int_0^\infty P(t\,;\epsilon)\,e^{-st}\,dt \\ &= \int_0^\epsilon \frac{1}{\epsilon}e^{-st}\,dt + \int_\epsilon^\infty 0\,e^{-st}\,dt \qquad P(t\,;\epsilon)=0 \text{ when } t>\epsilon \\ &= \frac{1}{\epsilon}\left[\frac{1-e^{-\epsilon s}}{s}\right].\end{aligned} \tag{6.30}$$

Now we may let $\epsilon \to 0$. This may be done by letting $\epsilon \to 0$ in Eq. (6.30), above. To do this we need to use the Taylor's series expansion of $e^{-\epsilon s}$ in terms of ϵ:

$$e^{-\epsilon s} = 1 - \epsilon s + \frac{(\epsilon s)^2}{2!} - \frac{(\epsilon s)^3}{3!} + \frac{(\epsilon s)^4}{4!}\cdots. \tag{6.31}$$

When Eq. (6.31) is substituted into Eq. (6.30) we get,

$$\begin{aligned}\mathcal{L}[P(t\,;\epsilon)] &= \frac{1}{\epsilon}\left[\frac{1-(1-\epsilon s+(\epsilon s)^2/2!-(\epsilon s)^3/3!\cdots)}{s}\right] \\ &= \frac{\epsilon s-(\epsilon s)^2/2!+(\epsilon s)^3/3!\cdots}{\epsilon s} \\ &= 1-\epsilon s/2!+(\epsilon s)^2/3!\cdots, \\ &\to 1 \qquad \text{as } \epsilon \to 0.\end{aligned} \tag{6.32}$$

Therefore,

$$\boxed{\mathcal{L}[\delta(t)] = 1.} \tag{6.33}$$

This value could also have been obtained using l'Hôpital's rule:

$$\mathcal{L}[\delta(t)] = \lim_{\epsilon\to 0}\frac{1-e^{-\epsilon s}}{\epsilon s} \overset{l'H}{=\!=} \lim_{\epsilon\to 0}\frac{s e^{-\epsilon s}}{s} = 1, \tag{6.34}$$

on taking derivatives of both the numerator and denominator with respect to ϵ.

Note: While there has been a lot of derivation here, it is only Eq. (6.33) that needs to be remembered.

6.6.3 The unit impulse and integration

The unit impulse also has the property that,

$$\begin{aligned} \int_{-\infty}^{\infty} g(t)\,\delta(t)\,dt &= \int_{-\infty}^{\infty} g(0)\,\delta(t)\,dt && \text{since } g(t)=g(0) \text{ where } \delta(t) \text{ is nonzero} \\ &= g(0)\int_{-\infty}^{\infty} \delta(t)\,dt && \text{elementary property of integrals} \\ &= g(0)\times 1 \quad = \quad g(0). \end{aligned} \tag{6.35}$$

In other words the integral of a function multiplied by the unit impulse is equivalent to picking out the value of $g(t)$ when $t = 0$. Bizarrely this is the easiest possible integral!

The apparent sleight of hand with the first equals sign above is motivated by the fact that the function, $g(t)$ and the value $g(0)$ are equal when $t = 0$ (see the red disk in Fig. 6.2a, below), and this implies that $g(t)\delta(t) = g(0)\delta(t)$ because $\delta(t) = 0$ when $t \neq 0$.

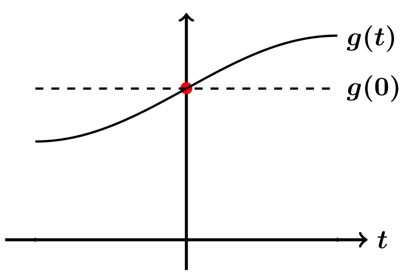


Figure 6.2a. Comparing $g(t)$ and $g(0)$.

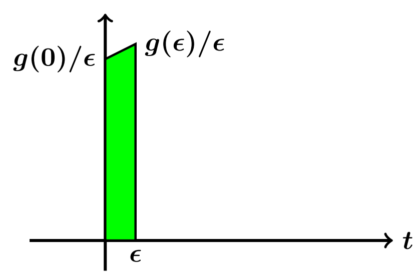


Figure 6.2b. Illustration of the trapezium rule.

An alternative proof may be provided by returning to the unit pulse, $P(t\,;\epsilon)$. When ϵ is very small, then we may approximate the integral of $g(t)\delta(t)$ using the trapezium rule (see Fig 6.2b, above):

$$\int_{-\infty}^{\infty} g(t)P(t\,;\epsilon)\;dt \quad\simeq\quad \underbrace{\left[\frac{1}{2}\left(\frac{g(\epsilon)}{\epsilon}+\frac{g(0)}{\epsilon}\right)\right]}_{\text{mean height}} \times \underbrace{\epsilon}_{\text{width}} \quad=\quad \tfrac{1}{2}[g(\epsilon)+g(0)]. \tag{6.36}$$

As $\epsilon \to 0$ this quantity tends towards $g(0)$ as well.

More generally, the unit impulse does not have to be located at $t = 0$. When it is centred at $t = a$ it is denoted by $\delta(t - a)$ and it is therefore infinite when $(t - a) = 0$. We now have the result,

$$\int_{-\infty}^{\infty} g(t)\,\delta(t-a)\,dt \quad=\quad \int_{-\infty}^{\infty} g(a)\,\delta(t-a)\,dt \quad=\quad g(a)\int_{-\infty}^{\infty} \delta(t-a)\,dt \quad=\quad g(a). \tag{6.37}$$

So the conclusion is that the integral of a function of t multiplied by a unit impulse is precisely equal to the value of the function at the point where the impulse is located.

Following on from this, we may find the Laplace Transform of $\delta(t-a)$:

$$\mathcal{L}[\delta(t-a)] = \int_0^\infty \delta(t-a)\, e^{-st}\, dt = e^{-as}, \tag{6.38}$$

and we may recover Eq. (6.35) when $a = 0$. Strictly speaking, this result applies only when $a \geq 0$. Should a be negative, then the impulse happens outside of the range of integration, and in that case the Laplace Transform is zero. We may state this mathematically as,

$$\mathcal{L}[\delta(t-a)] = \begin{cases} e^{-as} & (a \geq 0) \\ 0 & (a < 0) \end{cases} \tag{6.39}$$

Note: Again this has been a heavy subsubsection with multiple derivations. These derivations have been necessary in order that we may be convinced by the following results:

$$\boxed{\int_{-\infty}^{\infty} g(t)\, \delta(t-a)\, dt = g(a)} \qquad \text{and} \qquad \boxed{\mathcal{L}[\delta(t-a)] = e^{-as} \quad \text{when } a \geq 0.}$$

6.7 The solution of ODEs where the forcing term is a unit impulse

We shall have three examples of this type of ODE problem. The first is a 1st order ODE, while the others are of 2nd order.

Example 6.10: Solve $y' + ay = \delta(t)$ subject to $y(0) = 0$.

We do not yet have the technique to be able to find the Particular Integral that corresponds to a unit impulse, but we can use Laplace Transforms to solve this example. Again, using known results, we obtain the transformed version of the equation:

$$\begin{aligned} y' + ay = \delta(t) \quad &\Longrightarrow \quad \Big(sY - y(0)\Big) + aY = 1 \\ &\Longrightarrow \quad (s+a)Y = 1, \\ &\Longrightarrow \quad Y = \frac{1}{s+a}, \\ &\Longrightarrow \quad y = e^{-at} \qquad \text{using Eq. (6.4).} \end{aligned} \tag{6.40}$$

Note 1: The final solution, $y = e^{-at}$, is an example of what is known as the **impulse response** (or the **unit impulse response function**) for the system represented by the left hand side of the ODE. Crudely, it is how a system at rest then reacts to a unit impulse.

Note 2: The expression for Y is an example of what is known as the **Transfer Function** of the system. In a very real sense the Transfer Function encapsulates the full properties of the system. It is interesting to see that the reciprocal of the Transfer Function (i.e. $s + a$) is identical (account being taken for notation) with the Auxiliary equation, $\lambda + a = 0$, for the ODE. This is not a coincidence. There's a little more later.

Note 3: The exciting of a system by means of a unit impulse means that the given initial condition appears to have been violated. The original ODE has the initial condition, $y(0) = 0$, but the solution that we have derived satisfies $y(0) = 1$. Is this a contradiction?

Well, is it a contradiction? We'll need to analyse this in a little bit of detail, again just so that we can trust the solution that we have found. **Note:** that the rest of this page is for information only and serves solely to justify the Laplace Transform solution of Example 6.10.

Let us return to the unit pulse of duration, ϵ, and solve $y' + ay = P(t\,;\epsilon)$ instead. The objective here is to solve for the system's response to this pulse and then to let $\epsilon \to 0$ once more. Perhaps we need to rewrite the ODE as follows,

$$y' + ay = \begin{cases} \dfrac{1}{\epsilon} & (0 \le t \le \epsilon) \\ 0 & (\epsilon < t) \end{cases} \tag{6.41}$$

We'll minimise the detail of this analysis, but essentially (i) we solve for y in the range, $0 \le t \le \epsilon$, (ii) then determine the value of $y(\epsilon)$ which will then be used as an initial condition for (iii) the solution for y in the range, $\epsilon < t$. This solution is given by,

$$y = \begin{cases} \dfrac{1}{a\epsilon}(1 - e^{-at}) & (0 \le t \le \epsilon) \\ \dfrac{1}{a\epsilon}(1 - e^{-a\epsilon})e^{-a(t-\epsilon)} & (\epsilon < t) \end{cases} \tag{6.42}$$

where the solutions have been written in a way where it is easy to check that the two formulae for y agree at $t = \epsilon$. For the sake of illustration we have set $a = 1$ in Figure 6.3 where the solutions for three different pulse durations are shown.

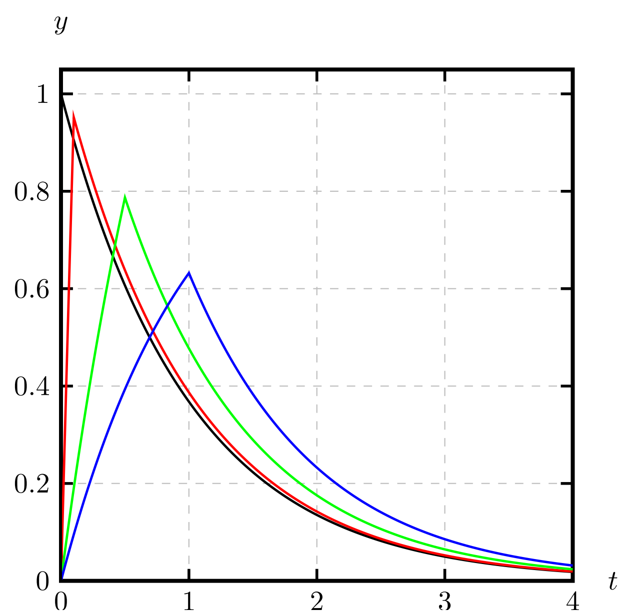


Figure 6.3. Solutions given by Eq. (6.42) for $\epsilon = 0$ (unit impulse) $\epsilon = 0.1$, $\epsilon = 0.5$ and $\epsilon = 1$.

In this Figure we see that the solutions corresponding to the various unit pulses converge towards that for the unit impulse as $\epsilon \to 0$. More specifically we see that the initial rise from $y = 0$ at $t = 0$ becomes increasingly steep in this limit. Equation (6.42) confirms this since $y'(0) = 1/\epsilon$. Although this is an unusual limit, the Laplace Transform solution given in Eq. (6.40) is indeed correct even though it appears to violate the given initial conditions; it is merely the case that there is an extremely rapid variation in y over the infinitesimally short duration of the impulse.

Example 6.11: Solve the equation $y'' + (a+b)y' + aby = \delta(t)$ subject to $y(0) = y'(0) = 0$.

On taking the Laplace Transform of the equation we get,

$$s^2Y + (a+b)sY + abY = 1 \qquad \text{noting that } y = y' = 0 \text{ at } t = 0$$

$$\Longrightarrow \qquad \Big[s^2 + (a+b)s + ab\Big]Y = 1$$

$$\Longrightarrow \qquad Y = \frac{1}{s^2 + (a+b)s + ab}$$

$$\Longrightarrow \qquad Y = \frac{1}{(s+a)(s+b)} \tag{6.43}$$

$$\Longrightarrow \qquad Y = \frac{1}{b-a}\left[\frac{1}{s+a} - \frac{1}{s+b}\right] \qquad \text{using partial fractions}$$

$$\Longrightarrow \qquad y = \left[\frac{e^{-at} - e^{-bt}}{b-a}\right] \qquad \text{taking the inverse Laplace Transform.}$$

In the light of the initial condition violation which we encountered in Example 6.10, let us check whether the same happens here. We may easily determine mathematically that $y(0) = 0$, and $y'(0) = 1$. Thus the **impulse response** for a second order ODE with zero initial conditions has a nonzero first derivative at $t = 0$, so it is the initial condition for the first derivatve that has been violated this time.

Although I shall not prove it, the general case is that the initial condition for the $(n-1)^{\rm st}$ derivative is violated when solving an $n^{\rm th}$ order ODE.

In the following Figure we show what the solution given by Eq. (6.43) looks like for the case, $a = 1$ and $b = 2$. The maximum value of y is 0.25 and this is attained when $t = \ln 2 = 0.69315$. The initial slope may also be found analytically and it is $y'(0) = 1$.

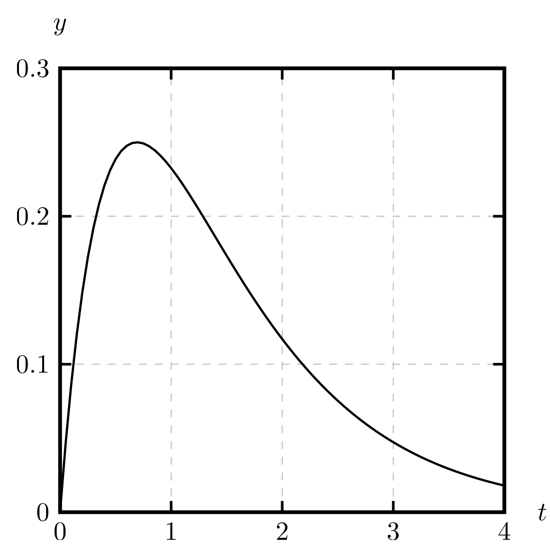


Figure 6.4. Solutions given by Eq. (6.43) when $a = 1$ and $b = 2$.

A final example involves an undamped mass/spring system.

Example 6.12: Solve the ODE, $my'' + ky = \delta(t)$, subject to $y(0) = y'(0) = 0$. Here m is mass and k, the spring stiffness.

We rearrange the equation for the ODE slightly to the form,

$$y'' + (k/m)y = (1/m)\delta(t), \tag{6.44}$$

and the result of taking the Laplace Transform is,

$$\Big(s^2 + \frac{k}{m}\Big)Y = \frac{1}{m}, \tag{6.45}$$

where the zero initial conditions have been accounted for. We now proceed as follows:

$$\begin{aligned}
& \Big(s^2 + \frac{k}{m}\Big)Y = \frac{1}{m} & \\
\Longrightarrow \quad & Y = \frac{1/m}{s^2 + k/m} & \text{which reminds us of the LT of a sine} \\
\Longrightarrow \quad & Y = \frac{1}{\sqrt{km}} \times \frac{\sqrt{k/m}}{s^2 + k/m} & \text{now use Ex. 6.2 with } a = \sqrt{k/m} \\
\Longrightarrow \quad & y = \frac{1}{\sqrt{km}} \sin\sqrt{\frac{k}{m}}\, t. &
\end{aligned} \tag{6.46}$$

This final solution certainly satisfies $y(0) = 0$, as did the solution in Example 6.11. However, we have

$$y' = \frac{1}{m}\cos\sqrt{\frac{k}{m}}\, t, \tag{6.47}$$

and therefore $y'(0) = 1/m$. So we have another violation of the $y'(0) = 0$ initial condition. In this case the physical meaning of this violation tells us that the initial velocity which is induced by the unit impulse decreases as the mass, m, increases. Clearly this is physically correct — think of the different responses of a ping pong ball and a cricket ball.

We may also say that the solution given in Eq. (6.46) satisfies $my'(0) = 1$ and, given that y' is the velocity, we can say that **the unit impulse imparts a unit momentum to the system.**

6.7.1 Some observations

Given that the presence of the unit impulse as a forcing term causes a violation of an initial condition, there is the scope to use our experience of solving these equations to write down an alternative versions of the ODEs and initial condition without the presence of the unit impulse. Thus for Example 6.10,

$$\begin{aligned}
& y' + ay = \delta(t), \qquad y(0) = 0, \\
\text{and} \quad & \\
& y' + ay = 0, \qquad y(0) = 1,
\end{aligned} \tag{6.48}$$

have the same solutions. Likewise for Example 6.11:

$$\begin{aligned}
& y'' + (a+b)y' + ab = \delta(t), \qquad y(0) = 0, \quad y'(0) = 0, \\
\text{and} \quad & \\
& y'' + (a+b)y' + ab = 0, \qquad y(0) = 0, \quad y'(0) = 1,
\end{aligned} \tag{6.49}$$

have the same solution. The more general case given in Example 6.12 (where the y'' term doesn't have a unit coefficient):

$$\begin{aligned} & my'' + ky = \delta(t), \qquad y(0) = 0, \quad y'(0) = 0, \\ \text{and} \qquad & \\ & my'' + ky = 0, \qquad\quad y(0) = 0, \quad my'(0) = 1, \end{aligned} \tag{6.50}$$

also have the same solution. As a final example which hasn't been covered above, the following two ODEs have identical solutions:

$$\begin{aligned} & my'' + ky = c\delta(t), \qquad y(0) = a, \quad y'(0) = b, \\ \text{and} \qquad & \\ & my'' + ky = 0, \qquad\quad y(0) = a, \quad y'(0) = b + c/m. \end{aligned} \tag{6.51}$$

Check carefully whether this last case makes sense, given the preceding analyses.

6.8 The Unit Step Function

6.8.1 Definition

This is the second of the two special functions that we'll consider and this one is sketched in Figure 6.5, below. It is denoted either by $u(t)$ or by $H(t)$; we will use $H(t)$. The former notation is because the step is a unit step, from 0 to 1. The latter notation comes from its alternative name, the Heaviside step function, named for the British physicist/engineer/mathematician, Oliver Heaviside, whose entry on Wikipedia is worth reading for many reasons!

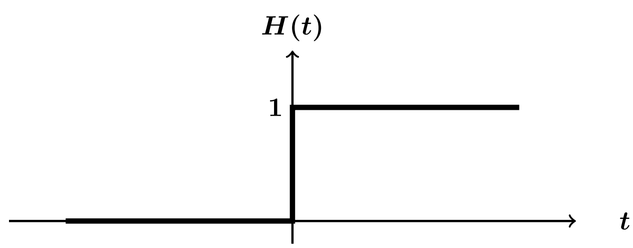


Figure 6.5. The unit step function.

Mathematically, the unit step function is defined as

$$H(t) = \begin{cases} 0 & (t < 0) \\ 1 & (t > 0). \end{cases} \tag{6.52}$$

The step rise does not need to occur at $t = 0$. Its equivalent at $t = a$ is denoted by $H(t - a)$ and this is shown in Figure 6.6.

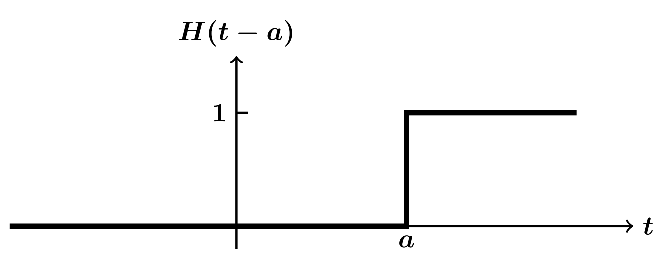


Figure 6.6. The unit step function at $t = a$.

This is defined by,

$$H(t - a) = \begin{cases} 0 & (t - a < 0) \\ 1 & (t - a > 0). \end{cases} \tag{6.53}$$

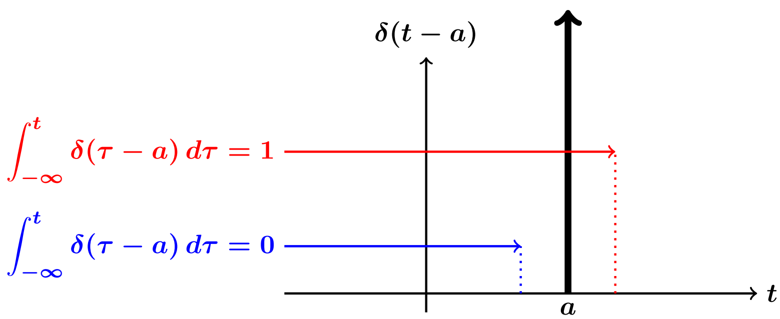
There is a very strong link between the unit impulse and the unit step function. The two main relationships may be expressed as,

$$H(t) = \int_{-\infty}^{t} \delta(\tau)\, d\tau \qquad \text{and} \qquad \delta(t) = \frac{dH(t)}{dt}, \tag{6.54}$$

or, more generally, as,

$$H(t-a) = \int_{-\infty}^{t} \delta(\tau - a)\, d\tau \qquad \text{and} \qquad \delta(t-a) = \frac{dH(t-a)}{dt.} \tag{6.55}$$

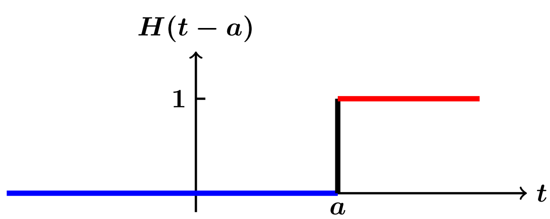
Figure 6.7 shows how the unit step function may be obtained by integrating the unit impulse. When $t < a$, which is represented by the blue line, the range of integration doesn't include the location of the unit impulse at $t = a$, and therefore the integral is zero. On the other hand, when $t > a$, which is represented by the red line, the range of integration includes the location of the unit impulse, and therefore the integral is 1.



$$\boxed{H(t-a) = \int_{-\infty}^{t} \delta(\tau - a)\, d\tau}$$

Figure 6.7. The unit impulse at $t = a$.

Figure 6.8 shows the unit step function and, in particular, we see that the slope is zero when $t \neq a$ (the red and the blue sections) and is infinite at $t = a$ (the black section). Pure mathematicians might have more to say about this but we will run with idea as being self-evident.



$$\boxed{\delta(t-a) = \frac{dH(t-a)}{dt}.}$$

Figure 6.8. The unit step function at $t = a$.

6.8.2 The Laplace Transform of the Unit Step Function

We will now consider the Laplace Transform of the unit step function because this has some use in the solution of ODEs. The Laplace Transform of $H(t-a)$ is given by

$$\begin{aligned}
\mathcal{L}[H(t-a)] &= \int_0^\infty H(t-a)\, e^{-st}\, dt \\
&= \int_0^a 0 \times e^{-st}\, dt + \int_a^\infty 1 \times e^{-st}\, dt \qquad \text{splitting the integral} \\
&= 0 + \int_a^\infty 1 \times e^{-st}\, dt \\
&= \left[-\frac{e^{-st}}{s}\right]_a^\infty = \frac{e^{-as}}{s}.
\end{aligned} \tag{6.56}$$

Note: that the use of $a = 0$ in the above shows that $\mathcal{L}[H(t)] = 1/s$. If this seems familiar then it is because we also showed that $\mathcal{L}[1] = 1/s$ in Example. 6.1. Although it appears that we have two functions with the same transform, the functions are identical within the range of integration of the Laplace Transform: $H(t) = 1$.

We will revisit both the unit impulse and the unit step function later.

6.9 The shift theorem in s

This is almost trivial! Here is a combined statement and proof of the theorem. Given that,

$$\mathcal{L}[f(t)] = \int_0^\infty f(t)\, e^{-st}\, dt = F(s), \tag{6.57}$$

then

$$\begin{aligned}
\mathcal{L}[f(t)e^{-at}] &= \int_0^\infty f(t)e^{-at}e^{-st}\, dt \qquad \text{by definition} \\
&= \int_0^\infty f(t)e^{-(s+a)t}\, dt \\
&= F(s+a).
\end{aligned} \tag{6.58}$$

If this proof seems to be too quick, then compare the role that $\boxed{s}$ plays in Eq. (6.57) with the role played by $\boxed{(s+a)}$ in Eq. (6.58). A simple statement of the theorem is,

$$\boxed{\mathcal{L}[f(t)] = F(s) \qquad \Longrightarrow \qquad \mathcal{L}[f(t)\, e^{-at}] = F(s+a).} \tag{6.59}$$

The simplicity of this theorem is demonstrated in the next two Examples.

Example 6.13: Find the Laplace Transform of te^{-at}.

We start by noting that we already know that $\mathcal{L}[t] = 1/s^2$ from Example 6.4. Therefore the s-shift theorem tells us that,

$$\mathcal{L}[te^{-at}] = 1/(s+a)^2. \tag{6.60}$$

This is the same answer as we found in Example 6.5, but there we undertook the integration explicitly rather than by using a previously-known result and the s-shift theorem.

Example 6.14: Find the Laplace Transform of $\cos bt\ e^{-at}$.

Example 6.3 shows that $\mathcal{L}[\cos bt] = s/(s^2 + b^2)$. So the s-shift theorem tells us that,

$$\mathcal{L}[\cos bt] = \frac{s}{s^2+b^2} \qquad \Longrightarrow \qquad \mathcal{L}[\cos bt\ e^{-at}] = \frac{s+a}{(s+a)^2+b^2}. \tag{6.61}$$

So every instance of s in the left hand equation is replaced by $(s + a)$ in the right hand equation.

Similarly, we may state that,

$$\mathcal{L}[\sin bt] = \frac{b}{s^2+b^2} \qquad \Longrightarrow \qquad \mathcal{L}[\sin bt\ e^{-at}] = \frac{b}{(s+a)^2+b^2}. \tag{6.62}$$

Both the left hand sides for both Eqs. (6.61) and (6.62) have been taken from Example 6.3.

We may also use this shift theorem to find inverse transforms and, indeed, this is its main use. The following is a typical example which arises from the solution of an ODE.

Example 6.15: Solve the ODE, $y'' + 4y' + 13y = 0$, subject to $y(0) = 1$ and $y'(0) = -4$.

On applying the Laplace Transform to the ODE we obtain,

$$\Big[s^2Y - y'(0) - sy(0)\Big] + 4\Big[sY - y(0)\Big] + 13Y = 0. \tag{6.63}$$

Using the given initial conditions this simplifies to,

$$(s^2 + 4s + 13)Y - s = 0, \tag{6.64}$$

and therefore,

$$Y = \frac{s}{s^2+4s+13}. \tag{6.65}$$

The determination of y, i.e. the inverse Laplace Transform of Y, will require us to coerce Y into a form which may then be inverted immediately. This will involve the completion of the square for the denominator, the s-shift theorem and the following Laplace Transforms for $\cos bt$ and $\sin bt$ from Example 6.3:

$$\mathcal{L}[\cos bt] = \frac{s}{s^2+b^2}, \qquad \mathcal{L}[\sin bt] = \frac{b}{s^2+b^2}.$$

First, we note that $s^2 + 4s + 13 = (s + 2)^2 + 9 = (s + 2)^2 + 3^2$. The presence of the $(s + 2)$ tells us that the s-shift theorem is soon to be used. So we have,

$$\begin{aligned} Y &= \frac{s}{s^2+4s+13} \\ &= \frac{s}{(s+2)^2+3^2} && \text{on completing the square} \\ &= \frac{s+2}{(s+2)^2+3^2} - \frac{2}{(s+2)^2+3^2} && \text{to get } (s+2) \text{ everywhere} \\ &= \frac{s+2}{(s+2)^2+3^2} - \tfrac{2}{3}\Big[\frac{3}{(s+2)^2+3^2}\Big] && \text{now for inversion.} \end{aligned} \tag{6.66}$$

The last step involved placing a 3 in the numerator of the fraction comprising the second quotient. Now every term is just an s-shift away from the Laplace Transforms of a cosine and a sine, respectuvely. Therefore the s-shift theorem tells us that

$$\mathcal{L}^{-1}\left[\frac{s+2}{(s+2)^2+3^2}\right] = e^{-2t}\cos 3t, \tag{6.67}$$

and

$$\mathcal{L}^{-1}\left[\tfrac{2}{3}\,\frac{3}{(s+2)^2+3^2}\right] = \tfrac{2}{3}e^{-2t}\sin 3t. \tag{6.68}$$

Therefore we may write the final solution of the ODE as,

$$y = e^{-2t}\cos 3t - \tfrac{2}{3}e^{-2t}\sin 3t = e^{-2t}\Big[\cos 3t - \tfrac{2}{3}\sin 3t\Big]. \tag{6.69}$$

6.10 The shift theorem in t

We shall begin with the following sketch of a function, $f(t)$ (black curve), and how it may be shifted sideways to the right but not retain any information about $f(t)$ when $t < 0$ (red curve). This uses $H(t-a)$ as a multiplier to filter out the unwanted information (the dotted line). So $f(t)$ has undergone a t-shift, and $f(t-a)H(t-a)$ contains exactly the same information as $f(t)$ does in the range, $0 \le t < \infty$.

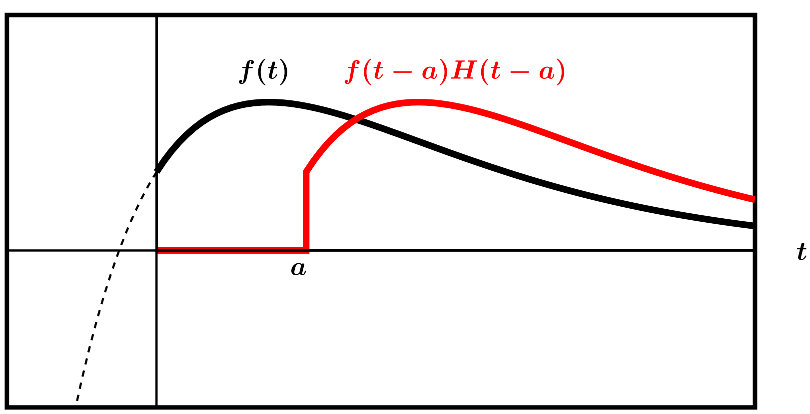


Figure 6.9. Showing a typical $f(t)$ and $f(t-a)H(t-a)$.

Now let us find the Laplace Transform of $f(t-a)H(t-a)$:

$$\begin{aligned}
\mathcal{L}[H(t-a)f(t-a)] &= \int_0^\infty f(t-a)\,H(t-a)\,e^{-st}\,dt \\
&= \int_a^\infty f(t-a)\,H(t-a)\,e^{-st}\,dt && \text{because the integrand is zero in } 0 \le t < a \\
&= \int_a^\infty f(t-a)\,e^{-st}\,dt && \text{because } H(t-a)=1 \text{ in } a \le t < \infty \\
&= \int_0^\infty f(\tau)e^{-s(\tau+a)}\,d\tau && \text{using } \tau = t-a\text{; also note the lower limit} \\
&= e^{-sa}\underbrace{\int_0^\infty f(\tau)\,e^{-s\tau}\,d\tau}_{F(s)} && e^{-sa} \text{ is a constant with respect to } \tau \\
&= e^{-sa}F(s).
\end{aligned} \tag{6.70}$$

Therefore we have shown that

$$\boxed{\mathcal{L}[H(t-a)f(t-a)] = e^{-sa}F(s),} \tag{6.71}$$

which is the Shift Theorem in t.

Example 6.16: Find $\mathcal{L}^{-1}\left[\dfrac{e^{-as}}{s^2}\right]$.

The presence of the exponential tells us that the t-shift theorem needs to be used. The presence of the $1/s^2$ tells us that the basic time-dependent function which lies beneath all of this is t, because $\mathcal{L}[t] = 1/s^2$. So therefore we apply the t-shift theorem and obtain,

$$\mathcal{L}[t] = \frac{1}{s^2} \qquad \Longrightarrow \qquad \mathcal{L}[(t-a)H(t-a)] = \frac{e^{-as}}{s^2}. \tag{6.72}$$

Therefore the answer to the original question is,

$$\mathcal{L}^{-1}\left[\frac{e^{-as}}{s^2}\right] = (t-a)H(t-a). \tag{6.73}$$

Example 6.17 Find the inverse Laplace Transform of $e^{-as}/(s+b)^2$.

This is an example of the use of both shift theorems. The exponential tells us that we need the t-shift theorem. The presence of the $(s+b)$ tells us that the s-shift is needed. From all this mess we can see that the underlying function is again, $1/s^2$. The only question now is, which is the better way to go? Should we use the s-shift theorem first or the t-shift theorem first? We'll do it both ways and everyone can then make up their own minds. But in both cases we lead off with the Laplace Transform of t.

If we start with the s-shift theorem, then we get,

$$\mathcal{L}[\underbrace{t}_{f(t)}] = \underbrace{\frac{1}{s^2}}_{F(s)}$$

$$\implies \quad \mathcal{L}[\underbrace{te^{-bt}}_{f(t)e^{-bt}}] = \underbrace{\frac{1}{(s+b)^2}}_{F(s+b)} \qquad \text{applying the } s\text{-shift theorem}$$

$$\implies \quad \mathcal{L}[\underbrace{te^{-bt}}_{f(t)}] = \underbrace{\frac{1}{(s+b)^2}}_{F(s)} \qquad \text{redefining the labelling}$$

$$\implies \quad \mathcal{L}[\underbrace{H(t-a)\times(t-a)e^{-b(t-a)}}_{H(t-a)f(t-a)}] = \underbrace{\frac{e^{-as}}{(s+b)^2}}_{e^{-as}F(s)} \qquad \text{applying the } t\text{-shift theorem} \tag{6.74}$$

where we have meticulously changed every t in te^{-bt} to $(t-a)$ in the last line.

On the other hand, if we start with the t-shift theorem, then we get,

$$\mathcal{L}[\underbrace{t}_{f(t)}] = \underbrace{\frac{1}{s^2}}_{F(s)}$$

$$\implies \quad \mathcal{L}[\underbrace{(t-a)H(t-a)}_{f(t-a)H(t-a)}] = \underbrace{\frac{e^{-as}}{s^2}}_{e^{-as}F(s)} \qquad \text{applying the } t\text{-shift theorem}$$

$$\implies \quad \mathcal{L}[\underbrace{(t-a)H(t-a)}_{f(t)}] = \underbrace{\frac{e^{-as}}{s^2}}_{F(s)} \qquad \text{relabelling} \tag{6.75}$$

$$\implies \quad \mathcal{L}[\underbrace{(t-a)e^{-bt}H(t-a)}_{f(t)e^{-bt}}] = \underbrace{\frac{e^{-a(s+b)}}{(s+b)^2}}_{F(s+b)} \qquad \text{applying the } s\text{-shift theorem}$$

$$\implies \quad \mathcal{L}[(t-a)e^{-b(t-a)}H(t-a)] = \frac{e^{-as}}{(s+b)^2} \qquad \text{multiplying both sides by } e^{ab}$$

and we obtain the same final line as before.

Hence,

$$\mathcal{L}^{-1}\left[\frac{e^{-as}}{(s+b)^2}\right] = H(t-a)\times(t-a)e^{-b(t-a)}. \tag{6.76}$$

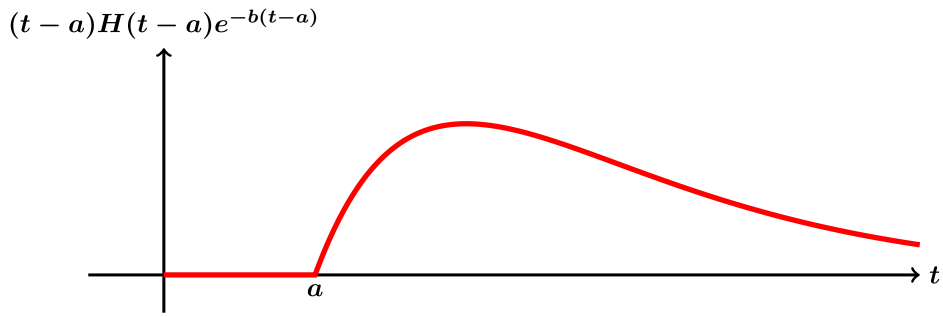


Figure 6.10. Sketch of $\mathcal{L}^{-1}[e^{-as}/(s+b)^2] \;=\; (t-a)H(t-a)e^{-b(t-a)}$

6.11 Convolution Theorem

Given the two functions $f(t)$ and $g(t)$ and their respective transforms, $F(s)$ and $G(s)$, then the convolution of f and g is defined by

$$f*g = \int_0^t f(\tau)g(t-\tau)\,d\tau = \int_0^t f(t-\tau)g(\tau)\,d\tau, \tag{6.77}$$

where it should be noted that each of these two integrals results in the same function of t. We may say that $f*g$ is **the convolution of f and g.**

It is interesting to note what happens to the arguments of f and g in the above integrals. In the first integral we see that, while the argument of $f(\tau)$ increases from 0 to t, the argument of $g(t-\tau)$ decreases from t to 0. This seems unusual but this behaviour arises elsewhere and for more obvious reasons. Perhaps the simplest example is the calculation of probabilities associated with rolling two standard dice. For example, the probability of throwing, say, a 9 in total is given by

$$\begin{aligned} P(9 \text{ with 2 dice}) \;&=\; P(3)\times P(6) \;+\; P(4)\times P(5) \;+\; P(5)\times P(4) \;+\; P(6)\times P(3) \\ &=\; \sum_{n=3}^{6} P(n)\times P(9-n). \end{aligned} \tag{6.78}$$

where I am assuming that my notation is self-explanatory. So we end up with a convolution sum where n increases while $9-n$ decreases.

However, it is more important to explain why it is even necessary to have the concept of convolution here at all. The reason is the following:

$$\mathcal{L}[f*g] = F(s)G(s), \tag{6.79}$$

i.e. that **the transform of the convolution is the product of the transforms.** I will not give a proof of this here, but it is relegated to an Appendix at the end of the Laplace Transform section should you be curious. This result may also be written in the form,

$$\mathcal{L}^{-1}[F(s)G(s)] \;=\; f*g, \tag{6.80}$$

which is the form that is generally needed in practice.

We will illustrate this by following the fate and fortune of the following two ODEs:

$$\boxed{\begin{aligned} y'+2y&=f(t)\\ y(0)&=0 \end{aligned}} \qquad \text{and} \qquad \boxed{\begin{aligned} z''+4z&=f(t)\\ z(0)&=z'(0)=0 \end{aligned}} \tag{6.81}$$

So we have two physical systems, one for $y(t)$ the other for $z(t)$, both of which are at rest for $t < 0$, and then each of which is set into motion solely by the forcing function, $f(t)$. These may be solved by first applying the Laplace Transform:

$$\boxed{(s+2)Y = F} \qquad \text{and} \qquad \boxed{(s^2+4)Z = F,} \tag{6.82}$$

These are solved easily for Y and Z:

$$\boxed{Y = F \times \frac{1}{s+2}} \qquad \text{and} \qquad \boxed{Z = F \times \frac{1}{s^2+4},} \tag{6.83}$$

both of which are products of functions of s and so we need the Convolution Theorem.

6.11.1 Some examples of the use of the Convolution Theorem

Before returning to Eqs. (6.82) and (6.83) we shall consider two Examples of the use of the Convolution Theorem.

Example 6.18: Given that $\mathcal{L}[t] = 1/s^2$, use the Convolution Theorem to find $\mathcal{L}^{-1}[1/s^4]$.

Clearly $1/s^4 = (1/s^2) \times (1/s^2)$ and therefore we may use the Convolution Theorem as given by Eq. (6.80). Formally, we shall let $F = G = 1/s^2$, which means that $f = g = t$. Equation (6.80) gives,

$$\begin{aligned}
\mathcal{L}^{-1}\left[\frac{1}{s^4}\right] = \mathcal{L}^{-1}[FG] \quad &= \quad f * g = \int_0^t f(\tau)\, g(t-\tau)\, d\tau \\
&= \quad t * t \\
&= \quad \int_0^t \tau(t-\tau)\, d\tau \\
&= \quad \left[\frac{\tau^2}{2}t - \frac{\tau^3}{3}\right]_0^t \\
&= \quad t^3/6.
\end{aligned} \tag{6.84}$$

Example 6.19 Use the Convolution Theorem to find the inverse Laplace Transform of

$$\frac{1}{s+1} \times \frac{1}{s+2}.$$

This particular function of s was obtained back in Example 6.7 when we were solving the ODE, $y' + 2y = e^{-t}$ subject to $y(0) = 0$. At that point we employed Partial Fractions to simplify this function of s before applying the Inverse Laplace Transform. Now we shall use the convolution theorem instead, and we really ought to obtain the same function of t.

We already know that

$$\mathcal{L}[e^{-t}] = \frac{1}{(s+1)} = F(s) \qquad \text{and} \qquad \mathcal{L}[e^{-2t}] = \frac{1}{(s+2)} = G(s),$$

and therefore we have $f(t) = e^{-t}$ and $g(t) = e^{-2t}$ to use in the Convolution Theorem. Then the required $\mathcal{L}^{-1}[FG] = f * g$. We'll use both versions of the convolution integral:

$$
\begin{aligned}
f*g &= e^{-t}*e^{-2t} & f*g &= e^{-t}*e^{-2t} \\
&= \int_0^t \underbrace{e^{-\tau}}_{f(\tau)} \times \underbrace{e^{-2(t-\tau)}}_{g(t-\tau)}\, d\tau & &= \int_0^t \underbrace{e^{-(t-\tau)}}_{f(t-\tau)} \times \underbrace{e^{-2\tau}}_{g(\tau)}\, d\tau \\
&= \int_0^t e^{-2t}e^{\tau}\, d\tau & &= \int_0^t e^{-t}e^{-\tau}\, d\tau \\
&= e^{-2t}\int_0^t e^{\tau}\, d\tau & &= e^{-t}\int_0^t e^{-\tau}\, d\tau \\
&= e^{-2t}\Big[e^{\tau}\Big]_0^t & &= e^{-t}\Big[-e^{-\tau}\Big]_0^t \\
&= e^{-2t}\Big[e^{t}-1\Big] & &= e^{-t}\Big[-e^{-t}+1\Big] \\
&= e^{-t}-e^{-2t}. & &= e^{-t}-e^{-2t}.
\end{aligned}
\tag{6.85}
$$

This is the same expression as was given in Eq. (6.19).

Note: In this Example we had a choice of which integral to use in the definition of the convolution integral in Eq. (6.77). Both have been applied and they are of equal difficulty, but there are some cases where one of the choices is either more useful or quicker than the other.

Example 6.20: Solve the ODE, $y' + 2y = f(t)$ subject to $y(0) = 0$ using the Convolution Theorem.

We have already started this problem back in Eq. (6.81), but we'll run it fully here. On taking the Laplace Transform (noting that $y(0) = 0$) we obtain,

$$
sY + 2Y = F \qquad \Longrightarrow \qquad Y = F \times \frac{1}{s+2}, \tag{6.86}
$$

which is a product and therefore a perfect target for the Convolution Theorem.

We may let $G(s) = 1/(s+2)$ and therefore $g(t) = e^{-2t}$. Now we face the choice of which integral to use. Here's the 'wrong' one:

$$
y(t) = f(t) * e^{-2t} = \int_0^{\tau} f(t-\tau)e^{-2\tau}\, d\tau. \tag{6.87}
$$

Whilst this a perfectly good expression for $y(t)$, it isn't as good as the other one. This second one is,

$$
y(t) = f(t) * e^{-2t} = \int_0^t f(\tau)e^{-2(t-\tau)}\, d\tau = e^{-2t}\int_0^t f(\tau)e^{2\tau}\, d\tau. \tag{6.88}
$$

On the face of it, this doesn't necessarily look better, but this is the form of solution which we would obtain by treating the original ODE as a first-order-linear ODE and using the Integrating Factor. For this ODE the

Integrating Factor is e^{2t}, and when the ODE is multiplied by it, we obtain,

$$\begin{aligned} e^{2t}(y'+2y) &= e^{2t}f(t) \\ \Longrightarrow \qquad (e^{2t}y)' &= e^{2t}f(t) && \text{exact derivative} \\ \Longrightarrow \qquad e^{2t}y &= \int_0^t e^{2\tau}f(\tau)\,d\tau && \text{using } y(0)=0 \\ \Longrightarrow \qquad y &= e^{-2t}\int_0^t e^{2\tau}f(\tau)\,d\tau. \end{aligned} \tag{6.89}$$

Note: If $f(t)=\delta(t)$, the unit impulse, then $F(s)=1$. Therefore Eq. (6.86) becomes

$$Y = \frac{1}{s+2}, \tag{6.90}$$

which is the **Transfer Function** for the system (i.e. for $y'+2y$), and its inverse Laplace Transform is

$$y = e^{-2t}, \tag{6.91}$$

is the **unit impulse response function**. This expression for y is obtained by substituting for $f(t)$ into Eq. (6.88).

Example 6.21: Solve the ODE, $z''+4z=f(t)$ subject to $z(0)=z'(0)=0$ using the Convolution Theorem.

Again we have already started this problem back in Eq. (6.81), and again we'll run it fully here.

On taking the Laplace Transform (noting that $z(0)=z'(0)=0$) we obtain,

$$(s^2+4)Z = F \qquad \Longrightarrow \qquad Z = F \times \frac{1}{s^2+4}. \tag{6.92}$$

If we let $G(s)=1/(s^2+4)$ then Eq. (6.7) gives its inverse as $g(t)=\frac{1}{2}\sin 2t$; check that one carefully. Hence the inverse Laplace Transform of Z is,

$$z = \tfrac{1}{2}\sin 2t * f(t) = \tfrac{1}{2}\int_0^t f(\tau)\sin 2(t-\tau)\,d\tau) = \tfrac{1}{2}\int_0^t f(t-\tau)\sin 2\tau\,d\tau. \tag{6.93}$$

I am not sure which is the better form of the convolution integral to use here, but both are fine. On the other hand, if we have $f(t)=\delta(t)$, then the first integral is better because it is slightly easier to use the result of integrating with the delta function. Once more, $f(t)=\delta(t) \Rightarrow F(s)=1$, and hence the **Transfer Function** is $Z=1/(s^2+4)$ and the **unit impulse response function** is $z=\frac{1}{2}\sin 2t$.

6.12 Solving systems of ODEs

We'll consider just one of these, and it will also involve the unit impulse. These become quite rapidly more difficult as the order of the system increases, and it will eventually be necessary to use numerical methods to tackle these. The example given here involves a pair of first order ODEs. There turns out to be more than one way of organising one's workings for this, and I will attempt to make each route as clear as possible.

Example 6.22: Solve the system of ODEs,

$$y'+y-z=0, \qquad z'+2y+4z=\delta(t), \qquad \text{subject to } y(0)=z(0)=0. \tag{6.94}$$

Method 1. This involves the elimination of one of the dependent variables to leave a second order ODE. So the elimination of z between the two yields, $y'' + 5y' + 6y = \delta(t)$, which has to be solved subect to $y(0) = y'(0) = 0$. Taking Laplace Transforms of this ODE, accounting for the initial conditions, yields,

$$(s^2 + 5s + 6)Y = 1 \qquad \Longrightarrow \qquad Y = \frac{1}{s^2 + 5s + 6} = \frac{1}{(s+2)(s+3)} = \frac{1}{s+2} - \frac{1}{s+3}. \tag{6.95}$$

Here we have used partial fractions to simplify the denominator, although the Convolution Theorem could also be used. Hence a simple Inverse Laplace Transform of these two components yields,

$$y = e^{-2t} - e^{-3t}. \tag{6.96}$$

One then needs to find z as well. Although one could also find a second order ODE for z which would then be transformed, this is a lot of work and very much a waste of time. Given that we have y, it is very much quicker to substititute this into the first 1st order ODE, namely $y' + y - z = 0$. to find z. Hence we get,

$$z = -e^{-2t} + 2e^{-3t}. \tag{6.97}$$

Method 2. This begins with taking the Laplace Transform of both ODEs simultaneously. This yields,

$$(s+1)Y - Z = 0, \qquad (s+4)Z + 2Y = 1. \tag{6.98}$$

At this point we may eliminate Z to obtain, not surprisingly,

$$Y = \frac{1}{(s+2)(s+3)}, \tag{6.99}$$

and from here on we follow the analysis of Method 1.

Method 3. This third method uses the language of matrix/vector equations. First we take Laplace Transforms to obtain Eq. (6.98), and then this is written in matrix/vector form.

$$\begin{pmatrix} s+1 & -1 \\ 2 & s+4 \end{pmatrix} \begin{pmatrix} Y \\ Z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{6.100}$$

Missing out a line of working, we may premultiply both sides by the inverse matrix, and this yields,

$$\begin{pmatrix} Y \\ Z \end{pmatrix} = \frac{1}{s^2 + 5s + 6} \begin{pmatrix} s+4 & 1 \\ -2 & s+1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad = \qquad \frac{1}{s^2 + 5s + 6} \begin{pmatrix} 1 \\ s+1 \end{pmatrix}, \tag{6.101}$$

where $s^2 + 5s + 6$ is the determinant of the matrix in Eq. (6.100). Now we need to use partial fractions:

$$Y = \frac{1}{s^2 + 5s + 6} = \frac{1}{s+2} - \frac{1}{s+3} \qquad \Longrightarrow \qquad y = e^{-2t} - e^{-3t} \tag{6.102}$$

and

$$Z = \frac{s+1}{s^2 + 5s + 6} = -\frac{1}{s+2} + \frac{2}{s+3} \qquad \Longrightarrow \qquad z = -e^{-2t} + 2e^{-3t}. \tag{6.103}$$

In the solutions, Eqs. (6.96) and (6.97), we notice that, while $y(0) = 0$, as required, we also have $z(0) = 1$ from this solution. This violation of the initial condition is to be expected because it is the z-equation that has the unit impulse as the forcing term.

6.13 Appendix: Proof of $\mathcal{L}[f * g] = F(s)G(s)$

This is for information and background only.

We start by writing down the definition of the Laplace Transform of $f * g$:

$$\mathcal{L}[f * g] = \int_0^\infty \Big[\int_0^t f(\tau)g(t-\tau)\,d\tau\Big] e^{-st}\,dt. \tag{6.104}$$

The proof proceeds by interchanging the order of integration, but this is not a rectangular region in (t,τ)–space and therefore it is not straightforward. In the inner integral, and for a given given value of t, τ varies between 0 and t. This is illustrated by the horizontal arrow in Fig. 6.11, and therefore the whole region of integration is given by the yellow shading.

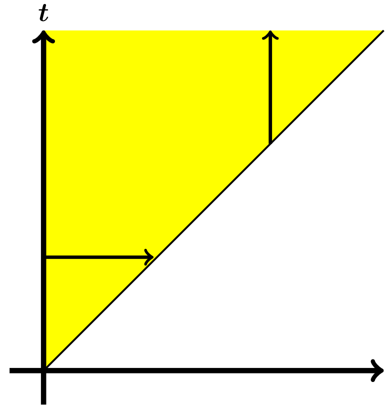


Figure 6.11. The yellow shading denotes the region of integration in Eq. (6.84).

When we interchange the order of integration, then we see from Fig. 6.11 that, for a fixed value of τ, t varies from τ to infinity (see the vertical arrow in Fig. 6.11). Hence Eq. (6.104) becomes,

$$\mathcal{L}[f * g] = \int_0^\infty \Big[\int_\tau^\infty f(\tau)g(t-\tau)e^{-st}\,dt\Big]\,d\tau. \tag{6.105}$$

Now the aim is to change the range of integration in the inner intergral so that the lower limit is zero. So we shall change variable from t to $\hat{t}$ where $\hat{t} = t - \tau$. Thus $d\hat{t} = dt$, and the lower limit in the inner integral will change from $t = \tau$ to $\hat{t} = 0$. So we get

$$\begin{aligned}
\mathcal{L}[f * g] &= \int_0^\infty \int_0^\infty f(\tau)g(\hat{t})e^{-s(\hat{t}+\tau)}\,d\hat{t}\,d\tau \\
&= \int_0^\infty \int_0^\infty \underbrace{f(\tau)e^{-s\tau}}_{\text{function of }\tau}\ \underbrace{g(\hat{t})e^{-s\hat{t}}}_{\text{function of }\hat{t}}\,d\hat{t}\,d\tau \\
&= \Big[\int_0^\infty f(\tau)e^{-s\tau}\,d\tau\Big] \times \Big[\int_0^\infty g(\hat{t})e^{-s\hat{t}}\,d\hat{t}\Big] \\
&= F(s)\,G(s).
\end{aligned} \tag{6.106}$$

Poetry in motion.

6.14 Standard results and theorems involving Laplace Transforms

I have listed below a set of results and theorems which are useful. Those theorems and transforms whose derivation could be in the exam are typeset in **red**. Not all of these results form part of the unit, so the items in the black font will either be quoted on the exam paper itself, or else will not be part of the exam.

Definition: $F(s) = \mathcal{L}[f(t)] = \int_0^\infty f(t)e^{-st}\, dt$

$\mathcal{L}\left[\frac{df}{dt}\right] = sF(s) - f(0)$

$\mathcal{L}\left[\frac{d^2 f}{dt^2}\right] = s^2 F(s) - f'(0) - sf(0)$

$\mathcal{L}\left[\frac{d^3 f}{dt^3}\right] = s^3 F(s) - f''(0) - sf'(0) - s^2 f(0)$

$\mathcal{L}[tf(t)] = -\frac{dF}{ds}$

$\mathcal{L}[t^2 f(t)] = \frac{d^2 F}{ds^2}$

$\mathcal{L}[t^n f(t)] = (-1)^n \frac{d^n F}{dx^n}.$

$\mathcal{L}\Big[\int_0^t f(x)dx\Big] = \frac{1}{s}F(s)$

Scaling theorem: $\mathcal{L}[f(at)] = \frac{1}{a}F\Big(\frac{s}{a}\Big)$

s-shift theorem: $\mathcal{L}[e^{-at}f(t)] = F(s+a)$

t-shift theorem: $\mathcal{L}[H(t-a)f(t-a)] = e^{-sa}F(s)$

Convolution theorem: $\mathcal{L}[f * g] = F(s)G(s)$ where $f * g = \int_0^t f(\tau)g(t-\tau)d\tau = \int_0^t f(t-\tau)g(\tau)d\tau$

Initial value theorem: $\lim_{s\to\infty} sF(s) = f(0)$

Final value theorem: $\lim_{s\to 0} sF(s) = \lim_{t\to\infty} f(t)$

6.15 Table of standard Laplace Transforms of functions

Each of the following could appear on the exam paper — all are derivable directly by applying the Laplace Transform definition, but some may also be derived using one of the shift theorems or the convolution theorem.

$f(t)$	$F(s)$	$f(t)$	$F(s)$
1	$\frac{1}{s}$	$e^{-at}\cos bt$	$\frac{s+a}{(s+a)^2+b^2}$
t	$\frac{1}{s^2}$	$\sinh bt$	$\frac{b}{s^2-b^2}$
t^2	$\frac{2}{s^3}$	$\cosh bt$	$\frac{s}{s^2-b^2}$
t^n	$\frac{n!}{s^{n+1}}$	$e^{-at}\sinh bt$	$\frac{b}{(s+a)^2-b^2}$
e^{at}	$\frac{1}{s-a}$	$e^{-at}\cosh bt$	$\frac{s+a}{(s+a)^2-b^2}$
e^{-at}	$\frac{1}{s+a}$	$t\sin bt$	$\frac{2bs}{(s^2+b^2)^2}$
$t^n e^{-at}$	$\frac{n!}{(s+a)^{n+1}}$	$t\cos bt$	$\frac{s^2-b^2}{(s^2+b^2)^2}$
$\sin bt$	$\frac{b}{s^2+b^2}$	$H(t)$	$\frac{1}{s}$
$\cos bt$	$\frac{s}{s^2+b^2}$	$H(t-a)$	$\frac{e^{-sa}}{s}$
$e^{-at}\sin bt$	$\frac{b}{(s+a)^2+b^2}$	$\delta(t-a)$	e^{-as} $(a>0)$

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME12002

Problem Sheet 1 — Complex Numbers

1. Simplify the following:

(a) j^3, **(b)** j^4, **(c)** j^5, **(d)** j^{10}, **(e)** j^{2023}, **(f)** $(1+j)(2+j)$, **(g)** $(1+j)(1-j)$,
(h) $(2+j)(1+3j)+(2-j)(1-3j)$, **(i)** $(2+j)(2+3j)-(2-j)(2-3j)$, **(j)** $2/(1-j)$,
(k) $(3+j)/(4+3j)$, **(l)** $(1+j)^2$, **(m)** $(1+j)^{100}$.

2. Find $(1+\sqrt{3}j)^3$, and hence find $(1+\sqrt{3}j)^{60}$.

3. Find the modulus and argument of each of the following complex numbers, and hence write them in complex exponential form.

(a) $2+3j$, **(b)** $-j$, **(c)** $40+9j$, **(d)** -5, **(e)** $1-100j$, **(f)** $1+\sqrt{3}j$.

4. Convert the following numbers from complex exponential form to Cartesian form and write in as simple form as possible.

(a) $e^{j\pi/3}$, **(b)** $4e^{2\pi j/3}$, **(c)** $\sqrt{2}e^{3\pi j/4}$.

5. Use de Moivre's theorem to find $\cos 4\theta$ and $\sin 4\theta$ in terms of $\cos\theta$ and $\sin\theta$. [Note that $\cos 4\theta$ may be written in terms of cosines only or sines only; find both forms.] Also find $\cos 5\theta$ and $\sin 5\theta$ in terms of $\cos\theta$ and $\sin\theta$.

Use your result for $\sin 5\theta$ to find an analytical expression for $\sin 36°$.

6. Evaluate the following roots of complex numbers:

(a) $(5+12j)^{1/2}$, **(b)** $(-16)^{1/4}$, **(c)** $1^{1/5}$, **(d)** $(-1)^{1/100}$,
(e) $(1+2j)^{2/7}$, **(f)** $(336+527j)^{1/4}$, **(g)** $(2j)^{1/2}$, **(h)** $(-15+8j)^{3/5}$.

7. Solve the following quadratic equations and plot the roots in the Argand diagram (or the complex plane). Find the modulus and argument of each root, and hence write them in complex exponential form.

(a) $x^2+2=0$ **(b)** $y^2+2jy-2=0$ **(c)** $z^2+2\sqrt{2}z+2-4j=0$.

8. When sketched in the complex plane, the three complex numbers, $z_1=1+j$, $z_2=3+j$ and $z_3=2$, clearly form a right-angled triangle. Multiply each of these numbers by j and sketch the results. How has the original triangle been transformed? Now multiply z_1, z_2 and z_3 by $1-j$; what is the nature of the transformation this time?

The following questions are for interest only and are definitely not examinable. If you are bored, then do have a go at them!

9. de Moivre's theorem is used to express $\cos n\theta$ and $\sin n\theta$ in terms of $\cos\theta$ and $\sin\theta$ when n is an integer. The equivalent for the hyperbolic sine and cosine is not quite as easy to write down. However, write $\cosh\theta$ in terms of e^{θ} and $e^{-\theta}$ and square both sides of the equation. Hence find $\cosh 2\theta$ in terms of $\cosh\theta$.

Further, can you think of a quick way of determining a simple expression for $\sinh 2\theta$ in terms of $\sinh\theta$ and $\cosh\theta$? Also, find an expression for $\sinh 3\theta$ in terms of powers of $\sinh\theta$, and an expression for $\cosh 3\theta$ in terms of powers of $\cosh\theta$.

10. Leibniz, he of the invention-of-calculus fame, came late to complex numbers. He found that

$$\sqrt{1+\sqrt{-3}}+\sqrt{1-\sqrt{-3}}=\sqrt{6},$$

but no-one was surprised! How does one prove this without finding the square roots of complex numbers?

11. Express $z=\frac{1}{2}(\sqrt{3}+1)+\frac{1}{2}(\sqrt{3}-1)j$ in complex exponential form. Now find the argument in terms of degrees, and hence find the first integer power of z for which it is a real value.

12. Some very weird ones. Not examinable. You may need to resort to some lateral thinking...

(a) $\ln(-1)$, **(b)** $\ln j$, **(c)** j^j, **(d)** $y=\cos^{-1}2$.

13. Definitely a maths department type of question, but let's fill up space. A Pythagorean triple is a set of three integers which satisfy, $n^2+m^2=q^2$. Now if a^2+b^2 is such a square, as is c^2+d^2, then the product, $(a^2+b^2)(c^2+d^2)$ may be written as the sum of the two squares, u^2+v^2, in two different ways. I've only just recently discovered this result and my initial reaction was disbelief! Your task is to prove it but I'll give a hint or two.

One may factorise (a^2+b^2) into complex factors (think complex conjugates). This means that $(a^2+b^2)(c^2+d^2)$ may be split into four complex factors. Now there are two different ways of pairing up these factors before they are multiplied. The rest is up to you.... Test out your result using

$$(3^2+4^2)(5^2+12^2).$$

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME12002

Problem Sheet 2 — Differentiation

1. Use the 'small increment' method (i.e. the approach using limits) described in the lecture notes to find the derivative of the following functions: [NOTE: **(i)** this type of question will *not* appear in the exam; **(ii)** part (d) begs the question – you'll see what I mean; **(iii)** there's a sneaky trick that you'll need to find for part (e) which is related in some way to finding z^{-1} in complex numbers.]

(a) x^3, (b) x^4, (c) x^{-1}, (d) $\sin x$, (e) $x^{1/2}$.

2. Find the derivatives of the following functions with respect to x:

(a) $4\sin x + 2x$, (b) $4e^{2x} + 5x^{-1}$, (c) $(bx)^{-1}$, (d) $-4 - 5x^{-2}$, (e) e^{3x-4}, (f) $\ln|2x^3|$, (g) $|x|$, (h) $\sin|x|$.

3. Find the derivatives of the following functions with respect to t:

(a) $t\sin t$, (b) $t^{-2}e^{3t}$, (c) $t\ln t - t$, (d) $te^{-t}\cos 2t$, (e) $\sin 2t \sinh 3t$, (f) $|t|\,\sin|t|$.

4. Differentiate the following problems with respect to t:

(a) e^{t^2}, (b) $\sqrt{1+t^2}$, (c) $(1+\sqrt{t})^2$, (d) $\sin[\sin(\sin t)]$, (e) $\tan(t^{1/2})$, (f) $e^{-\sin t^2}$, (g) $(\sin t)^{1/2}$, (h) $|t|^b$.

5. Using the quotient rule, differentiate the following with respect to x:

(a) $\tan(ax)$, (b) tanh(ax), (c) cosec (ax), (d) $e^x/(1+x)$, (e) e^{3x}/x^2, (f) $x/(1+x^2)$.

Use $\tan = \sin / \cos$, $\tanh = \sinh / \cosh$ and cosec $= 1/\sin$.

6. Find an expression for dy/dx in the following cases:

(a) $y^2 + y = x$, (b) $\sin(xy) = x$, (c) $\ln|y| = y - \cos x$.

7. Find the derivatives with respect to x of the following functions. You will need to use more than one of the above rules in some cases. Part (j) is rather lengthy.

(a) $\sin^{-1}(ax+b)$ (hint: let $y = \sin^{-1}(ax+b)$, find x in terms of y and then differentiate),

(b) $\sin^{-1}(\sin 2x)$, (c) $e^{x\sin x}$, (d) $(\sin x)e^{x^2}$,

(e) 2^x (hint: first show that $2^x = e^{x\ln 2}$), (f) x^x, (g) $\log_{10}|x|$,

(h) $\sinh^{-1}(ax+b)$, [N.b. this is an inverse sinh, not a reciprocal]

(i) $xe^{(x/\sqrt{1+x^2})}$, (j) $\sin(x^2)e^{x\sin x}/(1+x^2)$.

8. **[Lengthy and challenging.]** Find the first, second and third derivatives of $x^n e^{ax}$, where we can assume that $n > 3$. Can you write down a compact expression for the m^{th} derivative of this function?

9. **[Challenging.]** This one is the strangest one.... By considering a simple sketch it is easy to be convinced of the fact that $dy/dx = 1/(dx/dy)$. Use this result and the chain rule to find the appropriate formula for d^2y/dx^2 in terms of d^2x/dy^2. Check that your final formula is correct by applying it to $y = \ln x$ (for $x > 0$) and to $y = x^2$.

Department of Mechanical Engineering, University of Bath

ME12002 Engineering Mathematics S1 ME12002

Problem Sheet 3 — Differentiation — Critical points

1. Find the critical points of the following functions. Which are maxima and which are minima? Find the values of the functions at these points and sketch the functions.

(a) $f(x) = 2x^2 - 3$,

(b) $g(t) = t^2(t^2 - 1)$,

(c) $h(y) = ye^{-y}$,

(d) $F(x) = x^2 \sin^2 x$,

(e) $G(x) = x^4 + 2x^3 - 2x - 1$,

(f) $H(x) = x^2 e^{-x^2}$,

(g) $\theta(z) = 3z^4 - 2z^6$,

(h) $\Phi(\alpha) = (\alpha^2 - 4\alpha - 20)e^{-\alpha}$.

2. Find and classify any critical points of the function,

$$y = 10x^6 - 36x^5 + 45x^4 - 20x^3 + 1.$$

3. Find and classify any critical points of the function,

$$y = e^x/(x^2 + 1).$$

[Hint, this becomes very very messy after the first derivative. Therefore I would suggest that one should set $y' = (x - 1)^2 f(x)$ at that point before obtaining any higher derivatives.]

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME12002

Problem Sheet 4 — Integration

1. Find the following integrals:

(a) $\int x^3\,dx$, (b) $\int x^{-5}\,dx$, (c) $\int e^{-4t}\,dt$, (d) $\int \cosh 2y\,dy$,

(e) $\int_0^1 (2x+3)\,dx$, (f) $\int_{-1}^2 (4x^2-3x)\,dx$, (g) $\int_0^{\pi/4} (2\cos\alpha+\sin\alpha)\,d\alpha$.

2. Find the following integrals:

(a) $\int (x+3)^{-2}\,dx$, (b) $\int_0^1 2xe^{x^2}\,dx$ (hint: substitute $z=x^2$),

(c) $\int \frac{1}{x^2+a^2}\,dx$ (hint: substitute $x=a\tan\theta$),

(d) $\int_0^{\pi/2} \frac{\sin x}{1+\cos^2 x}\,dx$ (hint: substitute $z=\cos x$),

(e) $\int \frac{x}{1+x^2}\,dx$, (f) $\int \cos x\cos(\sin x)\,dx$, (g) $\int \theta^2\sin(\theta^3)\,d\theta$, (h) $\int_0^{\pi/2} \sqrt{\sin t}\,\cos t\,dt$,

(i) $\int \Big[f(x)\Big]^n f'(x)\,dx$, (j) $\int (ff''+f'f')\,dx$ (Try to find the integral by inspection),

(k) $\int_0^1 \sqrt{1-x^2}\,dx$, (l) $\int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx$,

(m) $\int_0^1 (1+\ln x)x^x\,dx$ [you will need to rewrite x^x in an alternative form (hint: take logs), and to assume that $\lim_{x\to 0} x^x = 1$].

3. Use partial fractions to simplify the integrand in the following integrals, and hence find those integrals:

(a) $\int \frac{2t+3}{t^2+3t+2}\,dt$, (b) $\int \frac{1}{t^2+3t+2}\,dt$, (c) $\int \frac{1}{t^3+t}\,dt$, (d) $\int \frac{t+3}{t^3+3t^2+2t}\,dt$,

(e) $\int \frac{2t+1}{t^2(t+1)}\,dt$, (f) $\int \frac{t^2-2}{(t+1)(t^2+2t+2)}\,dt$, (g) $\int \frac{1}{t(t^2+1)^2}\,dt$.

Please note that parts (f) and (g) are quite lengthy.

4. Find the following integrals which involve top-heavy quotients. Do this by first performing a kind of long division.

(a) $\int_0^1 \frac{t^2+4t+5}{t+1}\,dt$, (b) $\int \frac{t^3+t^2-3t-5}{t^2+3t+2}\,dt$,

(c) $\int \frac{3t^3+t^2-t-2}{t+1}\,dt$, (d) $\int_0^1 \frac{t^4(1-t)^4}{1+t^2}\,dt$.

Although it will take quite a while to solve part (d), you will know if the answer is correct because it should give you an "Oh, that's really interesting" moment.

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME12002

Problem Sheet 5 — Integration by Parts and some miscellanous ideas.

1 Obtain the following integrals.

(a) $\displaystyle\int x^5 \cos ax\,dx$, (b) $\displaystyle\int x^5 \sin ax\,dx$, (c) $\displaystyle\int x^5 e^{ax}\,dx$, (d) $\displaystyle\int x^5 e^{ajx}\,dx$.

(e) Can you find the integral, $\displaystyle\int x^5 e^{-ax}\,dx$, directly from the answer to part (c)?

(f) Have you seen the real part of the answer to Q1d before? And the imaginary part?

2. Evaluate the following integrals by integrating by parts twice.

(a) $\displaystyle\int \sin ax \sinh bx\,dx$, (b) $\displaystyle\int \sin ax \cosh bx\,dx$, (c) $\displaystyle\int_0^\infty e^{-ax}\cos\omega x\,dx$, (d) $\displaystyle\int_0^\infty e^{-ax}\sin\omega x\,dx$.

Can you think of an easy way of doing the last two integrals simultaneously which doesn't involve integration by parts? [Hint: consider the solution to Q1f.]

3. Evaluate

(a) $\displaystyle\int \Big[\ln|x|\Big]^2 dx$, (b) $\displaystyle\int \Big[\ln|x|\Big]^3 dx$, (c) $\displaystyle\int \Big[\ln|x|\Big]^n dx$, (d) $\displaystyle\int_0^1 \frac{\ln|x|}{x^{1/2}}\,dx$ (use: $x^{1/2}\ln x \to 0$ as $x \to 0^+$),

(e) $\displaystyle\int_0^2 x^3 \ln|x|\,dx$, (f) Evaluate Q3a by first using the substitution $x = e^y$.

4. The mandatory silly integral: $\displaystyle\int \sin(x)\ln\Big[\tan\Big(\frac{x}{2}+\frac{\pi}{4}\Big)\Big]\,dx.$

5. Use integration by parts once to obtain a formula for $I_n = \displaystyle\int_0^\infty x^n e^{-ax}\,dx$ in terms of I_{n-1} — such a formula is called a **Reduction Formula** or a **Recurrence Relation**. Find I_0 directly, and use the reduction formula to evaluate I_6. Check your answer by evaluating I_6 using integration by parts.

6. Something weird. Evaluate the following indefinite integral using one integration by parts where you'll choose to integrate f' first:

$$I = \int \frac{1}{f}\,f'\,dx.$$

Can you explain why the answer is incorrect? What happens if you choose to integrate between $x = a$ and $x = b$?

7. Also weird. Define $I(a)$ according to $I(a) = \displaystyle\int_0^\infty e^{-ax}\,dx$, and evaluate this integral. Clearly the value taken by $I(a)$ depends on a, and therefore we can differentiate it with respect to a. Do this and find $I'(a)$ both as an integral and as a function of a. Continue to differentiate in this way with the aim of eventually finding $\displaystyle\int_0^\infty x^4 e^{-ax}\,dx$.

Using this idea and the solution to Q2c, find $\displaystyle\int_0^\infty x e^{-ax}\cos\omega x\,dx$, the integral of the product of three functions.

8. Evaluate the mean and RMS of the following functions

(a) $f(t) = t^2 \quad (0 \le t \le 1)$,

(b) $f(t) = \sin t \qquad (0 \le t \le 2\pi)$,

(c) $f(t) = |\sin t| \qquad (0 \le t \le 2\pi)$,

(d) $f(t) = e^{-t} \qquad (0 \le t \le 1)$.

9. This is a set of miscellaneous questions and you may wish to wait until the revision period to tackle them.

I won't give too many hints about how to do them! Sorry. Some are substitutions but (e) should be done both as an integration by parts and by using a suitable multiple angle formula. Some will require the use of more than one method. Questions (g) and (h) are trick questions — it all depends on how quickly you can see the trick. As a guide (or perhaps a challenge!), when I saw (h) for the very first time (on a youtube video) I managed 5 seconds, and so it was immediately marked up as something that I just had to try out on you!

(a) $\displaystyle\int_0^{\pi^2} \sin\sqrt{x}\,dx$, (b) $\displaystyle\int_0^{\infty} e^{-\sqrt{x}}\,dx$,

(c) $\displaystyle\int_0^{\infty} e^{-x^{1/4}}\,dx$, (d) $\displaystyle\int_1^{e} \sin(\ln x)\,dx$,

(e) $\displaystyle\int_0^{\pi/2} \sin x\cos 5x\,dx$, (f) $\displaystyle\int \frac{x+3}{\sqrt{x^2+6x+10}}\,dx$,

(g) $\displaystyle\int_{-\infty}^{\infty} e^{-x^4}\sin 3x\,dx$, (h) $\displaystyle\int_0^{\pi/2} \frac{\sqrt{\sin x}}{\sqrt{\sin x}+\sqrt{\cos x}}\,dx.$

(i) $\displaystyle\int \frac{\sqrt{1-x}}{\sqrt{x}}\,dx$, (j) $\displaystyle\int \frac{1}{\sqrt{x+1}+\sqrt{x}}\,dx$, (k) $\displaystyle\int_0^{\infty} \frac{1}{1+e^{x}}\,dx$.

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME12002

Problem sheet 6 — Series – Binomial Series

1. Find the binomial expansions for **(a)** $(1+x)^{-2}$ and **(b)** $(1+x)^{-3}$, and express the final solution in summation form. Check your result for part (b) by differentiating the result for part (a). [You may need to alter the summation variable to get a perfect match.]

2. Following on from the previous question, differentiate the Binomial Series form of $(1+x)^{-3}$ to find the series for $(1+x)^{-4}$. Confirm this by using the general expression for the Binomial Series.

3. Find the binomial expansions for

$$\textbf{(a)}\ (1+x)^{-1/2}, \qquad \textbf{(b)}\ (1+x)^{1/2} \qquad \text{and} \qquad \textbf{(c)}\ (1+x)^{-3/2}.$$

Note that part (a) is in the lecture notes, but do resist the temptation to look before trying this out! Check your answers to (b) and (c) by the respective integration and differentiation of the result of (a).

4. Express $(1+x)^{-1/3}$ as a Binomial Series. Note that this one cannot be expressed in summation form....unless you decide to define your own notation! [Hint: do a google search for what are called double factorials.]

5. Find the Binomial Series representations for **(a)** $(1+2x)^{-1}$ and **(b)** $(3+2x)^{-1}$.

6. Obtain the binomial series for $(1+x^2)^{-1}$ and hence find a power series representation for $\tan^{-1} x$. Use this result to show that $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$.

7. **(a)** Use the chain rule (i.e. implicit differentiation) to find the derivative of $y = \tanh^{-1} x$. Use this result and the Binomial Series to find a power series representation for y.

(b) At some point in part (a) you will have an expression for x as an explicit function of y; replace the sinh and cosh terms by the correct exponentials, solve for e^y and hence find a different expression for y from the one given initially.

(c) Use the results of parts (a) and (b) to show that,

$$\frac{1}{2} + \frac{(\frac{1}{2})^3}{3} + \frac{(\frac{1}{2})^5}{5} + \frac{(\frac{1}{2})^7}{7} + \cdots = \frac{1}{2}\ln 3.$$

8. **[For fun only!]** Use the binomial expansion of $(1+x)^n$, where n is a positive integer, to show that

$$\sum_{i=0}^{n} \binom{n}{i} = 2^n \qquad \text{and} \qquad \sum_{i=0}^{n} (-1)^i \binom{n}{i} = 0.$$

9. **[For fun only!]** Write out the Binomial Expansions for both $(1+x)^n$ and $(1+x^{-1})^n$ where n is an integer. Use these results to show that

$$\sum_{r=0}^{n} \left[\binom{n}{r}\right]^2 = \binom{2n}{n}.$$

[Hint: where might the Binomial coefficient on the right hand side have come from?]

10. **[For fun only!]** Try to determine a general formula for the coefficient of $x^i y^j z^k$ in the expansion of $(x+y+z)^n$, where $i+j+k=n$. This expansion is called a Trinomial expansion.

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME120021

Problem sheet 7 — Series – Taylor's series, convergence, l'Hôpital's rule

1. Find a cubic polynomial, $f(x)$, which is such that
   (a) $f(1) = 2$, $f'(1) = -1$, $f''(1) = 0$ and $f'''(1) = 6$;
   (b) $f(0) = 3$, $f'(0) = 0$, $f''(0) = -6$ and $f'''(0) = 12$;
   (c) $f(2) = 7$, $f'(2) = 12$, $f''(2) = 18$ and $f'''(2) = 12$.
   In parts (a) and (c) rewrite the polynomials as finite power series in x.

2. **(A strange one!)** Find a way to use Taylor's series to rewrite $y = 1 + x + x^2 + x^3$ as a power series in $(x - 1)$. By this I mean that $y = \sum_{n=0}^{3} a_n (x-1)^n$.

3. Find the Taylor's series of $y = (1 + x)^{-2}$ about $x = 0$ and the one about $x = 1$. What are their radii of convergence?

4. Find the Taylor's series of $(1 + x)^{-1}$ about $x = 2$. Find its radius of convergence. It is possible to use the Binomial series to find this power series, but it may need a little bit of thinking...

5. Find the Maclaurin series for $\sin ax$ and use this to find the following limits:
$$\text{(a)} \quad \lim_{x\to 0} \frac{\sin(ax)}{x}, \qquad \text{(b)} \quad \lim_{x\to 0} \frac{\sin(ax) - ax}{x^3}.$$

6. Find the Maclaurin's series for both (a) $\cos(ax)$ and (b) $\cosh(ax)$. Use these results (not l'Hôpital's rule) to determine,
$$\lim_{x\to 0} \frac{\cos(ax) + \cosh(ax) - 2}{x^4}.$$
   You may, of course, use l'Hôpital's rule to confirm your solutions.

7. Find the Maclaurin series representation of $\ln(2 + x)$. What is its radius of convergence?

8. In the lecture notes we found a power series representation of e^{-x^2} by first finding a Maclaurin expansion of e^{-x} followed by the replacing of x by x^2. It was also mentioned that the direct determination of the Maclaurin series of e^{-x^2} is much more challenging. So here's your chance! Use the definition of the Maclaurin series to determine the power series of e^{-x^2} up to and including the x^6 term. If you have the stamina, go as far as the x^8 term!

9. Determine the convergence properties of the following numerical series.

(a) $\dfrac{2}{1^3} + \dfrac{3}{2^3} + \dfrac{4}{3^3} + \dfrac{5}{4^3} + \cdots$ (b) $\dfrac{1}{10} - \dfrac{2}{11} + \dfrac{3}{12} - \dfrac{4}{13} + \cdots$ (c) $\displaystyle\sum_{k=1}^{\infty} \frac{k^p}{k!} \quad (p > 0)$

(d) $\displaystyle\sum_{k=1}^{\infty} \frac{k!}{(2k)!}$ (e) $\displaystyle\sum_{k=1}^{\infty} \frac{k!\,k!}{(2k)!}$ (f) $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{n^n}$

In part (f) you will need to use the result that $\displaystyle\lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n = e$, which is to be proven in a later question.

10. Determine the radii of convergence of the following power series.

(a) $\displaystyle\sum_{n=1}^{\infty} x^n$ (b) $\displaystyle\sum_{n=1}^{\infty} nx^n$ (c) $\displaystyle\sum_{n=1}^{\infty} n^{100}x^n$

(d) $\cosh\sqrt{x} = 1 + \dfrac{x}{2!} + \dfrac{x^2}{4!} + \dfrac{x^3}{6!} + \cdots$ (e) $\cos x = 1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} \cdots$ (f) $\displaystyle\sum_{n=1}^{\infty} (-1)^n 2^n x^{2n}$

(g) $\displaystyle\sum_{n=1}^{\infty} n!\, x^n$ (h) $\displaystyle\sum_{n=1}^{\infty} \frac{n!\, x^n}{n^n}$ (i) $\displaystyle\sum_{n=0}^{\infty} \frac{(1+2n)}{(1+2^n)} x^{n/3}$

11. Use l'Hôpital's rule to find the following limits.

(a) $\displaystyle\lim_{x\to\frac{1}{2}} \frac{\cos \pi x}{2x-1}$ (b) $\displaystyle\lim_{x\to 1} \frac{\sin \pi x}{\sqrt{x^2-1}}$ (c) $\displaystyle\lim_{x\to 0} \frac{\cosh ax - 1}{x}$

(d) $\displaystyle\lim_{x\to\infty} \frac{x^2+2x+3}{3x^2+2x+1}$ (e) $\displaystyle\lim_{x\to 0} \frac{\sin x - x}{x^3}$ (f) $\displaystyle\lim_{x\to 0} \frac{\sin^2 x - x^2}{x^3}$

(g) $\displaystyle\lim_{x\to 4} \frac{x-4}{\sqrt{x^2+9}-5}$ (h) $\displaystyle\lim_{x\to 0} \frac{\sin(x^2)}{x\tan x}$ (i) $\displaystyle\lim_{x\to 0} \frac{\tan x - x}{x^3}$

(j) $\displaystyle\lim_{x\to 0} \frac{\cos(ax)+\cosh(ax)-2}{x^4}$ (k) $\displaystyle\lim_{x\to 0} \frac{\sinh x + \sin x - 2x}{x^5}$

12. This is not the type of question which I will give in the exam, but I would like to think that it is of interest. Your aim is to use l'Hôpital's rule to find the following three limits:

$$y = \lim_{x\to 0} x^x, \qquad y = \lim_{x\to\infty} x^{1/x}, \qquad \text{and} \qquad y = \lim_{n\to\infty}\Big(1+\frac{1}{n}\Big)^n.$$

None look like candidates for the use of l'Hôpital's rule, but if one takes the natural logarithm and then rearrange the resulting expressions appropriately, then you'll find that $\ln y$ does indeed take a suitable form. Note that, after you take the logarithm, then (i) the first two cases are of the form of ∞/∞, and (ii) the third case will eventually need the substitution $x = 1/n$ where the limit $x \to 0$ is taken. This third one arises when considering compound interest....more information in the solutions.

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME12002

Problem Sheet 8 — ODEs

1. What is the order of the following equations or systems of equations?

In each case rewrite them in first order form. **Do not** try to solve them!

Are these equations/systems linear or nonlinear, and do they constitute Initial Value Problems or Boundary Value Problems? (Primes denote derivatives with respect to t.)

(a) $y'' + ty = 0$ subject to $y(0) = 0$, $y'(0) = 1$.

(b) $y''' + y'' - 2yz = 0$, $z' = ty$ subject to $y(0) = 1$, $y'(0) = 0$, $y'(\infty) = 0$, $z(0) = 0$.

(c) $y'''' + 2(y + y'')^3 y' + y^5 = 1$, subject to $y = y' = y'' = y''' - 1 = 0$ at $t = 0$.

(d) $x'' + 2x - y = 0$, $y'' - x + 2y - z = 0$, $z'' - 3y + 2z = 0$,
subject to $x(0) = 1$, $x'(0) = 0$, $y(0) = y'(0) = 0$, $z(0) = z'(0) = 0$.

(e) $f' = g$, $g'' + fg' + f'g = 0$, subject to $f(0) = 0$, $g(0) = 1$, $g(\infty) = 0$.

2. Solve the following equations by direct integration of both sides.

(a) $y' = \cos t$ subject to $y(0) = 1$, (b) $y' = e^{2t} + 1$ subject to $y(1) = 1$.

3. Use separation of variables to find the solutions to the following ODEs.

(a) $\dfrac{dy}{dt} = \dfrac{4t}{y}$, $y(0) = 1$, (b) $\dfrac{dy}{dt} = 3t^2 y$, $y(0) = 1$,

(c) $\dfrac{dy}{dt} = t(1 + y^2)$ $y(0) = 1$, (d) $\dfrac{dy}{dt} = t^2(1 - y^2)$, $y(0) = 2$,

(e) $\dfrac{dy}{dt} = y - y^2$, $y(0) = 2$, (f) $t^2 \dfrac{dy}{dt} = y - t^3 y$, $y(1) = 1$,

(g) $\dfrac{d^2y}{dt^2} = \dfrac{1}{t}\dfrac{dy}{dt}$, $y(0) = 1$, $y'(1) = 2$.

4. Find the Integrating Factor and hence solve the following 1st order equations.

(a) $\dfrac{dy}{dt} + \dfrac{y}{t} = 1$, (b) $\dfrac{dy}{dt} - \dfrac{y}{t} = 1$, (c) $\dfrac{dy}{dt} + \dfrac{3y}{t} = t^{-2}$, (d) $\dfrac{dy}{dt} + 2ty = 2t$,

(e) $\dfrac{dy}{dt} + y \cot t = 1$, (f) $\dfrac{dy}{dt} + \dfrac{1 + 2t}{t} y = \dfrac{1}{t}$, (g) $\dfrac{dy}{dt} + 4t^3 y = t^3$

(h) $t\dfrac{dy}{dt} + (t + 1)y = t^2$.

5. The following differential equation

$$\frac{dy}{dt} = y^3 - y$$

falls into two different categories. First, it is of variables-separable type, and second it is an example of what is known as a Bernoulli equation.

(i) Use separation of variables, followed by partial fractions to find the solution subject to the initial condition that $y = 1/\sqrt{2}$ when $t = 0$.

(ii) Solve the equation for y again by first using the substitution, $y = z^{-1/2}$ where $z = z(t)$ is a new dependent variable — you will need to use the chain rule for this to find a formula for dz/dt in terms of dy/dt. This substitution should then give you a linear equation for z which may be solved.

6. The general form for Bernoulli's equation is

$$\frac{dy}{dt} + P(t)y^n + Q(t)y = 0,$$

where $P(t)$ and $Q(t)$ are given functions.

(i) Use the substitution $y = z^{\alpha}$, where z is a function of time and where α is constant to be found. After substitution, determine that value of α which reduces the equation for z into one of first order linear form.

(ii) You are now required to solve the ODE,

$$\frac{dy}{dt} + y - y^{-1} = 0, \qquad \text{subject to} \qquad y = 2 \text{ at } t = 0.$$

Use your general Bernoulli result to reduce the ODE to first order linear form and solve it. Check your answer for y by substituting it back into the above equation.

(iii) The above equation may also be solved using separation of variables. Please do it this way as well.

7. Another category of ODE could be called equidimensional. This is an example:

$$\frac{dy}{dx} = \frac{2y^2 + x^2}{2xy}.$$

The method of solution is to substitute $y(x) = x\,v(x)$ to form an ODE for $v(x)$. The resulting equation should then be solvable using separation of variables. Solve the above equation subject to the initial condition, $y = 1$ when $x = 1$. Then check that your solution satisfies the original ODE. [You may also attempt Q4a using the same idea.]

Please note that all of the above could form at least part of an exam question. I will not ask for the derivation required in Q6i. For equations of Bernoulli type and for those in equidimensional form I will give the required substitutions (as I have in Q5ii and Q7).

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME12002

Problem Sheet 9 — ODEs

Questions 1 and 2 contain equations most of which are of exam standard. Questions 3 and 4 are longer, and while they use some ideas (such as l'Hôpital's rule) which won't be examined in an ODE context, questions of this type may arise and have arisen in past exams. It is best to be guided by past exam papers in this regard.

1. First find the general solution of the following homogeneous equations. Then find the solution which satisfies $y(0) = 1$ and $y'(0) = 0$ (additionally $y''(0) = 0$ for third and fourth order equations and $y'''(0) = 0$ for fourth order equations).

(a) $\dfrac{d^2y}{dt^2} + 5\dfrac{dy}{dt} + 4y = 0$; (b) $\dfrac{d^2y}{dt^2} + 4\dfrac{dy}{dt} + 4y = 0$; (c) $\dfrac{d^2y}{dt^2} + 2\dfrac{dy}{dt} + 5y = 0$;

(d) $\dfrac{d^2y}{dt^2} - 4\dfrac{dy}{dt} + 29y = 0$; (e) $\dfrac{d^3y}{dt^3} + 2\dfrac{d^2y}{dt^2} + \dfrac{dy}{dt} + 2y = 0$;

(f) $\dfrac{d^3y}{dt^3} + \dfrac{d^2y}{dt^2} - 2y = 0$; (g) $\dfrac{d^3y}{dt^3} - 3\dfrac{d^2y}{dt^2} + 3\dfrac{dy}{dt} - y = 0$; (h) $\dfrac{d^4y}{dt^4} + 4y = 0$;

(i) $\dfrac{d^4y}{dt^4} + 5\dfrac{d^2y}{dt^2} + 4y = 0$; (j) $\dfrac{d^4y}{dt^4} + 2\dfrac{d^2y}{dt^2} + y = 0$.

2. Find the general solution of the following inhomogeneous equations.

(a) $\dfrac{d^2y}{dt^2} + 9y = f(t)$ where $f(t)$ takes the following forms: (i) e^{at}, (ii) t^3, (iii) $\cos at$, (iv) $\cos 3t$.

(b) $\dfrac{d^2y}{dt^2} + 2\dfrac{dy}{dt} + 2y = f(t)$ where $f(t)$ takes the following forms: (i) e^{at}, (ii) t^2, (iii) $\cos at$.

(c) $\dfrac{d^2y}{dt^2} - 7\dfrac{dy}{dt} + 12y = f(t)$ where $f(t)$ takes the following forms: (i) e^{2t}, (ii) e^{3t}, (iii) t^2 (iv) $\cos at$.

(d) $\dfrac{d^3y}{dt^3} + 3\dfrac{d^2y}{dt^2} + 3\dfrac{dy}{dt} + y = t^3e^{-t}$. (Use the standard way first, then use the substitution, $y(t) = z(t)e^{-t}$.)

3. Solve the equation

$$\frac{d^2y}{dt^2} + \frac{dy}{dt} - 6y = e^{2t}, \qquad y(0) = 0, \quad \left.\frac{dy}{dt}\right|_{t=0} = 0,$$

using standard CF/PI methods.

Now we will attempt to solve the same equation using a slightly different method. First solve,

$$\frac{d^2y}{dt^2} + \frac{dy}{dt} - 6y = e^{at}, \qquad y(0) = 0, \quad \left.\frac{dy}{dt}\right|_{t=0} = 0,$$

where $a \neq 2$. Now let $a \to 2$ in the answer, and use L'Hôpital's rule to recover the solution when $a = 2$.

4. In this question the equation,

$$\frac{d^2y}{dt^2} + 5\frac{dy}{dt} + 6y = te^{-2t},$$

will be solved in two different ways.

(a) Use the Complementary Function/Particular Integral approach.

(b) Use the substitution $y = z(t)e^{-2t}$ to simplify the equation. You should then be able to integrate the resulting equation once with respect to t. The final first order equation for z may then be solved using the CF/PI approach.

5. All of the above questions have resulted in either a general solution or else a solution which satisfies a given set of initial conditions. By contrast, this question concerns the solution of a second order ODE as a boundary value problem, and we shall consider two different types of BVP. The ODE is,

$$\frac{d^2y}{dt^2} - 3\frac{dy}{dt} + 2y = 2t + 3.$$

(a) Use the Complementary Function/Particular Integral approach to find the general solution.

(b) Find the specific solution for which $y(0) = 0$ and $y(1) = 0$ are satisfied.

(c) Now find the specific solution which has a unit periodicity. By this, I mean that the solution must satisfy the conditions, $y(0) = y(1)$ and $y'(0) = y'(1)$.

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME10305

Problem Sheet 9b (Extension) — ODE solutions

This problem sheet contains questions all of which are over and above what is required in the exams. You may therefore treat this sheet as **purely optional**. Nevertheless, everything that is given here may be completed using what has been taught in Maths 1 and Maths 2, with a few hints and nudges along the way.

1. One notation for dy/dt which is sometimes used in textbooks and research papers is Dy. In essence, d/dt and D are directly equivalent to one another and are simply alternative ways of writing down the same thing. Given this, one may try to determine the inverse of D in the following way. Given that

$$\frac{dy}{dt} = f(t) \qquad \Rightarrow \qquad y = c + \int f(t)\,dt,$$

then we may define D^{-1} as follows,

$$Dy = f \qquad \Rightarrow \qquad y = \frac{1}{D}f(t) \qquad = \qquad c + \int f(t)\,dt.$$

In other words, D^{-1} is equivalent to an indefinite integral plus an arbitrary constant.

(a) Now consider the differential equation, $(D+a)y = f(t)$. Rewrite this in the usual way (i.e. $dy/dt + ay = f(t)$) and use the integrating factor approach to find y, not forgetting the arbitrary constant. When this is done, identify which part of your solution forms the Complementary function and which the Particular Integral. What you have written is then the equivalent of

$$y = \frac{1}{D+a}f(t),$$

and it defines the meaning of $(D+a)^{-1}$.

(b) Let us extend the result of Q1a to the following differential equation,

$$\frac{d^2y}{dt^2} + (a+b)\frac{dy}{dt} + aby = f(t).$$

This may also be written as

$$D^2y + (a+b)Dy + aby = f(t), \qquad \text{or} \qquad (D+a)(D+b)y = f(t).$$

If we now set $z = (D+b)y$ then $(D+a)z = f(t)$.

First solve $(D+a)z = f(t)$ for z by applying the result of Q1a directly. Then solve $(D+b)y = z$ to find y. Keep your wits about you on this one — the final answer will involve a double integral.

(c) Now we will modify slightly the answer given in Q1b for the case when $a = b$, which (in the terminology of the lectures) is a repeated-λ case. You should find that some integrals will simplify slightly.

(d) Apply the formula found in Q1b to solve the two equations,

$$y'' + 3y' + 2y = e^t \qquad \text{and} \qquad y'' + 3y' + 2y = e^{-t}.$$

(e) Suppose that we are solving a third order ODE with $f(t)$ on the right hand side. If it is written in the form,

$$(D+a)(D+b)(D+c)y = f(t),$$

and given the form of the answer Q1b, can you guess what the solution is?

2. The aim for this question is to solve $y' + ay = 1$ subject to $y(0) = 0$ using Taylor's series. First, write down a general expression for the Taylor's series about $t = 0$ for the function $y(t)$ — this is *not* the solution because we don't yet know the value of all of the derivatives of y at $t = 0$. However, we may substitute the initial value of y into the governing equation to find $y'(0)$. Now differentiate the governing equation once; this will allow us to find $y''(0)$. Differentate again and hence find $y'''(0)$. The pattern should now be clear. Hence write down the Taylor's series of the solution. Can you identify it?

3. This question was devised while I was watching the film, Gravity, en route to India, with only a thin skin of aluminium between me and a quarter of an atmosphere of air pressure at -50°C while travelling ay 500mph six miles above the ground. I am not sure that I like disaster movies while flying! Suppose that Sandra Bullock and George Clooney are stranded in space, 20m apart and stationary relative to each another, i.e. 10m from their centre of gravity (I am assuming that they have the same mass!). How long will it take for gravitational attraction to cause the couple get close enough together that they may grasp each other's hand?

So the governing equation for $r(t)$ is, $x(t)$ is the distance of one of them from the other, how long will it take to reduce $x = 20$m to $x = 1$m as gravitational attraction draws them together? The governing equation, a nonlinear second order equation, is

$$m_1 \frac{d^2 r}{dt^2} = -\frac{m_1 m_2 G}{r^2},$$

where $m_1 = m_2 = 60$kg are their masses, and $G = 6.67408 \times 10^{-11}$N m^{2}kg^{-2} is the gravitational constant. Here, $r(t)$ is the distance between the two masses, m_1 and m_2. To be checked and corrected, and likewise the solution.

(a) Nothing in our lectures hints about how to solve this! However, d^2x/dt^2 is the same as dv/dt where $v = dx/dt$. Use the chain rule to show that

$$\frac{dv}{dt} = v\frac{dv}{dx}.$$

Use this substitution to solve for v in terms of x. Apply the initial condition, namely that at $t = 0$, $x = x_0$ and $v = 0$ (we'll keep the initial separation general for now).

(b) Now that we have v in terms of x, it is possible to solve this by first using the substitution, $x = x_0 \cos^2\theta$, to obtain an equation for θ in terms of t. This equation may be solved to find t in terms of θ. Don't let this worry you, for the whole point is that you need to find the time corresponding to a given distance. Now use $x_0 = 20$ and let $x = 1$ (it's probably best to find the corresponding value of θ here); what is the time? So how many days does it take for them to be reunited? (Cue suitable sad violin music...)

4. The Cauchy-Euler equation is a different class of linear ODE, and technically it is known as an equi-dimensional equation. The most general second order version is

$$x^2\frac{d^2y}{dx^2} + ax\frac{dy}{dx} + by = 0.$$

There are two ways of solving this equation, the first being to let $y = x^n$ (and then one will eventually be led to an indicial/auxiliary/characteristic equation for n) while the second is to change variables from x to ξ using $x = e^\xi$. Try to solve the equation

$$x^2\frac{d^2y}{dx^2} + 4x\frac{dy}{dx} + 2y = 0$$

using each of these two methods. [Note, when attempting the second, we are changing from dy/dx to $dy/d\xi$, and the chain rule will need to be used. Take care with the transformation of the second derivative — the product rule will be needed!]

[Continued overleaf]

Suppose now that we wish to solve

$$x^2\frac{d^2y}{dx^2}+5x\frac{dy}{dx}+4y=0.$$

The first method given above leads to a repeated value of n and then it isn't obvious how to proceed in this context. So adopt the second method, solve the equation, and this will show how one should proceed when using the otherwise quicker and simpler first method.

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME12002

Sheet 10 (preliminary) — Laplace Transforms

This problem sheet is an experiment to see if you can solve an ordinary differential equation using Laplace Transforms before I even begin lecturing on the topic. It's your choice if you wish to rise to the challenge. I'll try to walk you through it gently.

Here's the definition. If we have a function of time, $f(t)$, then its Laplace Transform is defined to be $F(s)$ where

$$F(s) = \mathcal{L}[f(t)] = \int_0^\infty f(t)e^{-st}\,dt,$$

where $\mathcal{L}[\quad]$ is just a mathematical shorthand for saying the words, 'Laplace Transform of...'. Don't worry about what all of this means — it's really weird — just go with the flow for now.

Q1. Use the definition of the Laplace Transform to find $\mathcal{L}[e^{-at}]$. This should be a nice quick integral, and your answer should come out to be $1/(s+a)$.

Q2. As the aim is to solve a differential equation we had better find an expression for $\mathcal{L}[dy/dt]$.

So let $Y(s) = \mathcal{L}[y(t)]$, and apply the Laplace Transform formula to dy/dt.

You will need one integration by parts to do this, and the final answer will involve both $Y(s)$ and the value of y at $t = 0$, i.e. $y(0)$. In the context of ODEs the value of $y(0)$ represents the initial condition.

Q3. Believe it or not, we are now in a position to solve an ODE. So use both of the above results to find the Laplace Transform of the ordinary differential equation,

$$\frac{dy}{dt} + 2y = e^{-t}, \qquad \text{subject to } y(0) = 2.$$

Once the equation (with the boundary condition) has been transformed, first rearrange it to find $Y(s)$ explicitly, then use partial fractions to simplify that expression, and then finally use the result in Q1 to find the function, $y(t)$, which corresponds to your $Y(s)$.

Q4. Perhaps you should solve the equation using the CF/PI method (i) to ensure that you can after such a long break(!) and (ii) to check that the Laplace Transform solution is correct.

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME12002

Sheet 10 — Laplace Transforms

It is normal in questions on Laplace Transforms to have ready access to the LTs of functions like sinusoids, exponentials and powers. Here, though, I will need you to derive these results.

1. Find the Laplace Transforms of the following functions using the definition of the Laplace Transform (rather than by looking up the result in a table):

(a) e^{3t} (b) e^{-3t} (c) $\cos\omega t$ (d) te^{-3t} (e) t^3 (f) $t\cos\omega t$ (g) $f'''(t)$

(h) The unit pulse: $f(t) = 1$ for $t < 1$, $f(t) = 0$ otherwise (i) $\cosh\omega t$ (j) $t^2 e^{-t}$

(k) $t^{-1/2}$ [Hint: set $x = (st)^{1/2}$ to transform the integral and use the result $\int_0^\infty e^{-x^2}\,dx = \sqrt{\pi}/2$.]

2. Use the Laplace Transform to solve the following equations:

$$\text{(a)} \quad \frac{dy}{dt} + 4y = 6, \qquad y(0) = 2.$$

$$\text{(b)} \quad \frac{d^2y}{dt^2} + 16y = 0, \qquad y(0) = 0, \quad \frac{dy}{dt}(0) = 1.$$

$$\text{(c)} \quad \frac{d^2y}{dt^2} + 4y = 29e^{-5t}, \qquad y(0) = 0, \quad \frac{dy}{dt}(0) = -3.$$

$$\text{(d)} \quad y''' + y'' + 4y' + 4y = 0, \qquad y(0) = 0, \quad y'(0) = 3, \quad y''(0) = -5.$$

[You may also practice on any of the linear constant coefficient equations from the ODEs section of the unit, but note that there may be some awkwardnesses due to the fact that (i) the questions weren't designed for nice LT solutions, (ii) many don't have initial conditions specified, (iii) some of the results derived in the 3rd and 4th Laplace Transform lectures may be of considerable use.]

3. Find the Laplace Transform of $z(t) = \displaystyle\int_0^t y(\tau)\,d\tau$. [Hint: recall that $z'(t) = y(t)$ here.]

4. Find the solution of the ODE, $y'' + 2y' + y = 2e^{-t}$, subject to $y(0) = y'(0) = 0$. [Hint: you may need to consult the solution to Q1j.]

5. Factorise the denominator of the following fractions into complex factors, and use partial fractions to find their Inverse Laplace Transforms: [**Note:** that I won't expect such complex factorisation in the exam.]

(a) $\dfrac{1}{s^2 + b^2}$ (b) $\dfrac{s}{s^2 + b^2}$ (c) $\dfrac{1}{s^2 + 2cs + c^2 + d^2}$ (d) $\dfrac{s + c}{s^2 + 2cs + c^2 + d^2}$.

These results may be used to solve the following equations:

(e) $y'' + 4y' + 5y = 0$, $y(0) = 0$, $y'(0) = 1$.

(f) $y'' + 2y' + 2y = e^{-t}$, $y(0) = 0$, $y'(0) = 0$.

6. Write down the values of the following integrals.

$$\int_{-\infty}^{\infty} \delta(t)e^{2t}\,dt, \qquad \int_{-\infty}^{\infty} \delta(t-1)e^{-t^2}\,dt, \qquad \int_{-\infty}^{\infty} \delta(t-2)\sin\pi t\,dt, \qquad \int_{0}^{\infty} \delta(t+2)t^3\,dt.$$

7. Find the Laplace Transforms of the following functions:

(a) $e^{e^t}\delta(t-1)$, (b) $\sum_{n=0}^{\infty}\delta(t-n) = \delta(t)+\delta(t-1)+\delta(t-2)+\delta(t-3)+\ldots.$

[Look out for the geometric series....]

8. Use the Laplace Transform to solve the following equations:

(a) $\dfrac{dy}{dt} + 3y = \delta(t), \qquad y(0) = 1.$

(b) $\dfrac{d^2y}{dt^2} + 3\dfrac{dy}{dt} + 2y = \delta(t), \qquad y(0) = 1, \quad \dfrac{dy}{dt}(0) = b,$ where b is a constant.

(c) $\dfrac{d^3y}{dt^3} - \dfrac{dy}{dt} = 3\delta(t), \qquad y(0) = 1, \quad \dfrac{dy}{dt}(0) = 0, \quad y''(0) = -1.$

9. Laplace Transforms are perfectly set up to solve Initial Value Problems, but let us try them out on a Boundary Value Problem. The aim, then, is to solve $y'' + y = 0$, subject to $y(0) = 1$ and $y(\frac{1}{2}\pi) = 1$. At the outset, let $y'(0) = c$ and carry out the analysis using this unknown constant. Eventually you will have the opportunity to find c.

Department of Mechanical Engineering, University of Bath

Engineering Mathematics S1 ME12002

Sheet 11 — Laplace Transforms

10. First sketch the following functions, and then Find their Laplace Transforms:

(a) $H(t-a)t^3$ (b) $\sum_{n=0}^{\infty}(-1)^n H(t-n) = H(t) - H(t-1) + H(t-2) - H(t-3) + \ldots$

(c) $H(a-t)$,

[In one case it may be possible to simplify the final answer...]

11. Use the s-shift theorem to find the Inverse Laplace Transform of:

(a) $\dfrac{1}{s+a}$ (b) $\dfrac{2}{(s+a)^3}$ (c) $\dfrac{b}{(s+a)^2+b^2}$ (d) $\dfrac{s}{(s+a)^2+b^2}$.

12. [This is an exam-style question.] Find the Laplace Transforms of both $\cos bt$ and $\sin bt$. Then use the s-shift theorem to write down the Laplace Transforms of $e^{-at}\cos bt$ and $e^{-at}\sin bt$. Hence solve the ODE,

$$y'' + 6y' + 25y = 0$$

subject to $y(0) = 1$ and $y'(0) = 5$.

13. Use the t-Shift Theorem to find the Inverse Laplace Transform of:

(a) $\frac{e^{-as}}{s^3}$ (b) $\frac{e^{-as}}{s+b}$ (c) $\frac{e^{-as}}{s^2+b^2}$ (d) $\frac{e^{-as}}{(s+c)^2+b^2}$.

14. Use the convolution theorem to find the Inverse Laplace Transform of:

(a) $\frac{1}{(s+a)^2}$ (b) $\frac{1}{(s+a)(s^2+b^2)}$ (c) $\frac{1}{(s^2+a^2)^2}$ (d) $e^{-as} \times \frac{1}{s^2}$.

15. Use the convolution theorem to find the solutions to the following equations:

(a) $y' + 3y = e^{-2t}, \quad y(0) = 0$;

(b) $y'' + 5y' + 6y = 0, \quad y(0) = 0, \quad y'(0) = 1$;

(c) $y'' + y = e^{-t}, \quad y(0) = 0, \quad y'(0) = 0$.

16. Solve the system of equations,

$$x' = 2x - 4y + \delta(t),$$
$$y' = 3x - 5y,$$

subject to the initial conditions, $x(0) = y(0) = 0$.

17. Solve the system of equations,

$$x'' + 2x - 2y = \delta(t),$$
$$y'' - x + 3y = 0,$$

subject to the initial conditions, $x(0) = x'(0) = y(0) = y'(0) = 0$. When the final solution has been obtained, determine which of the four initial conditions has been violated, but can you guess this in advance?

18. [This is a project-like question which combines quite a large number of mathematical results.]

The overall aim for this question is to solve the ODE,

$$y' + y = \text{Ш}(t) = \sum_{n=-\infty}^{\infty} \delta(t-n).$$

The unusual symbol, Ш, which I cannot typeset properly(!), is known as the Shah function, and the symbol itself is the Cyrillic character, sha, which mimics the shape of the function. In various contexts it is also known as (i) the Dirac comb, (ii) more picturesquely as the bed of nails function, and (iii) more functionally as an impulse train.

The solution is a periodic function which has a period of 1, but this can't be found simply using the Laplace Transform because that is an integral from $t = 0$ onwards, whereas the Shah function consists of impulses at all integer values of t, both positive and negative. Therefore we will do this by determining the eventual 'steady periodic state' that is achieved when t becomes large. Enjoy the ride!

(a) Find the Laplace Transform of e^{-t}, and then apply the t-shift theorem to find the inverse Laplace Transform of $e^{-ns}/(s+1)$.

(b) Use the Laplace Transform on the ODE,

$$y' + y = \sum_{n=0}^{\infty} \delta(t-n), \qquad y(0) = 0,$$

to find an expression for $Y(s)$, the transform of $y(t)$. Do not simplify this expression for Y by, say, summing the geometric series!

(c) Now use the result of part (a) to write down an expression for $y(t)$ in terms of a sum of terms involving unit step functions.

(d) The sum we have obtained for $y(t)$ is infinite in length; do make sure that you're happy with this idea! Now we let $t = n + \epsilon$ in your expression for y, where n is the first positive integer below t, and where $0 \le \epsilon < 1$. You should then be able to factor $e^{-\epsilon}$ out of the resulting mess(!), and then be able to sum the resulting geometric series to obtain a compact formula for y.

(e) Now find $\lim_{n\to\infty} y$. This will give the long-term formula for $y(t)$, but written in terms of ϵ — this will be a valid formula for $y(t)$ between two neighbouring integer values of t. In the ultimate steady periodic state, what are the maximum and minimum values of y?

(f) See if you can guess what $y(t)$ looks like.

(g) An easier way to solve the main problem is to concentrate on the interval of time, $0 \le t < 1$, and to solve for

$$y' + y = \delta(t), \qquad y(0) = c,$$

using Laplace Transforms. The value of c may be found by insisting that $y(1) = y(0)$.