CP3: Complex Numbers, Differential Equations and Linear Algebra

Toby Adkins

November 19, 2015

Contents

1. *Complex Numbers*

In this chapter, we aim to give an understanding of the basics of complex numbers, including:

- Representations of Complex Numbers
- De Moivre's Theorem
- Manipulations in the Argand Diagram
- Basic Functions of a Complex Variable
- Some examples

The material covered in this section is done so to a relatively shallow depth due to the fact that we use complex numbers merely as a tool for calculation, and further study of their properties is not required. As such, proofs of most of the formulae quoted in this section will not be included.

1.1 Representations of Complex Numbers

So *what is a complex number?* The field of complex numbers includes all real integers. The most basic way of representing a complex number is in the form:

$$z = x + iy \tag{1.1}$$

where x and y are real integers, and $i \equiv \sqrt{-1}$. The coefficient of i is known as the *imaginary part* of a complex number, while the remainder is known as the *real part.* In the case where $y = 0$, the complex number is purely real; all the real integers are just purely real complex numbers. In the case where $x = 0$, the complex number is purely imaginary.

The complex number z is related to the complex number $\bar{z}$:

$$\bar{z} = x - iy \tag{1.2}$$

This is known as the *complex conjugate* of z. In fact, the complex conjugate has a more general definition; it is the complex number that satisfies:

$$z\bar{z} = |z|^2 \tag{1.3}$$

where $|z|$ is the *modulus*, or magnitude, of z. This is because the modulus must always be a real number, and so the imaginary parts must cancel in the multiplication.

$$\begin{aligned} z\bar{z} &= (x + iy)(x - iy) \\ &= x^2 - ixy + ixy + y^2 \\ &= x^2 + y^2 \\ &= |z|^2 \end{aligned}$$

The definition in (1.3) holds true for all the representations of a complex number.

1.1.1 The Argand Diagram and Polar form

From the representation above, it is clear that z appears to represent some sort of coordinates of a point if we take account both of the real and imaginary part. This point can be sketched on what is known as the *Argand Diagram*, as shown in Figure (1.1). It is convention to represent the real part of z along the x-axis, and the imaginary part along the y-axis.

As we have already done work with plane polar coordinates, it should be pretty obvious to most of you that another way of representing a complex number is using this coordinate system.

$$z = r(\cos\theta + i\sin\theta) \tag{1.4}$$

$$r^2 = x^2 + y^2 \tag{1.5}$$

$$\theta = \tan^{-1}\left(\frac{y}{x}\right) \tag{1.6}$$

r gives the modulus of the complex number, and θ is known as it's *argument.* We often use "cis θ" to abbreviate $\cos\theta + i\sin\theta$. Note that conventionally, θ is defined in an anti-clockwise direction from 0 to 2π for a single revolution.

1.2 De Moivre's Theorem

De Moivre's Theorem simply states that, for a complex number z:

$$z^n = r^n(\cos n\theta + i\sin n\theta) = r^n e^{in\theta} \tag{1.7}$$

This is one of the most fundamental results in this entire chapter, and is used in nearly all of the mathematics based questions frequently asked in examinations.

1.2.1 Proving De Moivre's Theorem

We know from prior knowledge concerning exponentials that when you raise an exponential to a power, you simply multiply the exponent by that power. Thus, to prove De Moivre's theorem, we need to show equivalence between $\cos\theta + i\sin\theta$ and $e^{i\theta}$. This can be done in two ways:

- Taylor Expansion - We can consider the Taylor series expansion of $\cos\theta$ and $\sin\theta$. If you are unsure of what these are, please refer to the section on the Taylor series in the CP4 notes.

$$\begin{aligned}
\cos\theta &= 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots \\
\sin\theta &= \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots \\
\cos\theta + i\sin\theta &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots\right) \\
&= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \dots \\
&= e^{i\theta}
\end{aligned}$$

- Differential Equation - We can set up a differential equation in z and solve it with a general solution to prove the result. If you are unfamiliar on how to do this, please refer to the material in Chapter (2).

$$\begin{aligned}
z &= \cos\theta + i\sin\theta \\
\frac{dz}{d\theta} &= -\sin\theta + i\cos\theta \\
&= iz \\
\frac{dz}{d\theta} &= iz \\
\frac{dz}{z} &= i d\theta \\
\rightarrow z &= e^{i\theta}
\end{aligned}$$

  This is a much more elegant proof than the first as it does not require the use of approximations.

De Moivre's theorem leads directly to two very useful results:

$$\cos\theta = \frac{1}{2}\left(e^{i\theta} + e^{-i\theta}\right) \tag{1.8}$$

$$\sin\theta = \frac{1}{2i}\left(e^{i\theta} - e^{-i\theta}\right) \tag{1.9}$$

As we will see with some of the examples in Section (1.5), these can be used to prove results involving series of trigonometric functions.

1.3 Manipulations in the Argand Diagram

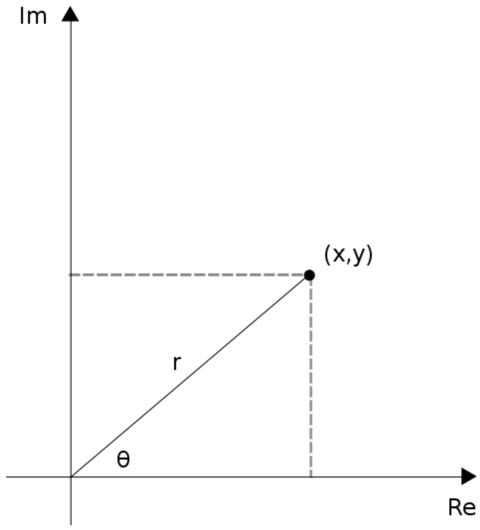

Figure 1.1: The Argand Diagram showing a complex number z

We can perform a number of operations on complex numbers in the Argand Diagram, including the following:

- Rotation - Conventionally, θ *increases* when moving in an anticlockwise direction. Hence, a rotation anticlockwise of z_θ by an angle α is analogous to multiplying by a complex number z_α with unit modulus:

$$\begin{aligned} z_\theta &= re^{i\theta} \\ z_\alpha &= e^{i\alpha} \\ z_\theta \cdot z_\alpha &= re^{i\theta} \cdot e^{i\alpha} \\ &= re^{i(\theta+\alpha)} \\ &= z_{\theta+\alpha} \end{aligned}$$

- Scaling - The is simply changing the modulus of a complex number by multiplying by a real scaling factor.

$$\begin{aligned} z &= re^{i\theta} \\ z' &= rb\, e^{i\theta} \end{aligned}$$

  for some real scalar b. For $|b| > 1$, we are *increasing* the modulus, while for $|b| < 1$, we are *decreasing* it.

- Mirroring - We can mirror a complex number in the real axis by instead representing it's complex conjugate. This is because complex conjugation reverses the sign of the imaginary part while keeping the real part the same.

$$\begin{aligned} z &= re^{i\theta} \\ z' &= re^{-i\theta} \end{aligned}$$

Like in the xy plane, we can also represent a number of curves in the Argand Diagram. If you are unsure of what a particular representation shows, the simply let $z = x + iy$ and simplify the resulting expression. This should give the equation of the curve in Cartesian Coordinates.

- Circle - A circle of radius r centred on some complex number z_o is given by:

$$|z - z_o| = r$$

- Ray - A ray from a point z_o outwards at some angle θ to the real axis is given by:

$$\arg(z - z_o) = \theta$$

  Note that this does not include the point z_o (we can show this using an open circle on z_o).

- Perpendicular bisector - The perpendicular bisector of the line joining z_1 and z_2 is given by:

$$|z - z_1| = |z - z_2|$$

- Spiral - A spiral or snail pattern is represented in the Argand diagram by:

$$z = te^{it}$$

  for some parameter t. This is because as the argument increases, so does the modulus (unlike for a circle). Try sketching this one yourself.

- Ellipse - An ellipse with foci z_1 and z_2 can be represented as:

$$|z - z_1| + |z - z_2| = r$$

  This is because the locus of an ellipse is the set of points that satisfy the property where the sum of their distance from two other points is equal to a constant value, in this case r.

- Hyperbola - A hyperbola with foci z_1 and z_2 in the complex plane can be represented as:

$$|z - z_1| - |z - z_2| = \pm r$$

As always, it is a good exercise to go through the process of finding the Cartesian forms of these representations for yourself, as some can prove to be quite algebraically intensive.

1.4 Basic Functions of a Complex Variable

In this section, we aim to cover some basic functions of complex numbers, or rather functions of a complex variable. Evidently, this will be a very basic outline, as more complex ideas will be covered in the notes entitled "Functions of a Complex Variable".

- Exponentials - In general, we want to let $z = x + iy$ in order to simplify these expressions.

$$\begin{aligned} e^z &= e^{x+iy} \\ &= e^x e^{iy} \\ &= e^x \cos y + i e^x \sin y \end{aligned}$$

- Logarithms - We want to let $z = re^{i(\theta+2k\pi)}$. We have to include this complex phase as the function may be periodic/multivalued. For some integer k:

$$\begin{aligned} ln(z) &= \log\left(re^{i\theta+2k\pi i}\right) \\ &= \log(r) + \log\left(e^{i\theta+2k\pi i}\right) \\ &= log(r) + i\theta + 2k\pi i \end{aligned}$$

- Trigonometric and Hyperbolic Functions - As could be guessed at by their similar form, there are relationships between the trigonometric and hyperbolic functions.

$$\begin{aligned} \cos x &= \frac{1}{2}\left(e^{ix} + e^{-ix}\right) \longleftrightarrow \cos ix = \cosh x \\ \sin x &= \frac{1}{2i}\left(e^{ix} - e^{-ix}\right) \longleftrightarrow \sin ix = i\sinh x \\ \cosh x &= \frac{1}{2}\left(e^{x} + e^{-x}\right) \longleftrightarrow \cosh ix = \cos x \\ \sinh x &= \frac{1}{2}\left(e^{x} - e^{-x}\right) \longleftrightarrow \sinh ix = i\sin x \end{aligned}$$

  There are also some useful "double-angle" identities that come from these results:

$$\begin{aligned} \sin(x+iy) &= \sin x \cosh y + i \cos x \sinh y \\ \cos(x+iy) &= \cos x \cosh y - i \sin x \sinh y \\ \sinh(x+iy) &= \sinh x \cos y + i \cosh x \sin y \\ \cosh(x+iy) &= \cosh x \cos y + i \sinh x \sin y \end{aligned}$$

- Inverse Hyperbolic Functions - Evidently we have to define the inverse of the hyperbolic functions for complex numbers.

$$\begin{aligned} \sinh^{-1} z &= \log\left(z + \sqrt{z^2+1}\right) \\ \cosh^{-1} z &= \log\left(z + \sqrt{z^2-1}\right) \\ \tanh^{-1} z &= \frac{1}{2}\log\left(\frac{1+z}{1-z}\right) \end{aligned}$$

These can be calculated by setting up a quadratic equation in e^{iw} and solving it. For $w = \sin^{-1} z$:

$$\begin{aligned}
\sin w &= z \\
z &= \frac{1}{2i}\left(e^{iw} - e^{-iw}\right) \\
0 &= e^{2iw} - 2ize^{iw} - 1 \\
e^{iw} &= iz \pm \sqrt{1 - z^2} \\
\to \sin^{-1} z &= -i \log\left(iz \pm \sqrt{1 - z^2}\right)
\end{aligned}$$

1.5 Some Examples

Following are some examples of questions using the concepts outlined in this chapter. It is advisable to work through the questions before looking at the solutions.

1. *Find the locus of points in the Argand Diagram satisfying the equation*

$$\arg\left(\frac{z-4}{z-1}\right) = \frac{\pi}{2}$$

We want to let $z = z + iy$ and then simplify to find the Cartesian equivalent.

$$\begin{aligned}\frac{x+iy-4}{x+iy-1} &= \frac{((x-4)+iy)((x-1)-iy)}{(x-1)^2+y^2} \\ &= \frac{x^2+y^2-5x+4+3iy}{(x-1)^2+y^2}\end{aligned}$$

However, we know that the real part of this fraction has to be zero for the argument of z to be $\pi/2$. Hence:

$$\begin{aligned} x^2+y^2-5x+4 &= 0 \\ \left(x-\frac{5}{2}\right)^2 - \frac{25}{4} + 4 + y^2 &= 0 \\ \rightarrow \left(x-\frac{5}{2}\right)^2 + y^2 &= \frac{9}{4}\end{aligned}$$

Thus, the locus of points is the semicircle above the real axis with centre $(5/2, 0)$ and radius $3/2$. Note that this does not include the end points.

2. *Show that*

$$\sum_{n=0}^{\infty} 2^{-n}\cos n\theta = \frac{1-\frac{1}{2}\cos\theta}{\frac{5}{4}-\cos\theta}$$

We can solve this by considering the complex exponential representation of $\cos n\theta$.

$$\begin{aligned}\sum_{n=0}^{\infty} 2^{-n}\cos n\theta &= \frac{1}{2}\sum_{n=0}^{\infty}\frac{1}{2^n}\left(e^{in\theta}+e^{-in\theta}\right) \\ &= \frac{1}{2}\sum_{n=0}^{\infty}\left(\frac{1}{2}e^{i\theta}\right)^n + \left(\frac{1}{2}e^{-i\theta}\right)^n\end{aligned}$$

We can sum these two infinite series:

$$\begin{aligned}\frac{1}{2}\sum_{n=0}^{\infty}\left(\frac{1}{2}e^{i\theta}\right)^n + \left(\frac{1}{2}e^{-i\theta}\right)^n &= \frac{1}{2}\left(\frac{1}{1-\frac{1}{2}e^{i\theta}} + \frac{1}{1-\frac{1}{2}e^{-i\theta}}\right) \\ &= \frac{1}{2}\left(\frac{8-2\left(e^{i\theta}+e^{-i\theta}\right)}{5-2\left(e^{i\theta}+e^{-i\theta}\right)}\right) \\ &= \frac{1}{2}\left(\frac{8-4\cos\theta}{5-4\cos\theta}\right) \\ &= \frac{1-\frac{1}{2}\cos\theta}{\frac{5}{4}-\cos\theta}\end{aligned}$$

We have thus shown the identity.

3. *Prove that*

$$\sum_{k=1}^{n}\binom{n}{k}\sin 2k\theta = 2^n \sin n\theta \cos^n\theta$$

Once again, we want to express the left-hand side as a series that we can more easily sum. Consider the following expression:

$$\begin{aligned}\sum_{k=1}^{n}\binom{n}{k}e^{i2k\theta} &= \binom{n}{1}e^{i2\theta} + \binom{n}{2}e^{i4\theta} + \cdots + \binom{n}{n}e^{i2n\theta} \\ &= \left(e^{i2\theta}+1\right)^n - 1 \\ &= \left(e^{i\theta}\right)^n\left(e^{i\theta}+e^{-i\theta}\right)^n - 1 \\ &= e^{in\theta}\left(2^n\cos^n\theta\right) - 1 \\ &= 2^n\cos^n\theta(\cos n\theta + i\sin n\theta) - 1 \\ &= i2^n\cos^n\theta\sin n\theta + 2^n\cos^n\theta\cos n\theta - 1\end{aligned}$$

If we take the imaginary part of both sides, we obtain the identity as required.

4. *Find the roots of the equation* $(z-1)^n + (z+1)^n = 0$*. Hence solve the equation* $x^3 + 15x^2 + 15x + 1 = 0$*.*

$$\begin{aligned}\frac{(z-1)^n}{(z+1)^n} &= -1 \\ \left(\frac{z-1}{z+1}\right)^n &= -1 \\ \frac{z-1}{z+1} &= e^{i\pi(2k+1)/n}\end{aligned}$$

for some integer $k = 0, 1, \ldots, n-1$. Let $\alpha = (2k+1)/n$.

$$\begin{aligned}z &= \frac{1+e^{i\alpha\pi}}{1-e^{i\alpha\pi}} \\ &= \frac{e^{i\frac{\pi}{2}\alpha}\left(e^{i\frac{\pi}{2}\alpha}+e^{-i\frac{\pi}{2}\alpha}\right)}{e^{i\frac{\pi}{2}\alpha}\left(e^{-i\frac{\pi}{2}\alpha}-e^{i\frac{\pi}{2}\alpha}\right)} \\ \to z &= i\cot\left(\frac{\pi}{2n}(2k+1)\right)\end{aligned}$$

In the original equation, let $n = 6$.

$$\begin{aligned}(z-1)^6 + (z+1)^6 &= 0 \\ 2\left(z^6 + 15z^4 + 15z^2 + 1\right) &= 0\end{aligned}$$

This means that z^2 is a solution to $x^3 + 15x^2 + 15x + 1 = 0$ for $n = 6$. Hence the roots of the polynomial are:

$$\begin{aligned}x &= -\cot^2\left(\frac{\pi}{12}\right),\ -\cot^2\left(\frac{\pi}{4}\right),\ -\cot^2\left(\frac{5\pi}{12}\right) \\ &= -(2+\sqrt{3})^2,\ -1,\ -(2-\sqrt{3})^2 \\ &= -(7+4\sqrt{3}),\ -1,\ -(7-4\sqrt{3})\end{aligned}$$

2. *Differential Equations*

This chapter aims to cover the basics of differential equation solving, including:

- Introduction and First Order ODE's
- Second Order ODE's
- Forced Oscillations and Resonance
- Simultaneous Differential Equations

Like with complex numbers, the mathematics of differential equations is mainly used as a tool for solving physics problems. That being said, it is recommended that students do a lot of practise on solving differential equations (DE's) so as to reach a stage where it is effectively a "handle-turning" process, and more time can be devoted to digesting the physical concepts at hand.

2.1 Ordinary Differential Equations

An ordinary differential equation (or ODE) is an equation containing a function or functions of one *independent* variable and its derivatives. For example, suppose that x is the independent variable in the following equation:

$$\frac{dy}{dx} = 1 + x^2$$

This means that y is the *dependent* variable as it's value depends on x. The *order* of the differential equation is equal to the highest derivative that appears in the equation; in this case, the equation is of first order. As solving ODE's invariably involved integration or trial solutions of some kind, always remember to include the arbitrary constant in the solution; initial conditions can always be imposed later.

2.1.1 First Order ODE's

There are a variety of types of first order ODE's with different methods of solving them, as shown below:

- Separable - For two arbitrary functions f and g, these are of the form:

  $$\frac{dy}{dx} = \frac{f(x)}{g(y)} \tag{2.1}$$

  To solve this, we simply re-arrange and integrate with respect to each variable.

  Solve the following equation subject to the condition that $y = 0$ at $x = 0$.

  $$\frac{dy}{dx} = \frac{xe^y}{1+x^2}$$

  It is evident that this equation is of the form in (2.1), and also because it conveniently appears right after we have introduced the concept.

  $$\begin{aligned} e^{-y}dy &= \frac{x}{1+x^2}dx \\ \int e^{-y}dy &= \int \frac{x}{1+x^2}dx \\ -e^{-y} &= \frac{1}{2}\log(1+x^2) + C_1 \end{aligned}$$

  Now we impose the boundary condition of $y = 0$ at $x = 0$:

  $$\begin{aligned} -1 &= \frac{1}{2}\log(1) + C_1 \\ \rightarrow C_1 &= -1 \\ e^{-y} &= -\frac{1}{2}\log(1+x^2) + 1 \\ y &= -\log\left(1 - \frac{1}{2}\log(1+x^2)\right) \end{aligned}$$

  Note that it is not always easy to express the solution in the form $y = \ldots$, so just simplify the equation as much as possible.

- Almost Separable - For two arbitrary functions a and b, these are of the form:

$$\frac{dy}{dx} = f(ax+by) \tag{2.2}$$

In order to solve these, we let $u = ax + by$ for a change of variables.

Solve the following equation

$$\frac{dy}{dx} = 2(2x+y)^2$$

As above, let $u = 2x + y$.

$$\begin{aligned}
\frac{du}{dx} &= 2 + \frac{dy}{dx} \\
\frac{du}{dx} - 2 &= 2u^2 \\
\frac{du}{dx} &= 2(u^2+1) \\
\frac{du}{u^2+1} &= 2dx \\
\tan^{-1} u &= 2x + C_2 \\
2x + y &= \tan(2x + C_2) \\
y &= \tan(2x + C_2) - 2x
\end{aligned}$$

- Homogeneous Equation - For an arbitrary function f, these are of the form:

$$\frac{dy}{dx} = f\left(\frac{y}{x}\right) \tag{2.3}$$

Similarly to almost separable equations, we make the substitution that $u = y/x$.

Solve the following equation

$$\frac{dy}{dx} + \frac{2x}{y} = 3$$

In this case, it is very clear that it is an homogeneous type equation, but be wary that it is not always this obvious. Let $u = y/x$.

$$\begin{aligned}
y &= ux \\
\frac{dy}{dx} &= u + x\frac{du}{dx}
\end{aligned}$$

Substituting this into the equation:

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

Integrating both sides:

$$\log(1-u) - 2\log(2-u) = \log x + C_3$$
$$\log\left(1-\frac{y}{x}\right) - 2\log\left(2-\frac{y}{x}\right) = \log x + C_3$$

Always remember to transform back to the original variables when stating the final answer!

- Integrating Factor - This is probably the most useful technique, and is used to solve equations of the form:

$$\frac{dy}{dx} + f(x)\,y = g(x) \tag{2.4}$$

for arbitrary functions f and g. We can solve this by multiplying throughout the whole equation by the *integrating factor* I:

$$I = e^{\int f(x)dx} \tag{2.5}$$

We then look for the 'reverse product rule' on the left-hand side of the equation. Let's take a look at an example to see what this means.

Solve the following equation

$$x(x-1)\frac{dy}{dx} + y = x(x-1)^2$$

This is not originally in the form that we require to apply the integrating factor, and so we need to first manipulate the equation.

$$x(x-1)\frac{dy}{dx} + y = x(x-1)^2$$
$$\frac{dy}{dx} + \frac{y}{x(x-1)} = (x-1)$$

Now we want to compute I:

$$\int \frac{1}{x(x-1)}dx = \log\left(\frac{1-x}{x}\right)$$
$$\rightarrow I = \frac{1-x}{x}$$

Multiplying throughout by I:

$$\left(\frac{1}{x}-1\right)\frac{dy}{dx} - \frac{y}{x^2} = -\frac{(x-1)^2}{x}$$
$$\frac{d}{dx}\left(y\frac{(1-x)}{x}\right) = -\frac{(x-1)^2}{x}$$
$$y\frac{(1-x)}{x} = \int \frac{2x-x^2-1}{x}dx$$
$$= 2x - \frac{x^2}{x} + \log x + C_4$$
$$y = \frac{2x^2}{(1-x)} - \frac{x^3}{2(1-x)} - \frac{x}{(1-x)}\log x + \frac{C_4 x}{(1-x)}$$

Always remember to add the constant of integration *immediately after integrating.*

- Bernoulli Equation - These at first glance do not appear to be linear equations as they are non-linear in the dependant variable, but can easily be solved with, you guessed it, a substitution. For arbitrary functions f and g, these are of the form:

$$\frac{dy}{dx} + f(x)\, y = g(x)\, y^n \tag{2.6}$$

In this case, we let $u = y^{1-n}$.

$$\begin{aligned} \frac{du}{dx} &= (1-n)y^{-n}\,\frac{dy}{dx} \\ \to \frac{du}{dx} + u(1-n)f(x) &= g(x)(1-n) \end{aligned}$$

We can then use the integrating factor method to solve the resultant equation.

Solve the following equation

$$\frac{dy}{dx} + \frac{y}{x} = 2x^{\frac{3}{2}}y^{\frac{1}{2}}$$

As stated above, let $u = y^{1/2}$.

$$\begin{aligned} \frac{du}{dx} &= \frac{1}{2}y^{-1/2}\,\frac{dy}{dx} \\ \frac{dy}{dx} &= 2y^{1/2}\,\frac{du}{dx} \\ \frac{du}{dx} + \frac{u}{2x} &= x^{3/2} \end{aligned}$$

Now we make use of the integrating factor technique.

$$\begin{aligned} I &= e^{\int \frac{1}{2x}dx} \\ &= e^{\frac{1}{2}\log x} \\ &= x^{1/2} \end{aligned}$$

Multiplying throughout by I:

$$\begin{aligned} x^{1/2}\frac{du}{dx} + \frac{u}{2}x^{-1/2} &= x^2 \\ \frac{d}{dx}\left(ux^{1/2}\right) &= x^2 \\ ux^{1/2} &= \frac{x^3}{3} + C_5 \\ y &= \left(\frac{x^{5/2}}{3} + \frac{C_5}{x^{5/2}}\right)^2 \end{aligned}$$

- Exact Equations - For some arbitrary functions p and q, these are of the form.

$$\frac{dy}{dx} = -\frac{p(x,y)}{q(x,y)} \tag{2.7}$$

These can be solved if and only if

$$\frac{\partial q}{\partial x} = \frac{\partial p}{\partial y} \tag{2.8}$$

This is in fact the condition for the equation to be an exact differential.

Solve the following equation

$$\frac{dy}{dx} = -\frac{6x + y + y^2}{x + 2xy}$$

Looking at (2.7), it is clear that:

$$\frac{\partial q}{\partial x} = \frac{\partial p}{\partial y} = 1 + 2y$$

Hence the equation is exact. Writing

$$df = \frac{df}{dx}dx + \frac{df}{dy}dy = (6x + y + y^2)dx + (x + 2xY + dy = 0$$

So, by inspection, the differential equation can be integrated to give

$$xy + xy^2 + 3x^2 = C_5$$

This is equal to a constant as the right-hand side of the differential equation is equal to zero, which integrates to give a constant. This makes more sense when regarded in the like of conservative forces and scalar potentials (see CP4 notes).

These are the six main types of first order differential equations that we will encounter. Always remember that if an equation does not initially appear in one of these forms, it will most likely be some re-arrangement of one of them.

2.2 Second Order ODE's

Second order differential equations are so called as they include the second derivative of the dependant variable with respect to the independent variable. They are of the form:

$$f(x)\,\frac{d^2y}{dx^2} + g(x)\,\frac{dy}{dx} + h(x)\,y = r(x) \tag{2.9}$$

for barbarity functions f, g, h and r. If $r(x) = 0$, the equation is *homogeneous*, where as if $r(x) \neq 0$, then it is *inhomogeneous.*

The fact that these equations are *linear* in the dependent variable and its derivatives is very important. This is because it allows us to use the *principle of superposition*; for an equation that has solutions $y_1(x)$ and $y_2(x)$, any linear combination of these two is also a solution. This means that to solve an inhomogeneous equation, we can solve the homogeneous and inhomogeneous parts separately, and their sum will be a solution to the entire equation. Note that in many cases there is not a *unique* solution, but instead a set of solutions to within an arbitrary set of constants.

2.2.1 Solving the Homogeneous Equation

To solve this, we want to use the substitution

$$y = Ce^{nx} \tag{2.10}$$

This will generate an axillary equation in n that can then be solved to find the general solution. Depending on the magnitude of n, this may lead to one of three different sets of solution behaviours. For the purposes of this subsection, we will assume that the equation is of the form

$$a\,\frac{d^2y}{dx^2} + b\,\frac{dy}{dx} + c\,y = r(x)$$

for constants a, b and c. The auxiliary equation is thus:

$$\begin{aligned} an^2 + bn + c &= 0 \\ \rightarrow n &= -\frac{b}{2a} \pm \sqrt{\left(\frac{b}{2a}\right)^2 - \frac{c}{a}} \\ &= -\frac{b}{2a} \pm \alpha \end{aligned}$$

The three solution sets are as follows:

1. Over-damped - This occurs when the solutions to n are real; for $b^2 > 4ac$. These are exponentially decaying solutions that usually disappear as $x \rightarrow \infty$, and are the result of strong damping.

$$y = e^{-bx/2a}\left(c_1 e^{\alpha x} + c_2 e^{-\alpha x}\right) \tag{2.11}$$

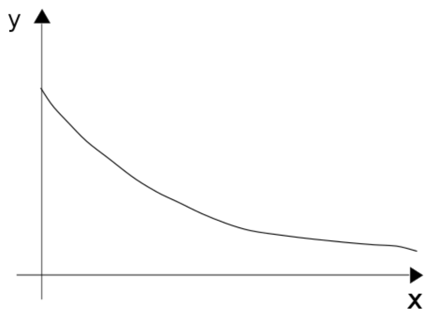

Figure 2.1: An example of an over-damped solution for y

2. Critical Damping - This occurs when the solutions to n are real and coincident; for $b^2 = 4ac$. This includes a linear term that is also modulated by the exponential decay amplitude.
$$y = e^{-bx/2a}(c_1 x + c_2) \tag{2.12}$$

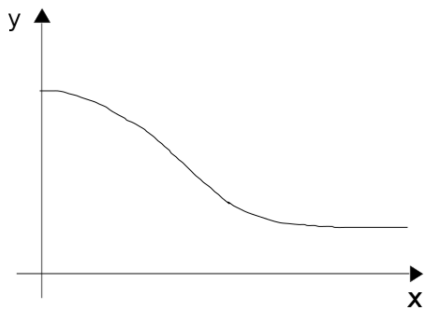

Figure 2.2: An example of a critically damped solution for y

3. Under-damping - This occurs when the solutions to n are complex; for $b^2 < 4ac$. The resultant solution will include sinusoidal/oscillatory terms in the decay.
$$y = e^{-bx/2a}(c_1 \cos(i\alpha x) + c_2 \sin(i\alpha x)) \tag{2.13}$$

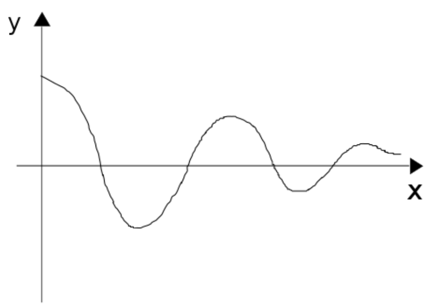

Figure 2.3: An example of an under-damped solution for y

2.2.2 Solving the Inhomogeneous Equation

Solving for the inhomogeneous solution is a little more of an 'art-form' than in the case of the homogeneous solution. In general, we have to find a trail solution that 'corresponds' to $r(x)$. For example, if $r(x) = \cos x$, then we would substitute the solution $y = c_1 \cos x + c_2 \sin x$ into the homogeneous part of the equation and equate coefficients to find the final solution. We can think of the trial solution as the most general form of the specific $r(x)$ given. Some trial solutions include:

$r(x)$	Trial Solution
A polynomial of degree n	A general polynomial of degree n
Sum of $\sin ax$ and $\cos ax$	$c_1 \cos ax + c_2 \sin ax$
Sum of $\sinh ax$ and $\cosh ax$	$c_1 e^{ax} + c_2 e^{-ax}$
e^{ax}	$c_1 e^{ax}$

If $r(x)$ is a sum of two such function types, then the trial solution should be a sum of those functions. If it is a product, then the trial solution should be a product (though pray it isn't as this can get quite algebraically intensive!).

Note that the solution to the homogeneous equation cannot appear in a solution to the inhomogeneous equation. For example, if e^{ax} appears in the solution to y_{hom}, then in our trial solution for the inhomogeneous function $r(x) = \cosh ax$, we have to use $y_{trial} = c_1 x e^{ax} + c_2 e^{-ax}$. Otherwise, we would just get a repeated solution.

Let us consider the following example. *A mass m is constrained to move in a straight line and is attached to a spring of length λm and a dashpot which produces a retarding force $-\alpha m v$, where v is the velocity of the mass. Find the steady state displacement of the mass when an amplitude modulated periodic force $Am \cos pt \sin \omega t$ with $p \ll \omega$ and $\alpha \ll \omega$ is applied to it.*

Show that for $\omega = \lambda$ the displacement of the amplitude modulated wave is approximately given by

$$-A \frac{\cos \omega t \sin(pt + \phi)}{\sqrt{4\omega^2 p^2 + \alpha^2 \omega^2}} \quad \text{where} \quad \cos \phi = \frac{2\omega p}{\sqrt{4\omega^2 p^2 + \alpha^2 \omega^2}}$$

At first glance, this looks like a very nasty question....and to tell you the truth, it really is. However, we recommend that you give it a go first, and come back here for hints if needed.

We first need to find the equation of motion of the mass, as this will give us the differential equation that we need to solve. By Newton's Second Law:

$$\begin{aligned} \sum F &= m\ddot{x} \\ m\ddot{x} &= -\lambda^2 m x - \alpha m v + Am \cos pt \sin \omega t \\ \rightarrow \frac{d^2x}{dt^2} + \alpha \frac{dx}{dt} + \lambda^2 x &= A \cos pt \sin \omega t \end{aligned}$$

As we are asked for the steady state displacement of the mass, we only have to find the solution to the inhomogeneous equation, as this will be the only remaining motion as $t \to \infty$. Consider

$$\begin{aligned} A \cos pt \sin \omega t &= \frac{A}{2} \left[\sin(p + \omega)t - \sin(p - \omega)t \right] \\ &= \mathrm{Im} \left(\frac{A}{2} e^{i(p+\omega)t} - \frac{A}{2} e^{i(p-\omega)t} \right) \end{aligned}$$

Thus, we want to trial solution to be $x = c_1 e^{i(p+\omega)t} + c_2 e^{i(p-\omega)t}$. Substituting this in, and going through some algebra, we obtain:

$$c_1 e^{i(p+\omega)t}\left[-(p+\omega)^2 + \alpha i(p+\omega) + \lambda^2\right] + c_2 e^{i(p-\omega)t}\left[-(p-\omega)^2 + \alpha i(p-\omega) + \lambda^2\right] = \frac{A}{2}e^{i(p+\omega)t} - \frac{A}{2}e^{i(p-\omega)t}$$

Equating coefficients:

$$\begin{aligned}
\frac{A}{2} &= c_1(\lambda^2 - (p+\omega)^2 + i\alpha(p+\omega)) \\
-\frac{A}{2} &= c_2(\lambda^2 - (p-\omega)^2 + i\alpha(p-\omega)) \\
c_1 &= \frac{A}{2(\lambda^2 - (p+\omega)^2 + i\alpha(p+\omega))} \\
c_2 &= \frac{A}{2(\lambda^2 - (p-\omega)^2 + i\alpha(p-\omega))}
\end{aligned}$$

Now, we can write c_1 and c_2 in complex exponential for that allows us to simplify the solutions somewhat.

$$\begin{aligned}
\text{Let}\quad r_1 e^{i\theta_1} &= \lambda^2 - (p+\omega)^2 + i\alpha(p+\omega) \\
r_1 &= \sqrt{\alpha^2(p+\omega)^2 + (\lambda^2 - (p+\omega)^2)^2} \\
\theta_1 &= \tan^{-1}\left(\frac{\alpha(p+\omega)}{\lambda^2 - (p+\omega)^2}\right) \\
\text{Let}\quad r_2 e^{i\theta_2} &= \lambda^2 - (p+\omega)^2 + i\alpha(p+\omega) \\
r_1 &= \sqrt{\alpha^2(p-\omega)^2 + (\lambda^2 - (p-\omega)^2)^2} \\
\theta_1 &= \tan^{-1}\left(\frac{\alpha(p-\omega)}{\lambda^2 - (p-\omega)^2}\right)
\end{aligned}$$

Thus our solution to the steady-state displacement is the imaginary part of the following:

$$x = \mathrm{Im}\left[\frac{A}{2}\left(\frac{e^{i((p+\omega)t-\theta_1)}}{r_1} - \frac{e^{i((p-\omega)t-\theta_2)}}{r_2}\right)\right] \tag{2.14}$$

Told you it was messy! We now have to apply the condition that both p and α are negligible in comparison to ω. In this limit, we find that

$$\begin{aligned}
r_1 = r_2 &\approx \sqrt{\alpha^2\omega^2 + (\lambda^2 - \omega^2)^2} \\
\theta_1 = -\theta_2 &\approx \tan^{-1}\left(\frac{\alpha\omega}{\lambda^2 - \omega^2}\right)
\end{aligned}$$

Substituting this back into (2.14), with a little simplification, we obtain:

$$x(t) = \frac{A\sin(\omega t - \theta_1)}{\sqrt{\alpha^2\omega^2 + (\lambda^2 - \omega^2)^2}}$$

Now we are asked the case with $\omega = \lambda$. We need to apply this condition *before* the approximations. Under this condition:

$$\begin{aligned}
r_1 &= \sqrt{\alpha^2(p+\omega)^2 + (2\omega p + p^2)^2} \\
r_2 &= \sqrt{\alpha^2(p+\omega)^2 + (2\omega p - p^2)^2} \\
\theta_1 &= \tan^{-1}\left(\frac{-\alpha(p+\omega)}{p(2\omega + p)}\right) \\
\theta_2 &= \tan^{-1}\left(\frac{\alpha(p-\omega)}{p(2\omega - p)}\right)
\end{aligned}$$

Now let us assume $p \ll \omega$ and $\alpha \ll \omega$:

$$r_1 = r_2 \approx \sqrt{4\omega^2 p^2 + \alpha^2\omega^2}$$
$$\theta_1 = \theta_2 \approx \tan^{-1}\left(-\frac{\alpha\omega}{2\omega p}\right) = \phi$$

We now want to find $\cos\phi$. Using the quadrant diagram below

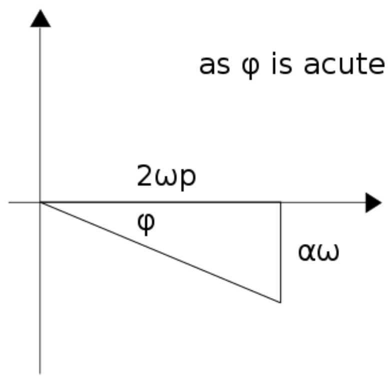

Figure 2.4: A quadrant diagram for an acute angle ϕ

we find that

$$\cos\phi = \frac{2\omega p}{\sqrt{4\omega^2 p^2 + \alpha^2\omega^2}}$$

Simplifying the expression for x based on these calculations:

$$\begin{aligned} x &= -\frac{A}{2}\left[\frac{\sin((p+\omega)t - \theta_1)}{r_1} + \frac{\sin((p-\omega)t - \theta_2)}{r_2}\right] \\ &= -\frac{A}{r_1}\left[\sin\left(\frac{(p+\omega)t - \phi + (p-\omega)t - \phi}{2}\right)\cdot\cos\left(\frac{(p+\omega)t - \phi - (p-\omega)t + \phi}{2}\right)\right] \\ \rightarrow x &= -A\frac{\cos\omega t\sin(pt+\phi)}{\sqrt{4\omega^2p^2 + \alpha^2\omega^2}} \end{aligned}$$

As with all questions, it helps to be very careful when making substitutions or approximations in order to preserve accuracy.

2.2.3 Euler-Cauchy Equation

Euler-Cauchy equations are of the form

$$ax^2\frac{d^y}{dx^2} + bx\frac{dy}{dx} + cy = r(x) \tag{2.15}$$

for constants a, b and c. This can be solved by letting $x = e^t$ to give a change of variables to an equation that we can solve by conventional techniques.

$$\begin{aligned}
\frac{dy}{dx} &= \frac{dt}{dx}\frac{dy}{dt} \\
&= \frac{1}{e^t}\frac{dy}{dt} \\
\frac{d^2y}{dx^2} &= \frac{d}{dx}\left(\frac{dy}{dt}\frac{dt}{dx}\right) \\
&= \frac{d}{dt}\left(\frac{dy}{dx}\right)\frac{dt}{dx} \\
&= \frac{d}{dt}\left(\frac{1}{e^t}\frac{dy}{dt}\right)\frac{1}{e^t} \\
&= \frac{1}{e^{2t}}\left(\frac{d^2y}{dt^2} - \frac{dy}{dt}\right)
\end{aligned}$$

Thus, the equation becomes

$$\begin{aligned}
a\left(\frac{d^2y}{dt^2} - \frac{dy}{dt}\right) + b\frac{dy}{dt} + cy &= r(e^t) \\
a\frac{d^2y}{dt^2} + (b-a)\frac{dy}{dt} + cy &= r(e^t)
\end{aligned}$$

which we can solve for $y(t)$ and then transform back to $y(x)$.

2.3 Forced Oscillations and Resonance

We jumped the gun slightly with the example in Section (2.2.2) as it actually uses some of the ideas covered in this section. However, now that you are familiar with the process of solving these types of driven equations, full derivations will not be outlined. If confused, refer to the aforementioned example, as it includes all the techniques to derive most of the results in this section.

Forced or driven oscillators typically have equations of motion of the form

$$\frac{d^2x}{dt^2} + \gamma\frac{dx}{dt} + {\omega_o}^2 x = \frac{F}{m}e^{i\omega t} \tag{2.16}$$

where γ is the damping constant, ω_o is the natural frequency (that results from a static force, letting $\ddot{x} = \dot{x} = 0$) and the right-hand side is the the sinusoidal driving force.

2.3.1 Solution

The transient solution ism given by finding the solution to the homogeneous equation. For $\gamma \neq 0$, this will have an exponential decay constant out the front, meaning it will decay to effectively zero over large time periods. If $\gamma = 0$, then we actually have simple harmonic motion.

The steady state solution is the solution to the inhomogeneous equation. In a similar way to in the example in Section (2.2.2), this can be done by letting $y = Ce^{i\omega t}$ and simplifying using complex exponential algebra. This leads to the solution of

$$x(t) = \frac{Fe^{i(\omega t - \phi)}}{m\sqrt{({\omega_o}^2 - \omega^2)^2 + \gamma^2\omega^2}} \tag{2.17}$$

$$\text{where} \quad \phi = \tan^{-1}\left(\frac{\gamma\omega}{{\omega_o}^2 - \omega^2}\right) \tag{2.18}$$

The resonant frequency for the displacement is found by differentiating the denominator and finding it's minimum value.

$$\begin{aligned} \frac{d}{d\omega}\left[({\omega_o}^2 - \omega^2)^2 + \gamma^2\omega^2\right]\Big|_{\omega=\omega_R} &= 0 \\ -4\omega({\omega_o}^2 - {\omega_R}^2) + 2\gamma^2\omega &= 0 \\ \rightarrow {\omega_R}^2 = {\omega_o}^2 - \frac{\gamma^2}{2} \end{aligned} \tag{2.19}$$

The resonant frequency for the velocity can be found by dividing through by ω, and then inspecting the values that minimise the denominator.

$$\begin{aligned} \dot{x} &= \frac{\omega F e^{i(\omega t - \phi)}}{m\sqrt{({\omega_o}^2 - \omega^2)^2 + \gamma^2\omega^2}} \\ &= \frac{Fe^{i(\omega t - \phi)}}{m\sqrt{({\omega_o}^2/\omega^2 - 1)^2 + \gamma^2}} \end{aligned}$$

Evidently, it is clear that $\omega = \omega_o$ for velocity resonance.

2.3.2 Power at Resonance

The power of the oscillator, assuming perfect coupling, is given by the time average of the instantaneous power at resonance.

$$\begin{aligned}\text{Power} &= \text{Force } \times \text{ Velocity}\\ P &= \frac{-F\omega m}{\sqrt{({\omega_o}^2-\omega^2)^2+\gamma^2\omega^2}}\ \sin(\omega t-\phi)\cos\omega t\end{aligned}$$

We want the instantaneous power, and so we need to find the time average of the sinusoidal terms.

$$\begin{aligned}\overline{\sin(\omega t-\phi)\cos\omega t} &= \overline{(\sin\omega t\cos\phi-\cos\omega t\sin\phi)\cos\omega t}\\ &= \overline{\sin\omega t\cos\omega t}\cos\phi-\overline{\cos^2\omega t}\sin\phi\end{aligned}$$

Averaging over a period:

$$\begin{aligned}\overline{\sin\omega t\cos\omega t} &= \frac{\omega}{2\pi}\int_0^{\frac{2\pi}{\omega}}\ \sin 2\omega t\ dt\\ &= 0\\ \overline{\cos^2\omega t} &= \frac{\omega}{2\pi}\int_0^{\frac{2\pi}{\omega}}\ \cos^2\omega t\ dt\\ &= \frac{1}{2}\end{aligned}$$

Hence, we arrive at the equation

$$\langle P\rangle = \frac{F\omega m}{2\sqrt{({\omega_o}^2-\omega^2)^2+\gamma^2\omega^2}}\ \sin\phi \tag{2.20}$$

$$= \frac{Fm\omega^2\gamma}{2(({\omega_o}^2-\omega^2)^2+\gamma^2\omega^2)} \tag{2.21}$$

Near resonance, $\omega\approx\omega_o$, meaning that the average power becomes:

$$\langle P\rangle_{Resonance} = \frac{F^2}{2m\gamma}$$

2.3.3 Quality Factor

The *Q factor* is a measure of the 'quality' of a resonant system. In general, a higher Q corresponds to a:

- Smaller resonant peak (see next section)
- Greater amplitude of oscillation
- Smaller decay constant, meaning that oscillations take longer to decay
- Good resonance

There are multiple definitions of Q factor that we can use when calculating it. These include:

$$\begin{aligned}Q &= 2\pi\cdot\frac{\text{Energy Stored}}{\text{Mean Power per Cycle}}\\ &= 2\pi\cdot\frac{\text{Amplitude at Resonance}}{\text{Max Displacement under static force}}\\ &= \frac{2\pi}{\text{Fractional Power Lost per Cycle}}\end{aligned}$$

The last of these can be easily calculated by Taylor expanding the exponential term that appears in the fractional power lost per cycle, assuming that the damping is small. Which definition used to calculate the Q factor will depend largely on the system at hand, especially whether it is forced or not.

In the case of the oscillator outlined at the start of this section, the Q factor is:

$$Q = \frac{\omega_o}{\gamma}$$

2.3.4 Width of Resonance

The width of the resonant peak corresponds to the quality of the resonator; the higher the Q factor, the smaller the resonant peak. Let ω_R be the resonant frequency.

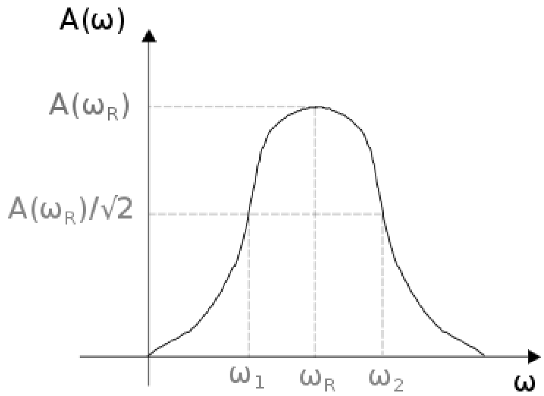

Figure 2.5: Finding the width of the resonant peak

The width of the resonance is defined as:

$$\Delta\omega = \omega_2 - \omega_1$$

We thus want to find an expression for $\omega_{1,2}$ in order to evaluate $\Delta\omega$. From (2.18):

$$A(\omega_1) = \frac{1}{\sqrt{2}}A(\omega_R)$$

$$\frac{F}{m\sqrt{({\omega_o}^2 - {\omega_1}^2)^2 + \gamma^2{\omega_1}^2}} = \frac{F}{\sqrt{2}m\sqrt{({\omega_o}^2 - {\omega_1}^2)^2 + \gamma^2{\omega_1}^2}}$$

$$\left({\omega_o}^2 - {\omega_1}^2\right)^2 + \gamma^2{\omega_1}^2 = 2\left[\left({\omega_o}^2 - {\omega_R}^2\right)^2 + \gamma^2{\omega_R}^2\right]$$

Recalling (2.19):

$$\left({\omega_o}^2 - {\omega_1}^2\right)^2 + \gamma^2{\omega_1}^2 = 2\gamma^2{\omega_o}^2 - \frac{\gamma^4}{2}$$

This can be solved as a quadratic in ${\omega_1}^2$, but the answer is not particularly enlightening. It it much easier to expand the equation by expressing ω_1 in terms of a relevant parameter. γ is a good choice as the damping will be small for a sharp resonance, and it has the right dimensions. Thus let

$$\omega_1 = \omega_o + a\gamma + O(\gamma^2)$$

for some constant a. Substitute this in and ignore terms of $O(\gamma^2)$ and above:

$$\left({\omega_o}^2 - {\omega_1}^2\right)^2 + \gamma^2{\omega_1}^2 = 2\gamma^2{\omega_o}^2 - \frac{\gamma^4}{2}$$
$$(\omega_o - \omega_1)^2(\omega_o + \omega_1)^2 + \gamma^2{\omega_1}^2 = 2\gamma^2{\omega_o}^2$$
$$4a^2\gamma^2{\omega_o}^2 + \gamma^2{\omega_o}^2 = 2\gamma^2{\omega_o}^2$$

Equating coefficients, we find that $a = \pm 1/2$. This means that:

$$\omega_1 = \omega_o - \frac{\gamma}{2}$$
$$\omega_2 = \omega_o + \frac{\gamma}{2}$$

Thus, the width of the resonance is:

$$\Delta\omega = \gamma \tag{2.22}$$

2.3.5 Phases

Evidently, there is a phase difference ϕ between the driving force and the response of the oscillator.

- Displacement Response: $\phi_x = \tan^{-1}\left(\frac{\gamma\omega}{{\omega_o}^2 - \omega^2}\right)$

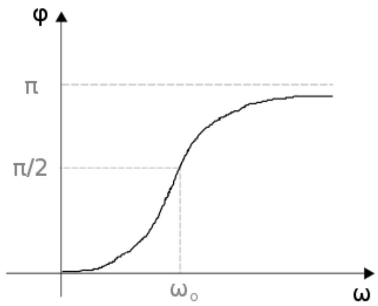

Figure 2.6: Displacement phase as a function of ω

- Velocity Response: $\phi_v = \tan^{-1}\left(\frac{\gamma\omega}{{\omega_o}^2 - \omega^2}\right) - \pi/2$

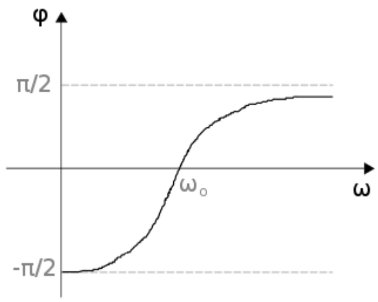

Figure 2.7: Velocity phase as a function of ω

2.4 Simultaneous Differential Equations

Simultaneous or coupled differential equations describe systems where there is no definitive dependent variable that can be isolated. In this case, we have to solve the system of equations. Note that we assume that the solutions to both variables will both 'vary together'; i.e they will have the same exponential terms. We can do this by following the following set of steps:

- Substitute the trial solutions $x_1 = C_1 e^{nt}$ and $x_2 = C_2 e^{nt}$
- Set the determinant of the resultant coefficient matrix to zero to find a solution to the homogeneous equation
- Find the amplitude ratios between the coefficients. It is useful to apply boundary conditions after this step to eliminate unnecessary algebra
- For the inhomogeneous solution, substitute the trial solutions into both equations and solve for coefficients as normal

As always, let's have a look at an example.

Solve the differential equations

$$2\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2\frac{dz}{dx} + 3y + z = e^{2x}$$
$$\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + \frac{dz}{dx} + 2y - z = 0$$

Is it possible to have a solution to these equations for which $y = z = 0$ when $x = 0$?

Let $y = Ae^{nx}$ and $z = Be^{nx}$. Taking linear combinations of the equations, we obtain the following coefficient equations

$$A(n^2+1) + B(n+2) = 0$$
$$A(n-1)(n-2) + B(n-1) = 0$$

Writing this as a matrix system and taking the determinant of the coefficient matrix:

$$\begin{pmatrix} n^2+1 & n+2 \\ (n-1)(n-2) & n-1 \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} = 0$$
$$\begin{vmatrix} n^2+1 & n+2 \\ (n-1)(n-2) & n-1 \end{vmatrix} \overset{!}{=} 0$$
$$(n^2+1)(n-1) - (n+1)(n-2)(n-1) = 0$$
$$5n - 5 = 0$$
$$\rightarrow n = 1$$

When $n = 1$,

$$A(2) + B(3) = 0$$
$$B = -\frac{2}{3}A$$

Hence, the homogeneous part of the solution is

$$y' = \alpha e^x$$
$$z' = -\frac{2}{3}\alpha e^x$$

for some constant α. For the inhomogeneous solution, we want to let $y = Ce^{2x}$ and $z = De^{2x}$.

$$\begin{aligned}
8C - 6C + 3C + 4D + 2D &= 1 \\
5C + 6D &= 1 \\
4C - 6C + 2D + 2C - D &= 0 \\
D &= 0 \\
C &= \frac{1}{5}
\end{aligned}$$

Hence, the final solutions are:

$$\begin{aligned}
y &= \alpha e^{x} + \frac{1}{5}e^{2x} \\
z &= -\frac{2}{3}\alpha e^{x}
\end{aligned}$$

It is evidently not possible to have solutions to this equation that satisfies $y = z = 0$ for $x = 0$, as z only takes non-zero values.

3. *Linear Algebra*

This chapter aims to cover the fundamental basics of linear algebra, including:

- Index Notation
- Vector Spaces
- Scalar and Vector Products
- Cardinal Geometry
- Matrices
- Determinants
- Matrix Inverse
- Eigenvalues, Eigenvectors and Quadratic Forms
- Systems of Linear Equations
- Matrix Types

Linear Algebra is in fact a very massive topic that goes well beyond the scope of this course. As such, this text will attempt to provide *some* of the theory underlying the more practical applications of linear algebra, but for a more full/in-depth explanation refer to Andre Lukas' type-set notes. Furthermore, it will not provide proofs of the more mathematically 'fundamental' concepts. Think of this as more of a distillation of the commonly used concepts and techniques. Throughout this chapter, we will use the superscript $\widehat{\underline{v}}$ to denote a unit vector; that is

$$\widehat{\underline{v}} = \frac{\underline{v}}{|\underline{v}|}$$

3.1 Index Notation

Index notation is an incredibly useful tool when it comes to linear algebra as it is a short-hand way of notating vector and matrix operations based on their elements. For example, suppose we have an n-dimensional row vector $\underline{v} = (v_1, v_2, v_3, \ldots, v_n)$. Then in index notation, this becomes v_i, where the subscript i denotes all of these elements within the vector; a row or column vector only has a single index as we only need one number to specify a unique element in the vector. For a matrix $\mathbf{A}$, then two indices are needed as you need to specify column position as well as row position; A_{ij}.

This brings with it some important properties and ideas. For example, if we have two vectors $\underline{v}$ and $\underline{u}$, which we will write in index notation as v_i and u_j, then the scalar product of these two vectors is only defined if $i = j$. This makes sense, as the vectors would have to be the same size for one to be able to calculate the product; else, what do you do with the entries left over when you have calculated all the other corresponding ones?

At this stage, we need to introduce two important entities that we use in index notation. There are as follows:

- Kronecker Delta - This is defined as:

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

  You can think of this as meaning "switches i to j" or vice versa; when multiplying a vector by this, we can simply switch the index.

- Levi-Cevita Tensor - This is defined as:

$$\epsilon_{i_1, i_2, \ldots, i_n} = \begin{cases} 1, & \text{if } (i_1, i_2, \ldots, i_n) \text{ is an even permeatation of } (1, 2, \ldots, n) \\ -1, & \text{if } (i_1, i_2, \ldots, i_n) \text{ is an odd permeatation of } (1, 2, \ldots, n) \\ 0, & \text{otherwise} \end{cases}$$

  Here we have defined the n-dimensional case; in this chapter, we will mostly work with the three dimensional case ϵ_{ijk}.

We can combine both of these entities to give us an 'index notation identity' that:

$$\epsilon_{ijk}\epsilon_{kmn} = \delta_{im}\delta_{jn} - \delta_{jm}\delta_{in} \tag{3.1}$$

As you are often told with identities such as this, if you stare at it long enough, it will make sense as to why it is what it is!

How is this useful? Well, we can use it to prove and simplify many vector operations and identities. For example, for two vectors $\underline{v}$ and $\underline{u}$, their scalar product is defined by

$$\underline{v} \cdot \underline{u} = v_i u_i$$

while the vector product is defined as

$$\underline{v} \times \underline{u} = \epsilon_{ijk} v_j u_k$$

Now, don't worry if you don't know what these operations actually are; we will define them fully later on in Section (3.3). For the meantime, let us use these 'index definitions' to prove some identities. In all of these cases, it is very important to keep track of the indices used. At any one time, there should only really be one "free" index that the summation takes place over.

- $\underline{c} \cdot (\underline{a} \times \underline{b}) = -\underline{c} \cdot (\underline{b} \times \underline{a})$

$$\begin{aligned} \underline{c} \cdot (\underline{a} \times \underline{b}) &= c_i (\underline{a} \times \underline{b})_i \\ &= c_i \epsilon_{ijk} a_j b_k \\ &= c_i \epsilon_{ikj} (-b_k) a_j \\ &= -c_i (\underline{b} \times \underline{a})_i \\ &= -\underline{c} \cdot (\underline{b} \times \underline{a}) \end{aligned}$$

  Here we have used the anti-cyclic property of the Levi-Civita tensor that introduces the minus sign when we change the order of the vector product (incidentally, this is one of the fundamental properties of the vector product due to its definition).

- $\underline{a} \times (\underline{b} \times \underline{c}) = \underline{b} \, (\underline{c} \cdot \underline{a}) - \underline{c} \, (\underline{b} \cdot \underline{a})$

$$\begin{aligned} \underline{a} \times (\underline{b} \times \underline{c}) &= \epsilon_{ijk} a_j (\underline{b} \times \underline{c})_k \\ &= \epsilon_{ijk} a_j \epsilon_{knm} b_m c_n \\ &= (\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}) a_j b_m c_n \\ &= a_j c_j b_i - b_j c_i a_j \\ &= \underline{b} \, (\underline{c} \cdot \underline{a}) - \underline{c} \, (\underline{b} \cdot \underline{a}) \end{aligned}$$

  We have used (3.1). Note how we have switched the indices on all the vectors when multiplying through by the Kronecker-Delta.

- $(\underline{a} \times \underline{b}) \cdot (\underline{c} \times \underline{d}) = (\underline{a} \cdot \underline{c}) \, (\underline{b} \cdot \underline{d}) - (\underline{a} \cdot \underline{d}) \, (\underline{b} \cdot \underline{c})$

$$\begin{aligned} (\underline{a} \times \underline{b}) \cdot (\underline{c} \times \underline{d}) &= (\underline{a} \times \underline{b})_k (\underline{c} \times \underline{d})_k \\ &= \epsilon_{ijk} a_j b_i \epsilon_{knm} c_n d_m \\ &= \epsilon_{ijk} \epsilon_{knm} a_j b_i c_n d_m \\ &= (\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}) a_j b_i c_n d_m \\ &= a_n c_n b_m d_m - a_m b_n c_n d_m \\ &= (\underline{a} \cdot \underline{c}) \, (\underline{b} \cdot \underline{d}) - (\underline{a} \cdot \underline{d}) \, (\underline{b} \cdot \underline{c}) \end{aligned}$$

We can prove a lot of identities such as these much more easily than otherwise using index notation, and so it is definitely worth becoming quite familiar with it's use and manipulations. We will be using it sporadically throughout this chapter when it is required; as always, if confused, come back to this section to refresh your memory.

3.2 Vector Space

A vector space $\mathcal{V}$ over $\mathcal{F}$ is a set with two operations:

1. Vector Addition - Adding vectors component by component: $\underline{u} + \underline{v} = u_i + v_i$
2. Scalar Multiplication - Multiplying the entire vector by a scalar: $\alpha\underline{u} = \alpha u_i$

The way to prove the existence of a vector space or sub-vector space is that it is closed under these two operations. This means that *any linear combination of vectors is also in the vector space.*

3.2.1 Span and Linear Independence

A *linear combination* of two vectors is the outcome of performing any linear operation (scalar multiplication and addition) on those two vectors. For example, $\alpha\underline{u} + \beta\underline{v}$ is an example of a linear combination of the vectors $\underline{u}$ and $\underline{v}$.

The *span* of a vector or sub-vector space is the *set of all linear combinations of those vectors.* The span of any vectors in a vector space itself forms a sub-vector space, as it is another vector set by definition:

$$\text{Span}(\underline{v}_1, \ldots, \underline{v}_k) = \left\{ \sum_{i=0}^{k} \alpha_i \underline{v}_i \;\middle|\; \alpha \in \mathcal{F} \right\} \tag{3.2}$$

where α_i are scalars. The vectors $\underline{v}_1, \ldots, \underline{v}_k$ are called *linearly independent* if and only if the only valid solution to

$$\sum_{i=0}^{k} \alpha_i \underline{v}_i = 0$$

is if all α_i's are zero. They are called *linearly dependent* if there is one non-trivial solution to this equation. Thus, to find whether a set of vectors are linearly dependent or not, write them as sum linear combination and attempt to find a solution.

Are $\underline{v}_1 = (1, 0, 1)$, $\underline{v}_2 = (2, 3, 1)$, $\underline{v}_3 = (1, 6, -1)$ *linearly independent?*

$$\alpha_1\underline{v}_1 + \alpha_2\underline{v}_2 + \alpha_3\underline{v}_3 = \begin{pmatrix} \alpha_1 + 2\alpha_2 + \alpha_3 \\ 3\alpha_2 + 6\alpha_3 \\ \alpha_1 + \alpha_2 - \alpha_3 \end{pmatrix} \stackrel{!}{=} 0$$

By inspection, we can see that a possible solution is $\alpha_1 = 3$, $\alpha_2 = -2$ and $\alpha_3 = 1$. Thus means that the three vectors are not linearly independent; the are in fact linearly dependent!

3.2.2 Basis and Dimension

The vectors $\underline{v}_1, \ldots, \underline{v}_k$ form a basis of $\mathcal{V}$ if and only if:

1. $\underline{v}_1, \ldots, \underline{v}_k$ are linearly independent
2. $\mathcal{V} = \text{Span}(\underline{v}_1, \ldots, \underline{v}_k)$

These two conditions essentially mean that we require that *every vector in the vector space can be written as a linear combination of the basis vectors.* Suppose that a non-zero vector $\underline{v}$ can be written as two different linear combinations of the basis vectors $\underline{u}_i$:

$$\begin{aligned}
\underline{v} = \sum_{i=0}^{n} \alpha_i \underline{u}_i &= \sum_{i=0}^{n} \beta_i \underline{u}_i \\
\sum_{i=0}^{n} (\alpha_i - \beta_i)\underline{u}_i &= 0 \\
\alpha_i - \beta_i &= 0 \\
\rightarrow \alpha_i &= \beta_i
\end{aligned}$$

We have just in fact shown that the linear combination of basis vectors is actually unique.

If $\underline{v}_1, \ldots, \underline{v}_k$ is a basis of $\mathcal{V}$, then we define the dimension of $\mathcal{V}$ as the 'size' of the set of basis vectors

$$\dim(\mathcal{V}) = k$$

For two vector spaces $\mathcal{V}$ and $\mathcal{W}$, we can say that

$$\dim(\mathcal{V} + \mathcal{W}) = \dim(\mathcal{V}) + \dim(\mathcal{W}) - \dim(\mathcal{V} \cap \mathcal{W}) \tag{3.3}$$

This is an intuitively obvious result, but can be proven by considering generalised vectors in each vector space.

Suppose that $\underline{u}_1, \ldots, \underline{u}_n$ are the basis for the sub-vector space $\mathcal{V} \cap \mathcal{W}$. Then $\dim(\mathcal{V} \cap \mathcal{W}) = n$. Suppose that there are further vectors $\underline{v}_1, \ldots, \underline{v}_j$ that form a basis of $\mathcal{V}$. Then $\dim(\mathcal{V}) = n + j$ as all the elements of $\mathcal{V} \cap \mathcal{W}$ must be contained within $\mathcal{V}$. Suppose that there are further vectors $\underline{w}_1, \ldots, \underline{w}_k$ that form a basis of $\mathcal{W}$. Then $\dim(\mathcal{W}) = n + k$ by similar logic.

The sub-vector space $\mathcal{V} + \mathcal{W}$ contains all the elements of $\mathcal{V}$ and $\mathcal{W}$, as well as their linear combinations. However, any of these combinations can be made out of a linear combination of the basis vectors of $\mathcal{V}$ or $\mathcal{W}$ or both.

Consider the vectors $\underline{V}_1$ and $\underline{W}_1$ defined by:

$$\begin{aligned}
\underline{V}_1 &= \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 + \ldots \\
\underline{W}_1 &= \beta_1 \underline{w}_1 + \beta_2 \underline{w}_2 + \ldots \\
\rightarrow \underline{V}_1 + \underline{W}_1 &= (\alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 + \ldots) + (\beta_1 \underline{w}_1 + \beta_2 \underline{w}_2 + \ldots)
\end{aligned}$$

It thus becomes clear that $\underline{V}_1 + \underline{W}_1$ is formed from a linear combination of the bases of $\mathcal{V}$ and $\mathcal{W}$. This means that $\mathcal{V} + \mathcal{W}$ has basis vectors

$$\underline{u}_1, \ldots, \underline{u}_n, \underline{v}_1, \ldots, \underline{v}_j, \underline{w}_1, \ldots, \underline{w}_k$$

It follows that $\dim(\mathcal{V}+\mathcal{W}) = n+k+j = (n+k)+(n+j)-n = \dim(\mathcal{V})+\dim(\mathcal{W})-\dim(\mathcal{V}\cap\mathcal{W})$ as required.

3.3 Scalar and Vector Products

These are the two most important vector operations covered in this chapter. Most students who have done a small amount of linear algebra should already be familiar with these. Note that all definitions will be defined for the three dimensional case.

3.3.1 Scalar Product

The scalar product, also known as the *dot product*, is defined as:

$$\underline{a} \cdot \underline{b} = a_i b_i = |\underline{a}||\underline{b}| \cos\theta = \langle a, b \rangle \tag{3.4}$$

where θ is the acute angle between the vectors. This will generate a scalar quantity from two vectors. We say that two vectors are *orthogonal* or perpendicular if $\underline{a} \cdot \underline{b} = 0$. Generally, a basis will be orthogonal, or else the vectors will not be linearly independent.

The scalar product *projects* the length of one vector onto another.

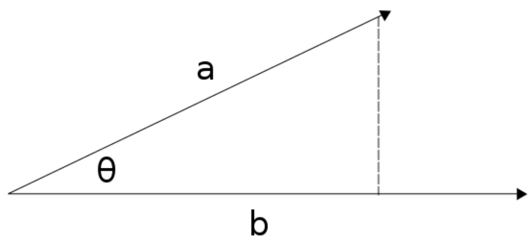

Figure 3.1: The projection property of the scalar product

The component along $\underline{b}$ is given by $\underline{a} \cdot \widehat{\underline{b}}$, where as the vector projection along $\underline{b}$ is given by $(\underline{a} \cdot \widehat{\underline{b}})\, \widehat{\underline{b}}$; the magnitude of the component multiplied given the direction of $\underline{b}$. We can use this property when writing a vector in terms of a basis. Suppose that $\underline{e}_1$, $\underline{e}_2$ and $\underline{e}_3$ are the three orthonormal vectors for some vector space. Then

$$\underline{v} = \alpha_1 \underline{e}_1 + \alpha_2 \underline{e}_2 + \alpha_3 \underline{e}_3$$

for constants α_i. We can thus find these constants by $\alpha_i = \underline{v} \cdot \underline{e}_i$, as this gives (in a sense) the 'amount' of the vector $\underline{v}$ in each of the unit directions.

3.3.2 Vector Product

The vector product, also known as the *cross product*, is defined as:

$$\underline{a} \times \underline{b} = \epsilon_{ijk} a_j b_k = |\underline{a}||\underline{b}| \sin\theta \tag{3.5}$$

where θ is the acute angle between the vectors. *This will generate a vector that is perpendicular to both of the vectors involved in the cross product.* Due to the nature of it's definition using the Levi-Civita tensor, it will reverse sign if the order of the operation is changed: $\underline{a} \times \underline{b} = -\underline{b} \times \underline{a}$.

We can compute the result of the vector product by computing the following determinant

$$\underline{a} \times \underline{b} = \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Do not worry if you don't know what a determinant is yet; this is coming up in Section (3.7).

3.3.3 Triple Scalar Product

The triple scalar product involves both the scalar and vector products, and is defined as

$$\langle \underline{a}, \underline{b}, \underline{a} \rangle = \underline{a} \cdot (\underline{b} \times \underline{c}) = \epsilon_{ijk} a_i b_j c_k \tag{3.6}$$

It's properties follow from the properties of the Levi-Civita Tensor:

– Linear in each argument

– Totally Antisymmetric

– Vanishes if two or more arguments are the same

If $\langle \underline{a}, \underline{b}, \underline{a} \rangle \neq 0$, then the three vectors $\underline{a}$, $\underline{b}$ and $\underline{c}$ are linearly independent. This is because the cross product $\underline{a} \times \underline{b}$ generates a vector perpendicular to both $\underline{b}$ and $\underline{c}$. If the subsequent dot product with $\underline{a}$ is non-zero, it means that $\underline{a}$ does not lie in the same plane as $\underline{b}$ and $\underline{c}$, and so the three vectors are linearly independent.

The triple scalar product also has some geometrical uses. For a figure defined by vectors $\underline{u}$, $\underline{v}$ and $\underline{w}$:

– Volume of Parallelepiped $= |\langle \underline{u}, \underline{v}, \underline{w} \rangle|$

– Volume of Tetrahedron $= \frac{1}{3!} | \langle \underline{u}, \underline{v}, \underline{w} \rangle |$

3.4 Cardinal Geometry

This section covers the techniques required to handle most of the geometrical questions that will be asked in this course. They will becomes easier to remember and then apply as one gets more familiar with the way the vector operations map to operations in three dimensional space.

3.4.1 Triangles

Suppose that a triangle in two dimensions has sides that are defined by the three vectors $\underline{a}$, $\underline{b}$ and $\underline{c}$. As these vectors must form a closed loop

$$\begin{aligned}\underline{a}+\underline{b}+\underline{c}&=0\\ \underline{a}\times(\underline{a}+\underline{b}+\underline{c})&=0\\ \underline{a}\times\underline{b}+\underline{a}\times\underline{c}&=0\\ &\rightarrow|\underline{a}\times\underline{b}|=|\underline{b}\times\underline{c}|=|\underline{a}\times\underline{c}|\end{aligned}$$

The area of the triangle is given by

$$\text{Area}=\frac{1}{2}|\underline{a}\times\underline{b}| \tag{3.7}$$

3.4.2 Lines in Three Dimensions

In two dimensions, we can define a line by it's gradient and some point that it passes through. We can do a similar thing in three dimensions, except we define the direction of the line, $\underline{p}$ and a point through which it passes $\underline{p}$. Parametrically, we can express thus as:

$$\underline{r}(t)=\underline{p}+t\underline{q} \tag{3.8}$$

This defines all the points on the line because as we vary the parameter t, the point specified moves along the direction of $\underline{q}$ with the constraint that at some point $(t=0)$ it must pass throught $\underline{p}$. This form can also be rearranged to express the line in terms of t:

$$\frac{x-p_x}{q_x}=\frac{y-p_y}{q_y}=\frac{z-p_z}{q_z} \tag{3.9}$$

We can find $\underline{q}$ by writing down two points on the line, or a line parallel to it, and taking their difference.

3.4.3 Planes

A plane consists of the locus of points that satisfy the condition that the line joining said points and another chosen point is perpendicular to some other vector. Thus, to define a plane, we just need a point and a *normal vector* that is perpendicular to the surface of the plane. This means the equation of a plane passing through $\underline{a}$ with normal vector $\underline{n}$ is

$$(\underline{r}-\underline{a})\,\underline{n}=0\rightarrow\underline{r}\cdot\underline{n}=d \tag{3.10}$$

We can also write this in 'expanded' form. Suppose that $\underline{n}=(a,b,c)$. Then, the equation of the plane becomes

$$ax+by+cz=d \tag{3.11}$$

We find $\underline{n}$ by finding two lines that lie in the plane (requiring three points at minimum) and taking their vector product as this generates a vector normal to both.

3.4.4 Geometric Techniques

This subsection is basically just a list of the manipulations that one should be familiar with when it comes to the geometry of lines and planes in three dimensions.

- Shortest Distance from a line to a point - Evidently, the shortest distance from a line to a point is in the direction perpendicular to the line. This means that we have to take the vector product with the unit vector $\hat{\underline{q}}$ for the line. For a general point $\underline{p}_o$:

$$d_{\text{min}} = \frac{|(\underline{p} - \underline{p}_o) \times \underline{q}|}{|\underline{q}|} \tag{3.12}$$

- Shortest Distance between two lines - Similarly, the shortest distance between two lines is the distance perpendicular to both. This will be in the direction of $\underline{q}_1 \times \underline{q}_2$ for the two lines (the cross product of the direction vectors). Thus, we just want find the projection of some line joining the two lines in this direction. Hence, the shortest distance is given by:

$$d_{\text{min}} = \frac{|(\underline{p}_1 - \underline{p}_2) \cdot (\underline{q}_1 \times \underline{q}_2)|}{|\underline{q}_1 \times \underline{q}_2|} \tag{3.13}$$

- Shortest Distance to a point from a plane - Again, this will be in the direction perpendicular to the surface of the plane, and so we have to find the projection of a line joining the given point $\underline{p}_o$ to a point in the plane $\underline{a}$ onto the normal to the plane $\underline{n}$.

$$d_{\text{min}} = (\underline{a} - \underline{p}_o)\, \hat{\underline{n}} \tag{3.14}$$

- Point of Intersection between a plane and a line - If it is a numerical question, we can simply express x, y, and z in parametric form using the equation of the line and substitute these into the equation for the plane. Solve for the free parameter (t), and substitute this back into the equation for the line to find the point of intersection.

  Suppose that a line is defined by $\underline{r} = \underline{a} + t\,\underline{b}$, and a plane by $(\underline{r} - \underline{c})\,\underline{d}$. Substituting the equation of the line into the equation of the plane:

$$(\underline{a} - \underline{c} + t\,\underline{b}) \cdot \underline{d} = 0$$
$$t = -\frac{\underline{d}\,(\underline{a} - \underline{c})}{\underline{b} \cdot \underline{d}}$$

  This is assuming that $\underline{b} \cdot \underline{d} \neq 0$, as otherwise the line and the plane would be parallel. This is a unique solution for t, meaning that there is a unique point of intersection of

$$\underline{p} = \frac{\underline{a}\,(\underline{b} \cdot \underline{d}) - \underline{d} \cdot (\underline{a} - \underline{c})}{\underline{b} \cdot \underline{d}} \tag{3.15}$$

- Line of Intersection between two planes - Assuming that two planes intersect, they will intersect along a line. The cross product of the normals to the plane $\underline{n}_1$ and $\underline{n}_2$ will give the direction of this line, and then one just has to find a point that lies on the intersection in order to have the equation of the line.

3.5 The Scalar Product

The scalar product between to 'entities' A and B is notated as $\langle A, B\rangle$. The reason why we say entities here is because the scalar product can be defined for more than just vectors or matrices. It must satisfy four conditions for it to be a scalar product:

$$\langle A, B\rangle = \langle B, A\rangle \tag{3.16}$$

$$\langle A, k\,B\rangle = k\,\langle A, B\rangle \tag{3.17}$$

$$\langle A, B + C\rangle = \langle A, B\rangle + \langle A, C\rangle \tag{3.18}$$

$$\langle A, A\rangle \geqslant 0 \tag{3.19}$$

If these properties are not satisfied, then we do not have ourselves a scalar product!

Does the expression

$$\langle f, g\rangle = \int_{-\infty}^{\infty} e^{-x^2} f(x)g(x)$$

define a scalar product? Consider the polynomial

$$p_a(x) = \frac{a_o + a_1 x + a_2 x^2 + \cdots + a_a x^a}{n_a}$$

Determine a general expression for n_a such that the polynomials p_a are normalised to one, so $\langle p_a, p_a\rangle = 1$.

First, we need to check whether it satisfies the conditions of the scalar product.

- Commutativity

$$\begin{aligned}\langle g, f\rangle &= \int_{-\infty}^{\infty} e^{-x^2} g(x)f(x) \\ &= \int_{-\infty}^{\infty} e^{-x^2} f(x)g(x) \\ &= \langle f, g\rangle\end{aligned}$$

- Linearity - Let $f(x) = \alpha a(x) + \beta b(x)$:

$$\begin{aligned}\langle \alpha a + \beta b, g\rangle &= \int_{-\infty}^{\infty} e^{-x^2} g(x)(\alpha a(x) + \beta b(x)) \\ &= \alpha \int_{-\infty}^{\infty} e^{-x^2} g(x)a(x) + \beta \int_{-\infty}^{\infty} e^{-x^2} g(x)b(x) \\ &= \alpha\,\langle a, g\rangle + \beta\,\langle b, g\rangle\end{aligned}$$

- Real

$$\begin{aligned}\langle f, f\rangle &= \int_{-\infty}^{\infty} e^{-x^2}\,(f(x))^2 \\ &\geqslant 0\end{aligned}$$

This is because the integrand, if non-zero, is positive definitive.

Thus, the expression does define a scalar product. Now for the next part:

$$
\begin{aligned}
\langle p_a, p_a \rangle &= \frac{1}{{n_a}^2} \int_{-\infty}^{\infty} e^{-x^2} (a_o + a_1 x + a_2 x^2 + \cdots + a_a x^a)^2 \\
&= \frac{1}{{n_a}^2} \int_{-\infty}^{\infty} e^{-x^2} \sum_{i,j=1}^{a} a_i x^i a_j x^j \\
&= \frac{1}{{n_a}^2} \int_{-\infty}^{\infty} e^{-x^2} \sum_{i,j=1}^{a} a_i a_j x^{i+j}
\end{aligned}
$$

Evidently, any terms where $i + j$ is odd will disappear, and hence let $i + j = n$ for even n.

$$
\begin{aligned}
I_n &= \int_{-\infty}^{\infty} e^{-x^2} x^n dx \\
&= -\frac{1}{2} \left(\left[x^{n-1} e^{-x^2} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} (n-1) x^{n-2} e^{-x^2} dx \right) \\
&= \frac{1}{2}(n-1) \int_{-\infty}^{\infty} x^{n-2} e^{-x^2} dx \\
&= \left(\frac{1}{2}\right)^n (n-1)!! \int_{-\infty}^{\infty} e^{-x^2} dx \\
\rightarrow I_n &= \left(\frac{1}{2}\right)^n (n-1)!! \sqrt{\pi}
\end{aligned}
$$

Thus

$$
\langle p_a, p_a \rangle = \frac{1}{{n_a}^2} \sum_{i,j=1}^{a} a_i a_j \left(\frac{1}{2}\right)^n (n-1)!! \sqrt{\pi}
$$

But we want to normalise.

$$
\rightarrow n_a = \sqrt{\sum_{i,j=1}^{a} a_i a_j (i+j-1)!! \sqrt{\pi} \left(\frac{1}{2}\right)^{i+j}}
$$

This is the value of n_a required in order for $\langle p_a, p_a \rangle = 1$.

3.6 Matrices

A matrix can be thought of as an array of elements, which can be integers, imaginary numbers or even operators. If we say a matrix $\mathbf{A}$ is an n by m matrix, this means that the matrix has n rows and m columns. We generally notate this as $n \times m$.

$$\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}$$

As can be seen in the matrix above, we refer to individual matrix elements by specifying their row and column numbers. Thus, like with vectors in index notation, we write a matrix $\mathbf{A}$ as A_{mn}.

At this stage, let us introduce an important matrix identity; the idea of the unit matrix I.

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \iddots & 0 \\ 0 & \iddots & \ddots & 0 \\ 0 & \dots & \dots & 1 \end{pmatrix} \tag{3.20}$$

That is, it is the matrix with 1's along the diagonal and zero's everywhere else. In index notation, it is written as δ_{ij} as it satisfies the properties of the Kronecker-Delta. For off-diagonal elements, $i \neq j$, meaning that the entry is zero. For on-diagonal elements, $i = j$, and so the entry is one. Multiplying a matrix by the unit matrix returns the original matrix; it can be thought of as multiplying by one.

3.6.1 Matrix Operations

There are a number of basic operations that we can perform on matrices, such as the following:

- Transpose - This involves 'flipping' the elements of the matrix elements about the diagonal; corresponding elements on either side of the diagonal are swapped.

$$(A_{ij})^T = A_{ji} \tag{3.21}$$

  This swaps the order of the indices in the index notation; order is very important here!

$$\mathbf{A} = \begin{pmatrix} 1 & 7 & -3 \\ -2 & 3 & 6 \\ 4 & 1 & 5 \end{pmatrix} \rightarrow \mathbf{A}^T = \begin{pmatrix} 1 & -2 & 4 \\ 7 & 3 & 1 \\ -3 & 6 & 5 \end{pmatrix}$$

- Complex Conjugate - Like with a normal complex number, we can apply the complex conjugate operation to an entire matrix; this takes the complex conjugate of each individual entry.

$$\overline{(B)}_{ij} = \overline{B_{ij}} \tag{3.22}$$

  Remember, this only operates on each individual element, and so we can take the conjugation operator inside the bracket.

$$\mathbf{B} = \begin{pmatrix} 1+i & -i & 4 \\ -i & -7+4i & i \\ 4 & i & 13 \end{pmatrix} \rightarrow \overline{\mathbf{B}} = \begin{pmatrix} 1-i & i & 4 \\ i & -7-4i & -i \\ 4 & -i & 13 \end{pmatrix}$$

- Hermitian Conjugate - This combines the previous two operations, and is usually notated by †.

$$\mathbf{C}^\dagger = \overline{\mathbf{C}}^T \tag{3.23}$$

3.6.2 Matrix Multiplication

As many of you may already know, we can multiply matrices together. Matrix multiplication is defined as:

$$C_{ik} = A_{ij}B_{jk} \tag{3.24}$$

Note that, *in general*, this is not commutative; that is, $\mathbf{AB} \neq \mathbf{BA}$. This means, again, that order is important. This definition also means that we can only multiply matrices if they are of the appropriate dimensions. For us to be able to multiply to matrices together, we require *the number of columns of the first to be equal to the number of rows of the second.* For example,

$$\mathbf{E} = \begin{pmatrix} 1 & -4 & 7 \\ 9 & 3 & -2 \end{pmatrix} \text{ and } \mathbf{F} = \begin{pmatrix} 3 & -2 \\ 4 & 6 \end{pmatrix}$$

are not able to be multiplied.

Why is this? It comes down to the way in which we perform matrix multiplication. For matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ such that $\mathbf{AB} = \mathbf{C}$, we take the dot product of the first row vector $\mathbf{A}$ with the first column vector $\mathbf{B}$, and this becomes the entry C_{11}. The next entry C_{12} is given by the product of the first row vector of $\mathbf{A}$ with the *second* column vector of $\mathbf{B}$. We progress along the columns of $\mathbf{B}$ with the first row of $\mathbf{A}$ until we run out of columns, and these entries form the first row of $\mathbf{C}$. We then use the second row vector of $\mathbf{A}$ and perform the same operation on the columns of $\mathbf{B}$ to give the second row of $\mathbf{C}$. We continue this until we have done all the row vectors of $\mathbf{A}$.

This might seem initially confusing, so let's take a look at an example. Suppose that we have the two matrices

$$\mathbf{A} = \begin{pmatrix} 1 & -4 & 7 \\ 9 & 3 & -2 \end{pmatrix} \text{ and } \mathbf{B} = \begin{pmatrix} 3 & -2 & 1 \\ 4 & 6 & 2 \\ 1 & 7 & 4 \end{pmatrix}$$

The product $\mathbf{AB}$ is defined, while $\mathbf{BA}$ is not due to the dimensions of each matrix. Let us compute the first of these. The dot product of the first row of $\mathbf{A}$ and first column of $\mathbf{B}$ is -6. Thus, this becomes the upper most left hand entry of the resultant matrix $\mathbf{C}$. The first row, second column: 23....We continue on this process until we obtain the result of:

$$\mathbf{C} = \mathbf{AB} = \begin{pmatrix} -6 & 23 & 21 \\ 37 & -14 & 7 \end{pmatrix}$$

Notice how we have multiplied a 2×3 matrix by a 3×3 matrix to generate a 2×3 matrix. We can think of the multiplication as 'collapsing' the inner two numbers; 2×3 and 3×3 $\rightarrow 2\times3$. This is always a good way to check whether your matrix multiplication is correct; if the dimensions of the resultant matrix are not correct, then something has gone wrong!

Now it might be a little more obvious as to why the matrices have to be of appropriate size; the dot product of two vectors is only defined if they are the same length, which places restrictions on the row and column sizes of the matrices.

3.6.3 Gaussian Elimination and Rank

The process of Gaussian elimination allows us to perform row operations on a matrix to reduce it to upper echelon form. For example, a matrix $\mathbf{A}$ in upper echelon form resembles

$$\mathbf{A} = \begin{pmatrix} 1 & -4 & 7 \\ 0 & 3 & 4 \\ 0 & 0 & 5 \end{pmatrix}$$

In general, upper echelon form is where the matrix only has entries above or on the diagonal, and the rest of the matrix is filled with zeros. We can do this because certain manipulations, such as swapping columns and rows, or multiplying a row/column by a constant does not change the properties of the matrix.

So how do we perform row reduction? Starting from the upper left-hand corner, we take the first non-zero entry, and divide that row by that entry to give us one. We then take this row away from the other rows to get a zero below the 'one' entry, remembering to modify the other entries in the other rows. Once we have obtained zero's in the first column, we move down and right by one, and apply the same process.

Consider the matrix $\mathbf{B}$ with rows R_1, R_2 and R_3.

$$\mathbf{B} = \begin{pmatrix} 1 & 4 & -3 \\ 3 & 2 & 7 \\ 6 & 4 & 1 \end{pmatrix}$$

We now want to perform row operations on this matrix. Note that we do not use an equality sign, a technically the matrices are not equal, but they exhibit the same properties.

$$\begin{pmatrix} 1 & 4 & -3 \\ 3 & -10 & 16 \\ 6 & 4 & 1 \end{pmatrix} \begin{matrix} -3 \\ + \\ \end{matrix}$$

$$\begin{pmatrix} 1 & 4 & -3 \\ 0 & -10 & 16 \\ 0 & -20 & 19 \end{pmatrix} \begin{matrix} -6 \\ \\ + \end{matrix}$$

$$\begin{pmatrix} 1 & 4 & -3 \\ 0 & 1 & -8/5 \\ 0 & -20 & 19 \end{pmatrix} \ \big| \cdot -1/10$$

$$\begin{pmatrix} 1 & 4 & -3 \\ 0 & 1 & -8/5 \\ 0 & 0 & -13 \end{pmatrix} \begin{matrix} \\ 20 \\ + \end{matrix}$$

$$\begin{pmatrix} 1 & 4 & -3 \\ 0 & 1 & -8/5 \\ 0 & 0 & 1 \end{pmatrix} \ \big| \cdot -1/13$$

Thus, in five steps, we have reduced the matrix $\mathbf{B}$ to upper echelon form.

A useful property to define at this stage is the *rank* of a matrix. The rank of a matrix gives the size of the image of the matrix; that is, it gives the dimensions of the space that the matrix projects into.

$$rk(\mathbf{M}) = \dim(\mathrm{Im}(\mathbf{M})) \tag{3.25}$$

For example, if a matrix represents the coefficients in a system of linear equations (see Section (3.10)), then the nature of the solution will depend on whether the matrix has a rank that is maximal or otherwise. *If the rank of a matrix is less than it's dimensions, then it does not have an inverse.* We can find the rank of a matrix by reducing it to upper echelon form; the rank is the number of 'steps' in this form. For example, the matrix **B** above is of rank 3. We will put this idea into practise later on in this chapter.

3.6.4 Matrix Trace

The trace is the sum of all the diagonal elements of a matrix.

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} A_{ii} \tag{3.26}$$

It has some important properties:

- $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$
- $\mathrm{tr}(k\mathbf{A}) = k\,\mathrm{tr}(\mathbf{A})$
- $\mathrm{tr}(\mathbf{A}^T) = \mathrm{tr}(\mathbf{A})$
- $\mathrm{tr}\left(\overline{\mathbf{A}}\right) = \overline{\mathrm{tr}(\mathbf{A})}$
- $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$

The trace is *base invariant*, meaning that all matrices it is equal to the sum of the eigenvalues, but more on this later!

3.7 Determinants

A determinant maps vectors $\underline{a}_1, \ldots, \underline{a}_n$ to a number number $\det(\underline{a}_1, \ldots, \underline{a}_n)$ such that:

$$\det(\ldots, \alpha\underline{a} + \beta\underline{b}, \ldots) = \alpha \det(\ldots, \underline{a}, \ldots) + \beta \det(\ldots, \underline{b}, \ldots) \tag{3.27}$$

$$\det(\ldots, \underline{a}, \ldots, \underline{b}, \ldots) = -\det(\ldots, \underline{b}, \ldots, \underline{a}, \ldots) \tag{3.28}$$

$$\det(\underline{e}_1, \ldots, \underline{e}_n) = 1 \tag{3.29}$$

$$\det(\ldots, \underline{a}_i, \ldots, \underline{a}_i, \ldots) = 0 \tag{3.30}$$

Notice how is obeys similar symmetry properties to the Levi-Civita tensor; this is because one can define the determinant using this tensor, but we will not do this here. An alternative, computational definition of the determinant is

$$\det \mathbf{A} = \sum_{i,j=1}^{n} (-1)^{i+j} A_{ij} \det \mathbf{A}_{(i,j)} \tag{3.31}$$

where $\mathbf{A}_{(i,j)}$ is the matrix $\mathbf{A}$ without row i and column j. This is known as the co-factor matrix. The determinant is only defined *if the rank of the matrix is maximal.*

3.7.1 Properties of the Determinant

The determinant has a number of properties. For an $n \times n$ square matrix $\mathbf{A}$ and invertible matrix $\mathbf{P}$, and scalar λ:

- $\det(\mathbf{A}^T) = \det \mathbf{A}$
- $\det(\mathbf{AB}) = \det \mathbf{A} \det \mathbf{B}$
- $\det(\mathbf{PAP}^{-1}) = \det \mathbf{A}$
- $\det(\mathbf{A}^{-1}) = 1/\det \mathbf{A}$
- $\det(\lambda \mathbf{A}) = \lambda^n \det \mathbf{A}$

3.7.2 Calculating the Determinant

We can calculate the determinant by the cofactor method. For a 2×2 matrix, the determinant is simply given by:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \tag{3.32}$$

For a 3×3 matrix, this becomes:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & h \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \tag{3.33}$$

Evidently, this process can be repeated for larger and large matrices, but the calculation soon becomes horrendously complicated, and so we start to use computation at this point.

3.8 Matrix Inverse

A matrix inverse is a matrix $\mathbf{A}^{-1}$ that is related to the matrix $\mathbf{A}$ such that:

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I} \tag{3.34}$$

A matrix only has a well defined inverse if it is a quadratic matrix. A quadratic matrix is only invertible if the rank of said matrix is maximal, meaning that all rows are linearly independent (as otherwise the determinant is undefined).

Unlike with integers, the inverse is not just simply the inversion of all the entries in the matrix (except for a diagonal matrix); this should not reverse the action of the matrix. It is in fact a little more complicated. There are two main methods that we can use to calculate the determinant:

1. The Co-factor Method - For a matrix $\mathbf{A}$, let $\mathbf{C}$ be the co-factor matrix defined by

   $$C_{ij} = (-1)^{i+j} \det A_{(i,j)}$$

   Then, the inverse of $\mathbf{A}$ is calculated by:

   $$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \mathbf{C}^T \tag{3.35}$$

2. Row Operations - We can write the matrix as an augmented matrix with the identity matrix on the right-hand side. If we row reduce the original matrix on the left-hand side to the identity matrix, applying the same operations to the right-hand side, we will obtain the inverse on the right-hand side.

Find the inverse of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & -2 \\ 1 & -3 & 0 \end{pmatrix}$$

by row reduction. Check your answer using the cofactor method.

Writing the system as an augmented matrix:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 2 & 1 & -2 & 0 & 1 & 0 \\ 1 & -3 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\begin{array}{cc} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 1 & -3 & 0 & 0 & 0 & 1 \end{array}\right) & \begin{array}{l} -3 \\ + \\ \end{array} \end{array}$$

$$\begin{array}{cc} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & -3 & 1 & -1 & 0 & 1 \end{array}\right) & \begin{array}{l} -1 \\ \\ + \end{array} \end{array}$$

$$\begin{array}{cc} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -7 & 3 & 1 \end{array}\right) & \begin{array}{l} \\ -3 \\ + \end{array} \end{array}$$

$$\begin{array}{cc} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -6 & 3 & 1 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -7 & 3 & 1 \end{array}\right) & \begin{array}{l} + \\ \\ 1 \end{array} \end{array}$$

Thus, the inverse of the matrix $\mathbf{A}$ is

$$\mathbf{A}^{-1} = \begin{pmatrix} -6 & 3 & 1 \\ -2 & 1 & 0 \\ -7 & 3 & 1 \end{pmatrix}$$

Now let us check this using the co-factor matrix method. It is trivial to find that the determinant is 1. The co-factor matrix is:

$$\mathbf{C} = \begin{pmatrix} -6 & -2 & -7 \\ 3 & 1 & 3 \\ 1 & 0 & 1 \end{pmatrix}$$

The result follows by taking the transpose.

3.9 Eigenvalues, Eigenvectors and Quadratic Forms

A quadratic matrix $\mathbf{A}$ has a scalar eigenvalue λ with corresponding eigenvector $\underline{v}$ if these obey the equation

$$\mathbf{A}\underline{v} = \lambda\underline{v} \tag{3.36}$$

We can find the eigenvalues of the matrix by calculating the characteristic polynomial given by

$$\mathcal{X}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) \tag{3.37}$$

and setting it equal to zero. The characteristic polynomial has a number of properties:

- It is basis independent

$$\begin{aligned}
\mathcal{X}_{\mathbf{PAP}^{-1}}(\lambda) &= \det(\mathbf{PAP}^{-1} - \lambda\mathbf{I}) \\
&= \det(\mathbf{P}(\mathbf{A} - \lambda\mathbf{I})\mathbf{P}^{-1}) \\
&= \det\mathbf{P}\det(\mathbf{A} - \lambda\mathbf{I})\det\mathbf{P}^{-1} \\
&= \det(\mathbf{A} - \lambda\mathbf{I})
\end{aligned}$$

- The coefficients are basis independent

$$\begin{aligned}
\mathcal{X}(\lambda) &= \prod_{i=1}^{n}(A_{ii} - \lambda) + Q(\lambda^{n-2}) \\
&= (-1)^n\lambda^n + (-1)^{n-1}\lambda^{n-1}(\sum_{i=1}^{n} A_{ii}) + \cdots + \det\mathbf{A} \\
c_n &= (-1)^n \\
c_{n-1} &= (-1)^{n-1}\,\mathrm{tr}\mathbf{A} \\
c_o &= \det\mathbf{A}
\end{aligned}$$

To find the eigenvectors, we have to apply the condition that

$$(\mathbf{A} - \lambda\mathbf{I})\underline{v} = 0 \tag{3.38}$$

A matrix can be diagonalised if it has a set of eigenvectors $\underline{v}_1, \ldots, \underline{v}_n$ that form a basis. We define

$$\mathbf{P} = (\underline{v}_1, \ldots, \underline{v}_n)$$

i.e a matrix with the eigenvectors as column vectors. It is invertible because it is of maximal rank owing to the fact that $\underline{v}_1, \ldots, \underline{v}_n$ are linearly independent because they form a basis. We can thus diagonalise a matrix by

$$\mathbf{P}^{-1}\mathbf{AP} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \tag{3.39}$$

where $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix with the eigenvalues along the diagonal. This means that as soon as we have found the eigenvalues of a matrix, we know it's diagonalised form.

Lastly, two matrices $\mathbf{A}$ and $\mathbf{B}$ are simultaneously diagonalisable if and only if

$$[\mathbf{A}, \mathbf{B}] = 0 \tag{3.40}$$

$$\mathbf{AB} - \mathbf{BA} = 0 \tag{3.41}$$

Let $\widehat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ and $\widehat{\mathbf{B}} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P}$.

$$
\begin{aligned}
[\mathbf{A}, \mathbf{B}] &= \mathbf{P}\widehat{\mathbf{A}}\mathbf{P}^{-1}\mathbf{P}\widehat{\mathbf{B}}\mathbf{P}^{-1} - \mathbf{P}\widehat{\mathbf{B}}\mathbf{P}^{-1}\mathbf{P}\widehat{\mathbf{A}}\mathbf{P}^{-1} \\
&= \mathbf{P}\widehat{\mathbf{A}}\widehat{\mathbf{B}}\mathbf{P}^{-1} - \mathbf{P}\widehat{\mathbf{B}}\widehat{\mathbf{A}}\mathbf{P}^{-1} \\
&= \mathbf{P}\left(\widehat{\mathbf{A}}\widehat{\mathbf{B}} - \widehat{\mathbf{B}}\widehat{\mathbf{A}}\right)\mathbf{P}^{-1} \\
&= \mathbf{P}\left[\widehat{\mathbf{A}}, \widehat{\mathbf{B}}\right]\mathbf{P}^{-1} \\
&= 0
\end{aligned}
$$

The other side of the proof is much more complicated, and will not be covered here.

Diagonalise the matrix

$$
\mathbf{A} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}
$$

over complex numbers.

First, let us find the eigenvalues.

$$
\begin{aligned}
\mathcal{X}(\lambda) &= \begin{vmatrix} \cos\phi - \lambda & -\sin\phi \\ \sin\phi & \cos\phi - \lambda \end{vmatrix} \\
&= (\cos\phi - \lambda)(\cos\phi - \lambda) + \sin^2\phi \\
&= \cos^2\phi + \sin^2\phi - 2\cos\phi\lambda + \lambda^2 \\
0 &\overset{!}{=} \lambda^2 - 2\cos\phi\lambda + 1 \\
\lambda &= \cos\phi \pm i\sin\phi
\end{aligned}
$$

For $\lambda_1 = \cos\phi + i\sin\phi$:

$$
\begin{aligned}
(\mathbf{A} - \lambda_1\mathbf{I})\underline{v}_1 &= \begin{pmatrix} -i\sin\phi & -\sin\phi \\ \sin\phi & -i\sin\phi \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \overset{!}{=} 0 \\
-ix - y &= 0 \\
x - iy &= 0 \\
x &= iy \\
\rightarrow \underline{v}_1 &= \frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \end{pmatrix}
\end{aligned}
$$

For $\lambda_2 = \cos\phi - i\sin\phi$:

$$
\begin{aligned}
(\mathbf{A} - \lambda_1\mathbf{I})\underline{v}_1 &= \begin{pmatrix} i\sin\phi & -\sin\phi \\ \sin\phi & i\sin\phi \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \overset{!}{=} 0 \\
ix - y &= 0 \\
x + iy &= 0 \\
x &= -iy \\
\rightarrow \underline{v}_1 &= \frac{1}{\sqrt{2}}\begin{pmatrix} -i \\ 1 \end{pmatrix}
\end{aligned}
$$

Thus, we can automatically write down the diagonalised matrix:

$$
\widehat{\mathbf{A}} = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}
$$

3.9.1 Quadratic Forms

Quadratic forms are polynomial expressions of the form

$$Q(x) = \underline{x}^T \mathbf{A} \underline{x} = c \tag{3.42}$$

where $\mathbf{A}$ is a symmetric matrix.

$$\begin{aligned}
\left(\underline{x}^T \mathbf{A'} \underline{x}\right)^T &= \underline{x} \mathbf{A'}^T \underline{x}^T \\
2Q(x) &= \underline{x}^T \mathbf{A'} \underline{x} + \underline{x} \mathbf{A'}^T \underline{x}^T \\
&= \underline{x}^T \left(\mathbf{A'} + \mathbf{A'}^T\right) \underline{x} \\
Q(x) &= \frac{1}{2} \underline{x}^T \left(\mathbf{A'} + \mathbf{A'}^T\right) \underline{x}
\end{aligned}$$

Thus $\mathbf{A}$ can always be written as a symmetric matrix as

$$\mathbf{A} = \frac{1}{2}\left(\mathbf{A'} + \mathbf{A'}^T\right)$$

If we obtain the eigenvalues of this matrix, we can deduce the form of the surfaces or curves described by Q(x).

Condition on λ_i	2D	3D
All equal, same sign as c	circle	sphere
All same sign as c	ellipse	ellipsoid
λ_i with both signs	hyperbola	hyperboloid

The lengths of the semi-major axis of these are given by:

$$l_i = \sqrt{\frac{c}{\lambda_i}} \tag{3.43}$$

A curve in two-dimensional space is defined by all $\underline{x} = (x, y)^T$ which solve the equation $x^2 + 3y^2 - 2xy = 1$. Show that the curve is an ellipse and determine the length of it's two axes.

Clearly

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix}$$

Finding the eigenvalues:

$$\begin{aligned}
\begin{vmatrix} 1-\lambda & -1 \\ -1 & 3-\lambda \end{vmatrix} &\overset{!}{=} 0 \\
(1-\lambda)(3-\lambda) - 1 &= 0 \\
\lambda^2 - 4\lambda + 2 &= 0 \\
&\rightarrow \lambda = 2 \pm \sqrt{2}
\end{aligned}$$

Using (3.43), we can thus intermediately state that:

$$\begin{aligned}
\text{Major Axis} &= \frac{1}{\sqrt{2+\sqrt{2}}} \\
\text{Minor Axis} &= \frac{1}{\sqrt{2-\sqrt{2}}}
\end{aligned}$$

3.10 Systems of Linear Equations

As we have seen in the previous chapter, we can solve systems of equations by the use of matrices. For example, the system of linear equations given by

$$\begin{array}{ccccccc} a_1x & + & a_2y & + & a_3z & = & b_1 \\ a_4x & + & a_5y & + & a_6z & = & b_2 \\ a_7x & + & a_8y & + & a_9z & = & b_3 \end{array}$$

can be written in matrix from as

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_5 \\ a_6 & a_7 & a_9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Thus, for a vector of variables $\underline{x}$ and a vector of solutions $\underline{b}$, then we have

$$\mathbf{A} \cdot \underline{x} = \underline{b} \tag{3.44}$$

For a homogeneous equation (that is, $\underline{b} = 0$), $\det \mathbf{A} = 0$ in order for a set of solutions to exist.

For an inhomogeneous equation (that is, $\underline{b} \neq 0$), $\det \mathbf{A} \neq 0$ for a solution. The nature of the solution will depend on the rank of the coefficient matrix $\mathbf{A}$. For a three-variable system:

- $\text{rk}(\mathbf{A}) = 3 \rightarrow$ unique solution (single point)
- $\text{rk}(\mathbf{A}) = 2 \rightarrow$ a line (one free parameter)
- $\text{rk}(\mathbf{A}) = 1 \rightarrow$ a plane (two free parameters)

There are multiple methods to solve a system of linear equations. These include:

- Row Reduction - Use row reduction on the augmented coefficient matrix with the solutions on the right-hand side, and then solve the 'reduced' system of linear equations. This is generally the preferred method as it handles free parameters and discontinuities more easily than the other methods
- Matrix Inverse - Assuming that the matrix is invertible, we can simply calculate $\underline{x} = \mathbf{A}^{-1}\underline{b}$.
- Cramer's Method - This involves replacing each of the columns of $\mathbf{A}$ by the solutions of the equation and calculating the determinant. Let $\mathbf{B}_i$ be the matrix $\mathbf{A}$ with column i replaced by $\underline{b}$.

$$x_i = \frac{\det \mathbf{B}_i}{\det \mathbf{A}} \tag{3.45}$$

$$\underline{x} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} \det \mathbf{B}_1 \\ \det \mathbf{B}_2 \\ \det \mathbf{B}_3 \end{pmatrix} \tag{3.46}$$

- Diagonalisation - We can diagonalise the coefficient matrix, which will give us a one-to-one equality between coefficient modified variables and the set of solutions. Let $\mathbf{P} = (\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n)$, where the vectors $\underline{v}_i$ are the normalised eigenvectors of the coefficient matrix $\mathbf{A}$.

$$\frac{d\underline{x}}{dt} + \mathbf{A}\underline{x} = \underline{b}$$

$$\mathbf{P}^{-1} \frac{d\underline{x}}{dt} \mathbf{P} + \widehat{\mathbf{A}}\underline{x} = \mathbf{P}^{-1}\underline{b}\, \mathbf{P}$$

The linear system

$$\begin{aligned} x+y+z &= 1 \\ x+2y+4z &= \eta \\ x+4y+10z &= \eta^2 \end{aligned}$$

depends on the parameter η. Show that the rank of the coefficient matrix is two. Explicitly solve the system for cases where a solution exists.

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & \eta \\ 1 & 4 & 10 & \eta^2 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \eta-1 \\ 0 & 3 & 9 & \eta^2-1 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \eta-1 \\ 0 & 0 & 0 & (\eta^2-1)-3(\eta-1) \end{array}\right)$$

Thus, the rank of the coefficient matrix is 2. For a solution to exist

$$\begin{aligned} (\eta^2-1)-3(\eta-1) &\stackrel{!}{=} 0 \\ \eta^2-3\eta+2 &= 0 \\ (\eta-2)(\eta-1) &= 0 \\ &\to \eta = 2 \text{ or } 1 \end{aligned}$$

The solutions will be parametrised by a free variable as the rank of the matrix is one less than maximal.

Let $\eta = 1$ and $\underline{x} = (x, y, t)$. Here we have chosen the z coordinate as our free parameter. Hence:

$$\begin{aligned} x+y+t &= 1 \\ y+3t &= 0 \\ y &= -3t \\ x &= 1+2t \end{aligned}$$

The solutions will be in the form of the line

$$\underline{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$$

Similarly, when $\eta = 2$, the solutions are

$$\underline{x} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$$

Note that the solutions have the same direction vector; they are parallel lines passing through different points.

3.11 Matrix Types

There are many different types of matrices, all of which have different properties. Some relevant ones have been included below.

3.11.1 Hermitian and Symmetric Matrices

Hermitian matrices are those that obey the property

$$\mathbf{H} = \mathbf{H}^\dagger \tag{3.47}$$

Symmetric matrices obey the same property, except they are restricted to just the real case (i.e just the transpose). *Their eigenvalues are real and their eigenvectors are orthonormal.* Using the scalar product:

$$\begin{aligned}\lambda \langle \underline{v}, \underline{v} \rangle &= \langle \underline{v}, \mathbf{H}\underline{v} \rangle \\ &= \langle \mathbf{H}\underline{v}, \underline{v} \rangle \\ &= \langle \lambda\underline{v}, \underline{v} \rangle \\ &= \overline{\lambda} \langle \underline{v}, \underline{v} \rangle\end{aligned}$$

This means that $\lambda = \overline{\lambda}$, meaning that the eigenvalues must be real. Consider two eigenvalues and corresponding eigenvectors of $\mathbf{H}$:

$$\begin{aligned}\mathbf{H}\underline{v}_1 &= \lambda_1 \underline{v}_1 \\ \mathbf{H}\underline{v}_2 &= \lambda_2 \underline{v}_2 \\ (\lambda_1 - \lambda_2) \langle \underline{v}_1, \underline{v}_2 \rangle &= \langle \lambda_1 \underline{v}_1, \underline{v}_2 \rangle - \langle \underline{v}_1, \lambda_2 \underline{v}_2 \rangle \\ &= \langle \underline{v}_1, \mathbf{H}\underline{v}_2 \rangle - \langle \underline{v}_1, \mathbf{H}\underline{v}_2 \rangle \\ &= 0\end{aligned}$$

As $\lambda_1 \neq \lambda_2$, $\langle \underline{v}_1, \underline{v}_2 \rangle \stackrel{!}{=} 0$. This means that the eigenvectors are orthogonal.

3.11.2 Unitary and Orthogonal Matrices

Orthogonal matrices are those that satisfy

$$\mathbf{O}\mathbf{O}^T = \mathbf{I} \tag{3.48}$$

$$\mathbf{O}^T = \mathbf{O}^{-1} \tag{3.49}$$

The column vectors of such matrices are mutually orthogonal.

$$\begin{aligned}\mathbf{O} &= (\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_n) \\ \mathbf{O}\mathbf{O}^T &= (\underline{c}_1 \cdot \underline{c}_1, \underline{c}_2 \cdot \underline{c}_2, \ldots, \underline{c}_n \cdot \underline{c}_n) = \mathbf{I} \\ c_i c_j &= \delta_{ij}\end{aligned}$$

For all $i \neq j$, the product is zero; thus the columns are mutually orthogonal.

Unitary matrices are very similar to orthogonal matrices, except they are defined over complex fields as well.

$$\mathbf{U}\mathbf{U}^\dagger = \mathbf{I} \tag{3.50}$$

$$\mathbf{U}^\dagger = \mathbf{U}^{-1} \tag{3.51}$$

The eigenvalues of unitary matrices are complex phases with magnitude one.

$$\begin{aligned}
\mathbf{U}^\dagger \underline{v} &= \lambda \underline{v} \\
\mathbf{U} \underline{v} &= \lambda \underline{v} \\
\mathbf{U}^\dagger \mathbf{U} \underline{v} &= \mathbf{U}^\dagger \lambda \underline{v} \\
\underline{v} &= \lambda \underline{v} \\
\rightarrow \lambda &= e^{i\phi}
\end{aligned}$$

3.11.3 Projection Matrices

Projection matrices map a vector space onto a sub-vector space. If one projects a vector into a sub-vector space, and then attempts to apply the projection again, one should obtain the same result. This means that for projection matrices, we can write that:

$$\mathbf{P}^2 = \mathbf{P} \tag{3.52}$$

The projector will create a projection for each non zero element of the diagonal (all off-diagonal entries must be zero by this definition), and so the size of the vector space is given by $\mathrm{tr}(\mathbf{P})$.

$$\begin{aligned}
\mathbf{P}^2 \underline{v} &= \lambda \underline{v} \\
\mathbf{P} \underline{v} &= \lambda \underline{v} \\
\mathbf{P}^2 \underline{v} &= \mathbf{P} \lambda \underline{v} \\
\mathbf{P}(\lambda \underline{v}) &= \lambda \underline{v} \\
\lambda(\lambda \underline{v}) &= \lambda \underline{v} \\
\lambda(\lambda - 1) &= 0
\end{aligned}$$

Thus, projection matrices have eigenvalues $\lambda = 0$ or 1.

3.11.4 Rotational Matrices

Rotational matrices are those that leave the scalar product invariant. This makes sense for it to be a rotation, as a rotation leaves the size of something the same, but just changes it's orientation.

$$\begin{aligned}
\langle \mathbf{R}\underline{v}, \mathbf{R}\underline{w} \rangle &= \langle \underline{v}, \underline{w} \rangle \\
&= R_{ij} v_j R_{ik} v_k \\
&= (R_{ji})^T R_{ik} v_j w_k \\
\rightarrow \langle \underline{v}, \underline{w} \rangle &= \mathbf{R}^T \mathbf{R} \langle \underline{v}, \underline{w} \rangle
\end{aligned}$$

Thus, the product of the matrix and it's transpose must be equal to the unit matrix in order to preserve size.

$$\begin{aligned}
\mathbf{R}^T \mathbf{R} &= \mathbf{I} \\
\det(\mathbf{R}^T \mathbf{R}) &= 1 \\
\det(\mathbf{R})^2 &= 1 \\
\det(\mathbf{R}) &= \pm 1
\end{aligned}$$

If we impose these two conditions on a general matrix two-dimensional matrix, we can calculate the rotational matrix in two dimensions.

$$
\begin{aligned}
\mathbf{R} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\
\mathbf{R}^T\mathbf{R} &\overset{!}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} \\
&= \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix} \\
\det(\mathbf{R}) &\overset{!}{=} \pm 1 \\
ad - bc &= \pm 1
\end{aligned}
$$

Solving the resultant equations simultaneously, we obtain the rotational matrix:

$$\mathbf{R}(\theta) = \begin{pmatrix} \cos\theta & \mp\sin\theta \\ \sin\theta & \pm\cos\theta \end{pmatrix} \tag{3.53}$$

This satisfies the properties that:

$$
\begin{aligned}
\mathbf{R}(\theta_1)\mathbf{R}(\theta_2) &= \mathbf{R}(\theta_1+\theta_2) \\
\mathbf{R}(\alpha\theta_1+\beta\theta_2) &= \alpha\mathbf{R}(\theta_1)+\beta\mathbf{R}(\theta_2)
\end{aligned}
$$

Thus, we can find composite rotations by multiplying two rotational matrices together.

To find a the rotational matrices in three dimensions, one *can* go through the same process with a general 3×3 matrix, but this is quite messy. Instead, we can decompose the rotation into a rotation around the three axes. This means that all we need to do is fix a single axis, and the apply the rotation to a particular plane. This allows us to obtain the three rotational matrices of:

$$
\begin{aligned}
\mathbf{R}_x(\theta) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \\
\mathbf{R}_y(\theta) &= \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \\
\mathbf{R}_z(\theta) &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}
$$

3.11.5 The Exponential Matrix

The *exponential matrix* is an entity defined by

$$\exp(\mathbf{A}) = \sum_{n=0}^{\infty}\frac{\mathbf{A}^n}{n!} \tag{3.54}$$

For certain matrices, this can represent a rotation. For example, consider

$$\exp(\mathbf{M}\phi) \;\text{ for }\; \mathbf{M} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Let us look at the positive powers of $\mathbf{M}$.

$$\begin{aligned}
\mathbf{M}^2 &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \\
&= -\mathbf{I} \\
\exp(\mathbf{M}\phi) &= \mathbf{I} + \mathbf{M}\phi - \frac{1}{2!}\mathbf{I}\phi^2 - \frac{1}{3!}\mathbf{M}\phi^3 + \frac{1}{4!}\mathbf{I}\phi^4 + \dots \\
&= \mathbf{M}\left(\phi - \frac{\phi^3}{3!} + \frac{\phi^5}{5!} + \dots\right) + \mathbf{I}\left(1 - \frac{\phi^2}{2!} + \frac{\phi^4}{4!} + \dots\right) \\
&= \mathbf{M}\sin\phi + \mathbf{I}\cos\phi \\
&= \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}
\end{aligned}$$

It thus represents a rotation for $\exp(\mathbf{M}\phi)$.

For two matrices $\mathbf{A}$ and $\mathbf{B}$, $[\mathbf{A}, \mathbf{B}] = 0$ implies that they can be diagonalised by the same matrix. Let $\widehat{\mathbf{A}} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}$ and $\widehat{\mathbf{B}} = \mathbf{T}\mathbf{B}\mathbf{T}^{-1}$.

$$\begin{aligned}
\exp(\widehat{\mathbf{A}} + \widehat{\mathbf{B}}) &= \sum_{n=0}^{\infty} \frac{\widehat{A_{ii}} + \widehat{B_{ii}}^n + \widehat{B_{ii}}^n}{n!} \\
&= \sum_{n=0}^{\infty} \frac{\widehat{A_{ii}}^n}{n!} + \frac{\widehat{B_{ii}}^n}{n!} \\
&= \exp\widehat{\mathbf{A}}\exp\widehat{\mathbf{B}} \\
\exp(\mathbf{T}\mathbf{A}\mathbf{T}^{-1}) &= 1 + \mathbf{T}\mathbf{A}\mathbf{T}^{-1} + \frac{\mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{T}\mathbf{A}\mathbf{T}^{-1}}{2!} + \dots \\
&= \mathbf{T}\left(1 + \mathbf{A} + \frac{\mathbf{A}}{2!} + \dots\right)\mathbf{T}^{-1} \\
&= \mathbf{T}(\exp\mathbf{A})\mathbf{T}^{-1} \\
\exp(\widehat{\mathbf{A}} + \widehat{\mathbf{B}}) &= \mathbf{T}(\exp\mathbf{A})\mathbf{T}^{-1} \cdot \mathbf{T}(\exp\mathbf{B})\mathbf{T}^{-1} \\
\mathbf{T}(\exp(\mathbf{A} + \mathbf{B}))\mathbf{T}^{-1} &= \mathbf{T}(\exp\mathbf{A}\exp\mathbf{B})\mathbf{T}^{-1} \\
\to \exp(\mathbf{A} + \mathbf{B}) &= \exp\mathbf{A}\exp\mathbf{B}
\end{aligned}$$